

PRACTICE QUESTIONS

AWS

DEA-C01

AWS Certified Data Engineer - Associate

SHARED BY [u/Glad-Kiwi-6631](#) on aiexamcerts

EDITION Free Set 1

QUESTIONS 10

SHARED October 15, 2026

EXAM AWS DEA-C01, AWS Certified Data Engineer - Associate

Shared by u/Glad-Kiwi-6631 on aiexamcerts

10 practice questions for the AWS DEA-C01 exam, AWS Certified Data Engineer - Associate, each followed by its answer and a short explanation. More free sets and the full practice exam: www.aiexamcerts.com/aws/dea-c01. Spotted a mistake? Email support@aiexamcerts.com with the exam code DEA-C01 and the question number.

1. Raw S3 objects must be kept for 90 days, then moved to a colder storage class, and deleted after one year. Which feature should be configured?
- A S3 Replication rule
 - B S3 Object Lock in compliance mode
 - C S3 Lifecycle configuration
 - D S3 Requester Pays

Answer: C

S3 Lifecycle rules transition objects between storage classes and expire them after set ages, which covers both the move and the deletion. Replication copies objects to another bucket and does not age them out. Object Lock blocks deletion for a retention period, which works against the required expiry.

2. Amazon CloudWatch _____ are time-ordered data points that alarms and dashboards read.

Answer: metrics

CloudWatch stores metrics as time-ordered series of data points, identified by namespace, name, and dimensions. Alarms evaluate those series against thresholds. Logs hold event records and traces hold request paths, so neither fits this definition.

3. Which AWS service uses machine learning to discover and classify sensitive data, such as personally identifiable information, stored in Amazon S3?
- A AWS Config
 - B Amazon Inspector
 - C AWS Security Hub
 - D Amazon Macie

Answer: D

Amazon Macie analyzes S3 buckets and objects with machine learning and pattern matching to find sensitive data such as PII. AWS Config evaluates resource configuration, not object contents. Inspector scans workloads for vulnerabilities, and Security Hub aggregates findings from other services.

4. Amazon Managed Workflows for Apache Airflow (MWAA) runs Apache Airflow environments as a managed service.
- A True
 - B False

Answer: A (True)

MWAA provisions and operates Airflow environments, so you supply DAG files and AWS handles the scheduler, workers, and web server. Teams often confuse it with Glue workflows, which are a separate orchestration feature for Glue jobs and crawlers.

5. Which two controls protect data at rest in an Amazon Redshift environment? (Choose two.)
- A Enable AWS KMS encryption for the cluster
 - B Require SSL for client connections
 - C Place the cluster in a public subnet with a public IP
 - D Encrypt UNLOAD output in S3 with a KMS key
 - E Grant redshift:* to every analyst role

Answer: A, D

Cluster encryption with KMS protects the data stored on disk, and UNLOAD can write exported files to S3 encrypted with a KMS key. Requiring SSL protects data in transit, not at rest. Public placement and broad permissions increase exposure rather than protecting stored data.

6. An ingestion pipeline sends small JSON records through Firehose to Amazon S3 and produces thousands of tiny objects. Which Firehose setting most directly reduces the object count?
- A Increase the buffer size and buffer interval
 - B Decrease the buffer size to 1 MiB
 - C Turn off compression on the output
 - D Add a second delivery stream for each producer

Answer: A

Firehose writes one object per buffer flush, so a larger buffer size or longer buffer interval means fewer, larger files. A smaller buffer does the opposite and creates even more objects. Compression changes object size per object, not how many objects are written, and extra delivery streams multiply the problem.

7. Which AWS Glue component runs Spark or Python code to transform data?

- A Glue crawler
- B Glue job
- C Glue trigger
- D Glue workflow

Answer: B

A Glue job is the unit that runs ETL code, either Spark or Python shell, against your sources. Crawlers discover metadata and do not transform data. Triggers and workflows start or group jobs and crawlers, but they run no transformation code themselves.

8. Continuous logging for AWS Glue jobs sends driver and executor output to Amazon CloudWatch _____ groups.

Answer: log

Glue writes driver and executor output to CloudWatch Logs log groups, which can be searched and used in metric filters and alarms. CloudWatch Metrics holds numeric time series and is a different feature. Continuous logging is what makes logs visible while a long job is still running.

9. A DynamoDB table must support queries on customer_status, a non-key attribute, without scanning the whole table. What should be added?

- A A DynamoDB Accelerator (DAX) cluster in front of the table
- B A Time to Live attribute
- C A global secondary index on customer_status
- D A second table with the same partition key

Answer: C

A global secondary index lets you query by a different partition key, here customer_status, and it can be created on an existing table. DAX speeds up reads of items already known by key and does not add a query path. TTL deletes expired items and a second table would need its own sync logic.

10. A pipeline must alert an on-call engineer when a Step Functions execution fails. Which configuration receives the state change event?

- A An Amazon Athena workgroup with query logging
- B An AWS CloudTrail trail for the state machine
- C An Amazon Redshift WLM queue for the workflow
- D An EventBridge rule for failed executions

Answer: D

Step Functions emits execution status change events to EventBridge, so a rule matching a failed status can route the event to SNS or another target. CloudTrail records API calls and is not the place to subscribe to execution status. Athena workgroups and Redshift WLM queues have no part in alerting on workflow failures.