



DAS-C01^{Q&As}

AWS Certified Data Analytics - Specialty (DAS-C01)

Pass Amazon DAS-C01 Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

<https://www.pass4itsure.com/das-c01.html>

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by Amazon
Official Exam Center

-  **Instant Download** After Purchase
-  **100% Money Back** Guarantee
-  **365 Days** Free Update
-  **800,000+** Satisfied Customers



**QUESTION 1**

A company uses Amazon Redshift as its data warehouse. A new table has columns that contain sensitive data. The data in the table will eventually be referenced by several existing queries that run many times a day.

A data analyst needs to load 100 billion rows of data into the new table. Before doing so, the data analyst must ensure that only members of the auditing group can read the columns containing sensitive data.

How can the data analyst meet these requirements with the lowest maintenance overhead?

- A. Load all the data into the new table and grant the auditing group permission to read from the table. Load all the data except for the columns containing sensitive data into a second table. Grant the appropriate users read-only permissions to the second table.
- B. Load all the data into the new table and grant the auditing group permission to read from the table. Use the GRANT SQL command to allow read-only access to a subset of columns to the appropriate users.
- C. Load all the data into the new table and grant all users read-only permissions to non-sensitive columns. Attach an IAM policy to the auditing group with explicit ALLOW access to the sensitive data columns.
- D. Load all the data into the new table and grant the auditing group permission to read from the table. Create a view of the new table that contains all the columns, except for those considered sensitive, and grant the appropriate users read-only permissions to the table.

Correct Answer: D

QUESTION 2

An analytics team uses Amazon OpenSearch Service for an analytics API to be used by data analysts. The OpenSearch Service cluster is configured with three master nodes. The analytics team uses Amazon Managed Streaming for Apache Kafka (Amazon MSK) and a customized data pipeline to ingest and store 2 months of data in an OpenSearch Service cluster. The cluster stopped responding, which is regularly causing timeout requests. The analytics team discovers the cluster is handling too many bulk indexing requests.

Which actions would improve the performance of the OpenSearch Service cluster? (Choose two.)

- A. Reduce the number of API bulk requests on the OpenSearch Service cluster and reduce the size of each bulk request.
- B. Scale out the OpenSearch Service cluster by increasing the number of nodes.
- C. Reduce the number of API bulk requests on the OpenSearch Service cluster, but increase the size of each bulk request
- D. Increase the number of master nodes for the OpenSearch Service cluster
- E. Scale down the pipeline component that is used to ingest the data into the OpenSearch Service cluster.

Correct Answer: AB

QUESTION 3



A mortgage company has a microservice for accepting payments. This microservice uses the Amazon DynamoDB encryption client with AWS KMS managed keys to encrypt the sensitive data before writing the data to DynamoDB. The finance team should be able to load this data into Amazon Redshift and aggregate the values within the sensitive fields. The Amazon Redshift cluster is shared with other data analysts from different business units.

Which steps should a data analyst take to accomplish this task efficiently and securely?

- A. Create an AWS Lambda function to process the DynamoDB stream. Decrypt the sensitive data using the same KMS key. Save the output to a restricted S3 bucket for the finance team. Create a finance table in Amazon Redshift that is accessible to the finance team only. Use the COPY command to load the data from Amazon S3 to the finance table.
- B. Create an AWS Lambda function to process the DynamoDB stream. Save the output to a restricted S3 bucket for the finance team. Create a finance table in Amazon Redshift that is accessible to the finance team only. Use the COPY command with the IAM role that has access to the KMS key to load the data from S3 to the finance table.
- C. Create an Amazon EMR cluster with an EMR_EC2_DefaultRole role that has access to the KMS key. Create Apache Hive tables that reference the data stored in DynamoDB and the finance table in Amazon Redshift. In Hive, select the data from DynamoDB and then insert the output to the finance table in Amazon Redshift.
- D. Create an Amazon EMR cluster. Create Apache Hive tables that reference the data stored in DynamoDB. Insert the output to the restricted Amazon S3 bucket for the finance team. Use the COPY command with the IAM role that has access to the KMS key to load the data from Amazon S3 to the finance table in Amazon Redshift.

Correct Answer: B

QUESTION 4

A healthcare company ingests patient data from multiple data sources and stores it in an Amazon S3 staging bucket. An AWS Glue ETL job transforms the data, which is written to an S3-based data lake to be queried using Amazon Athena. The company wants to match patient records even when the records do not have a common unique identifier.

Which solution meets this requirement?

- A. Use Amazon Macie pattern matching as part of the ETLjob
- B. Train and use the AWS Glue PySpark filter class in the ETLjob
- C. Partition tables and use the ETL job to partition the data on patient name
- D. Train and use the AWS Glue FindMatches ML transform in the ETLjob

Correct Answer: D

The FindMatches transform enables you to identify duplicate or matching records in your dataset, even when the records do not have a common unique identifier and no fields match exactly.

Reference: <https://docs.aws.amazon.com/glue/latest/dg/machine-learning.html>

QUESTION 5

A company is hosting an enterprise reporting solution with Amazon Redshift. The application provides reporting capabilities to three main groups: an executive group to access financial reports, a data analyst group to run long-running ad-hoc queries, and a data engineering group to run stored procedures and ETL processes. The executive



team requires queries to run with optimal performance. The data engineering team expects queries to take minutes.

Which Amazon Redshift feature meets the requirements for this task?

- A. Concurrency scaling
- B. Short query acceleration (SQA)
- C. Workload management (WLM) D. Materialized views

Correct Answer: D

Materialized views:

Reference: <https://aws.amazon.com/redshift/faqs/>

[DAS-C01 PDF Dumps](#)

[DAS-C01 Exam Questions](#)

[DAS-C01 Braindumps](#)