



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 technology company is creating a dashboard that will visualize and analyze time-sensitive data. The data will come in through Amazon Kinesis Data Firehose with the buffer interval set to 60 seconds. The dashboard must support near-realtime data.

Which visualization solution will meet these requirements?

- A. Select Amazon OpenSearch Service (Amazon Elasticsearch Service) as the endpoint for Kinesis Data Firehose. Set up an OpenSearch Dashboards (Kibana) using the data in Amazon OpenSearch Service (Amazon ES) with the desired analyses and visualizations.
- B. Select Amazon S3 as the endpoint for Kinesis Data Firehose. Read data into an Amazon SageMaker Jupyter notebook and carry out the desired analyses and visualizations.
- C. Select Amazon Redshift as the endpoint for Kinesis Data Firehose. Connect Amazon QuickSight with SPICE to Amazon Redshift to create the desired analyses and visualizations.
- D. Select Amazon S3 as the endpoint for Kinesis Data Firehose. Use AWS Glue to catalog the data and Amazon Athena to query it. Connect Amazon QuickSight with SPICE to Athena to create the desired analyses and visualizations.

Correct Answer: A

Reference: <https://aws.amazon.com/blogs/big-data/ingest-streaming-data-into-amazon-elasticsearch-service-within-the-privacy-of-your-vpc-with-amazon-kinesis-data-firehose/>

QUESTION 2

A company has an electronic healthcare system that contains data of patients. The data is consolidated from multiple systems and is stored in an Amazon S3 bucket in .csv format. The company has created an AWS Glue Data Catalog. The dataset contains duplicate data, and no unique keys exist to identify a patient. Fields do not match exactly across the systems.

A data analytics specialist must design a solution to identify and remove duplicates. The solution must minimize the amount of human intervention and code that are required.

The data analytics specialist starts by using labeled data to teach the FindMatches machine learning (ML) transform.

What must the data analytics specialist do next to meet these requirements?

- A. Identify matches in the dataset by using an AWS Glue ETL job with Spark distinct(). Review the output by using Amazon Redshift Spectrum.
- B. Identify matches in the dataset by using an AWS Glue ETL job with Spark distinct(). Create a Data Catalog of transformed results. Review the output by using Amazon Athena.
- C. Identify matches in the dataset by using an AWS Glue ETL job that has a transform type of "find matching records." Create a Data Catalog of transformed results. Review the output by using Amazon Athena.
- D. Identify matches in the dataset by using an AWS Glue ETL job that has a transform type of "find matching records." Review the output by using Amazon Redshift Spectrum.



Correct Answer: C

QUESTION 3

A retail company uses Amazon Aurora MySQL for its operational data store and Amazon Redshift for its data warehouse. The MySQL database resides in a different VPC than the Redshift cluster. Data analysts need to query data in both MySQL and Amazon Redshift to provide business insights. The company wants the lowest network latency between the two VPCs.

Which combination of solutions meet these requirements? (Choose two.)

- A. Set up VPC peering between the MySQL VPC and the Redshift VPC.
- B. Set up a transit gateway to connect the MySQL VPC with the Redshift VPC.
- C. Use a Redshift federated query to retrieve live data from the MySQL database. Create a late-binding view to combine the MySQL data with the Redshift data.
- D. Use Amazon Redshift Spectrum to retrieve live data from the MySQL database. Create a late-binding view to combine the MySQL data with the Redshift data.
- E. Use the Redshift COPY command to constantly copy live data from MySQL to the Redshift cluster. Create a late-binding view to combine the MySQL data with the Redshift data.

Correct Answer: AC

QUESTION 4

A company uses Amazon OpenSearch Service (Amazon Elasticsearch Service) to store and analyze its website clickstream data. The company ingests 1 TB of data daily using Amazon Kinesis Data Firehose and stores one day's worth of data in an Amazon OpenSearch Service (Amazon Elasticsearch Service) cluster.

The company has very slow query performance on the Amazon OpenSearch Service (Amazon Elasticsearch Service) index and occasionally sees errors from Kinesis Data Firehose when attempting to write to the index. The Amazon OpenSearch Service (Amazon Elasticsearch Service) cluster has 10 nodes running a single index and 3 dedicated master nodes. Each data node has 1.5 TB of Amazon EBS storage attached and the cluster is configured with 1,000 shards. Occasionally, JVMMemoryPressure errors are found in the cluster logs.

Which solution will improve the performance of Amazon OpenSearch Service (Amazon Elasticsearch Service)?

- A. Increase the memory of the Amazon OpenSearch Service (Amazon Elasticsearch Service) master nodes.
- B. Decrease the number of Amazon OpenSearch Service (Amazon Elasticsearch Service) data nodes.
- C. Decrease the number of Amazon OpenSearch Service (Amazon Elasticsearch Service) shards for the index.
- D. Increase the number of Amazon OpenSearch Service (Amazon Elasticsearch Service) shards for the index.

Correct Answer: C

Reference: <https://www.blumatador.com/docs/troubleshooting/aws-elasticsearch-jvm->

