

# **2026 Data-Engineer-Associate Exam Sample Online: AWS Certified Data Engineer - Associate (DEA-C01) - The Best Amazon Data-Engineer-Associate Standard Answers**



BTW, DOWNLOAD part of PDFTorrent Data-Engineer-Associate dumps from Cloud Storage: <https://drive.google.com/open?id=1DJGpc0UHnd0rqzYDjTjGeFvohHLBfs4A>

PDFTorrent's Data-Engineer-Associate exam dumps comprise a brief and succinct set of exam questions that provides authentic, updated and the most relevant information on each syllabus contents that may be the part of your Data-Engineer-Associate exam paper. The Data-Engineer-Associate dumps have been verified and approved by the skilled professional. Hence, there is no question of irrelevant or substandard information. The feedback of our customers evaluates Data-Engineer-Associate Brain Dumps as the top dumps that helped their overcome all their exam worries rather enabled them to ace it with brilliant success.

Are you worried about where to find reliable and valid Data-Engineer-Associate practice exam cram? Please stop hunting with aimless, Amazon Data-Engineer-Associate free study dumps will help you and solve your problems. If you still have doubts, you can download Data-Engineer-Associate free demo to have a try. If you have any questions about Data-Engineer-Associate Study Tool, please contact us by email or chat with our online customer service, we will always here to answers your questions. Our Data-Engineer-Associate test practice will enhance your professional skills and expand your knowledge, which will ensure you a define success in our Data-Engineer-Associate actual test.

**>> Data-Engineer-Associate Exam Sample Online <<**

## **Data-Engineer-Associate Exam Torrent: AWS Certified Data Engineer - Associate (DEA-C01) - Data-Engineer-Associate Prep Torrent & Data-Engineer-Associate Test Braindumps**

The PDF version of the PDFTorrent Amazon Data-Engineer-Associate prep material is easily accessible. This format is ideal for someone who is constantly on the move, as you can prepare for your AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) exam whether you are using your smartphone, tablet, or laptop. You can study anywhere, at any time, without

having to worry about installing anything. Furthermore, you can study with a hard copy by printing all of your AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) PDF questions. We offer regular updates in PDF format to improve AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) questions according to changes in the exam.

## Amazon AWS Certified Data Engineer - Associate (DEA-C01) Sample Questions (Q211-Q216):

### NEW QUESTION # 211

A company stores daily records of the financial performance of investment portfolios in .csv format in an Amazon S3 bucket. A data engineer uses AWS Glue crawlers to crawl the S3 data.

The data engineer must make the S3 data accessible daily in the AWS Glue Data Catalog.

Which solution will meet these requirements?

- A. Create an IAM role that includes the AmazonS3FullAccess policy. Associate the role with the crawler. Specify the S3 bucket path of the source data as the crawler's data store. Allocate data processing units (DPUs) to run the crawler every day. Specify a database name for the output.
- B. Create an IAM role that includes the AmazonS3FullAccess policy. Associate the role with the crawler. Specify the S3 bucket path of the source data as the crawler's data store. Create a daily schedule to run the crawler. Configure the output destination to a new path in the existing S3 bucket.
- C. Create an IAM role that includes the AWSGlueServiceRole policy. Associate the role with the crawler. Specify the S3 bucket path of the source data as the crawler's data store. Allocate data processing units (DPUs) to run the crawler every day. Configure the output destination to a new path in the existing S3 bucket.
- **D. Create an IAM role that includes the AWSGlueServiceRole policy. Associate the role with the crawler. Specify the S3 bucket path of the source data as the crawler's data store. Create a daily schedule to run the crawler. Specify a database name for the output.**

**Answer: D**

Explanation:

To make the S3 data accessible daily in the AWS Glue Data Catalog, the data engineer needs to create a crawler that can crawl the S3 data and write the metadata to the Data Catalog. The crawler also needs to run on a daily schedule to keep the Data Catalog updated with the latest data. Therefore, the solution must include the following steps:

Create an IAM role that has the necessary permissions to access the S3 data and the Data Catalog. The AWSGlueServiceRole policy is a managed policy that grants these permissions<sup>1</sup>.

Associate the role with the crawler.

Specify the S3 bucket path of the source data as the crawler's data store. The crawler will scan the data and infer the schema and format<sup>2</sup>.

Create a daily schedule to run the crawler. The crawler will run at the specified time every day and update the Data Catalog with any changes in the data<sup>3</sup>.

Specify a database name for the output. The crawler will create or update a table in the Data Catalog under the specified database. The table will contain the metadata about the data in the S3 bucket, such as the location, schema, and classification.

Option B is the only solution that includes all these steps. Therefore, option B is the correct answer.

Option A is incorrect because it configures the output destination to a new path in the existing S3 bucket. This is unnecessary and may cause confusion, as the crawler does not write any data to the S3 bucket, only metadata to the Data Catalog.

Option C is incorrect because it allocates data processing units (DPUs) to run the crawler every day. This is also unnecessary, as DPUs are only used for AWS Glue ETL jobs, not crawlers.

Option D is incorrect because it combines the errors of option A and C. It configures the output destination to a new path in the existing S3 bucket and allocates DPUs to run the crawler every day, both of which are irrelevant for the crawler.

Reference:

1: AWS managed (predefined) policies for AWS Glue - AWS Glue

2: Data Catalog and crawlers in AWS Glue - AWS Glue

3: Scheduling an AWS Glue crawler - AWS Glue

[4]: Parameters set on Data Catalog tables by crawler - AWS Glue

[5]: AWS Glue pricing - Amazon Web Services (AWS)

### NEW QUESTION # 212

A data engineer needs to create a new empty table in Amazon Athena that has the same schema as an existing table named old-table.

Which SQL statement should the data engineer use to meet this requirement?

- A.
- B.
- C.
- D.

**Answer: B**

Explanation:

\* Problem Analysis:

\* The goal is to create a new empty table in Athena with the same schema as an existing table (old\_table).

\* The solution must avoid copying any data.

\* Key Considerations:

\* CREATE TABLE AS (CTAS) is commonly used in Athena for creating new tables based on an existing table.

\* Adding the WITH NO DATA clause ensures only the schema is copied, without transferring any data.

\* Solution Analysis:

\* Option A: Copies both schema and data. Does not meet the requirement for an empty table.

\* Option B: Inserts data into an existing table, which does not create a new table.

\* Option C: Creates an empty table but does not copy the schema.

\* Option D: Creates a new table with the same schema and ensures it is empty by using WITH NO DATA.

\* Final Recommendation:

\* Use D. CREATE TABLE new\_table AS (SELECT \* FROM old\_table) WITH NO DATA to create an empty table with the same schema.

:

Athena CTAS Queries

CREATE TABLE Statement in Athena

### NEW QUESTION # 213

A data engineer wants to orchestrate a set of extract, transform, and load (ETL) jobs that run on AWS. The ETL jobs contain tasks that must run Apache Spark jobs on Amazon EMR, make API calls to Salesforce, and load data into Amazon Redshift.

The ETL jobs need to handle failures and retries automatically. The data engineer needs to use Python to orchestrate the jobs.

Which service will meet these requirements?

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

**Answer: A**

Explanation:

The data engineer needs to orchestrate ETL jobs that include Spark jobs on Amazon EMR, API calls to Salesforce, and loading data into Redshift. They also need automatic failure handling and retries. Amazon Managed Workflows for Apache Airflow (Amazon MWAA) is the best solution for this requirement.

Option A: Amazon Managed Workflows for Apache Airflow (Amazon MWAA)

Apache Airflow is designed for complex job orchestration, allowing users to define workflows (DAGs) in Python. MWAA manages Airflow and its integrations with other AWS services, including Amazon EMR, Redshift, and external APIs like Salesforce. It provides automatic retry handling, failure detection, and detailed monitoring, which fits the use case perfectly.

Option B (AWS Step Functions) can orchestrate tasks but doesn't natively support complex workflow definitions with Python like Airflow does.

Option C (AWS Glue) is more focused on ETL and doesn't handle the orchestration of external systems like Salesforce as well as Airflow.

Option D (Amazon EventBridge) is more suited for event-driven architectures rather than complex workflow orchestration.

Reference:

Amazon Managed Workflows for Apache Airflow

Apache Airflow on AWS

### NEW QUESTION # 214

A company needs to set up a data catalog and metadata management for data sources that run in the AWS Cloud. The company will use the data catalog to maintain the metadata of all the objects that are in a set of data stores. The data stores include structured

sources such as Amazon RDS and Amazon Redshift. The data stores also include semistructured sources such as JSON files and .xml files that are stored in Amazon S3.

The company needs a solution that will update the data catalog on a regular basis. The solution also must detect changes to the source metadata.

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

- A. Use Amazon DynamoDB as the data catalog. Create AWS Lambda functions that will connect to the data catalog. Configure the Lambda functions to gather the metadata information from multiple sources and to update the DynamoDB data catalog. Schedule the Lambda functions to run periodically.
- B. Use Amazon Aurora as the data catalog. Create AWS Lambda functions that will connect to the data catalog. Configure the Lambda functions to gather the metadata information from multiple sources and to update the Aurora data catalog. Schedule the Lambda functions to run periodically.
- C. Use the AWS Glue Data Catalog as the central metadata repository. Extract the schema for Amazon RDS and Amazon Redshift sources, and build the Data Catalog. Use AWS Glue crawlers for data that is in Amazon S3 to infer the schema and to automatically update the Data Catalog.
- **D. Use the AWS Glue Data Catalog as the central metadata repository. Use AWS Glue crawlers to connect to multiple data stores and to update the Data Catalog with metadata changes. Schedule the crawlers to run periodically to update the metadata catalog.**

**Answer: D**

Explanation:

This solution will meet the requirements with the least operational overhead because it uses the AWS Glue Data Catalog as the central metadata repository for data sources that run in the AWS Cloud. The AWS Glue Data Catalog is a fully managed service that provides a unified view of your data assets across AWS and on-premises data sources. It stores the metadata of your data in tables, partitions, and columns, and enables you to access and query your data using various AWS services, such as Amazon Athena, Amazon EMR, and Amazon Redshift Spectrum. You can use AWS Glue crawlers to connect to multiple data stores, such as Amazon RDS, Amazon Redshift, and Amazon S3, and to update the Data Catalog with metadata changes.

AWS Glue crawlers can automatically discover the schema and partition structure of your data, and create or update the corresponding tables in the Data Catalog. You can schedule the crawlers to run periodically to update the metadata catalog, and configure them to detect changes to the source metadata, such as new columns, tables, or partitions<sup>12</sup>.

The other options are not optimal for the following reasons:

A: Use Amazon Aurora as the data catalog. Create AWS Lambda functions that will connect to the data catalog. Configure the Lambda functions to gather the metadata information from multiple sources and to update the Aurora data catalog. Schedule the Lambda functions to run periodically. This option is not recommended, as it would require more operational overhead to create and manage an Amazon Aurora database as the data catalog, and to write and maintain AWS Lambda functions to gather and update the metadata information from multiple sources. Moreover, this option would not leverage the benefits of the AWS Glue Data Catalog, such as data cataloging, data transformation, and data governance.

C: Use Amazon DynamoDB as the data catalog. Create AWS Lambda functions that will connect to the data catalog. Configure the Lambda functions to gather the metadata information from multiple sources and to update the DynamoDB data catalog. Schedule the Lambda functions to run periodically. This option is also not recommended, as it would require more operational overhead to create and manage an Amazon DynamoDB table as the data catalog, and to write and maintain AWS Lambda functions to gather and update the metadata information from multiple sources. Moreover, this option would not leverage the benefits of the AWS Glue Data Catalog, such as data cataloging, data transformation, and data governance.

D: Use the AWS Glue Data Catalog as the central metadata repository. Extract the schema for Amazon RDS and Amazon Redshift sources, and build the Data Catalog. Use AWS Glue crawlers for data that is in Amazon S3 to infer the schema and to automatically update the Data Catalog. This option is not optimal, as it would require more manual effort to extract the schema for Amazon RDS and Amazon Redshift sources, and to build the Data Catalog. This option would not take advantage of the AWS Glue crawlers' ability to automatically discover the schema and partition structure of your data from various data sources, and to create or update the corresponding tables in the Data Catalog.

References:

- 1: AWS Glue Data Catalog
- 2: AWS Glue Crawlers
- 3: Amazon Aurora
- 4: AWS Lambda
- 5: Amazon DynamoDB

## NEW QUESTION # 215

A company receives a data file from a partner each day in an Amazon S3 bucket. The company uses a daily AWS Glue extract, transform, and load (ETL) pipeline to clean and transform each data file. The output of the ETL pipeline is written to a CSV file

named Dairy.csv in a second S3 bucket.

Occasionally, the daily data file is empty or is missing values for required fields. When the file is missing data, the company can use the previous day's CSV file.

A data engineer needs to ensure that the previous day's data file is overwritten only if the new daily file is complete and valid. Which solution will meet these requirements with the LEAST effort?

- A. Invoke an AWS Lambda function to check the file for missing data and to fill in missing values in required fields.
- B. Use AWS Glue Studio to change the code in the ETL pipeline to fill in any missing values in the required fields with the most common values for each field.
- C. Run a SQL query in Amazon Athena to read the CSV file and drop missing rows. Copy the corrected CSV file to the second S3 bucket.
- **D. Configure the AWS Glue ETL pipeline to use AWS Glue Data Quality rules. Develop rules in Data Quality Definition Language (DQDL) to check for missing values in required files and empty files.**

**Answer: D**

Explanation:

Problem Analysis:

The company runs a daily AWS Glue ETL pipeline to clean and transform files received in an S3 bucket.

If a file is incomplete or empty, the previous day's file should be retained.

Need a solution to validate files before overwriting the existing file.

Key Considerations:

Automate data validation with minimal human intervention.

Use built-in AWS Glue capabilities for ease of integration.

Ensure robust validation for missing or incomplete data.

Solution Analysis:

Option A: Lambda Function for Validation

Lambda can validate files, but it would require custom code.

Does not leverage AWS Glue's built-in features, adding operational complexity.

Option B: AWS Glue Data Quality Rules

AWS Glue Data Quality allows defining Data Quality Definition Language (DQDL) rules.

Rules can validate if required fields are missing or if the file is empty.

Automatically integrates into the existing ETL pipeline.

If validation fails, retain the previous day's file.

Option C: AWS Glue Studio with Filling Missing Values

Modifying ETL code to fill missing values with most common values risks introducing inaccuracies.

Does not handle empty files effectively.

Option D: Athena Query for Validation

Athena can drop rows with missing values, but this is a post-hoc solution.

Requires manual intervention to copy the corrected file to S3, increasing complexity.

Final Recommendation:

Use AWS Glue Data Quality to define validation rules in DQDL for identifying missing or incomplete data.

This solution integrates seamlessly with the ETL pipeline and minimizes manual effort.

Implementation Steps:

Enable AWS Glue Data Quality in the existing ETL pipeline.

Define DQDL Rules, such as:

Check if a file is empty.

Verify required fields are present and non-null.

Configure the pipeline to proceed with overwriting only if the file passes validation.

In case of failure, retain the previous day's file.

AWS Glue Data Quality Overview

Defining DQDL Rules

AWS Glue Studio Documentation

## NEW QUESTION # 216

.....

You feel tired when you are preparing hard for Amazon Data-Engineer-Associate exam, do you know what other candidates are doing? Look at the candidates in IT certification exam around you. Why are they confident when you are nervous about the exam? Is your ability below theirs? Of course not. Have you wondered why other IT people can easily pass Amazon Data-Engineer-

Associate test? The answer is to use PDFTorrent Amazon Data-Engineer-Associate questions and answers which can help you sail through the exam with no mistakes. Don't believe it? Do you feel it is amazing? Have a try. You can confirm quality of the exam dumps by experiencing free demo. Hurry up and click PDFTorrent.com

**Data-Engineer-Associate Standard Answers:** <https://www.pdf torrent.com/Data-Engineer-Associate-exam-prep-dumps.html>

Our company has established a long-term partnership with those who have purchased our Data-Engineer-Associate exam questions, It can be said that choosing Data-Engineer-Associate study engine is your first step to pass the exam, Amazon Data-Engineer-Associate Exam Sample Online Among which the PDF version is the most popular one, because it is universally acknowledged that the PDF version is convenient for you to read as well as printing. Of course, there are many benefits after you have chosen our Data-Engineer-Associate exam prep, such as good jobs, high salaries and promising futures, to name but a few.

Following us, find out the secrets how you can pass your exam or test with our Data-Engineer-Associate PDF & test engine dumps and become specialized experts in this area, Opening Existing Documents.

Our company has established a long-term partnership with those who have purchased our Data-Engineer-Associate Exam Questions, It can be said that choosing Data-Engineer-Associate study engine is your first step to pass the exam.

## High Pass-Rate Amazon Data-Engineer-Associate Exam Sample Online Offer You The Best Standard Answers | AWS Certified Data Engineer - Associate (DEA-C01)

Among which the PDF version is the most popular one, because Data-Engineer-Associate it is universally acknowledged that the PDF version is convenient for you to read as well as printing.

Of course, there are many benefits after you have chosen our Data-Engineer-Associate exam prep, such as good jobs, high salaries and promising futures, to name but a few, Provides you a guaranteed pathway of exam success!

- Data-Engineer-Associate 100% Accuracy  Data-Engineer-Associate Download Pdf  Test Data-Engineer-Associate Result  Easily obtain free download of 「 Data-Engineer-Associate 」 by searching on 「 [www.dumpsquestion.com](http://www.dumpsquestion.com) 」  Data-Engineer-Associate 100% Accuracy
- Trustable Data-Engineer-Associate Exam Sample Online - Win Your Amazon Certificate with Top Score  Download ▶ Data-Engineer-Associate ◀ for free by simply entering ✨: [www.pdfvce.com](http://www.pdfvce.com)  ✨  website  Valid Data-Engineer-Associate Test Notes
- New Data-Engineer-Associate Exam Questions  New Data-Engineer-Associate Exam Questions  Valid Test Data-Engineer-Associate Fee  Download  Data-Engineer-Associate  for free by simply searching on ✨: [www.examdiscuss.com](http://www.examdiscuss.com)  ✨  Free Data-Engineer-Associate Download
- Data-Engineer-Associate Download Pdf  New Data-Engineer-Associate Exam Questions  Data-Engineer-Associate Latest Dumps Sheet  Simply search for ⇒ Data-Engineer-Associate ⇐ for free download on 【 [www.pdfvce.com](http://www.pdfvce.com) 】  Reliable Data-Engineer-Associate Exam Book
- Free PDF 2026 Amazon High Pass-Rate Data-Engineer-Associate Exam Sample Online  Search for { Data-Engineer-Associate } and easily obtain a free download on 「 [www.dumpsmaterials.com](http://www.dumpsmaterials.com) 」  Data-Engineer-Associate Valid Test Cram
- Top Data-Engineer-Associate Dumps  New Data-Engineer-Associate Exam Questions  Data-Engineer-Associate Valid Test Cram  The page for free download of { Data-Engineer-Associate } on  [www.pdfvce.com](http://www.pdfvce.com)  will open immediately  Valid Data-Engineer-Associate Test Notes
- 2026 Data-Engineer-Associate Exam Sample Online | Trustable AWS Certified Data Engineer - Associate (DEA-C01) 100% Free Standard Answers  Search for ▷ Data-Engineer-Associate ◁ and obtain a free download on ✓ [www.pass4test.com](http://www.pass4test.com)  ✓  Reliable Data-Engineer-Associate Test Tips
- Pass Guaranteed Quiz Data-Engineer-Associate - High-quality AWS Certified Data Engineer - Associate (DEA-C01) Exam Sample Online  ✨: [www.pdfvce.com](http://www.pdfvce.com)  ✨  is best website to obtain { Data-Engineer-Associate } for free download  Customizable Data-Engineer-Associate Exam Mode
- Data-Engineer-Associate Valid Test Cram  Data-Engineer-Associate Valid Exam Experience  Valid Data-Engineer-Associate Test Notes  Easily obtain  Data-Engineer-Associate  for free download through  [www.easy4engine.com](http://www.easy4engine.com)   Data-Engineer-Associate Valid Test Cram
- Customizable Data-Engineer-Associate Exam Mode  Data-Engineer-Associate Valid Exam Experience  New Data-Engineer-Associate Exam Questions  Search for ✓ Data-Engineer-Associate  ✓  and easily obtain a free download on ➤ [www.pdfvce.com](http://www.pdfvce.com)   Valid Test Data-Engineer-Associate Fee
- New Data-Engineer-Associate Exam Questions  Data-Engineer-Associate Well Prep  Data-Engineer-Associate Download Pdf  Search for { Data-Engineer-Associate } and obtain a free download on 《 [www.pdfdumps.com](http://www.pdfdumps.com) 》   Test Data-Engineer-Associate Result

