



**Free Questions for DAS-C01 by ebraindumps**

**Shared by Rodgers on 06-06-2022**

**For More Free Questions and Preparation Resources**

**Check the Links on Last Page**

# Question 1

---

**Question Type:** MultipleChoice

---

A company uses Amazon Redshift as its data warehouse. A new table includes some columns that contain sensitive data and some columns that contain non-sensitive data. The data in the table eventually will be referenced by several existing queries that run many times each day.

A data analytics specialist must ensure that only members of the company's auditing team can read the columns that contain sensitive data. All other users must have read-only access to the columns that contain non-sensitive data.

Which solution will meet these requirements with the LEAST operational overhead?

## Options:

---

- A-** Grant the auditing team permission to read from the table. Load the columns that contain non-sensitive data into a second table. Grant the appropriate users read-only permissions to the second table.
- B-** Grant all users read-only permissions to the columns that contain non-sensitive data. Use the GRANT SELECT command to allow the auditing team to access the columns that contain sensitive data.
- C-** Grant all users read-only permissions to the columns that contain non-sensitive data. Attach an IAM policy to the auditing team with an explicit Allow action that grants access to the columns that contain sensitive data.
- D-** Grant the auditing team permission to read from the table. Create a view of the table that includes the columns that contain non-sensitive data. Grant the appropriate users read-only permissions to that view.

**Answer:**

---

B

**Explanation:**

---

<https://aws.amazon.com/jp/about-aws/whats-new/2020/03/announcing-column-level-access-control-for-amazon-redshift/>

## Question 2

---

**Question Type:** MultipleChoice

---

A company is building an analytical solution that includes Amazon S3 as data lake storage and Amazon Redshift for data warehousing. The company wants to use Amazon Redshift Spectrum to query the data that is stored in Amazon S3.

Which steps should the company take to improve performance when the company uses Amazon Redshift Spectrum to query the S3 data files? (Select THREE )

Use gzip compression with individual file sizes of 1-5 GB

### Options:

---

- A- Use a columnar storage file format
- B- Partition the data based on the most common query predicates
- C- Split the data into KB-sized files.
- D- Keep all files about the same size.
- E- Use file formats that are not splittable

### Answer:

---

B, C, D

## Question 3

---

### Question Type: MultipleChoice

---

An advertising company has a data lake that is built on Amazon S3. The company uses AWS Glue Data Catalog to maintain the metadata

a. The data lake is several years old and its overall size has increased exponentially as additional data sources and metadata are stored in the data lake. The data lake administrator wants to implement a mechanism to simplify permissions management between Amazon S3 and the Data Catalog to keep them in sync

Which solution will simplify permissions management with minimal development effort?

### Options:

---

- A- Set AWS Identity and Access Management (IAM) permissions for AWS Glue
- B- Use AWS Lake Formation permissions
- C- Manage AWS Glue and S3 permissions by using bucket policies
- D- Use Amazon Cognito user pools.

### Answer:

---

B

## Question 4

---

**Question Type:** MultipleChoice

---

A gaming company is collecting clickstream data into multiple Amazon Kinesis data streams. The company uses Amazon Kinesis Data Firehose delivery streams to store the data in JSON format in Amazon S3. Data scientists use Amazon Athena to query the most recent data and derive business insights. The company wants to reduce its Athena costs without having to recreate the data pipeline. The company prefers a solution that will require less management effort.

Which set of actions can the data scientists take immediately to reduce costs?

### Options:

---

- A-** Change the Kinesis Data Firehose output format to Apache Parquet Provide a custom S3 object YYYYMMDD prefix expression and specify a large buffer size For the existing data, run an AWS Glue ETL job to combine and convert small JSON files to large Parquet files and add the YYYYMMDD prefix Use ALTER TABLE ADD PARTITION to reflect the partition on the existing Athena table.
- B-** Create an Apache Spark Job that combines and converts JSON files to Apache Parquet files Launch an Amazon EMR ephemeral cluster daily to run the Spark job to create new Parquet files in a different S3 location Use ALTER TABLE SET LOCATION to reflect the new S3 location on the existing Athena table.
- C-** Create a Kinesis data stream as a delivery target for Kinesis Data Firehose Run Apache Flink on Amazon Kinesis Data Analytics on the stream to read the streaming data, aggregate it and save it to Amazon S3 in Apache Parquet format with a custom S3 object YYYYMMDD prefix Use ALTER TABLE ADD PARTITION to reflect the partition on the existing Athena table
- D-** Integrate an AWS Lambda function with Kinesis Data Firehose to convert source records to Apache Parquet and write them to Amazon S3 In parallel, run an AWS Glue ETL job to combine and convert existing JSON files to large Parquet files Create a custom S3 object YYYYMMDD prefix Use ALTER TABLE ADD PARTITION to reflect the partition on the existing Athena table.

### Answer:

---

D

## Question 5

---

**Question Type:** MultipleChoice

---

A data analyst runs a large number of data manipulation language (DML) queries by using Amazon Athena with the JDBC driver. Recently, a query failed after It ran for 30 minutes. The query returned the following message

Java.sql.SQLException: Query timeout

The data analyst does not immediately need the query results However, the data analyst needs a long-term solution for this problem

Which solution will meet these requirements?

### Options:

---

- A-** Split the query into smaller queries to search smaller subsets of data.
- B-** In the settings for Athena, adjust the DML query timeout limit
- C-** In the Service Quotas console, request an increase for the DML query timeout
- D-** Save the tables as compressed .csv files

### Answer:

---

A

## Question 6

---

**Question Type: MultipleChoice**

---

A company uses Amazon Kinesis Data Streams to ingest and process customer behavior information from application users each day. A data analytics specialist notices that its data stream is throttling. The specialist has turned on enhanced monitoring for the Kinesis data stream and has verified that the data stream did not exceed the data limits. The specialist discovers that there are hot shards.

Which solution will resolve this issue?

**Options:**

---

- A-** Use a random partition key to ingest the records.
- B-** Increase the number of shards. Split the size of the log records.
- C-** Limit the number of records that are sent each second by the producer to match the capacity of the stream.
- D-** Decrease the size of the records that are sent from the producer to match the capacity of the stream.

**Answer:**

---

A

## Question 7

---

**Question Type: MultipleChoice**

---



A software company wants to use instrumentation data to detect and resolve errors to improve application recovery time. The company requires API usage anomalies, like error rate and response time spikes, to be detected in near-real time (NRT) The company also requires that data analysts have access to dashboards for log analysis in NRT

Which solution meets these requirements'?

### Options:

---

**A-** Use Amazon Kinesis Data Firehose as the data transport layer for logging data Use Amazon Kinesis Data Analytics to uncover the NRT API usage anomalies Use Kinesis Data Firehose to deliver log data to Amazon OpenSearch Service (Amazon Elasticsearch Service) for search, log analytics, and application monitoring Use OpenSearch Dashboards (Kibana) in Amazon OpenSearch Service (Amazon Elasticsearch Service) for the dashboards.

**B-** Use Amazon Kinesis Data Analytics as the data transport layer for logging data. Use Amazon Kinesis Data Streams to uncover NRT monitoring metrics. Use Amazon Kinesis Data Firehose to deliver log data to Amazon OpenSearch Service (Amazon Elasticsearch Service) for search, log analytics, and application monitoring Use Amazon QuickSight for the dashboards

**C-** Use Amazon Kinesis Data Analytics as the data transport layer for logging data and to uncover NRT monitoring metrics Use Amazon Kinesis Data Firehose to deliver log data to Amazon OpenSearch Service (Amazon Elasticsearch Service) for search, log analytics, and application monitoring Use OpenSearch Dashboards (Kibana) in Amazon OpenSearch Service (Amazon Elasticsearch Service) for the dashboards

**D-** Use Amazon Kinesis Data Firehose as the data transport layer for logging data Use Amazon Kinesis Data Analytics to uncover NRT monitoring metrics Use Amazon Kinesis Data Streams to deliver log data to Amazon OpenSearch Service (Amazon Elasticsearch Service) for search, log analytics, and application monitoring Use Amazon QuickSight for the dashboards.

**Answer:**

---

C

## Question 8

---

**Question Type:** MultipleChoice

---

A data engineer is using AWS Glue ETL jobs to process data at frequent intervals. The processed data is then copied into Amazon S3. The ETL jobs run every 15 minutes. The AWS Glue Data Catalog partitions need to be updated automatically after the completion of each job.

Which solution will meet these requirements MOST cost-effectively?

**Options:**

---

- A-** Use the AWS Glue Data Catalog to manage the data catalog. Define an AWS Glue workflow for the ETL process. Define a trigger within the workflow that can start the crawler when an ETL job run is complete.
- B-** Use the AWS Glue Data Catalog to manage the data catalog. Use AWS Glue Studio to manage ETL jobs. Use the AWS Glue Studio feature that supports updates to the AWS Glue Data Catalog during job runs.
- C-** Use an Apache Hive metastore to manage the data catalog. Update the AWS Glue ETL code to include the `enableUpdateCatalog` and `partitionKeys` arguments.

**D-** Use the AWS Glue Data Catalog to manage the data catalog Update the AWS Glue ETL code to include the enableUpdateCatalog and partitionKeys arguments.

**Answer:**

---

A

## Question 9

---

**Question Type: MultipleChoice**

---

An ecommerce company ingests a large set of clickstream data in JSON format and stores the data in Amazon S3. Business analysts from multiple product divisions need to use Amazon Athena to analyze the dat

a. The company's analytics team must design a solution to monitor the daily data usage for Athena by each product division. The solution also must produce a warning when a divisions exceeds its quota

Which solution will meet these requirements with the LEAST operational overhead?

**Options:**

---

**A-** Use a CREATE TABLE AS SELECT (CTAS) statement to create separate tables for each product division Use AWS Budgets to track Athena usage Configure a threshold for the budget Use Amazon Simple Notification Service (Amazon SNS) to send notifications when

thresholds are breached.

**B-** Create an AWS account for each division Provide cross-account access to an AWS Glue Data Catalog to all the accounts. Set an Amazon CloudWatch alarm to monitor Athena usage. Use Amazon Simple Notification Service (Amazon SNS) to send notifications.

**C-** Create an Athena workgroup for each division Configure a data usage control for each workgroup and a time period of 1 day Configure an action to send notifications to an Amazon Simple Notification Service (Amazon SNS) topic

**D-** Create an AWS account for each division Configure an AWS Glue Data Catalog in each account Set an Amazon CloudWatch alarm to monitor Athena usage Use Amazon Simple Notification Service (Amazon SNS) to send notifications.

**Answer:**

---

C

## Question 10

---

**Question Type:** MultipleChoice

---

A company is reading data from various customer databases that run on Amazon RDS. The databases contain many inconsistent fields For example, a customer record field that is place\_id in one database is location\_id in another database. The company wants to link customer records across different databases, even when many customer record fields do not match exactly

Which solution will meet these requirements with the LEAST operational overhead?

### Options:

---

- A-** Create an Amazon EMR cluster to process and analyze data in the databases. Connect to the Apache Zeppelin notebook, and use the FindMatches transform to find duplicate records in the data.
- B-** Create an AWS Glue crawler to crawl the databases. Use the FindMatches transform to find duplicate records in the data. Evaluate and tune the transform by evaluating performance and results of finding matches.
- C-** Create an AWS Glue crawler to crawl the data in the databases. Use Amazon SageMaker to construct Apache Spark ML pipelines to find duplicate records in the data.
- D-** Create an Amazon EMR cluster to process and analyze data in the databases. Connect to the Apache Zeppelin notebook, and use Apache Spark ML to find duplicate records in the data. Evaluate and tune the model by evaluating performance and results of finding duplicates.

### Answer:

---

B

## Question 11

---

### Question Type: MultipleChoice

---

A company using Amazon QuickSight Enterprise edition has thousands of dashboards, analyses, and datasets. The company struggles to manage and assign permissions for granting users access to various items within QuickSight. The company wants to make it easier to

implement sharing and permissions management.

Which solution should the company implement to simplify permissions management?

**Options:**

---

- A-** Use QuickSight folders to organize dashboards, analyses, and datasets Assign individual users permissions to these folders
- B-** Use QuickSight folders to organize dashboards analyses, and datasets Assign group permissions by using these folders.
- C-** Use AWS IAM resource-based policies to assign group permissions to QuickSight items
- D-** Use QuickSight user management APIs to provision group permissions based on dashboard naming conventions

**Answer:**

---

C

**To Get Premium Files for DAS-C01 Visit**

**<https://www.p2pexams.com/products/das-c01>**

**For More Free Questions Visit**

**<https://www.p2pexams.com/amazon/pdf/das-c01>**

