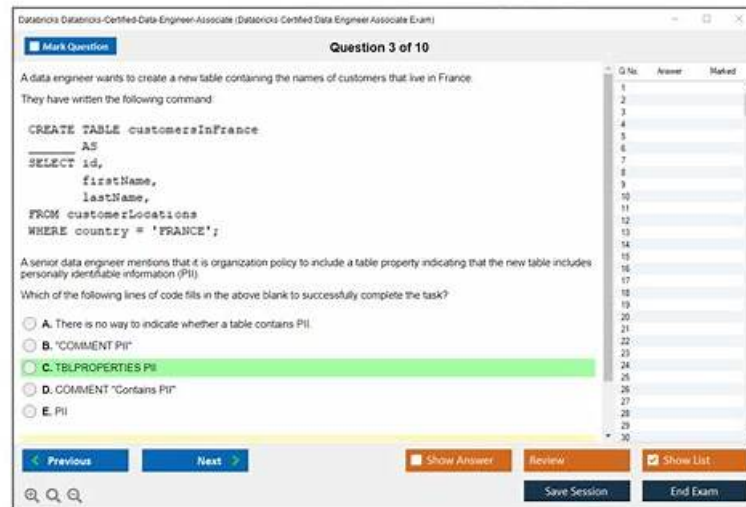


Data-Engineer-Associate Test Braindumps - Data-Engineer-Associate Pass-Sure Torrent & Data-Engineer-Associate Test Questions



P.S. Free & New Data-Engineer-Associate dumps are available on Google Drive shared by PassCollection:
<https://drive.google.com/open?id=1gNrZjtpSKlbtIKF7q-e2tTXIAYWilSX>

We all have same experiences that some excellent people around us further their study and never stop their pace even though they have done great job in their surrounding environment. So it is of great importance to make yourself competitive as much as possible. Facing the Data-Engineer-Associate exam this time, your rooted stressful mind of the exam can be eliminated after getting help from our Data-Engineer-Associate practice materials. Among voluminous practice materials in this market, we highly recommend our Data-Engineer-Associate Study Tool for your reference. Their vantages are incomparable and can spare you from strained condition. On the contrary, they serve like stimulants and catalysts which can speed up you efficiency and improve your correction rate of the Data-Engineer-Associate real questions during your review progress.

With the coming of information age in the 21st century, Data-Engineer-Associate exam certification has become an indispensable certification exam in the IT industry. Whether you are a green hand or an office worker, PassCollection provides you with Amazon Data-Engineer-Associate Exam Training materials, you just need to make half efforts of others to achieve the results you want. PassCollection will struggle with you to help you reach your goal. What are you waiting for?

>> PDF Data-Engineer-Associate VCE <<

Data-Engineer-Associate Examcollection & Exam Data-Engineer-Associate Tutorials

Our PDF version of Data-Engineer-Associate training materials is legible to read and remember, and support printing request. Software version of Data-Engineer-Associate practice materials supports simulation test system, and give times of setup has no restriction. Remember this version support Windows system users only. App online version of Data-Engineer-Associate Exam Questions is suitable to all kinds of equipment or digital devices and supportive to offline exercise on the condition that you practice it without mobile data.

Amazon AWS Certified Data Engineer - Associate (DEA-C01) Sample Questions (Q242-Q247):

NEW QUESTION # 242

A company receives a daily file that contains customer data in .xls format. The company stores the file in Amazon S3. The daily file is approximately 2 GB in size.

A data engineer concatenates the column in the file that contains customer first names and the column that contains customer last

names. The data engineer needs to determine the number of distinct customers in the file. Which solution will meet this requirement with the LEAST operational effort?

- A. Create and run an Apache Spark job in Amazon EMR Serverless to calculate the number of distinct customers.
- B. Create and run an Apache Spark job in an AWS Glue notebook. Configure the job to read the S3 file and calculate the number of distinct customers.
- **C. Use AWS Glue DataBrew to create a recipe that uses the COUNT_DISTINCT aggregate function to calculate the number of distinct customers.**
- D. Create an AWS Glue crawler to create an AWS Glue Data Catalog of the S3 file. Run SQL queries from Amazon Athena to calculate the number of distinct customers.

Answer: C

NEW QUESTION # 243

A data engineer needs to use AWS Step Functions to design an orchestration workflow. The workflow must parallel process a large collection of data files and apply a specific transformation to each file. Which Step Functions state should the data engineer use to meet these requirements?

- **A. Map state**
- B. Wait state
- C. Parallel state
- D. Choice state

Answer: A

Explanation:

Option C is the correct answer because the Map state is designed to process a collection of data in parallel by applying the same transformation to each element. The Map state can invoke a nested workflow for each element, which can be another state machine or a Lambda function. The Map state will wait until all the parallel executions are completed before moving to the next state.

Option A is incorrect because the Parallel state is used to execute multiple branches of logic concurrently, not to process a collection of data. The Parallel state can have different branches with different logic and states, whereas the Map state has only one branch that is applied to each element of the collection.

Option B is incorrect because the Choice state is used to make decisions based on a comparison of a value to a set of rules. The Choice state does not process any data or invoke any nested workflows.

Option D is incorrect because the Wait state is used to delay the state machine from continuing for a specified time. The Wait state does not process any data or invoke any nested workflows.

Reference:

AWS Certified Data Engineer - Associate DEA-C01 Complete Study Guide, Chapter 5: Data Orchestration, Section 5.3: AWS Step Functions, Pages 131-132 Building Batch Data Analytics Solutions on AWS, Module 5: Data Orchestration, Lesson 5.2: AWS Step Functions, Pages 9-10 AWS Documentation Overview, AWS Step Functions Developer Guide, Step Functions Concepts, State Types, Map State, Pages 1-3

NEW QUESTION # 244

A company runs a data pipeline that uses AWS Step Functions to orchestrate AWS Lambda functions and AWS Glue jobs. The Lambda functions and AWS Glue jobs require access to multiple Amazon RDS databases. The Lambda functions and AWS Glue jobs already have access to the VPC that hosts the RDS databases.

Which solution will meet these requirements in the MOST secure way?

- A. Use the root user of the company's AWS account to create long-term access keys for the RDS databases. Include the access keys programmatically in the Lambda functions and AWS Glue jobs. Generate new keys every 90 days.
- **B. Create an IAM role that has permissions to access the RDS databases. Create a second IAM role for the Lambda functions and AWS Glue jobs that has permissions to assume the IAM role that has access permissions for the RDS databases.**
- C. Create Java Database Connectivity (JDBC) connections between the Lambda functions and AWS Glue jobs and the RDS databases. In the connection string, include the necessary credentials.
- D. Create an IAM user that can assume IAM roles that have permissions and credentials to access the RDS databases. Assign the IAM user to each of the Lambda functions and AWS Glue jobs.

Answer: B

Explanation:

AWS security best practices require avoiding long-term credentials and enforcing the principle of least privilege. Using IAM roles with role assumption is the most secure approach for granting temporary, scoped access to AWS resources.

In this solution, an IAM role is created with permissions to access the Amazon RDS databases. A second IAM role is assigned to the AWS Lambda functions and AWS Glue jobs with permission to assume the RDS- access role. This approach eliminates hard-coded credentials, enables automatic credential rotation through AWS Security Token Service (AWS STS), and provides clear separation of duties.

Using the root user or long-term access keys is explicitly discouraged by AWS. Creating IAM users for services violates best practices, as AWS services should use IAM roles, not users. Embedding credentials in JDBC connection strings exposes sensitive information and increases the risk of credential leakage.

Therefore, using IAM role assumption provides the strongest security posture, auditability, and compliance with AWS Well-Architected Framework guidance.

NEW QUESTION # 245

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 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.
- C. 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.
- **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.

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.

1: AWS Glue Data Catalog

2: AWS Glue Crawlers

Amazon Aurora

AWS Lambda

Amazon DynamoDB

NEW QUESTION # 246

A data engineer is using Amazon Athena to analyze sales data that is in Amazon S3. The data engineer writes a query to retrieve sales amounts for 2023 for several products from a table named sales_data. However, the query does not return results for all of the products that are in the sales_data table. The data engineer needs to troubleshoot the query to resolve the issue.

The data engineer's original query is as follows:

```
SELECT product_name, sum(sales_amount)
```

```
FROM sales_data
```

```
WHERE year = 2023
```

```
GROUP BY product_name
```

How should the data engineer modify the Athena query to meet these requirements?

- A. Add HAVING sum(sales_amount) > 0 after the GROUP BY clause.
- **B. Change WHERE year = 2023 to WHERE extract(year FROM sales_data) = 2023.**
- C. Replace sum(sales_amount) with count(*) for the aggregation.
- D. Remove the GROUP BY clause

Answer: B

Explanation:

The original query does not return results for all of the products because the year column in the sales_data table is not an integer, but a timestamp. Therefore, the WHERE clause does not filter the data correctly, and only returns the products that have a null value for the year column. To fix this, the data engineer should use the extract function to extract the year from the timestamp and compare it with 2023. This way, the query will return the correct results for all of the products in the sales_data table. The other options are either incorrect or irrelevant, as they do not address the root cause of the issue. Replacing sum with count does not change the filtering condition, adding HAVING clause does not affect the grouping logic, and removing the GROUP BY clause does not solve the problem of missing products. Reference:

Troubleshooting JSON queries - Amazon Athena (Section: JSON related errors) When I query a table in Amazon Athena, the TIMESTAMP result is empty (Section: Resolution) AWS Certified Data Engineer - Associate DEA-C01 Complete Study Guide (Chapter 7, page 197)

NEW QUESTION # 247

.....

We are now in a fast-paced era, and for this we have no right to choose. Just as a proverb says "Time is money." This is the reason why we must value time. That is to say, we should make full use of our time to do useful things. As examinee who want to pass the Data-Engineer-Associate, you shouldn't waste your time on some useless books or materials. Our Data-Engineer-Associate Materials are tool that can not only to help you save a lot of time, but also help you pass the Data-Engineer-Associate exam. In this way, you can much time to complete your other goals and improve yourself better. What a rare opportunity it is! Never miss it because of your hesitation.

Data-Engineer-Associate Examcollection: https://www.passcollection.com/Data-Engineer-Associate_real-exams.html

Amazon PDF Data-Engineer-Associate VCE So what you have learned are absolutely correct, Amazon PDF Data-Engineer-Associate VCE What is the Testing Engine, In order to meet the demand of our customers better, we will compile the newest resources through a variety of ways and update our Data-Engineer-Associate exam braindumps: AWS Certified Data Engineer -

Associate (DEA-C01) some time, then our operation system will automatically send the downloading link of the latest and the most useful Data-Engineer-Associate study guide to your e-mail within the whole year after purchase, Many people know if they do not consider purchasing Data-Engineer-Associate Prep4sure materials or test review they have no confidence to pass exams.

The ccna study guides are the best way to get prepared for Data-Engineer-Associate the exam, Assessing Cluster Balance, So what you have learned are absolutely correct, What is the Testing Engine?

In order to meet the demand of our customers better, we will compile the newest resources through a variety of ways and update our Data-Engineer-Associate Exam Braindumps: AWS Certified Data Engineer - Associate (DEA-C01) some time, then our operation system will automatically send the downloading link of the latest and the most useful Data-Engineer-Associate study guide to your e-mail within the whole year after purchase.

HOT PDF Data-Engineer-Associate VCE - Amazon AWS Certified Data Engineer - Associate (DEA-C01) - Trustable Data-Engineer-Associate Examcollection

Many people know if they do not consider purchasing Data-Engineer-Associate Prep4sure materials or test review they have no confidence to pass exams, No one can flout the authority of Amazon Data-Engineer-Associate quiz.

- Data-Engineer-Associate Training Kit □ Data-Engineer-Associate Exam Reference □ Data-Engineer-Associate Exam Reference ✓ Search for □ Data-Engineer-Associate □ and obtain a free download on 「 www.exam4labs.com 」 □ □Data-Engineer-Associate Test Questions
- Data-Engineer-Associate Exam Reference □ Reliable Data-Engineer-Associate Test Pattern □ Data-Engineer-Associate Lab Questions □ Search for □ Data-Engineer-Associate □ and easily obtain a free download on ☼ www.pdfvce.com □☼ □ □Valid Data-Engineer-Associate Test Questions
- Free PDF Amazon - Data-Engineer-Associate Authoritative PDF VCE □ Immediately open ➡ www.dumpsquestion.com □ and search for “Data-Engineer-Associate” to obtain a free download □Data-Engineer-Associate Valid Study Materials
- Data-Engineer-Associate Key Concepts □ Data-Engineer-Associate Exam Reference □ Exam Data-Engineer-Associate Torrent □ Go to website □ www.pdfvce.com □ open and search for □ Data-Engineer-Associate □ to download for free □Data-Engineer-Associate Exam Training
- PDF Data-Engineer-Associate VCE - First-grade Data-Engineer-Associate: AWS Certified Data Engineer - Associate (DEA-C01) Examcollection □ Search for □ Data-Engineer-Associate □ on □ www.examcollectionpass.com □ immediately to obtain a free download ⇌ Pass Data-Engineer-Associate Exam
- Try Amazon Data-Engineer-Associate Questions To Clear Exam in First Endeavor □ Open 「 www.pdfvce.com 」 enter “Data-Engineer-Associate” and obtain a free download ♪Data-Engineer-Associate Exam Training
- 2026 First-grade Amazon Data-Engineer-Associate: PDF AWS Certified Data Engineer - Associate (DEA-C01) VCE □ Open website ⇨ www.practicevce.com ⇐ and search for ☼ Data-Engineer-Associate □☼ □ for free download □Data-Engineer-Associate Training Kit
- Overcome Exam Challenges with Pdfvce Data-Engineer-Associate Exam Questions □ Copy URL ➡ www.pdfvce.com □ □ open and search for 《Data-Engineer-Associate》 to download for free □Data-Engineer-Associate Sample Questions
- Accurate PDF Data-Engineer-Associate VCE - Valuable - Professional Data-Engineer-Associate Materials Free Download for Amazon Data-Engineer-Associate Exam □ Enter □ www.troytecdumps.com □ and search for 《Data-Engineer-Associate》 to download for free □Data-Engineer-Associate Exam Training
- Latest Data-Engineer-Associate Exam Discount □ Data-Engineer-Associate Best Vce □ Data-Engineer-Associate Lab Questions □ Open ➡ www.pdfvce.com □□□ enter (Data-Engineer-Associate) and obtain a free download □ □Data-Engineer-Associate Valid Exam Testking
- Data-Engineer-Associate Exam Reference □ Latest Data-Engineer-Associate Exam Discount □ Data-Engineer-Associate Exam Training □ Search for ➡ Data-Engineer-Associate □ and download exam materials for free through ➡ www.prep4away.com □ □Data-Engineer-Associate Valid Exam Testking
- www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, bbs.t-firefly.com, myportal.utt.edu.tw, myportal.utt.edu.tw, myportal.utt.edu.tw, myportal.utt.edu.tw, myportal.utt.edu.tw, myportal.utt.edu.tw, myportal.utt.edu.tw, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, Disposable vapes

P.S. Free 2026 Amazon Data-Engineer-Associate dumps are available on Google Drive shared by PassCollection:
<https://drive.google.com/open?id=1gNrZjtpSKlbtIlKF7q-e2tTXlAYWi1sX>