Question 1

The term “data repositories” **exclusively** refers to RDBMes and NoSQL databases that are used to collect, organize, and isolate data for analytics.

True

False

Correct

The term “data repositories” includes not just RDBMSes and NoSQL databases, it also includes data warehouses, data marts, and data lakes.

Question 2

In use cases for RDBMS, what is one of the reasons that relational databases are so well suited for OLTP applications?

Offer easy backup and restore options

Minimize data redundancy

Support the ability to insert, update, or delete small amounts of data

Allow you to make changes in the database even while a query is being executed

Correct

This is one of the abilities of RDBMSs that make them very well suited for OLTP applications.

Question 3

Which NoSQL database type stores each record and its associated data within a single document and also works well with Analytics platforms?

Document-based

Key-value store

Column-based

Graph-based

Correct

Document-based NoSQL databases store each record and its associated data within a single document and work well with Analytics platforms.

Question 4

What type of data repository is used to isolate a subset of data for a particular business function, purpose, or community of users?

Data Mart

Data Warehouse

Data Lake

Data Pipeline

Correct

A data mart is a sub-section of the data warehouse used to isolate a subset of data for a particular business function, purpose, or community of users.

Question 5

_____ is ideal for data lakes where transformations on data are applied after raw data is loaded into the data lake.

Data Pipeline

ETL (Extract-Transform-Load) Process

Batch Processing

Correct

ELT is useful for processing large sets of unstructured and non-relational data, which makes it ideal for use in data lakes, generally used for storing large amounts of structured, semi-structured, and unstructured data in their native format.

Question 6

Which one of these statements explains what data integration is?

Data Integration is the process of loading data into a data repository

Data Integration includes extracting, transforming, merging, and delivering quality data for analytical purposes

Data Integration is the process of extracting data

Data Integration is the process of applying business logic to source data

Correct

Data Integration extracts and combines disparate source data into a unified view so that data consumers can query and analyze the integrated data.

Question 1

Data Marts and Data Warehouses have typically been relational, but the emergence of what technology has helped to let these be used for non-relational data?

NoSQL

Data Lake

SQL

ETL

Correct

The emergence of NoSQL technology has made it possible for data marts and data warehouses to be used for both relational and non-relational data.

Question 2

What is one of the most significant advantages of an RDBMS?

Is ACID-Compliant

Can store only structured data

Requires source and destination tables to be identical for migrating data

Enforces a limit on the length of data fields

Correct

ACID-Compliance is one of the significant advantages of an RDBMS.

Question 3

Which one of the NoSQL database types uses a graphical model to represent and store data, and is particularly useful for visualizing, analyzing, and finding connections between different pieces of data?

Graph-based

Column-based

Document-based

Key value store

Correct

Graph-based NoSQL databases use a graphical model to represent and store data and are used for visualizing, analyzing, and finding connections between different pieces of data.

Question 4

Which of the data repositories serves as a pool of raw data and stores large amounts of structured, semi-structured, and unstructured data in their native formats?

Relational Databases

Data Lakes

Data Warehouses

Data Marts

Correct

A Data Lake can store large amounts of structured, semi-structured, and unstructured data in their native format, classified and tagged with metadata.

Question 5

While data integration combines disparate data into a unified view of the data, a data pipeline covers the entire data movement journey from source to destination systems, and ETL is a process within data integration.

True

False

Correct

A data pipeline covers the entire journey of data from source to destination. Data integration is performed within a data pipeline, while ETL is a process within data integration.