DBCLOUD
EXAMS

QUESTIONS & ANSWERS

verified



Technology Provider: Amazon
Exam: AWS Certified Data Engineer - Associate DEA-C01
Number of questions in the database: 377
Exam Version: NOV.2026

Topic 1 - Exam A

Question #1*Topic 1*

A data engineer is configuring an AWS Glue job to read data from an Amazon S3 bucket. The data engineer has set up the necessary AWS Glue connection details and an associated IAM role. However, when the data engineer attempts to run the AWS Glue job, the data engineer receives an error message that indicates that there are problems with the Amazon S3 VPC gateway endpoint. The data engineer must resolve the error and connect the AWS Glue job to the S3 bucket.
Which solution will meet this requirement?

- A. Update the AWS Glue security group to allow inbound traffic from the Amazon S3 VPC gateway endpoint.
- B. Configure an S3 bucket policy to explicitly grant the AWS Glue job permissions to access the S3 bucket.
- C. Review the AWS Glue job code to ensure that the AWS Glue connection details include a fully qualified domain name.
- D. Verify that the VPC's route table includes inbound and outbound routes for the Amazon S3 VPC gateway endpoint.

Correct Answer: *D*

Respuesta por mayor votación de la comunidad

D (88%)

6%

Question #2*Topic 1*

A retail company has a customer data hub in an Amazon S3 bucket. Employees from many countries use the data hub to support company-wide analytics. A governance team must ensure that the company's data analysts can access data only for customers who are within the same country as the analysts.
Which solution will meet these requirements with the LEAST operational effort?

- A. Create a separate table for each country's customer data. Provide access to each analyst based on the country that the analyst serves.
- B. Register the S3 bucket as a data lake location in AWS Lake Formation. Use the Lake Formation row-level security features to enforce the company's access policies.

https://dbcloud.academy/

- C. Move the data to AWS Regions that are close to the countries where the customers are. Provide access to each analyst based on the country that the analyst serves.
- D. Load the data into Amazon Redshift. Create a view for each country. Create separate IAM roles for each country to provide access to data from each country. Assign the appropriate roles to the analysts.

Correct Answer: *B*

Respuesta por mayor votación de la comunidad

B (92%)

8%

Question #3*Topic 1*

A media company wants to improve a system that recommends media content to customer based on user behavior and preferences. To improve the recommendation system, the company needs to incorporate insights from third-party datasets into the company's existing analytics platform.
The company wants to minimize the effort and time required to incorporate third-party datasets.
Which solution will meet these requirements with the LEAST operational overhead?

- A. Use API calls to access and integrate third-party datasets from AWS Data Exchange.
- B. Use API calls to access and integrate third-party datasets from AWS DataSync.
- C. Use Amazon Kinesis Data Streams to access and integrate third-party datasets from AWS CodeCommit repositories.
- D. Use Amazon Kinesis Data Streams to access and integrate third-party datasets from Amazon Elastic Container Registry (Amazon ECR).

Correct Answer: *A*

Respuesta por mayor votación de la comunidad

A (94%)

6%

Question #4*Topic 1*

A financial company wants to implement a data mesh. The data mesh must support centralized data governance, data analysis, and data access control. The company has decided to use AWS Glue for data catalogs and extract, transform, and load (ETL) operations.
Which combination of AWS services will implement a data mesh? (Choose two.)

- A. Use Amazon Aurora for data storage. Use an Amazon Redshift provisioned cluster for data analysis.
- B. Use Amazon S3 for data storage. Use Amazon Athena for data analysis.
- C. Use AWS Glue DataBrew for centralized data governance and access control.

https://dbcloud.academy/

- D. Use Amazon RDS for data storage. Use Amazon EMR for data analysis.
- E. Use AWS Lake Formation for centralized data governance and access control.

Correct Answer: *BE*

Respuesta por mayor votación de la comunidad

BE (100%)

Question #5*Topic 1*

A data engineer maintains custom Python scripts that perform a data formatting process that many AWS Lambda functions use. When the data engineer needs to modify the Python scripts, the data engineer must manually update all the Lambda functions.
The data engineer requires a less manual way to update the Lambda functions.
Which solution will meet this requirement?

- A. Store a pointer to the custom Python scripts in the execution context object in a shared Amazon S3 bucket.
- B. Package the custom Python scripts into Lambda layers. Apply the Lambda layers to the Lambda functions.
- C. Store a pointer to the custom Python scripts in environment variables in a shared Amazon S3 bucket.
- D. Assign the same alias to each Lambda function. Call reach Lambda function by specifying the function's alias.

Correct Answer: *B*

Respuesta por mayor votación de la comunidad

B (100%)

Question #6*Topic 1*

A company created an extract, transform, and load (ETL) data pipeline in AWS Glue. A data engineer must crawl a table that is in Microsoft SQL Server. The data engineer needs to extract, transform, and load the output of the crawl to an Amazon S3 bucket. The data engineer also must orchestrate the data pipeline.
Which AWS service or feature will meet these requirements MOST cost-effectively?

- A. AWS Step Functions
- B. AWS Glue workflows
- C. AWS Glue Studio
- D. Amazon Managed Workflows for Apache Airflow (Amazon MWAA)

Correct Answer: *B*

https://dbcloud.academy/

Respuesta por mayor votación de la comunidad

B (100%)

Question #7*Topic 1*

A financial services company stores financial data in Amazon Redshift. A data engineer wants to run real-time queries on the financial data to support a web-based trading application. The data engineer wants to run the queries from within the trading application.
Which solution will meet these requirements with the LEAST operational overhead?

- A. Establish WebSocket connections to Amazon Redshift.
- B. Use the Amazon Redshift Data API.
- C. Set up Java Database Connectivity (JDBC) connections to Amazon Redshift.
- D. Store frequently accessed data in Amazon S3. Use Amazon S3 Select to run the queries.

Correct Answer: *B*

Respuesta por mayor votación de la comunidad

B (100%)

Question #8*Topic 1*

A company uses Amazon Athena for one-time queries against data that is in Amazon S3. The company has several use cases. The company must implement permission controls to separate query processes and access to query history among users, teams, and applications that are in the same AWS account.
Which solution will meet these requirements?

- A. Create an S3 bucket for each use case. Create an S3 bucket policy that grants permissions to appropriate individual IAM users. Apply the S3 bucket policy to the S3 bucket.
- B. Create an Athena workgroup for each use case. Apply tags to the workgroup. Create an IAM policy that uses the tags to apply appropriate permissions to the workgroup.
- C. Create an IAM role for each use case. Assign appropriate permissions to the role for each use case. Associate the role with Athena.
- D. Create an AWS Glue Data Catalog resource policy that grants permissions to appropriate individual IAM users for each use case. Apply the resource policy to the specific tables that Athena uses.

Correct Answer: *B*

Respuesta por mayor votación de la comunidad

B (100%)

https://dbcloud.academy/