

**Q. 1**

During the process of creating a notebook, you need to specify the pool that needs to be attached to the notebook that is, a SQL or Spark pool.

Correct or Incorrect: Notebook cells are individual blocks of code or text that runs as a group. If you want to skip cells within the group, a simple skip notation in the cell is all that is required.

 Incorrect Correct[Report Error](#)**TEST MO**

1	[ ]
8	[ ]
15	[ ]
22	[ ]
29	[ ]
36	[ ]
43	[ ]
50	[ ]
57	[ ]

[◀ PREV](#)[NEXT ▶](#)

Type here to search



**Q. 2**

Scenario: Dr. Karl Malus works for the Power Broker Corporation (PBC) founded by Curtiss Jackson, using technology to service various countries and their military efforts. You have been contracted by the company to assist Dr. Malus with their Microsoft Azure Databricks projects.

The team plans to create an Azure Databricks workspace that has a tiered structure. The workspace will contain the following three workloads:

- A workload for data engineers who will use Python and SQL.
- A workload for jobs that will run notebooks that use Python, Scala, and SQL.
- A workload that data scientists will use to perform ad hoc analysis in Scala and R.

The enterprise architecture team at PBC identifies the following standards for Databricks environments:

- The data engineers must share a cluster.
- The job cluster will be managed by using a request process whereby data scientists and data engineers provide packaged notebooks for deployment to the cluster.
- All the data scientists must be assigned their own cluster that terminates automatically after 120 minutes of inactivity. Currently, there are three data scientists.

Required:

The team needs to create the Databricks clusters for the workloads.

Solution: The team creates a Standard cluster for each data scientist, a High Concurrency cluster for the data engineers, and a High Concurrency cluster for the jobs.

Does this meet the requirement?

Correct

Incorrect

Report Error

< PREV

NEXT >

Type here to search



TEST MO
1
8
15
22
29
36
43
50
57

**Q. 3**

Azure offers several types of storage for data, the one chosen should depend on the needs of the users. Each data store has a different price structure.

When you want to store data but don't need to query it, which would be the most cost efficient choice?

- Azure Storage
- Azure Databricks
- Azure Data Lake Storage
- Azure Data Factory
- Azure Data Catalogue
- Azure Stream Analytics

[Report Error](#)

TEST MODE	
1	[ ]
8	[ ]
15	[ ]
22	[ ]
29	[ ]
36	[ ]
43	[ ]
50	[ ]
57	[ ]

[◀ PREV](#)[NEXT ▶](#)

Type here to search



**Q. 4**

When queries are submitted, a dedicated SQL pool query optimizer tries to determine which access paths to the data will result in the least amount of effort to retrieve the data required to resolve the query. It is a cost-based optimizer, and compares the cost of various query plans, and then chooses the plan with the lowest cost.

Correct or Incorrect: It is important to update the statistics after you load data or update large ranges of data, so that queries can benefit from the updated statistics information.

 Correct Incorrect[Report Error](#)

TEST MO	
1	1
8	1
15	1
22	2
29	2
36	2
43	2
50	2
57	2

[◀ PREV](#)[NEXT ▶](#)

Type here to search



**Q. 5**

Azure offers a service to detect anomalies in account activities. These anomalies generate security alerts which are integrated with Azure Security Centre, and are also sent via email to subscription administrators, with details of suspicious activity and recommendations on how to investigate and remediate threats.

Which of the below is the name of this service?

- Azure Shield for Storage
- Azure Defender for Storage
- Azure Storage Account Security Feature
- Azure Armour for Storage
- Encryption in transit

[Report Error](#)

TEST MODE	
1	[ ]
8	[ ]
15	[ ]
22	[ ]
29	[ ]
36	[ ]
43	[ ]
50	[ ]
57	[ ]

[◀ PREV](#)[NEXT ▶](#)

Type here to search



**Q. 6**

The first step in deploying Azure Synapse Analytics is to deploy an Azure Synapse Analytics workspace. A shared Hive-compatible metadata system allows tables defined on files in the data lake to be seamlessly consumed by either Spark or Hive.

SQL and Spark can directly explore and analyze which types of files stored in the data lake? (Select all that apply)

- TSV
- CSV
- XLSX
- Parquet
- JSON
- TXT

[Report Error](#)

TEST MO
1
8
15
22
29
36
43
50
57

[◀ PREV](#)[NEXT ▶](#)

Type here to search



**Q. 7**

A DataFrame creates a data structure and it's one of the core data structures in Spark. In Spark, it is seen as a distributed collection of data that is organized into columns that have names.

Synapse Spark can be used to read and transform objects into a flat structure through data frames. Synapse SQL serverless can be used to query such objects directly and return those results as a regular table. With Synapse Spark, you can transform nested structures into columns and array elements into multiple rows.

The steps show the techniques involved to deal with complex data types have been shuffled.

- a. Flatten nested schema Use the function to flatten the nested schema of the data frame (df) into a new data frame.
- b. Define a function for flattening We define a function to flatten the nested schema.
- c. Flatten child nested Schema Use the function you create to flatten the nested schema of the data frame into a new data frame.
- d. Explode Arrays Transform the array in the data frame into a new dataframe where you also define the column that you want to select.

Which is the correct technique sequence to deal with complex data types?

- b,a,c,d
- c,b,d,a
- b,a,d,c
- a,c,b,d

[Report Error](#)

TEST MO	
1	[ ]
8	[ ]
15	[ ]
22	[ ]
29	[ ]
36	[ ]
43	[ ]
50	[ ]
57	[ ]

[◀ PREV](#)

[NEXT ▶](#)

Type here to search



**Q. 8**

Scenario: A customer of Ultron Electronics is attempting to use a \$300 store credit for the full amount of a new purchase. They are trying to double-spend their credit by creating two transactions at the exact same time using the entire store credit. The customer is making two transactions using two different devices.

The database behind the scenes is an ACID-compliant transactional database.

What would be the result?

- One order would be processed and use the in-store credit, and the other order would update the remaining inventory for the items in the basket, but would not complete the order.
- Both orders would be processed and use the in-store credit.
- One order would be processed and use the in-store credit, and the other order would not be processed.
- None of the listed options.

[Report Error](#)

TEST MO
1
8
15
22
29
36
43
50
57

[PREV](#)

[NEXT >](#)

Type here to search



**Q. 9**

Azure Synapse Analytics is a high performing Massively Parallel Processing (MPP) engine that is built with loading and querying large datasets in mind. There are times though when performance expectations are not met, and it is necessary then to know what aspects of the table structures and architecture can be reviewed and adapted to maximize query performance.

What is the following code intended to accomplish?

```
1.PowerShell  
2.Set-AzSqlDatabase -ResourceGroupName "resourcegroupname" -DatabaseName "mySampleDataWarehouse" -ServerName "sqlpoolservername" -  
RequestedServiceObjectiveName "DW300c"
```

- None of the listed options.
- Address the issue of poor load performance.
- Address the issue of low concurrency.
- Address the issue of poor query performance.
- Address the issue of poor response time.

[Report Error](#)

TEST MO	
1	[ ]
8	[ ]
15	[ ]
22	[ ]
29	[ ]
36	[ ]
43	[ ]
50	[ ]
57	[ ]

[◀ PREV](#)[NEXT ▶](#)

Type here to search



**Q. 10**

Identify the missing word(s) in the following sentence within the context of Microsoft Azure.

A(n) [?] schema may be defined at query time.

- Hybrid data type
- Structured data type
- Unstructured data type
- Azure Cosmos DB data type

[Report Error](#)

TEST MO	
1	[ ]
8	[ ]
15	[ ]
22	[ ]
29	[ ]
36	[ ]
43	[ ]
50	[ ]
57	[ ]

[PREV](#)

[NEXT >](#)

Type here to search



**Q. 11**

Azure Storage accounts are the base storage type within Azure. Azure Storage offers a very scalable object store for data objects and file system services in the cloud. It can also provide a messaging store for reliable messaging, or it can act as a NoSQL store.

Which of the following are Azure Storage configuration options? (Select all that apply)

- Azure Blob
- Azure Data Lake
- Azure Table
- Azure Queue
- Azure Database Server
- Azure Files

[Report Error](#)

TEST MO	
1	1
8	1
15	1
22	2
29	2
36	2
43	2
50	2
57	2

[◀ PREV](#)[NEXT ▶](#)

Type here to search



**Q. 12**

When talking about the Azure Databricks workspace, we refer to two different things.

- The first reference is the logical Azure Databricks environment in which clusters are created, data is stored (via DBFS), and in which the server resources are housed.
  - The second reference is the more common one used within the context of Azure Databricks.
- The first step to using Azure Databricks is to create and deploy a Databricks workspace, which is the logical environment. You can do this in the Azure portal.

There are a number of required values to create your Azure Databricks workspace.

Which are they? (Select five)

 Workspace Name Pricing Tier Resource Group Subscription Location Cluster Mode[Report Error](#)**TEST MO**

1	[ ]
8	[ ]
15	[ ]
22	[ ]
29	[ ]
36	[ ]
43	[ ]
50	[ ]
57	[ ]

[◀ PREV](#)[NEXT ▶](#)

Type here to search



**Q. 13**

Scenario: O'Shaughnessy's is a fast food restaurant. The chain has stores nationwide and is rivalled by Big Belly Burgers. You have been hired by the company to advise on moving data from an Azure Data Lake Storage Gen2 container to a database in an Azure Synapse Analytics dedicated SQL pool with an Azure Data Factory pipeline.

Given: Data in the container is stored in the following folder structure:

- /in/{YYYY}/{MM}/{DD}/{HH}/{mm}
- The earliest folder is /in/2021/01/01/00/00
- The latest folder is /in/2021/01/15/01/45.

Required: Configure a pipeline trigger to meet the following

- Existing data must be loaded.
- Data must be loaded every 30 minutes.
- Late-arriving data of up to two minutes must be included in the load for the time at which the data should have arrived.

When configuring the pipeline trigger, which of the following properties should be used?

TEST MO
1
8
15
22
29
36
43
50
57

Prefix: /in/, Event: Blob created

Recurrence: 30 minutes, Start time: 2021-01-01T00:00, Delay 2 minutes

Recurrence: 30 minutes, Start time: 2021-01-01T00:00

Recurrence: 32 minutes, Start time: 2021-01-15T01:45

Report Error

< PREV

NEXT >

Type here to search



**Q. 14**

Knowing now the different concepts of spark it is imperative to understand how it fits in with the different Data services on Azure.  
Which of the following is best described by:

"Organizations that don't have existing spark implementations yet, get the functionality to spin up a spark cluster to meet data engineering needs without the overhead of the other Spark Platforms. Data Engineers, Data scientist, Data Platform Experts, and Data Analyst can come together within Synapse Analytics where the Spark cluster is spun up quickly to meet the needs.

It provides scale in an efficient way for Spark Clusters and integrates with the one stop shop Data warehousing platform of Synapse."

- HDI
- Azure Databricks
- Spark Pools in Azure Synapse Analytics
- Apache Spark

[Report Error](#)

TEST MO	
1	1
8	1
15	1
22	2
29	2
36	2
43	2
50	2
57	2

[◀ PREV](#)[NEXT ▶](#)

**Q. 15**

Scenario: A teammate is working on solution for transferring data between a dedicated SQL Pool and a serverless Apache Spark Pool using the Azure Synapse Apache Spark Pool to Synapse SQL connector.

When could SQL Auth be used for this connection?

- Never, it is not necessary to use SQL Auth when transferring data between a SQL or Spark Pool.
- When you need a token-based authentication to a dedicated SQL outside of the Synapse Analytics workspace.
- None of the listed options.
- Always, anytime you want to transfer data between the SQL and Spark Pool.

[Report Error](#)

TEST MO
1
8
15
22
29
36
43
50
57

[◀ PREV](#)[NEXT ▶](#)

Type here to search



**Q. 16**

Identify the missing word(s) in the following sentence within the context of Microsoft Azure.

Many business application architectures separate transactional and analytical processing into separate systems with data stored and processed on separate infrastructures. These infrastructures are commonly referred to as OLTP (online transaction processing) systems working with operational data, and OLAP (online analytical processing) systems working with historical data, with each system optimized for their specific task.

Azure Cosmos DB provides ... [?]

- Both a transactional store optimized for transactional workloads and an analytical store optimized for analytical workloads and a fully managed autosync process to keep the data within these stores in sync.
- None of the listed options.
- A transactional store optimized for transactional workloads and a fully managed autosync process to keep the data within these stores in sync.
- An analytical store optimized for analytical workloads and a fully managed autosync process to keep the data within these stores in sync.

[Report Error](#)

TEST MO
1
8
15
22
29
36
43
50
57

[◀ PREV](#)

[NEXT ▶](#)

Type here to search



**Q. 17**

Identify the missing word(s) in the following sentence within the context of Microsoft Azure.

Azure Synapse Analytics can work by acting as the one stop shop to meet all of your analytical needs in an integrated environment if you do not have an analytical environment in place already.

[?] is a single web UI that allows you to:

- Explore your data estate.
- Develop TSQL scripts and notebooks to interact with the analytical engines.
- Build data integration pipelines for managing data movement.
- Monitor the workloads within the service.
- Manage the components of the service.

Azure Portal

Azure Synapse Studio

Azure Monitor

Azure DevOps

Azure Pipelines

Azure Designer

Report Error

**TEST MO**

1	[...]
8	[...]
15	[...]
22	[...]
29	[...]
36	[...]
43	[...]
50	[...]
57	[...]

< PREV

NEXT >

Type here to search



**Q. 19 When creating a new cluster in the Azure Databricks workspace, what happens behind the scenes?**

- Azure Databricks provisions a dedicated VM that processes all jobs, based on your VM type and size selection.
- None of the listed options.
- Azure Databricks creates a cluster of driver and worker nodes, based on your VM type and size selections.
- When an Azure Databricks workspace is deployed, you are allocated a pool of VMs. Creating a cluster draws from this pool.

[Report Error](#)

TEST MO
1
8
15
22
29
36
43
50
57

[PREV](#)[NEXT >](#)

Type here to search



**Q. 20**

Identify the missing word(s) in the following sentence within the context of Microsoft Azure.

When working with large data sets, it can take a long time to run the sort of queries that clients need. These queries can't be performed in real time, and often require algorithms such as MapReduce that operate in parallel across the entire data set. The results are then stored separately from the raw data and used for querying.

One drawback to this approach is that it introduces latency. If processing takes a few hours, a query may return results that are several hours old. Ideally, you would like to get some results in real time (perhaps with some loss of accuracy), and combine these results with the results from the batch analytics.

The Lambda architecture is a big data processing architecture that addresses this problem by combining both batch- and real-time processing methods. It features an append-only immutable data source that serves as system of record. Timestamped events are appended to existing events (nothing is overwritten). Data is implicitly ordered by time of arrival.

The [?] is a vast improvement upon the traditional Lambda architecture. At each stage, we enrich our data through a unified pipeline that allows us to combine batch and streaming workflows through a shared filestore with ACID-compliant transactions.

- Data Sea architecture
- Anaconda architecture
- No-SQL architecture
- Delta Lake architecture
- Serverless architecture
- Data Lake architecture

[Report Error](#)

[PREV](#)

[NEXT >](#)

Type here to search



TEST MO
1
8
15
22
29
36
43
50
57

**Q. 21**

As great as data lakes are at inexpensively storing our raw data, they also bring with them performance challenges:

- Too many small or very big files - more time opening & closing files rather than reading contents (worse with streaming).
- Partitioning also known as "poor man's indexing"- breaks down if you picked the wrong fields or when data has many dimensions, high cardinality columns.
- No caching - cloud storage throughput is low (cloud object storage is 20-50MB/s/core vs 300MB/s/core for local SSDs).

As a solution to the challenges with Data Lakes noted above, Delta Lake is a file format that can help you build a data lake comprised of one or many tables in Delta Lake format. Delta Lake integrates tightly with Apache Spark, and uses an open format that is based on Parquet.

Two of the core features of Delta Lake are performing UPSERTS and Time Travel operations.

What does the Time Travel operation do?

- All of these
- Writing complex temporal queries.
- Providing snapshot isolation for a set of queries for fast changing tables.
- Re-creating analyses, reports, or outputs (for example, the output of a machine learning model). This could be useful for debugging or auditing, especially in regulated industries.
- Because Delta Lake is version controlled, you have the option to query past versions of the data using a single file storage system.

Report Error

TEST MO
1
8
15
22
29
36
43
50
57

< PREV

NEXT >

Type here to search



## Q. 22 How can parameters be passed into an Azure Databricks notebook from Azure Data Factory?

- Use the new API endpoint option on a notebook in Databricks and provide the parameter name.
- Deploy the notebook as a web service in Databricks, defining parameter names and types.
- Render the notebook to an API endpoint in Databricks, defining parameter names and types.
- Use notebook widgets to define parameters that can be passed into the notebook.

[Report Error](#)

TEST MO
1
8
15
22
29
36
43
50
57

[◀ PREV](#)[NEXT ▶](#)

Type here to search



**Q. 23**

Identify the missing word(s) in the following sentence within the context of Microsoft Azure.

Azure Synapse Analytics is an integrated analytics platform, which combines data warehousing, big data analytics, data integration, and visualization into a single environment. Azure Synapse Analytics empowers users of all abilities to gain access and quick insights across all of their data, enabling a whole new level of performance and scale.

Diagnostic analytics deals with answering the question [?].

- "When will the modification made meet my goals?"
- "What is likely to happen in the future based on previous trends and patterns?"
- "What is happening in my business?"
- "Why is it happening?"

[Report Error](#)

TEST MO	
1	1
8	8
15	15
22	22
29	29
36	36
43	43
50	50
57	57

[◀ PREV](#)

[NEXT ▶](#)

Type here to search



Q. 24 Which Azure Synapse Studio hub would you go to create Notebooks?

- Data
- Integrate
- Manage
- None of the listed options
- Create

Report Error

TEST MO
1
8
15
22
29
36
43
50
57

< PREV

NEXT >

Type here to search



**Q. 25**

While Agile, CI/CD, and DevOps are different, they support one another.

Which is best described by:

"Focuses on software-defined life cycles highlighting tools that emphasize automation."

- DevOps
- Agile
- CI/CD
- SDLC

Report Error

TEST MO
1
8
15
22
29
36
43
50
57

< PREV

NEXT >

Type here to search



Q. 26 Correct or Incorrect : Azure Synapse functionality requires integration with Azure Data Factory, Azure Databricks and Power BI.

Incorrect

Correct

Report Error

TEST MO	
1	
8	
15	
22	
29	
36	
43	
50	
57	

< PREV

NEXT >

Type here to search



**Q. 27**

Scenario: Jungle.com uses Azure Cosmos DB to store sales orders and customer profile data from their eCommerce site. The NoSQL document store provided by the Azure Cosmos DB provides the familiarity of managing their data using SQL syntax, while being able to read and write the files at a massive, global scale.

While Jungle.com is happy with the capabilities and performance of Azure Cosmos DB, they are concerned about the cost of executing a large volume of complex analytical queries needed to fulfill their operational reporting requirements.

They want to efficiently access all their operational data stored in Cosmos DB without needing to increase the Azure Cosmos DB throughput and associate cost. They have looked at options for extracting data from their containers to the data lake as it changes, through the Azure Cosmos DB change feed mechanism.

The problem with this approach is the extra service and code dependencies and long-term maintenance of the solution. They could perform bulk exports from a Synapse Pipeline, but then they won't have the most up-to-date information at any given moment.

Which would be the best action to take?

- Enable Azure VPN Gateway for Cosmos DB and enable the analytical store on their Azure Cosmos DB containers.
- Enable Azure Synapse Link for Cosmos DB and enable the analytical store on their Azure Cosmos DB containers.
- Enable Azure VNet Peering for Cosmos DB and enable the analytical store on their Azure Cosmos DB containers.
- Enable Azure Dedicated Connect for Cosmos DB and enable the analytical store on their Azure Cosmos DB containers.

[Report Error](#)

TEST MO
1
8
15
22
29
36
43
50
57

[PREV](#)

[NEXT >](#)

Type here to search 



**Q. 28**

Identify the missing word(s) in the following sentence within the context of Microsoft Azure.

As a Data Engineer, you can transfer and move data in several ways. The most common tool is Azure Data Factory which provides robust resources and nearly 100 enterprise connectors. Azure Data Factory also allows you to transform data by using a wide variety of languages.

Azure has opened the way for technologies that can handle unstructured data at an unlimited scale. T

his change has shifted the paradigm for loading and transforming data from [?].

 ETL ? MTD MTD ? RPO ELT ? ETL ETL ? ELT ETL ? RTO RPO ? RTO[Report Error](#)

TEST MO
1
8
15
22
29
36
43
50
57

[◀ PREV](#)[NEXT ▶](#)

**Q. 29**

You can use either the REST API or the Azure client library to programmatically access a storage account.

What is the primary advantage of using the client library?

- Localization
- Convenience
- Availability
- Cost

[Report Error](#)

TEST MO
1
8
15
22
<b>29</b>
36
43
50
57

[PREV](#)[NEXT >](#)

Type here to search



**Q. 30**

Activities within Azure Data Factory define the actions that will be performed on the data and there are three categories including:

- Data movement activities
- Data transformation activities
- Control activities

Pipelines in Data Factory are defined in JSON format as follows:

```
1.JSON
2.{  
3.  "name": "PipelineName",  
4.  "properties":  
5.  {  
6.    "description": "pipeline description",  
7.    "activities":  
8.    {  
9.    },  
10.   "parameters": {  
11.   },  
12.   }  
13. }
```

Which of the JSON properties are required? (Select all that apply)

 name activities description parameters[Report Error](#)[◀ PREV](#)[NEXT ▶](#)**TEST MODE**

1	2
8	9
15	16
22	23
29	<b>30</b>
36	37
43	44
50	51
57	58

**Q. 31**

The following are the facets of Azure Databricks security:

- Data Protection
- IAM/Auth
- Network
- Compliance

Which of the following comprise Data Protection within Azure Databricks security? (Select five)

 Vault Secrets VNet Peering TLS Managed Keys AAD ACLs[Report Error](#)**TEST MO**

1	[ ]
8	[ ]
15	[ ]
22	[ ]
29	[ ]
36	[ ]
43	[ ]
50	[ ]
57	[ ]

[PREV <](#)[NEXT >](#)

Q. 32 Which of the following services allow customers to store semi-structured datasets in Azure.

Azure File Storage

Azure SQL Database

Azure Content Delivery Network (CDN)

Azure Blob Storage

Azure Table Storage

Azure Cosmos DB

Report Error

TEST MO

1	[ ]
8	[ ]
15	[ ]
22	[ ]
29	[ ]
36	[ ]
43	[ ]
50	[ ]
57	[ ]

< PREV

NEXT >

Type here to search



**Q. 33**

Identify the missing word(s) in the following sentence within the context of Microsoft Azure.

[?] is a data flow object that can be added to the canvas designer as an activity in an Azure Data Factory pipeline to perform code free data preparation. It enables individuals who are not conversant with the traditional data preparation technologies such as Spark or SQL Server, and languages such as Python and T-SQL to prepare data at cloud scale iteratively.

- Data Expression Script Builder
- Mapping Data Flow
- Data Expression Orchestrator
- Data Stream Expression Builder
- Power Query
- Data Flow Expression Builder

[Report Error](#)

TEST MO
1
8
15
22
29
36
43
50
57

[◀ PREV](#)[NEXT >](#)

Type here to search

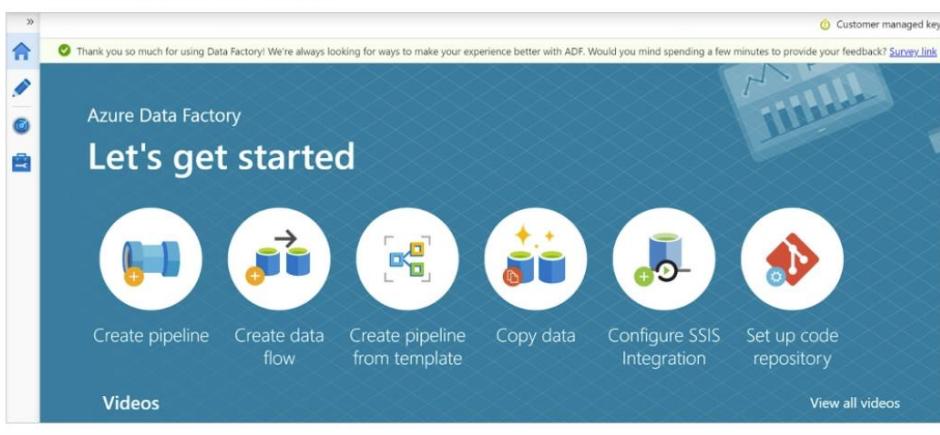


Q. 34

Scenario: Your team is working on a project using the Azure Data Factory authoring tool.

A junior team member comes to you and asks "Where can I find the Copy Data activity?"

Which of the below is the correct location?



Batch Service

◀ PREV

NEXT ▶

Type here to search



^

TEST MODE : pr

1	2	3
8	9	10
15	16	17
22	23	24
29	30	31
36	37	38
43	44	45
50	51	52
57	58	59

## Let's get started



Create pipeline



Create data flow



Create pipeline from template



Copy data



Configure SSIS Integration



Set up code repository

Videos

[View all videos](#)

Batch Service

Data Explorer

Move & Transform

Databricks

Azure Function

[Report Error](#)

[PREV](#)

[NEXT](#)

Type here to search



TEST MODE : pr

1	2	3
8	9	10
15	16	17
22	23	24
29	30	31
36	37	38
43	44	45
50	51	52
57	58	59

**Q. 35**

Identify the missing word(s) in the following sentence within the context of Microsoft Azure.

In Azure Data Factory, a(n) [?] is a logical grouping of activities that together perform a task.

- Pipeline
- Linked Service
- Sink
- Orchestration
- Activity

[Report Error](#)

**TEST MODE**

1	2
8	9
15	16
22	23
29	30
36	37
43	44
50	51
57	58

[◀ PREV](#)

[NEXT ▶](#)

Type here to search



**Q. 36**

Scenario: Honest Eddie's Car Dealership is an establishment in South Carolina USA, which is dedicated to the purchase and sale of cars and light trucks. Currently the IT team is looking into distribution options for a product dimension table.

Which distribution option should Eddie's IT team use for a product dimension table that will contain 1,000 records in Synapse Analytics?

- DISTRIBUTION = REPLICATE
- DISTRIBUTION = HEAP([ProductId])
- DISTRIBUTION = ROUND\_ROBIN
- DISTRIBUTION = HASH([ProductId])

[Report Error](#)

**TEST MODE**

1	2
8	9
15	16
22	23
29	30
36	37
43	44
50	51
57	58

[◀ PREV](#)

[NEXT ▶](#)

Type here to search



**Q. 37**

Authentication is the process of validating credentials as you access resources in a digital infrastructure. This ensures that you can validate that an individual, or a service that wants to access a service in your environment can prove who they are. Azure Synapse Analytics provides several different methods for authentication.

Which are valid authentication methods in Azure Synapse Analytics? (Select all that apply)

- Azure Key Vault
- Managed identity
- SQL Authentication
- Azure Active Directory
- MFA
- All of these

[Report Error](#)

**TEST MODE**

1	2
8	9
15	16
22	23
29	30
36	<b>37</b>
43	44
50	51
57	58

[◀ PREV](#)

[NEXT ▶](#)

Type here to search



^

Q. 38 If you are performing analytics on the data, set up the storage account as an Azure Data Lake Storage Gen2 account by setting the Hierarchical Namespace option to which of the following?

- Ticked/Checked
- Enabled
- Auto-scale
- Disabled
- ON
- OFF

Report Error

TEST MODE

1	2
8	9
15	16
22	23
29	30
36	37
43	44
50	51
57	58

◀ PREV

NEXT ▶

Type here to search



**Q. 39**

Scenario: You are working as a consultant at Advanced Idea Mechanics (A.I.M.) who is a privately funded think tank organized of a group of brilliant scientists whose sole dedication is to acquire and develop power through technological means. Their goal is to use this power to overthrow the governments of the world. They supply arms and technology to radicals and subversive organizations in order to foster a violent technological revolution of society while making a profit.

The company has 10,000 employees. Most employees are located in Europe. The company supports teams worldwide.

AIM has two main locations: a main office in London, England, and a manufacturing plant in Berlin, Germany.

During events, 100 engineers set up a remote portable office by using a VPN to connect the datacentre in the London office. The portable office is set up and torn down in approximately 20 different countries each year.

AIM runs Microsoft SQL Server in an on-premises virtual machine (VM).

Required:

- Migration of the database to Azure SQL Database
- Synchronize users from Active Directory to Azure Active Directory (Azure AD)
- Configure Azure SQL Database to use an Azure AD user as administrator

Which of the following should be configured?

- For each Azure SQL Database server, set the Access Control to administrator.
- For each Azure SQL Database, set the Access Control to administrator.
- For each Azure SQL Database server, set the Active Directory to administrator.
- For each Azure SQL Database, set the Active Directory administrator role.

[Report Error](#)

[◀ PREV](#)

[NEXT ▶](#)

Type here to search



TEST MO	
1	1
8	1
15	1
22	1
29	1
36	1
43	1
50	1
57	1

**Q. 40**

Scenario: The organization you work at has data which is specific to a country or region due to regulatory control requirements.

When considering Azure Storage Accounts, which option meets the data diversity requirement?

- Enable virtual networks for the proprietary data and not for the public data. This will require separate storage accounts for the proprietary and public data.
- Locate the organization's data in a data centre with the strictest data regulations to ensure that regulatory requirement thresholds have been met. In this way, only one storage account will be required for managing all data, which will reduce data storage costs.
- None of the listed options.
- Locate the organization's data in a data centre in the required country or region with one storage account for each location.

**TEST MODE**

1	2
8	9
15	16
22	23
29	30
36	37
43	44
50	51
57	58

[Report Error](#)

[PREV](#)

[NEXT](#)

Type here to search



**Q. 41**

Azure provides many ways to store your data and there are several tools that create a storage account.

Which aspects guide a user's decision on the tool used to create a storage account? (Select two)

- Location restrictions of the data centre
- If the user needs automation
- The datatype being stored in the account
- Tool cost
- If the user wants a GUI

[Report Error](#)

**TEST MODE**

1	2
8	9
15	16
22	23
29	30
36	37
43	44
50	51
57	58

[◀ PREV](#)

[NEXT ▶](#)

Type here to search



**Q. 42**

Azure Synapse SQL pools support placing complex data processing logic into Stored procedures

Correct or Incorrect: Multiple users and client programs can perform operations on underlying database objects through a procedure, even if the users and programs do not have direct permissions on those underlying objects.

Correct

Incorrect

[Report Error](#)

**TEST MODE**

1	2
8	9
15	16
22	23
29	30
36	37
43	44
50	51
57	58

[◀ PREV](#)

[NEXT ▶](#)

Type here to search



^

Q. 43 What type of process are the driver and the executors?

- C++ processes
- Python processes
- JavaScript
- Java processes

Report Error

TEST MODE

1	2
8	9
15	16
22	23
29	30
36	37
43	44
50	51
57	58

◀ PREV

NEXT ▶

Type here to search



**Q. 44 How do you disable Azure Synapse Link for Azure Cosmos DB?**

- Delete the Azure Cosmos DB container
- Set the Azure Synapse Link option to disable on the Azure Cosmos DB instance.
- Set the Azure Synapse Link option to disable on the Azure Cosmos DB container.
- Delete the Azure Cosmos DB account

[Report Error](#)

**TEST MODE**

1	2
8	9
15	16
22	23
29	30
36	37
43	<b>44</b>
50	51
57	58

[◀ PREV](#)

[NEXT ▶](#)

Type here to search



rep  
raining

Q. 45 Correct or Incorrect: The Apache Spark history server can be used to debug and diagnose completed only.

Correct

Incorrect

Report Error

TEST MODE

1	2
8	9
15	16
22	23
29	30
36	37
43	44
50	51
57	58

◀ PREV

NEXT ▶

Type here to search



Q. 46 Which is the correct syntax for overwriting data in Azure Synapse Analytics from a Databricks notebook?

- df.write.format("com.databricks.spark.sqldw").update() option("...").option("...") save()
- df.write.format("com.databricks.spark.sqldw").mode("overwrite").option("...").option("...") save()
- df.write.mode("overwrite") option("...") option("...") save()
- df.write.format("com.databricks.spark.sqldw").overwrite() option("...").option("...") save()

Report Error

TEST MODE

1	2
8	9
15	16
22	23
29	30
36	37
43	44
50	51
57	58

◀ PREV

NEXT ▶

Type here to search



^

**Q. 47**

Identify the missing word(s) in the following sentence within the context of Microsoft Azure.

[?] provides one-click setup, streamlined workflows, an interactive workspace for Spark-based applications plus it adds capabilities to Apache Spark, including fully managed Spark clusters and an interactive workspace.

- Azure SQL Data warehouse
- Azure Storage Explorer
- Azure Cosmos DB
- Azure Data Catalogue
- Azure Data Factory
- Azure Databricks

[Report Error](#)

**TEST MODE**

1	2
8	9
15	16
22	23
29	30
36	37
43	44
50	51
57	58

[◀ PREV](#)

[NEXT ▶](#)

Type here to search



**Q. 48**

When working with Azure Data Factory, a dataset is a named view of data that simply points or references the data you want to use in your activities as inputs and outputs.

A dataset in Data Factory can be defined in a JSON format for programmatic creation as follows:

Which of the JSON properties are required? (Select all that apply)

```
1. JSON
2. {
3.   "name": "<name of dataset>",
4.   "properties": {
5.     "type": "<type of dataset: AzureBlob, AzureSql etc...>",
6.     "linkedServiceName": {
7.       "referenceName": "<name of linked service>",
8.       "type": "LinkedServiceReference",
9.     },
10.    "schema": [
11.      {
12.        "name": "<Name of the column>",
13.        "type": "<Name of the type>"
14.      }
15.    ],
16.    "typeProperties": {
17.      "<type specific property>": "<value>",
18.      "<type specific property 2>": "<value 2>",
19.    }
20.  }
21. }
```

 type typeProperties structure name**TEST MODE :**

1	2	[1]
8	9	[1]
15	16	[1]
22	23	[2]
29	30	[2]
36	37	[2]
43	44	[2]
50	51	[2]
57	58	[2]

[◀ PREV](#)[NEXT ▶](#)

Type here to search



**Q. 49**

Scenario: You have been contracted by Wayne Enterprises, a company owned by Bruce Wayne with market value of over twenty seven million dollars. Bruce founded Wayne Enterprises shortly after he created the Wayne Foundation and he became the president and chairman of the company.

Bruce has come to you because his IT team needs advice on the configuration and synchronization of data between an on-premises Microsoft SQL Server database to Azure SQL Database.

Recently, ad-hoc and reporting queries are being overutilized on the on-premises production instance and your expert advise is required on the following points.

Requirements:

- Execute an initial data synchronization to Azure SQL Database (minimize downtime)
- Execute bi-directional data synchronization after initial synchronization

A synchronization solution must be created and implemented and Bruce and the team look to you as the Azure expert.

Which synchronization method should you advise the team to use?

Transactional replication

SQL Server Agent job

Azure SQL Data Sync

Data Migration Assistant

Backup and restore

**TEST MODE :**

1	2	[1]
8	9	[1]
15	16	[1]
22	23	[2]
29	30	[2]
36	37	[2]
43	44	[2]
50	51	[2]
57	58	[2]

< PREV

NEXT >

Type here to search



Q. 50 When doing a write stream command, what does the outputMode("append") option do?

- The append outputMode allows records to be added to the output sink.
- The append mode replaces existing records and updates aggregates.
- The append mode allows records to be updated and changed in place.
- The append outputMode allows records to update to the output log.

Report Error

TEST MODE

1	2
8	9
15	16
22	23
29	30
36	37
43	44
50	51
57	58

◀ PREV

NEXT ▶

Type here to search



^

**Q. 51**

Synapse Studio comes with an integrated notebook experience. The notebooks in Synapse studio, are a web interface that enables you to create, edit, or transform data in the files. It is based on a live code experience, including visualizations and narrative text.

Correct or Incorrect: You can access data in the primary storage account directly. There's no need to provide the secret keys.

Incorrect

Correct

[Report Error](#)

**TEST MODE**

1	2
8	9
15	16
22	23
29	30
36	37
43	44
50	51
57	58

[◀ PREV](#)

[NEXT ▶](#)

Type here to search



^

**Q. 52**

Connectors are Azure Data Factory objects that enable your Linked Services and Datasets to connect to a wide variety of data sources and sinks. These can include connections to Azure resources and third-party connectors such as Amazon S3 or Google cloud. There are nearly 100 connectors that are available.

Which are among the file formats supported? (Select all that apply)

XLSX format

CSV format

All of these

Binary format

Parquet format

JSON format

**TEST MODE**

1 2

8 9

15 16

22 23

29 30

36 37

43 44

50 51

57 58

Report Error

◀ PREV

NEXT ▶

Type here to search



^

**Q. 53**

Unlike fully logged operations, which use the transaction log to keep track of every row change, minimally logged operations keep track of extent allocations and meta-data changes only. Therefore, minimal logging involves logging only the required information to roll back the transaction after a failure or for an explicit request.

Which is used to roll back the transaction after a failure or for an explicit request?

- ROLLBACK POINT
- ROLLBACK TRAN
- ROLLBACK TO MIN
- ROLLBACK TO SAVEPOINT

[Report Error](#)

**TEST MODE**

1	2
8	9
15	16
22	23
29	30
36	37
43	44
50	51
57	58

[◀ PREV](#)

[NEXT ▶](#)

Type here to search



**Q. 54**

Scenario: Queen Consolidated was overtaken by Raymond Carson Palmer and re-branded as Palmer Technologies. Now that Ray is overseeing the operations at Palmer, Ray has decided to implement better applications.

You are working as a consultant with Palmer, and in a meeting with Ray and his IT team discussing Azure Data Factory. The team plans to use Azure Data Factory to prepare data to be queried by Azure Synapse Analytics serverless SQL pools.

Files will be initially ingested into an Azure Data Lake Storage Gen2 account as 10 small JSON files. Each file will contain the same data attributes and data from a subsidiary of Palmer. The team needs to move the files to a different folder and transform the data.

Required:

- Provide the fastest possible query times.
- Automatically infer the schema from the underlying files.

As the Azure expert, the team looks to you for advice on how they should configure the Data Factory copy activity with respect to the copy behaviour.

Which of the following should you advise them to use?

Flatten hierarchy

Append Files

Merge Files

Preserve hierarchy

Report Error

◀ PREV

NEXT ▶

Type here to search



**TEST MODE :**

1	2	1
8	9	1
15	16	1
22	23	2
29	30	2
36	37	2
43	44	2
50	51	2
57	58	2

**Q. 55**

Identify the missing word(s) in the following sentence within the context of Microsoft Azure.

Azure Data Factory is the cloud-based [?] and data integration service that allows you to create data-driven workflows for orchestrating data movement and transforming data at scale. You can build complex [?] processes that transform data visually with data flows or by using compute services such as Azure HDInsight Hadoop, Azure Databricks, and Azure Synapse Analytics. (Select all that apply)

ETL

ELT

OLTP

OLAP

CI/CD

[Report Error](#)

**TEST MODE**

1 2

8 9

15 16

22 23

29 30

36 37

43 44

50 51

57 58

[◀ PREV](#)

[NEXT ▶](#)

Type here to search



^

**Q. 56**

Correct or Incorrect: Concurrency and the allocation of resources across connected users are also a factor that can limit the load performance into Azure Synapse Analytics SQL pools.

To optimize the load execution operations, recommendations are to reduce or minimize the number of simultaneous load jobs that are running or assigning higher resource classes that reduce the number of active running tasks.

Incorrect

Correct

[Report Error](#)

**TEST MODE**

1	2
8	9
15	16
22	23
29	30
36	37
43	44
50	51
57	58

[◀ PREV](#)

[NEXT ▶](#)

Type here to search



**Q. 57**

Scenario: You are working on a project using Azure Synapse Studio and want to configure a private endpoint. You open up Azure Synapse Studio, go to the manage hub, and see that the private endpoints is greyed out.

Why is the option not available?

- Azure Synapse Studio does not support the creation of private endpoints.
- A managed virtual network has not been created.
- There are service interruptions which must be troubleshoot first.
- A conditional access policy has to be defined first.

[Report Error](#)

**TEST MODE**

1	2
8	9
15	16
22	23
29	30
36	37
43	44
50	51
57	58

[◀ PREV](#)

[NEXT ▶](#)

Type here to search



**Q. 58**

Scenario: Your team has deployed a factory to production and realizes there's a bug that needs to be fixed right away, but you can't deploy the current collaboration branch.

What is the best action to take?

- Create a rollback to a savepoint
- Deploy a hotfix
- Utilize a workhole
- Deploy a timeshift
- None of the listed options

[Report Error](#)

**TEST MODE**

1	2
8	9
15	16
22	23
29	30
36	37
43	44
50	51
57	<b>58</b>

[◀ PREV](#)

[NEXT ▶](#)

Type here to search



Q. 59 What does Azure Data Lake Storage (ADLS) Passthrough enable?

- Automatically mounting ADLS accounts to the workspace that are added to the managed resource group.
- Commands running on a configured cluster can read and write data in ADLS without configuring service principal credentials.
- Blocking ADLS resources through a mount point when credential passthrough is enabled.
- User security groups that are added to ADLS are automatically created in the workspace as Databricks groups.

Report Error

TEST MODE

1	2
8	9
15	16
22	23
29	30
36	37
43	44
50	51
57	58

◀ PREV

NEXT ▶

Type here to search



**Q. 60**

One of the key management features that you have at your disposal within Azure Synapse Analytics, is the ability to scale the compute resources for SQL or Spark pools to meet the demands of processing your data. Compute is separate from storage, which enables you to scale compute independently of the data in your system. This means you can scale up and scale down the compute power to meet your needs.

Apache Spark pools for Azure Synapse Analytics uses an Autoscale feature that automatically scales the number of nodes in a cluster instance up and down.

Autoscale continuously monitors the Spark instance and collects which of the following metrics? (Select five)

- Total Pending CPU
- Total seeds on each individual Node
- Total Free CPU
- Total Free Memory
- Used Memory per Node
- Total Pending Memory

[Report Error](#)

**TEST MODE**

1	2
8	9
15	16
22	23
29	30
36	37
43	44
50	51
57	58

[PREV](#)

[SUBMIT TEST >](#)

Type here to search



^

**Q. 18**

Although you have the opportunity to ingest data at the source directly into a data warehouse, it is more typical to store the source data within a staging area, which is also referred to as a landing zone. This typically is a neutral storage area that sits between the source systems and the data warehouse.

Which are main reasons for adding a staging area into the architecture of a modern data warehouse? (Select all that apply)

- To reduce contention on source systems.
- To rerun failed data warehouse loads from a staging area
- To join data together from different source systems
- Enables you to deal with the ingestion of source systems on different schedules.
- Data storage cost savings
- Higher availability of data

[Report Error](#)**TEST MO**

1	[ ]
8	[ ]
15	[ ]
22	[ ]
29	[ ]
36	[ ]
43	[ ]
50	[ ]
57	[ ]

[◀ PREV](#)[NEXT ▶](#)

Type here to search

