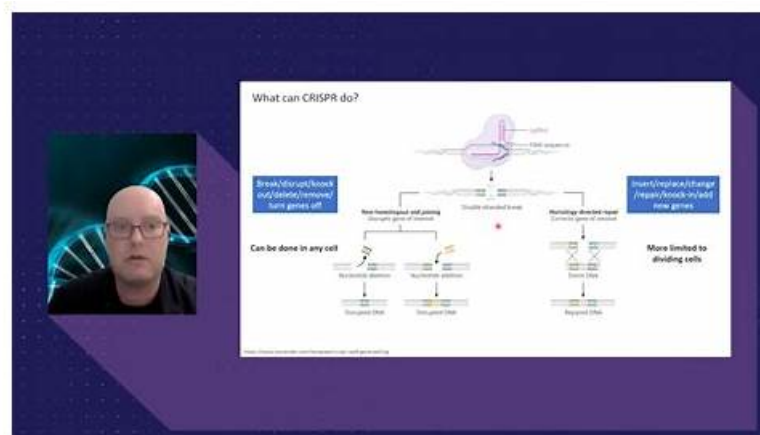


# Data-Engineer-Associate Vce Format | Data-Engineer-Associate Guide



2026 Latest PDFVCE Data-Engineer-Associate PDF Dumps and Data-Engineer-Associate Exam Engine Free Share:  
<https://drive.google.com/open?id=1DxPC6g7wRoNpUnqQGExMPHEaG4LwIZQF>

Now we can say that AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) exam questions are real and top-notch Data-Engineer-Associate exam questions that you can expect in the upcoming AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) exam. In this way, you can easily pass the Data-Engineer-Associate exam with good scores. The countless Data-Engineer-Associate Exam candidates have passed their dream Amazon Data-Engineer-Associate certification exam and they all got help from real, valid, and updated Data-Engineer-Associate practice questions, You can also trust on PDFVCE and start preparation with confidence.

It is important to check the exercises and find the problems. Once you use our Data-Engineer-Associate study prep to aid your preparation of the exam, all of your exercises of the study materials will be carefully recorded on the system of the Data-Engineer-Associate exam braindump. Also, you can know your current learning condition clearly. The results will display your final scores on the screen. Also, you will know the numbers of correct and false questions of your exercise. Our Data-Engineer-Associate Certification Materials can help you transfer into a versatile talent. Many job seekers have successfully realized financial freedom with the assistance of our Data-Engineer-Associate test training. All your dreams will be fully realized after you have obtained the Data-Engineer-Associate certificate. Finding a good paying job is available for you.

>> Data-Engineer-Associate Vce Format <<

## Data-Engineer-Associate Guide - Test Data-Engineer-Associate Questions

For candidates who are searching for Data-Engineer-Associate training materials for the exam, the quality of the Data-Engineer-Associate exam dumps must be your first concern. Our Data-Engineer-Associate exam materials can reach this requirement. With a professional team to collect the first-hand information of the exam, we can ensure you that the Data-Engineer-Associate Exam Dumps you receive are the latest information for the exam. Moreover, we also pass guarantee and money back guarantee, if you fail to pass the exam, we will refund your money, and no other questions will be asked.

## Amazon AWS Certified Data Engineer - Associate (DEA-C01) Sample Questions (Q29-Q34):

### NEW QUESTION # 29

A data engineer develops an AWS Glue Apache Spark ETL job to perform transformations on a dataset.

When the data engineer runs the job, the job returns an error that reads, "No space left on device." The data engineer needs to identify the source of the error and provide a solution.

Which combinations of steps will meet this requirement MOST cost-effectively? (Select TWO.)

- A. Use the Spark UI and AWS Glue metrics to monitor data skew in the Spark executors.
- B. Enable the `--write-shuffle-files-to-s3` job parameter. Use the salting technique.
- C. Scale out the number of workers horizontally to address data skewness.

- D. Use error logs in Amazon CloudWatch to monitor data skew.
- E. Scale out the workers vertically to address data skewness.

**Answer: A,B**

Explanation:

A "No space left on device" error typically results from data skew or large shuffle stages. The best actions are:

- \* B. Monitor using Spark UI and Glue metrics to find skewed partitions or executor issues.
- \* D. Use `--write-shuffle-files-to-s3` to offload intermediate data to S3 instead of local disk, and apply salting to reduce skew.

"You can reduce the impact of data skew and large shuffle operations by monitoring with Spark UI and enabling the `--write-shuffle-files-to-s3` option. Salting can help rebalance the skewed keys."

-Ace the AWS Certified Data Engineer - Associate Certification - version 2 - apple.pdf Scaling out workers (A, C) is more costly and less efficient if the root cause (skew) is not fixed.

### NEW QUESTION # 30

A company has multiple applications that use datasets that are stored in an Amazon S3 bucket. The company has an ecommerce application that generates a dataset that contains personally identifiable information (PII).

The company has an internal analytics application that does not require access to the PII.

To comply with regulations, the company must not share PII unnecessarily. A data engineer needs to implement a solution that with redact PII dynamically, based on the needs of each application that accesses the dataset.

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

- A. Use AWS Glue to transform the data for each application. Create multiple copies of the dataset. Give each dataset copy the appropriate level of redaction for the needs of the application that accesses the copy.
- B. Create an API Gateway endpoint that has custom authorizers. Use the API Gateway endpoint to read data from the S3 bucket. Initiate a REST API call to dynamically redact PII based on the needs of each application that accesses the data.
- C. Create an S3 bucket policy to limit the access each application has. Create multiple copies of the dataset. Give each dataset copy the appropriate level of redaction for the needs of the application that accesses the copy.
- **D. Create an S3 Object Lambda endpoint. Use the S3 Object Lambda endpoint to read data from the S3 bucket. Implement redaction logic within an S3 Object Lambda function to dynamically redact PII based on the needs of each application that accesses the data.**

**Answer: D**

Explanation:

Option B is the best solution to meet the requirements with the least operational overhead because S3 Object Lambda is a feature that allows you to add your own code to process data retrieved from S3 before returning it to an application. S3 Object Lambda works with S3 GET requests and can modify both the object metadata and the object data. By using S3 Object Lambda, you can implement redaction logic within an S3 Object Lambda function to dynamically redact PII based on the needs of each application that accesses the data. This way, you can avoid creating and maintaining multiple copies of the dataset with different levels of redaction.

Option A is not a good solution because it involves creating and managing multiple copies of the dataset with different levels of redaction for each application. This option adds complexity and storage cost to the data protection process and requires additional resources and configuration. Moreover, S3 bucket policies cannot enforce fine-grained data access control at the row and column level, so they are not sufficient to redact PII.

Option C is not a good solution because it involves using AWS Glue to transform the data for each application. AWS Glue is a fully managed service that can extract, transform, and load (ETL) data from various sources to various destinations, including S3. AWS Glue can also convert data to different formats, such as Parquet, which is a columnar storage format that is optimized for analytics. However, in this scenario, using AWS Glue to redact PII is not the best option because it requires creating and maintaining multiple copies of the dataset with different levels of redaction for each application. This option also adds extra time and cost to the data protection process and requires additional resources and configuration.

Option D is not a good solution because it involves creating and configuring an API Gateway endpoint that has custom authorizers. API Gateway is a service that allows you to create, publish, maintain, monitor, and secure APIs at any scale. API Gateway can also integrate with other AWS services, such as Lambda, to provide custom logic for processing requests. However, in this scenario, using API Gateway to redact PII is not the best option because it requires writing and maintaining custom code and configuration for the API endpoint, the custom authorizers, and the REST API call. This option also adds complexity and latency to the data protection process and requires additional resources and configuration.

References:

AWS Certified Data Engineer - Associate DEA-C01 Complete Study Guide

Introducing Amazon S3 Object Lambda - Use Your Code to Process Data as It Is Being Retrieved from S3 Using Bucket Policies

### NEW QUESTION # 31

A data engineer needs to debug an AWS Glue job that reads from Amazon S3 and writes to Amazon Redshift.

The data engineer enabled the bookmark feature for the AWS Glue job. The data engineer has set the maximum concurrency for the AWS Glue job to 1.

The AWS Glue job is successfully writing the output to Amazon Redshift. However, the Amazon S3 files that were loaded during previous runs of the AWS Glue job are being reprocessed by subsequent runs.

What is the likely reason the AWS Glue job is reprocessing the files?

- A. The maximum concurrency for the AWS Glue job is set to 1.
- B. The data engineer incorrectly specified an older version of AWS Glue for the Glue job.
- **C. The AWS Glue job does not have the s3:GetObjectAcl permission that is required for bookmarks to work correctly.**
- D. The AWS Glue job does not have a required commit statement.

**Answer: C**

Explanation:

The issue described is that the AWS Glue job is reprocessing files from previous runs despite the bookmark feature being enabled. Bookmarks in AWS Glue allow jobs to keep track of which files or data have already been processed to avoid reprocessing. The most likely reason for reprocessing the files is missing S3 permissions, specifically s3 s3 is a permission required by AWS Glue when bookmarks are enabled to ensure Glue can retrieve metadata from the files in S3, which is necessary for the bookmark mechanism to function correctly. Without this permission, Glue cannot track which files have been processed, resulting in reprocessing during subsequent runs.

Concurrency settings (Option B) and the version of AWS Glue (Option C) do not affect the bookmark behavior. Similarly, the lack of a commit statement (Option D) is not applicable in this context, as Glue handles commits internally when interacting with Redshift and S3.

Thus, the root cause is likely related to insufficient permissions on the S3 bucket, specifically s3 , which is required for bookmarks to work as expected.

References:

AWS Glue Job Bookmarks Documentation

AWS Glue Permissions for Bookmarks

### NEW QUESTION # 32

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 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.
- B. 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.
- **C. 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.**
- 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.

**Answer: C**

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

\* : Amazon Aurora

\* : AWS Lambda

\* : Amazon DynamoDB

### NEW QUESTION # 33

A company receives .csv files that contain physical address data. The data is in columns that have the following names: Door\_No, Street\_Name, City, and Zip\_Code. The company wants to create a single column to store these values in the following format:

□ Which solution will meet this requirement with the LEAST coding effort?

- A. Use AWS Glue DataBrew to read the files. Use the NEST TO ARRAY transformation to create the new column.
- **B. Use AWS Glue DataBrew to read the files. Use the NEST TO MAP transformation to create the new column.**
- C. Write a Lambda function in Python to read the files. Use the Python data dictionary type to create the new column.
- D. Use AWS Glue DataBrew to read the files. Use the PIVOT transformation to create the new column.

**Answer: B**

Explanation:

The NEST TO MAP transformation allows you to combine multiple columns into a single column that contains a JSON object with key-value pairs. This is the easiest way to achieve the desired format for the physical address data, as you can simply select the columns to nest and specify the keys for each column. The NEST TO ARRAY transformation creates a single column that contains an array of values, which is not the same as the JSON object format. The PIVOT transformation reshapes the data by creating new columns from unique values in a selected column, which is not applicable for this use case. Writing a Lambda function in Python requires more coding effort than using AWS Glue DataBrew, which provides a visual and interactive interface for data transformations. References:

7 most common data preparation transformations in AWS Glue DataBrew (Section: Nesting and unnesting columns) NEST TO MAP - AWS Glue DataBrew (Section: Syntax)

## NEW QUESTION # 34

.....

If you prefer to prepare for your Data-Engineer-Associate exam on paper, we will be your best choice. Data-Engineer-Associate PDF version is printable, and you can print them into hard one and take some notes on them if you like, and you can study them anytime and anywhere. In addition, Data-Engineer-Associate Pdf Version have free demo for you to have a try, so that you can have deeper understanding of what you are going to buy. Data-Engineer-Associate exam dumps are edited by skilled experts, and therefore the quality can be guaranteed. And you can use them at ease.

**Data-Engineer-Associate Guide:** <https://www.pdfvce.com/Amazon/Data-Engineer-Associate-exam-pdf-dumps.html>

Amazon Data-Engineer-Associate Vce Format It's a very short time, no worry to cost your delivery to get it, We are sure that we offer the best professional Data-Engineer-Associate dumps torrent materials, If you are workers in IT field, holding a Data-Engineer-Associate certification (with the help of Data-Engineer-Associate prep + test bundle) will be an outstanding advantages over others when you are facing promotion or better jobs opportunities especially for companies which have business with Amazon or sell Amazon products, Also you will find that most of our Amazon Data-Engineer-Associate exam bootcamp have 85% similarity or above with the real questions of real test after you purchase our products.

But, to answer your question directly, I thought one of my favorite things in Data-Engineer-Associate this book from a workflow perspective is the whole chapter on tapeless media, The secret to system success is the correct organization of its data.

## **Data-Engineer-Associate Vce Format - Amazon Data-Engineer-Associate Guide: AWS Certified Data Engineer - Associate (DEA-C01) Pass Certainly**

It's a very short time, no worry to cost your delivery to get it, We are sure that we offer the best professional Data-Engineer-Associate Dumps Torrent materials, If you are workers in IT field, holding a Data-Engineer-Associate certification (with the help of Data-Engineer-Associate prep + test bundle) will be an outstanding advantages over others when you are facing promotion or better jobs opportunities especially for companies which have business with Amazon or sell Amazon products.

Also you will find that most of our Amazon Data-Engineer-Associate exam bootcamp have 85% similarity or above with the real questions of real test after you purchase our products.

If you think I am exaggerating, you can try it for yourself.

- Accurate Amazon Data-Engineer-Associate Vce Format Are Leading Materials - Fantastic Data-Engineer-Associate Guide  
□ The page for free download of □ Data-Engineer-Associate □ on ► [www.pass4test.com](http://www.pass4test.com) ◀ will open immediately □  
□ Data-Engineer-Associate New Exam Braindumps
- Data-Engineer-Associate Authorized Certification □ Latest Data-Engineer-Associate Test Notes □ Certification Data-Engineer-Associate Training □ Open [ [www.pdfvce.com](http://www.pdfvce.com) ] enter □ Data-Engineer-Associate □ and obtain a free download □ Data-Engineer-Associate Valid Test Questions
- 2026 Useful Data-Engineer-Associate Vce Format | 100% Free AWS Certified Data Engineer - Associate (DEA-C01) Guide □ Download ► Data-Engineer-Associate □ for free by simply entering 【 [www.pdfdumps.com](http://www.pdfdumps.com) 】 website □ Prep Data-Engineer-Associate Guide
- Pass Guaranteed Quiz Amazon - Data-Engineer-Associate Newest Vce Format □ Easily obtain free download of ► Data-Engineer-Associate □ by searching on [ [www.pdfvce.com](http://www.pdfvce.com) ] □ Data-Engineer-Associate Current Exam Content
- Latest Data-Engineer-Associate Demo □ Latest Data-Engineer-Associate Test Notes ✓ □ Reliable Data-Engineer-Associate Braindumps Free □ Simply search for □ Data-Engineer-Associate □ for free download on ► [www.examdisscuss.com](http://www.examdisscuss.com) □ □ Data-Engineer-Associate Certification Practice
- Pass Guaranteed Quiz Data-Engineer-Associate - AWS Certified Data Engineer - Associate (DEA-C01) High Hit-Rate Vce Format □ Search for □ Data-Engineer-Associate □ and download exam materials for free through ► [www.pdfvce.com](http://www.pdfvce.com) ◀ □ Detailed Data-Engineer-Associate Study Plan
- Pass Guaranteed Quiz Data-Engineer-Associate - AWS Certified Data Engineer - Associate (DEA-C01) High Hit-Rate Vce Format □ Download ✓ Data-Engineer-Associate □ ✓ □ for free by simply searching on □ [www.vce4dumps.com](http://www.vce4dumps.com) □ □ □ Data-Engineer-Associate Valid Test Questions
- Reliable Data-Engineer-Associate Test Price □ Data-Engineer-Associate Braindump Pdf □ Detailed Data-Engineer-Associate Study Plan □ Enter □ [www.pdfvce.com](http://www.pdfvce.com) □ and search for ► Data-Engineer-Associate ◀ to download for free □ □ Certification Data-Engineer-Associate Training
- Practice Data-Engineer-Associate Exam □ Data-Engineer-Associate Current Exam Content □ Dumps Data-Engineer-Associate Free Download □ Simply search for ( Data-Engineer-Associate ) for free download on ► [www.pdfdumps.com](http://www.pdfdumps.com) ◀ □ Prep Data-Engineer-Associate Guide
- Pass Guaranteed Quiz Data-Engineer-Associate - AWS Certified Data Engineer - Associate (DEA-C01) High Hit-Rate Vce Format □ Simply search for □ Data-Engineer-Associate □ for free download on □ [www.pdfvce.com](http://www.pdfvce.com) □ □ Data-

## Engineer-Associate Current Exam Content

- [illegible]

DOWNLOAD the newest PDFVCE Data-Engineer-Associate PDF dumps from Cloud Storage for free:  
<https://drive.google.com/open?id=1DxPC6g7wRoNpUnqQGExMPHsG4LwZQF>