

Knowledge Check

Total points 6

1. Question 1

High-availability systems must be available most of the time. Service-level agreements (SLAs) specify your organization's availability expectations. System uptime can be expressed as three nines, four nines, or five nines. These expressions indicate system uptimes of 99.9 percent, 99.99 percent, or 99.999 percent. How many downtime hours per year will be covered with an SLA of 99.9% uptime?

1 / 1 point

- ☐ .88 Hours
- ☐ 24 Hours
- ☒ 8.76 Hours
- ☐ .09 Hours

Correct

An SLA of 99.9% uptime covers downtime of 8.76 hours per year.

2. Question 2

Which of the following terms describes the final cost of owning a given technology and includes the following costs:

Hardware

Software licensing

Labor (installation, upgrades, maintenance)

Datacenter overhead (power, telecommunications, building, heating and cooling)

1 / 1 point

- ☐ RPO
- ☐ MTD
- ☒ TCO
- ☐ RTO

Correct

The total cost of ownership (TCO) is the purchase price of an asset plus the costs of operation. Assessing the total cost of ownership represents taking a bigger picture look at what the product is and what its value is over time.

3. Question 3

Azure supports both structured and unstructured data. Unstructured data in Azure would commonly be held in which of the following?

1 / 1 point

- ☐ Azure SQL Database
- ☒ Azure Cosmos DB

☐ Azure SQL Data Warehouse

Correct

For unstructured (NoSQL) databases, each data element can have its own schema at query time. Data can be stored as a file in Azure Blob storage or as NoSQL data in Azure Cosmos DB or Azure HDInsight.

4. Question 4

The term *lift and shift* is used when migrating physical or virtualized servers from an on-premises environment to a Microsoft Azure virtual machine cloud-based environment without the need to rearchitect the application.

Which of the following are benefits of carrying out a Lift and Shift operation?

Select all options that apply.

1 / 1 point

☒ Higher Availability

Correct

The lift-and-shift strategy provides immediate benefits, including higher availability, lower operational costs, and the ability to transfer workloads from one datacenter to another

☒ Lower operational costs

Correct

The lift-and-shift strategy provides immediate benefits, including higher availability, lower operational costs, and the ability to transfer workloads from one datacenter to another.

☐ Take advantage of all Azure Features.

5. Question 5

The schema of what data type can be defined at query time.

1 / 1 point

☐ Structured data

☐ Azure Cosmos DB

☒ Unstructured data

Correct

The schema of unstructured data is typically defined at query time. This means that data can be loaded onto a data platform in its native format.

6. Question 6

Which cloud technical requirement is met by duplicating customer content for redundancy and meeting service-level agreements (SLAs) in Azure?

1 / 1 point

☐ Maintainability

☒ High availability

☐ Multilingual support

Correct

High availability duplicates customer content for redundancy and meets SLAs in Azure.

Knowledge check

Total points 6

1. Question 1

As data processing techniques change with technology, new roles are starting to appear. These roles provide specialized skills to help streamline the data engineering process. Which of the following roles have been identified as new roles in modern data Projects?

1 / 1 point



Data Engineer

Correct

Data Engineer has been identified as a new role with specialized skills to help streamline the data engineering process.



Database Administrator



Artificial Intelligence Engineer

Correct

Artificial Intelligence Engineer has been identified as a new role with specialized skills to help streamline the data engineering process



Data Scientist

Correct

Data Scientist has been identified as a new role with specialized skills to help streamline the data engineering process.

2. Question 2

The role of a Data Engineer includes which of the following tasks?
Select all options that apply.

1 / 1 point



Manage, monitor, and ensure the security and privacy of data to satisfy business needs.

Correct

The Azure data engineer focuses on data-related tasks in Azure. Including Managing, monitoring, and ensure the security and privacy of data to satisfy business needs.



Using services and tools to ingest, egress, and transform data from multiple sources.

Correct

The Azure data engineer focuses on data-related tasks in Azure. Primary responsibilities include using services and tools to ingest, egress, and transform data from multiple sources.



Perform advanced analytics to extract value from data



Working with services such as Cognitive Services, Cognitive Search, and Bot Framework.

3. Question 3

Data Engineers may sometimes perform ETL process when processing data. The extract may come from many sources including databases, files, and streams. As part of the Extract process which of the following information must be supplied?

0.333333333333333 / 1 point



Define the Transformation

This should not be selected

Data Engineers define any data transformation operations after extracting the data, which can include splitting, combining, deriving, adding, removing, or pivoting columns and is part of the Transformation process.



Define the Data Source



Define the Data

Correct

During the extraction process, Data Engineers must define the Data to be processed.

4. Question 4

In Microsoft Azure, Data Engineers will use several tools to perform ETL processes.

Which of the following tools will commonly be used to perform ETL processes in Azure?

1 / 1 point



Azure Cosmos DB



Azure Data Factory



Azure Synapse Analytics

Correct

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

5. Question 5

Which role works with Azure Cognitive Services, Cognitive Search, and the Bot Framework?

1 / 1 point



AI engineer



Data scientist



Data engineer

Correct

Artificial intelligence (AI) engineers work with AI services such as Cognitive Services, Cognitive Search, and the Bot Framework.

6. Question 6

Which Azure data platform is commonly used to process data in an ELT framework?

1 / 1 point



Azure Data Lake Storage



Azure Databricks



Azure Data Factory

Correct

Azure Data Factory is a cloud-integration service that orchestrates the movement of data between various data stores.

Knowledge check

Total points 6

1. Question 1

Azure provides many data platform technologies to meet the needs of common data varieties. In relational database systems like Microsoft SQL Server which of the following types of data structures is used?

1 / 1 point

- ☐ Semi-structured Data
- ☐ Unstructured Data
- ☒ Structured Data

Correct

In relational database systems like Microsoft SQL Server, Azure SQL Database, and Azure SQL Data Warehouse, Structured Data is used as the data structure.

2. Question 2

Nonstructured data is stored in nonrelational systems, commonly called unstructured or NoSQL systems. Which of the following are types of NoSQL databases? Select all options that apply.

1 / 1 point

- ☒ Key-value store

Correct

Key Value store stores key-value pairs of data in a table structure. They are a type of NoSQL database.

- ☒ Column Store Database

Correct

Column Store Databases store data based on columns rather than rows. They are a type of NoSQL database. Columns can be defined at the query's runtime, allowing flexibility in the data that's returned performantly.

- ☐ Relational database

- ☒ Graph database

Correct

Graph database finds relationships between data points by using a structure that's composed of vertices and edges. They are a type of NoSQL database.

- ☒ Document databases.

Correct

Document databases store documents that are tagged with metadata to aid document searches. They are a type of NoSQL database.

3. Question 3

You need to provision a data store that will store but not query data. You also want to avail of the cheapest storage option. Which of the following is most suitable in this scenario?

1 / 1 point

- ☐ Azure Queue:

☐ Azure Files

☒ Azure Blob

Correct

Azure Blob is a scalable object store for text and binary data and is the cheapest option for this type of storage.

4. Question 4

Azure Data Lake Storage Gen 2 is designed to store massive amounts of data for big-data analytics. Which of the following are features of Azure Data Lake?

Select all options that apply.

1 / 1 point

☒ Hadoop compatibility

Correct

One of the features of Data Lake Storage includes Hadoop compatibility.

☒ Security support for access control lists (ACLs)

Correct

One of the features of Data Lake Storage includes support for access control lists.

☒ Unlimited scalability

Correct

One of the features of Data Lake Storage includes unlimited scalability.

☐ Limited support for Zone Redundant storage

5. Question 5

Which data platform technology is a globally distributed, multi model database that can perform queries in less than a second?

1 / 1 point

☐ Azure SQL Data Warehouse

☐ Azure SQL Database

☒ Azure Cosmos DB

Correct

Azure Cosmos DB is a globally distributed, multi model database that can offer sub second query performance.

6. Question 6

Which Azure service is the best choice to store documentation about a data source?

1 / 1 point

☐ Azure Data Factory

☐ Azure Data Lake Storage

☒ Azure Data Catalog

Correct

Azure Data Catalog is a central place where an organization's users can contribute their knowledge. Together, they build a community of data sources that the organization owns.

Knowledge check

Total points 4

1. Question 1

Synapse SQL offers both serverless and dedicated resource models to work with both descriptive and diagnostic analytical scenarios. To provide predictable performance and cost you should create which of the following.

1 / 1 point



Serverless SQL endpoints



Dedicated SQL pools

Correct

For predictable performance and cost, create dedicated SQL pools to reserve processing power for data stored in SQL tables.

2. Question 2

Azure Synapse Pipelines leverages the capabilities of Azure Data Factory and is a cloud-based solution for which of the following types of data ingestion and transformation?

1 / 1 point



ELT



ETL

Correct

Azure Synapse Pipelines leverages the capabilities of Azure Data Factory and is the cloud-based ETL and data integration service that allows you to create data-driven workflows for orchestrating data movement and transforming data at scale.

3. Question 3

Azure Stream Analytics is a perfect solution to implement when your data is which of the following types?

1 / 1 point



Real-time Analysis



Batch process

Correct

If your organizations' challenge is to respond to data event in real-time or perform a post mortem of large batches of data in a continuous time bound stream, then Azure Stream Analytics is for your organization

4. Question 4

Data Engineers configure ingestion components of Azure Stream Analytics by configuring data inputs from sources including Azure Event Hubs, Azure IoT Hub, or Azure Blob storage.

Azure Event Hub provides which of the following features.

Select all that apply!

1 / 1 point



Azure Storage



Partitioned consumer model

Correct

Event Hubs uses a partitioned consumer model to scale out your data stream and is integrated into the big data and analytics services of Azure including Databricks, Stream Analytics, ADLS, and HDInsight.



Big data streaming service

Correct

Event hubs provide a big data streaming service of Azure. It is designed for high throughput data streaming scenarios where customers may send billions of requests per day.

Test Prep

Latest Submission Grade 93.75%

1. Question 1

Moving resources such as servers and services from an on-premises environment to a cloud-based solution will have a benefit on which of the following.

1 / 1 point



Operational expenditure only



Both capital and operational expenditure



Capital expenditure only

Correct

One benefit of Cloud environments requires no capital investment. T- You pay for a service or product as you use it i.e., pay-as-you-go pricing. Moving servers and services to the cloud also reduces operational costs.

2. Question 2

Which of the following data processing frameworks are used by data engineers to ingest data from an on-premises database to an on-premises data warehouse?

1 / 1 point



Online Transactional Processing (OLTP)



Extract, Transform, and Load (ETL)



Extract, Load, and Transform (ELT)

Correct

Extract, Transform and Load (ETL) is a typical process for ingesting data from an on-premises database to an on-premises data warehouse.

3. Question 3

Unstructured data differs from Structured Data in many features.

Which of the following are features of Unstructured data?

Select all options that apply.

1 / 1 point

☒ Schema-on-read.

Correct

Schema on read is a feature of unstructured data

☐ Schema-on-write

☐ Predefined format

☒ Native format

Correct

Native Format is a feature of unstructured data.

☐ Commonly stored in data warehouses

☒ Commonly stored in data lakes

Correct

Storage in Data Lakes is a feature of unstructured data

4. Question 4

Data Engineers may sometimes perform ETL process when processing data. The extract may come from many sources including databases, files, and streams. These data sources can have unique data formats including which of the following?

Choose all that apply.

0.5 / 1 point

☒

Partially structured

This should not be selected

Partially structured is not a term used in Azure. The nearest term would be semi-structured. Try going back and reviewing **The tasks of a data engineer** lesson.

☐ Structured

☒ Semi-structured

Correct

Each data source has unique data formats one of which and can be structured, semi-structured, or unstructured.

☒ Unstructured

Correct

Each data source has unique data formats one of which and can be structured, semi-structured, or unstructured.

5. Question 5

The Extract Load and Transform (ELT) process has which of the following benefits.

Select all options that apply.

1 / 1 point

☒ Data can be stored in its original format.

Correct

A benefit of ELT is that you can store data in its original format, be it JSON, XML, PDF, or images.

☒ Reduces the time required to load the data into a destination system.

Correct

In an ELT process, data is extracted and loaded in its native format. This change reduces the time required to load the data into a destination system. The change also limits resource contention on the data sources.

☐ Data will be transformed prior to storage in the destination

6. Question 6

A Data Engineers scope of work includes the process of getting, ingesting, transforming, validating, and cleaning up data to meet business requirements. Which of the following terms describe this process?

1 / 1 point

☐ Data Whaling

☐ Data Ranching

☒ Data Wrangling

Correct

A data engineer's scope of work goes well beyond looking after a database and the server where it's hosted. Data engineers must also get, ingest, transform, validate, and clean up data to meet business requirements.

7. Question 7

Which data processing framework will a data engineer use to ingest data onto cloud data platforms in Azure?

1 / 1 point

☒ Extract, load, and transform (ELT)

☐ Online transaction processing (OLTP)

☐ Extract, transform, and load (ETL)

Correct

ELT is a typical process for ingesting data from an on-premises database into the cloud.

8. Question 8

Data Engineers may sometimes perform ETL process when processing data. The extract may come from many sources including databases, files, and streams. As part of the Extract process which of the following information must be supplied?

Select all options that apply.

1 / 1 point

☐ Define the Transformation

☒ Define the Data

Correct

This is an initial step to specify the source data.

☒ Define the Data Source

Correct

During the extraction process, Data Engineers must define the Data Source.

Test prep

Latest Submission Grade 90.63%

1. Question 1

Azure Cosmos DB is a globally distributed, multimodal database and can be deployed using several API models. Which of the following APIs are supported in Cosmos DB?

Select all options that apply.

1 / 1 point



MongoDB

Correct

MongoDB API is a supported model with Azure Cosmos DB.



Cassandra

Correct

Cassandra API is a supported model with Azure Cosmos DB.



SQL

Correct

SQL API is a supported model with Azure Cosmos DB.



Ansible

2. Question 2

What type of data is typically stored in a relational database such as SQL Server or Azure SQL?

1 / 1 point



Structured



Semi-structured



Unstructured

Correct

Structured data is typically stored in a relational database such as SQL Server or Azure SQL Database. Azure SQL Database is a service that runs in the cloud.

3. Question 3

Which of the following services allow customers to store semi-structured datasets in Microsoft Azure?

Select all options that apply.

1 / 1 point



Azure SQL Datawarehouse



Azure Cosmos DB

Correct

Cosmos DB supports semi-structured (NoSQL) database services in Azure.



Azure File storage

Correct

Azure File Storage is primarily for unstructured data but can store semi-structured data as well.

☐ Azure SQL Database

☒ Azure Blob Storage

Correct

Azure Blob Storage is primarily for unstructured data but can store semi-structured data as well.

4. Question 4

Azure Synapse Analytics is an integrated analytics platform and brings together which of the following features?

Select all options that apply.

1 / 1 point

☐ Data backup

☒ Data integration

Correct

Azure Synapse Analytics is an integrated analytics platform, which combines data warehousing, big data analytics, data integration, and visualization into a single environment.

☒ Big data analytics

Correct

Azure Synapse Analytics is an integrated analytics platform, which combines data warehousing, big data analytics, data integration, and visualization into a single environment.

☒ Visualization

Correct

Azure Synapse Analytics is an integrated analytics platform, which combines data warehousing, big data analytics, data integration, and visualization into a single environment.

☒ Data warehousing

Correct

Azure Synapse Analytics is an integrated analytics platform, which combines data warehousing, big data analytics, data integration, and visualization into a single environment.

5. Question 5

Which of the following is a globally distributed multi-model database?

1 / 1 point

☐ Azure Synapse Analytics

☐ Azure SQL Database

☒ Azure Cosmos DB

Correct

Azure Cosmos DB is a globally distributed, multi-model database that can offer sub second query performance.

6. Question 6

Which Azure service is the best choice to store documentation about a data source?

1 / 1 point

☐ Azure Data Lake Storage



Azure Data Catalog



Azure Data Factory

Correct

Azure Data Catalog is a central place where an organization's users can contribute their knowledge. Together, they build a community of data sources that the organization owns.

7. Question 7

Which of the following are types of NoSQL database?

Select all options that apply.

1 / 1 point



Graph Database

Correct

Graph database finds relationships between data points by using a structure that's composed of vertices and edges. They are a type of NoSQL database.



Key-value store

Correct

Key Value store stores key-value pairs of data in a table structure. They are a type of NoSQL database.



B: Relational Database

8. Question 8

You have been asked by management to investigate a database solution in Azure that will satisfy the following.

Ability to build applications with guaranteed low latency and high availability anywhere, at any scale, and the capability to migrate existing MongoDB, and other NoSQL workloads to the cloud.

Which of the following solutions will be most appropriate?

0.25 / 1 point



Azure Database for MySQL



Azure Cosmos DB



SQL Server on Virtual Machines

This should not be selected

This will allow you to migrate your SQL workloads to Azure while maintaining complete SQL Server compatibility and operating system-level access. Try going back and reviewing Core Microsoft Azure data platform technologies Lesson.



Azure SQL Database

Feedback. This will provide a managed, intelligent SQL only solution. Try going back and reviewing Core Microsoft Azure data platform technologies Lesson.

This should not be selected

Course Practice Exam

Latest Submission Grade 95%

1. Question 1

Which of the following is a cloud integration service that orchestrates the movement of data between various data stores?

1 / 1 point

- ☐ Azure HDInsight
- ☐ Azure Synapse Analytics
- ☒ Azure Data Factory

Correct

Azure Data Factory (ADF) is a cloud integration service that orchestrates that movement of data between various data stores. Data Engineers can create data-driven workflows in the cloud for orchestrating and automating data movement and data transformation.

2. Question 2

Azure HDInsight provides technologies for ingesting, processing, and analyzing big data to support batch processing, data warehousing, IoT, and Data Science. Which of the following technologies integrated into HDInsight provides a NoSQL database built upon Hadoop commonly used for search engines and includes automatic failover?

1 / 1 point

- ☐ Apache Storm
- ☒ Apache HBase
- ☐ Apache Kafka

Correct

Apache HBase is a NoSQL database built upon Hadoop commonly used for search engines and includes automatic failover.

3. Question 3

As part of an ETL process, data in date-formats may be modified to match the destination format. At what step in the ETL process is this carried out?

1 / 1 point

- ☒ Transformation
- ☐ Extraction
- ☐ Load

Correct

In the ETL process, data is modified during the transformation step.

4. Question 4

Which one of the following is a benefit of using the Azure SQL Database cloud-based Platform-as-a-Service (PaaS) service rather than on-premises database management systems?

1 / 1 point



Increased scalability



Increased day to day management costs



Increased functionality

Correct

PaaS increases scalability as resources can be scaled up and scaled out on demand.

5. Question 5

Which Cosmos DB API works with Graph Databases?

1 / 1 point



Gremlin



Table



MongoDB



CassandraDB

Correct

The Gremlin API enables you to perform graph queries over data.

6. Question 6

Which of the following services in Microsoft Azure are used to process data?

Select all options that apply.

0.5 / 1 point



Azure Blob Storage



Event Hub

This should not be selected

Try going back and reviewing **Services on the Microsoft Azure Data Platform**.



Stream Analytics Job

Correct

Feedback: An Azure Stream Analytics job consists of an input, query, and an output.



Synapse Analytics

7. Question 7

Which service in Azure consists of input and output pipelines, and can be used to process streaming data?

1 / 1 point



Synapse Analytics



Azure Databricks



Stream Analytics Job

Correct

An Azure Stream Analytics job consists of an input, query, and output.

8. Question 8

Data Engineers configure ingestion components of Azure Stream Analytics by configuring data inputs from sources including Azure Event Hubs, Azure IoT Hub, or Azure Blob storage.

Which of the following is a feature of Azure Event Hub?

1 / 1 point

- ☐ Partitioned consumer model
- ☐ Azure Storage
- ☒ Big data streaming service

Correct

Event hubs provide a big data streaming service of Azure. It is designed for high throughput data streaming scenarios where customers may send billions of requests per day.

9. Question 9

Nonstructured data is stored in nonrelational systems, commonly called unstructured or NoSQL systems. Which of the following are types of NoSQL databases?

Select all options that apply.

1 / 1 point

- ☒ Key-value store

Correct

A Key Value store can store key-value pairs of data in a table structure.

- ☒ Document databases.

Correct

A document database can store documents that are tagged with metadata to aid document searches.

- ☐ Relational database

- ☒ Column database

Correct

Column Databases store data based on columns rather than rows. Columns can be defined at the query's runtime, allowing flexibility in the data that's returned performantly.

- ☒ Graph database

Correct

Graph database finds relationships between data points by using a structure that's composed of vertices and edges.

10. Question 10

Which of the following is not a data job role?

1 / 1 point

- ☐ Data Analyst
- ☐ Data Administrator
- ☒ Systems Administrator

Correct

Systems administrators are responsible for maintaining and managing components such as networks, virtual machines, and other physical devices in a datacenter.