DP-900: Microsoft Azure Data Fundamentals
Sample Questions

Last updated: 3/26/2022

PLEASE COMPLETE THIS SURVEY (https://aka.ms/samplequestions)

Microsoft is exploring the possibility of providing sample questions as an exam preparation resource, and we would like your feedback. While we prefer that you complete the survey after taking the exam, you may complete it at any time. Thank You!

User Guide

These sample questions are intended to provide an overview of the style, wording, and difficulty of the questions that you are likely to experience on this exam. These questions are **not** the same as what you will see on the exam nor is this document illustrative of the length of the exam or its complexity (e.g., you may see additional question types, multiple case studies, and possibly labs). These questions are **examples** only to provide insight into what to expect on the exam and help you determine if additional preparation is required.

In the first section, you will find the questions without answers so that you can test your knowledge. In the second section, the answer, a rationale, and a URL that will link you to additional information is provided immediately below each question.

Contents

Questions

Question # 1 (Sentence completion)

Select the answer that correctly completes the sentence.

Objects in which things about data should be captured and stored are called: _____.

A. tables
B. entities
C. rows
D. columns

Question # 2 (Sentence completion)

Select the answer that correctly completes the sentence.

You need to process data that is generated continuously and near real-time responses are required.

You should use _____.

A. batch processing
B. scheduled data processing
C. buffering and processing
D. streaming data processing

Question # 3 (Matching)

You are evaluating data processing approaches.

Match the data processing approaches on the left to the requirements on the right.

Data processing approaches	**Descriptions**
A. Extract, Transform, Load (ETL)	_____ 1. Optimize data privacy.
B. Extract, Load, Transform (ELT)	_____ 2. Provide support for Azure Data Lake.
	_____ 3. Manage large volumes of data.

Question # 4 (Multiple choice)

Select the answer that correctly completes the sentence.

The technique that provides recommended actions that you should take to achieve a goal or target is called _____ analytics.

A. descriptive
B. diagnostic
C. predictive
D. prescriptive

Question # 5 (Matching)

Match the data processing objects on the left to the requirements on the right.

Data processing objects	**Descriptions**
A. Tables	_____ 1. Create relationships.
B. Indexes	_____ 2. Improve processing speed for data searches.
C. Views	_____ 3. Store instances of entities as rows.
D. Keys	_____ 4. Display data from predefined queries.

Question # 6 (Sentence completion)

Select the answer that correctly completes the sentence.

The process of splitting an entity into more than one table to reduce data redundancy is called: _____.

A. deduplication
B. denormalization
C. normalization
D. optimization

Question # 7 (Sentence completion)

Select the answer that correctly completes the sentence.

Azure SQL Database is an example of _____ -as-a-service.

A. platform
B. infrastructure
C. software
D. application

Question # 8 (Matching)

You need to query an Azure SQL database.

Match the query tools on the left to the correct scenarios on the right.

Query Tools	**Descriptions**
A. Azure Data Studio	_____ 1. Query data while working within a Visual Studio project.
B. Azure Query editor	_____ 2. Query data located in a non-Microsoft platform.
C. SQL Server Data Tools	_____ 3. Query data from within the Azure portal.

Question # 9 (Sentence completion)

Select the answer that correctly completes the sentence.

The act of increasing or decreasing the resources that are available for a service is called: _____.

A. computing
B. provisioning
C. networking
D. scaling

Question # 10 (Matching)

You are creating queries to retrieve data from an Azure SQL database.

Match the SQL clauses or functions on the left to the requirements on the right.

SQL clauses	**Descriptions**
A. JOIN	_____ 1. Filter records.
B. WHERE	_____ 2. Combine rows from multiple tables.
C. SUM	_____ 3. Calculate the total value of a numeric column.
D. COUNT	_____ 4. Determine the number of rows retrieved.

Question # 11 (Multiple choice)

What are three characteristics of non-relational data? Each correct answer presents a complete solution.

A. Forced schema on data structures
B. Flexible storage of ingested data
C. Entities are self-describing
D. Entities may have different fields
E. Each row has the exact same columns

Question # 12 (Sentence completion)

Select the answer that correctly completes the sentence.

You have data that consists of JSON-based documents.

You need to store the data in an Azure environment that supports efficient non-key, field-based searching.

You should use _____ as the data store.

A. Azure Table Storage
B. Azure Blob Storage
C. Azure File Storage
D. Azure Cosmos DB

Question # 13 (Multiple Choice)

You need to create a graph database.

Which Azure data service should you use?

A. Azure Table
B. Azure Cosmos DB
C. Azure Blob
D. Azure File

Question # 14 (Sentence completion)

Select the answer that correctly completes the sentence.

You use Azure Table Storage as a non-relational data store.

You need to optimize data retrieval. You should use _____ as query criteria.

A. only partition keys
B. only row keys
C. partition keys and row keys
D. only properties

Question # 15 (Multiple Choice)

You need to use JavaScript Object Notation (JSON) files to provision Azure storage.

What should you use?

A. Azure portal
B. Azure command-line interface (CLI)
C. Azure PowerShell
D. Azure Resource Manager (ARM) templates

Question # 16 (Multiple Choice)

For which reason should you deploy a data warehouse?

A. Record daily sales transactions.
B. Perform sales trend analyses.
C. Print sales orders.
D. Search status of sales orders.

Question # 17 (Multiple Choice)

Which two Azure data services support Apache Spark clusters? Each correct answer presents a complete solution.

A. Azure Synapse Analytics
B. Azure Cosmos DB
C. Azure Databricks
D. Azure Data Factory

Question # 18 (Multiple Choice)

You design a data ingestion and transformation solution by using Azure Data Factory service.

You need to get data from an Azure SQL database.

Which two resources should you use? Each correct answer presents part of the solution.

A. Linked service
B. Copy data activity
C. Dataset
D. Azure Databricks notebook

Question # 19 (Multiple Choice)

Which Azure Data Factory component should you use to represent data that you want to ingest for processing?

A. Linked services
B. Datasets
C. Pipelines
D. Notebooks

Question # 20 (Multiple Choice)

You are designing reports by using Microsoft Power BI.

For which three scenarios can you create Power BI reports as paginated reports? Each correct answer presents a complete solution.

A. a report that has a table visual with an ability to print all the data in the table
B. a report that has a table visual with an ability to see all the data in the table
C. a report with a repeatable header and footer
D. a report that is formatted to fit well on a printed page
E. a report that uses only Power BI visuals

Questions and Answers

Question # 1 (Sentence completion)

Select the answer that correctly completes the sentence.

Objects in which things about data should be captured and stored are called: _____.

A. tables
B. entities
C. rows
D. columns

Answer:	B. Entities
Objective:	1.1 Describe types of core data workloads
Rationale:	An entity is a thing in which information needs to be known or held. A table is the object that stores a collection of entities. A row represents a single instance of an entity. A column defines a specific property of an entity.
URL:	https://docs.microsoft.com/en-us/learn/modules/describe-concepts-of-relational-data/2-explore-characteristics

Question # 2 (Sentence completion)

Select the answer that correctly completes the sentence.

You need to process data that is generated continuously and near real-time responses are required.

You should use _____.

A. batch processing
B. scheduled data processing
C. buffering and processing
D. streaming data processing

Answer:	D. Streaming data processing
Objective:	1.1 Describe types of core data workloads
Rationale:	When data is generated on a continual basis and insights must you must be able see data insights immediately, process data as it arrives rather than storing and processing data later as a group. Batch processing, scheduled processing, and buffering all collect and store data for later processing.
URL:	https://docs.microsoft.com/en-us/learn/modules/explore-core-data-concepts/4-describe-difference

Question # 3 (Matching)

You are evaluating data processing approaches.

Match the data processing approaches on the left to the requirements on the right.

Data processing approaches A. Extract, Transform, Load (ETL) B. Extract, Load, Transform (ELT)	**Descriptions** _____ 1. Optimize data privacy. _____ 2. Provide support for Azure Data Lake. _____ 3. Manage large volumes of data.
Answer:	Extract, Transform, Load (A) matches description 1: Optimize data privacy. Extract, Load, Transform (B) matches description 2: Provide support for Azure Data Lake, and description 3: Manage large volumes of data.
Objective:	1.2 Describe data analytics core concepts
Rationale:	Extract, Transform, Load (ETL) is the correct approach when you need to filter sensitive data before loading the data into an analytical model. It is suitable for simple data models that do not require Azure Data Lake support. Extract, Load, Transform (ELT) is the correct approach because it supports Azure Data Lake as the data store and manages large volumes of data.
URL:	https://docs.microsoft.com/en-us/learn/modules/explore-concepts-of-data-analytics/2-describe-data-ingestion-process

Question # 4 (Multiple choice)

Select the answer that correctly completes the sentence.

The technique that provides recommended actions that you should take to achieve a goal or target is called _____ analytics.

A. descriptive
B. diagnostic
C. predictive
D. prescriptive

Answer:	D. Prescriptive
Objective:	Exam objective number and text
Rationale:	Prescriptive analytics helps you define actions (prescription) that you should perform to achieve what you need (or overcome a problem).

	Descriptive analytics help you understand what has happened by looking at historical data. You can further analyze information by using diagnostic analytics. Predictive analytics helps to forecast the future values using historical values.
URL:	https://docs.microsoft.com/en-us/learn/modules/explore-concepts-of-data-analytics/4-explore

Question # 5 (Matching)

Match the data processing objects on the left to the requirements on the right.

Data processing objects	**Descriptions**
A. Tables B. Indexes C. Views D. Keys	_____ 1. Create relationships. _____ 2. Improve processing speed for data searches. _____ 3. Store instances of entities as rows. _____ 4. Display data from predefined queries.

Answer:	Keys (D) matches description 1: Create relationships. Indexes (B) matches description 2: Improve processing speed for data searches. Tables (A) matches description 3: Store instances of entities as rows. Views (C) matches description 4: Display data from predefined queries.
Objective:	2.1 Describe relational data workloads
Rationale:	Tables use rows to store data or instances of entities. Indexes are a subset of data used to speed up data searches. Views are virtual tables based on the results of a query. You use primary keys and foreign keys to create relationships between tables.
URL:	https://docs.microsoft.com/en-us/learn/modules/describe-concepts-of-relational-data/2-explore-characteristics https://docs.microsoft.com/en-us/learn/modules/describe-concepts-of-relational-data/3-explore-structures

Question # 6 (Sentence completion)

Select the answer that correctly completes the sentence.

The process of splitting an entity into more than one table to reduce data redundancy is called: _____.

A. deduplication
B. denormalization
C. normalization
D. optimization

Answer:	C. Normalization
Objective:	2.2 Describe relational Azure Data Services
Rationale:	Normalization is the process of splitting a logical entity into multiple tables and designing the relationship between tables. Normalization is often used to improve query completion speeds for transaction processing databases. Deduplication is the process of removing duplicate data from tables. Denormalization is the process of combining tables. This process is often used to store data in databases which are used for data mining and reporting. Optimization is the process of modifying fields and database structure to improve overall performance.
URL:	https://docs.microsoft.com/en-us/learn/modules/describe-concepts-of-relational-data/2-explore-characteristics

Question # 7 (Sentence completion)

Select the answer that correctly completes the sentence.

Azure SQL Database is an example of _____ -as-a-service.

A. platform
B. infrastructure
C. software
D. application

Answer:	A. Platform
Objective:	2.2 Describe relational Azure Data Services
Rationale:	Azure SQL Database is an example of platform-as-a-service. Infrastructure-as-a-service includes technologies such as virtual machines and virtual networks. Software-as-a-service describes a method of software delivery in which users license software online by subscription. Microsoft offers subscriptions to license Office products like Microsoft Excel.

	Application-as-a-service describes applications that are hosted in the cloud and offered to users by subscription. This is like software-as-a-service but often relates to custom line-of-business solutions. There are no Azure data services that use the software-as-a-service or application-as-a-service model.
URL:	https://docs.microsoft.com/en-us/learn/modules/explore-relational-data-offerings/2-azure-data-services

Question # 8 (Matching)

You need to query an Azure SQL database.

Match the query tools on the left to the correct scenarios on the right.

Query Tools	**Descriptions**
A. Azure Data Studio B. Azure Query editor C. SQL Server Data Tools	_____ 1. Query data while working within a Visual Studio project. _____ 2. Query data located in a non-Microsoft platform. _____ 3. Query data from within the Azure portal.
Answer:	SQL Server Data Tools (C) matches description 1: Query data while working within a Visual Studio project. Azure Data Studio (A) matches description 2: Query data located in a non-Microsoft platform. Azure Query editor (B) matches description 3: Query data from within the Azure portal.
Objective:	2.3 Identify basic management tasks for relational data
Rationale:	Azure Data Studio is a cross-platform database tool that you can use with both on-premises and cloud data platforms on Windows, MacOS and Linux. Azure Query editor is available in the Azure portal. You can use this tool for querying Azure SQL databases. SQL Server Data Tools is available in Visual Studio. You can use this tool to connect to and query on-premises and cloud data services.
URL:	https://docs.microsoft.com/en-us/learn/modules/query-relational-data/3-sql-database?ns-enrollment-type=LearningPath&ns-enrollment-id=learn.wwl.azure-data-fundamentals-explore-relational-data https://docs.microsoft.com/en-us/sql/azure-data-studio/download-azure-data-studio?view=sql-server-ver15 https://docs.microsoft.com/en-us/sql/ssms/download-sql-server-management-studio-ssms?view=sql-server-ver15

Question # 9 (Sentence completion)

Select the answer that correctly completes the sentence.

The act of increasing or decreasing the resources that are available for a service is called: _____.

A. computing
B. provisioning
C. networking
D. scaling

Answer:	D. Scaling
Objective:	2.3 Identify basic management tasks for relational data
Rationale:	Scaling is the act of increasing or decreasing the resources used by a service. Computing is the act of processing data. Provisioning is the act of running series of tasks to create and configure a service. Networking is the act of providing connectivity to a data resource.
URL:	https://docs.microsoft.com/en-us/learn/modules/explore-provision-deploy-relational-database-offerings-azure/2-describe-provision-relational-data-services

Question # 10 (Matching)

You are creating queries to retrieve data from an Azure SQL database.

Match the SQL clauses or functions on the left to the requirements on the right.

SQL clauses	**Descriptions**
A. JOIN B. WHERE C. SUM D. COUNT	_____ 1. Filter records. _____ 2. Combine rows from multiple tables. _____ 3. Calculate the total value of a numeric column. _____ 4. Determine the number of rows retrieved.
Answer:	WHERE (B) matches description 1: Filter records. JOIN (A) matches description 2: Combine rows from multiple tables. SUM (C) matches description 3: Calculate the total value of a numeric column. COUNT (D) matches description 4: Determine the number of rows retrieved.
Objective:	2.4 Describe query techniques for data using SQL language
Rationale:	Use the JOIN clause for combining rows retrieved from multiple tables based on a condition. Use the WHERE clause for filtering out rows. Use the SUM function for calculating the total sum of a numeric column. Use the COUNT function returns the number of rows that matches a specified criterion.
URL:	https://docs.microsoft.com/en-us/learn/modules/query-relational-data/6-exercise-perform-query?ns-enrollment-type=LearningPath&ns-enrollment-id=learn.wwl.azure-data-fundamentals-explore-relational-data

Question # 11 (Multiple choice)

What are three characteristics of non-relational data? Each correct answer presents a complete solution.

A. Forced schema on data structures
B. Flexible storage of ingested data
C. Entities are self-describing
D. Entities may have different fields
E. Each row has the exact same columns

Answer:	B. Flexible storage of ingested data. AND C. Entities are self-describing. AND D. Entities may have different fields.

Objective:	3.1 Describe non-relational data workloads
Rationale:	Characteristics of non-relational data are flexible storage of ingested data, self-describing entities, and the ability to have different fields for entities. Forced schemas are a feature of relational tables. Forcing each row to maintain the same columns describes a relational table.
URL:	https://docs.microsoft.com/en-us/learn/modules/explore-concepts-of-non-relational-data/2-explore-characteristics

Question # 12 (Sentence completion)

Select the answer that correctly completes the sentence.

You have data that consists of JSON-based documents.

You need to store the data in an Azure environment that supports efficient non-key, field-based searching.

You should use _____ as the data store.

A. Azure Table Storage
B. Azure Blob Storage
C. Azure File Storage
D. Azure Cosmos DB

Answer:	D. Azure Cosmos DB
Objective:	3.1 Describe non-relational data workloads
Rationale:	Azure Cosmos DB is a non-relational document DB that supports high latency for both reading and writing. It adds indexes automatically on elements. Azure Table Storage is a NoSQL key-value storage that supports semi-structured data with dynamic column. This technique optimizes both data retrieval and writing but performance is not good when searching on non-partition key and non-key values. Azure File Storage and Blob Storage are for files and BLOBs, not for searchable JSON documents.
URL:	https://docs.microsoft.com/en-us/learn/modules/explore-non-relational-data-offerings-azure/5-explore-azure-cosmos-database?ns-enrollment-type=LearningPath&ns-enrollment-id=learn.wwl.azure-data-fundamentals-explore-non-relational-data

Question # 13 (Multiple Choice)

You need to create a graph database.

Which Azure data service should you use?

A. Azure Table
B. Azure Cosmos DB
C. Azure Blob
D. Azure File

Answer:	B. Azure Cosmos DB
Objective:	3.2 Describe non-relational data offerings on Azure
Rationale:	Only Azure Cosmos DB supports creating graph databases. Azure Table Storage, Azure Blob Storage, and Azure File Storage do not support graph databases.
URL:	https://docs.microsoft.com/en-us/learn/modules/explore-non-relational-data-offerings-azure/5-explore-azure-cosmos-database

Question # 14 (Sentence completion)

Select the answer that correctly completes the sentence.

You use Azure Table Storage as a non-relational data store.

You need to optimize data retrieval. You should use _____ as query criteria.

A. only partition keys
B. only row keys
C. partition keys and row keys
D. only properties

Answer:	C. partition keys and row keys
Objective:	3.2 Describe non-relational data offerings on Azure
Rationale:	You can optimize data read performance with Azure Table Storage when the search completes by using both a partition key and arow key. Azure Table Storage does not support search on properties.
URL:	https://docs.microsoft.com/en-us/learn/modules/explore-non-relational-data-offerings-azure/2-explore-azure-table-storage?ns-enrollment-type=LearningPath&ns-enrollment-id=learn.wwl.azure-data-fundamentals-explore-non-relational-data

Question # 15 (Multiple Choice)

You need to use JavaScript Object Notation (JSON) files to provision Azure storage.

What should you use?

A. Azure portal
B. Azure command-line interface (CLI)
C. Azure PowerShell
D. Azure Resource Manager (ARM) templates

Answer:	D. Azure Resource Manager (ARM) templates
Objective:	3.3 Identify basic management tasks for non-relational data
Rationale:	Azure Resource Manager templates use JSON to create text files to standardize Azure Storage provisioning. Azure portal, Azure CLI, and Azure PowerShell do not use JSON text files.
URL:	https://docs.microsoft.com/en-us/learn/modules/explore-provision-deploy-non-relational-data-services-azure/2-describe-provision-non-relational-data-services

Question # 16 (Multiple Choice)

For which reason should you deploy a data warehouse?

A. Record daily sales transactions.
B. Perform sales trend analyses.
C. Print sales orders.
D. Search status of sales orders.

Answer:	Insert correct answer(s)
Objective:	Exam objective number and text
Rationale:	You typically create data warehouses to support analytical queries that involve high volume data and generating aggregated values. Recording daily sales, printing reports, and searching sales orders are typical operations seen in online data processing (OLTP) databases and not in data warehouses.
URL:	https://docs.microsoft.com/en-us/learn/modules/examine-components-of-modern-data-warehouse/1-introduction https://docs.microsoft.com/en-us/learn/modules/examine-components-of-modern-data-warehouse/2-describe-warehousing

Question # 17 (Multiple Choice)

Which two Azure data services support Apache Spark clusters? Each correct answer presents a complete solution.

A. Azure Synapse Analytics
B. Azure Cosmos DB
C. Azure Databricks
D. Azure Data Factory

Answer:	A. Azure Synapse Analytics AND C. Azure Databricks
Objective:	4.2 Describe the components of a modern data warehouse
Rationale:	Azure Synapse Analytics and Azure Databricks both support using Apache Spark clusters to process data. Azure Cosmos DB is for non-relational data. You can use Azure Data Factory for data integration and migration.
URL:	https://docs.microsoft.com/en-us/learn/modules/explore-provision-deploy-non-relational-data-services-azure/2-describe-provision-non-relational-data-services

Question # 18 (Multiple Choice)

You design a data ingestion and transformation solution by using Azure Data Factory service.

You need to get data from an Azure SQL database.

Which two resources should you use? Each correct answer presents part of the solution.

A. Linked service
B. Copy data activity
C. Dataset
D. Azure Databricks notebook

Answer:	A. Linked service AND B. Dataset
Objective:	4.3 Describe data ingestion and processing on Azure
Rationale:	To get data from a source, you need to create a linked service for Azure Data Factory. The linked service contains details about the data source including the server name and credentials. You must also define a dataset to describe the expected data structure. A dataset stores data retrieved from a data source. You can use the Copy data activity for transferring data from one source to another (destination). You can use an Azure Databricks notebook for processing data using spark clusters with given instructions.

URL:	https://docs.microsoft.com/en-us/learn/modules/explore-data-ingestion-azure/2-describe-common-practices-for-data-loading?ns-enrollment-type=LearningPath&ns-enrollment-id=learn.wwl.azure-data-fundamentals-explore-data-warehouse-analytics

Question # 19 (Multiple Choice)

Which Azure Data Factory component should you use to represent data that you want to ingest for processing?

A. Linked services
B. Datasets
C. Pipelines
D. Notebooks

Answer:	B. Datasets
Objective:	4.3 Describe data ingestion and processing on Azure
Rationale:	A dataset represents the data that you want to ingest for processing. A dataset can also represent output data from a process. You use a pipeline to perform tasks and processes. You use a linked service to connect to a source or destination. A notebook can contain cells that read data, process data, and write the results out to a data store.
URL:	https://docs.microsoft.com/en-us/learn/modules/explore-data-ingestion-azure/2-describe-common-practices-for-data-loading

Question # 20 (Multiple Choice)

You are designing reports by using Microsoft Power BI.

For which three scenarios can you create Power BI reports as paginated reports? Each correct answer presents a complete solution.

A. a report that has a table visual with an ability to print all the data in the table
B. a report that has a table visual with an ability to see all the data in the table
C. a report with a repeatable header and footer
D. a report that is formatted to fit well on a printed page
E. a report that uses only Power BI visuals

Answer:	A. a report that has a table visual with an ability to print all the data in the table AND C. a report with a repeatable header and footer AND D. a report with a repeatable header and footer
Objective:	4.4 Describe data visualization in Power BI

Rationale:	When a Power BI report that has a table visual contains multiple rows, printed, only records that can are displayed will be printed. All records print if you design the report by using Report Builder as a paginated report, all records print. Only paginated report supports repeatable headers and footers. You cannot create paginated reports by using Power BI visuals. You must use Report Builder instead.
URL:	https://docs.microsoft.com/en-us/learn/modules/get-started-with-power-bi/1-introduction?ns-enrollment-type=LearningPath&ns-enrollment-id=learn.wwl.azure-data-fundamentals-explore-data-warehouse-analytics