Data-Engineer-Associate Reliable Exam Braindumps, Discount Data-Engineer-Associate Code

BTW, DOWNLOAD part of Real4Prep Data-Engineer-Associate dumps from Cloud Storage: https://drive.google.com/open?id=1oY_VTKUWYrHZG8COhtvO9A5G8noDShRy

One strong point of our APP online version is that it is convenient for you to use our Data-Engineer-Associate exam dumps even though you are in offline environment. In other words, you can prepare for your Data-Engineer-Associate exam with under the guidance of our Data-Engineer-Associate Training Materials anywhere at any time. Just take action to purchase we would be pleased to make you the next beneficiary of our Data-Engineer-Associate exam practice. Trust us and you will get what you are dreaming!

You can use this Data-Engineer-Associate simulation software without an internet connection after installation. Tracking and reporting features of our Amazon Data-Engineer-Associate practice exam software makes it easier for you to identify and overcome mistakes. Customization feature of this format allows you to change time limits and questions numbers of mock exams.

>> Data-Engineer-Associate Reliable Exam Braindumps <<

Discount Data-Engineer-Associate Code & Data-Engineer-Associate Study Plan

Getting an authoritative IT certification will make a great difference to your career like Data-Engineer-Associate exam tests. The

difficulty and profession of real questions need much time and energy to prepare, which can be solved by our Data-Engineer-Associate dumps torrent. The latest training materials are tested by IT experts and certified trainers who studied Data-Engineer-Associate Exam Questions for many years. The high quality of our vce braindumps are the guarantee of high passing score.

Amazon AWS Certified Data Engineer - Associate (DEA-C01) Sample Questions (Q27-Q32):

NEW QUESTION # 27
A data engineer has a one-time task to read data from objects that are in Apache Parquet format in an Amazon S3 bucket. The data engineer needs to query only one column of the data.
Which solution will meet these requirements with the LEAST operational overhead?

- A. Use S3 Select to write a SQL SELECT statement to retrieve the required column from the S3 objects.
- B. Prepare an AWS Glue DataBrew project to consume the S3 objects and to query the required column.
- C. Run an AWS Glue crawler on the S3 objects. Use a SQL SELECT statement in Amazon Athena to query the required column.
- D. Configure an AWS Lambda function to load data from the S3 bucket into a pandas dataframe- Write a SQL SELECT statement on the dataframe to query the required column.

Answer: A

Explanation:
Option B is the best solution to meet the requirements with the least operational overhead because S3 Select is a feature that allows you to retrieve only a subset of data from an S3 object by using simple SQL expressions.
S3 Select works on objects stored in CSV, JSON, or Parquet format. By using S3 Select, you can avoid the need to download and process the entire S3 object, which reduces the amount of data transferred and the computation time. S3 Select is also easy to use and does not require any additional services or resources.
Option A is not a good solution because it involves writing custom code and configuring an AWS Lambda function to load data from the S3 bucket into a pandas dataframe and query the required column. This option adds complexity and latency to the data retrieval process and requires additional resources and configuration.Moreover, AWS Lambda has limitations on the execution time, memory, and concurrency, which may affect the performance and reliability of the data retrieval process.
Option C is not a good solution because it involves creating and running an AWS Glue DataBrew project to consume the S3 objects and query the required column. AWS Glue DataBrew is a visual data preparation tool that allows you to clean, normalize, and transform data without writing code. However, in this scenario, the data is already in Parquet format, which is a columnar storage format that is optimized for analytics.
Therefore, there is no need to use AWS Glue DataBrew to prepare the data. Moreover, AWS Glue DataBrew adds extra time and cost to the data retrieval process and requires additional resources and configuration.
Option D is not a good solution because it involves running an AWS Glue crawler on the S3 objects and using a SQL SELECT statement in Amazon Athena to query the required column. An AWS Glue crawler is a service that can scan data sources and create metadata tables in the AWS Glue Data Catalog. The Data Catalog is a central repository that stores information about the data sources, such as schema, format, and location.
Amazon Athena is a serverless interactive query service that allows you to analyze data in S3 using standard SQL. However, in this scenario, the schema and format of the data are already known and fixed, so there is no need to run a crawler to discover them. Moreover, running a crawler and using Amazon Athena adds extra time and cost to the data retrieval process and requires additional services and configuration.
References:
AWS Certified Data Engineer - Associate DEA-C01 Complete Study Guide
S3 Select and Glacier Select - Amazon Simple Storage Service
AWS Lambda - FAQs
What Is AWS Glue DataBrew? - AWS Glue DataBrew
Populating the AWS Glue Data Catalog - AWS Glue
What is Amazon Athena? - Amazon Athena

NEW QUESTION # 28
A financial company recently added more features to its mobile app. The new features required the company to create a new topic in an existing Amazon Managed Streaming for Apache Kafka (Amazon MSK) cluster.
A few days after the company added the new topic, Amazon CloudWatch raised an alarm on the RootDiskUsed metric for the MSK cluster.
How should the company address the CloudWatch alarm?

- A. Specify the Target-Volume-in-GiB parameter for the existing topic.
- B. Expand the storage of the MSK broker. Configure the MSK cluster storage to expand automatically.
- C. Update the MSK broker instance to a larger instance type. Restart the MSK cluster.
- D. Expand the storage of the Apache ZooKeeper nodes.

Answer: B

Explanation:
The RootDiskUsed metric for the MSK cluster indicates that the storage on the broker is reaching its capacity. The best solution is to expand the storage of the MSK broker and enable automatic storage expansion to prevent future alarms.
* Expand MSK Broker Storage:
* AWS Managed Streaming for Apache Kafka (MSK) allows you to expand the broker storage to accommodate growing data volumes. Additionally, auto-expansion of storage can be configured to ensure that storage grows automatically as the data increases.

NEW QUESTION # 29
A company is building a data lake for a new analytics team. The company is using Amazon S3 for storage and Amazon Athena for query analysis. All data that is in Amazon S3 is in Apache Parquet format.
The company is running a new Oracle database as a source system in the company's data center. The company has 70 tables in the Oracle database. All the tables have primary keys. Data can occasionally change in the source system. The company wants to ingest the tables every day into the data lake.
Which solution will meet this requirement with the LEAST effort?

- A. Create an AWS Glue connection to the Oracle database. Create an AWS Glue bookmark job to ingest the data incrementally and to write the data to Amazon S3 in Parquet format.
- B. Create an Apache Sqoop job in Amazon EMR to read the data from the Oracle database. Configure the Sqoop job to write the data to Amazon S3 in Parquet format.
- C. Create an AWS Database Migration Service (AWS DMS) task for ongoing replication. Set the Oracle database as the source. Set Amazon S3 as the target. Configure the task to write the data in Parquet format.
- D. Create an Oracle database in Amazon RDS. Use AWS Database Migration Service (AWS DMS) to migrate the on-premises Oracle database to Amazon RDS. Configure triggers on the tables to invoke AWS Lambda functions to write changed records to Amazon S3 in Parquet format.

Answer: C

Explanation:
The company needs to ingest tables from an on-premises Oracle database into a data lake on Amazon S3 in Apache Parquet format. The most efficient solution, requiring the least manual effort, would be to use AWS Database Migration Service (DMS) for continuous data replication.
Option C: Create an AWS Database Migration Service (AWS DMS) task for ongoing replication. Set the Oracle database as the source. Set Amazon S3 as the target. Configure the task to write the data in Parquet format.AWS DMS can continuously replicate data from the Oracle database into Amazon S3, transforming it into Parquet format as it ingests the data. DMS simplifies the process by providing ongoing replication with minimal setup, and it automatically handles the conversion to Parquet format without requiring manual transformations or separate jobs. This option is the least effort solution since it automates both the ingestion and transformation processes.
Other options:
Option A (Apache Sqoop on EMR) involves more manual configuration and management, including setting up EMR clusters and writing Sqoop jobs.
Option B (AWS Glue bookmark job) involves configuring Glue jobs, which adds complexity. While Glue supports data transformations, DMS offers a more seamless solution for database replication.
Option D (RDS and Lambda triggers) introduces unnecessary complexity by involving RDS and Lambda for a task that DMS can handle more efficiently.
References:
AWS Database Migration Service (DMS)
DMS S3 Target Documentation

NEW QUESTION # 30
A data engineer maintains a materialized view that is based on an Amazon Redshift database. The view has a column named load_date that stores the date when each row was loaded.
The data engineer needs to reclaim database storage space by deleting all the rows from the materialized view.

Which command will reclaim the MOST database storage space?

A.

```
DELETE FROM materialized_view_name where 1=1
```

B.

```
TRUNCATE materialized_view_name
```

C.

```
VACUUM table_name where load_date<=current_date
materializedview
```

D.

```
DELETE FROM materialized_view_name where load_date<=current_date
```


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

Answer: A

Explanation:
To reclaim the most storage space from a materialized view in Amazon Redshift, you should use a DELETE operation that removes all rows from the view. The most efficient way to remove all rows is to use a condition that always evaluates to true, such as 1=1. This will delete all rows without needing to evaluate each row individually based on specific column values like load_date.
* Option A: DELETE FROM materialized_view_name WHERE 1=1;This statement will delete all rows in the materialized view and free up the space. Since materialized views in Redshift store precomputed data, performing a DELETE operation will remove all stored rows.
Other options either involve inappropriate SQL statements (e.g., VACUUM in option C is used for reclaiming storage space in tables, not materialized views), or they don't remove data effectively in the context of a materialized view (e.g., TRUNCATE cannot be used directly on a materialized view).
References:
* Amazon Redshift Materialized Views Documentation
* Deleting Data from Redshift

NEW QUESTION # 31
A company uses AWS Step Functions to orchestrate a data pipeline. The pipeline consists of Amazon EMR jobs that ingest data from data sources and store the data in an Amazon S3 bucket. The pipeline also includes EMR jobs that load the data to Amazon Redshift.
The company's cloud infrastructure team manually built a Step Functions state machine. The cloud infrastructure team launched an EMR cluster into a VPC to support the EMR jobs. However, the deployed Step Functions state machine is not able to run the EMR jobs.
Which combination of steps should the company take to identify the reason the Step Functions state machine is not able to run the EMR jobs? (Choose two.)

- A. Check the retry scenarios that the company configured for the EMR jobs. Increase the number of seconds in the interval between each EMR task. Validate that each fallback state has the appropriate catch for each decision state. Configure an Amazon Simple Notification Service (Amazon SNS) topic to store the error messages.
- B. Use AWS CloudFormation to automate the Step Functions state machine deployment. Create a step to pause the state machine during the EMR jobs that fail. Configure the step to wait for a human user to send approval through an email message. Include details of the EMR task in the email message for further analysis.
- C. Query the flow logs for the VPC. Determine whether the traffic that originates from the EMR cluster can successfully reach

the data providers. Determine whether any security group that might be attached to the Amazon EMR cluster allows connections to the data source servers on the informed ports.

- D. Check for entries in Amazon CloudWatch for the newly created EMR cluster. Change the AWS Step Functions state machine code to use Amazon EMR on EKS. Change the IAM access policies and the security group configuration for the Step Functions state machine code to reflect inclusion of Amazon Elastic Kubernetes Service (Amazon EKS).
- E. Verify that the Step Functions state machine code has all IAM permissions that are necessary to create and run the EMR jobs. Verify that the Step Functions state machine code also includes IAM permissions to access the Amazon S3 buckets that the EMR jobs use. Use Access Analyzer for S3 to check the S3 access properties.

Answer: C,E

Explanation:
To identify the reason why the Step Functions state machine is not able to run the EMR jobs, the company should take the following steps:
Verify that the Step Functions state machine code has all IAM permissions that are necessary to create and run the EMR jobs. The state machine code should have an IAM role that allows it to invoke the EMR APIs, such as RunJobFlow, AddJobFlowSteps, and DescribeStep. The state machine code should also have IAM permissions to access the Amazon S3 buckets that the EMR jobs use as input and output locations. The company can use Access Analyzer for S3 to check the access policies and permissions of the S3 buckets12.
Therefore, option B is correct.
Query the flow logs for the VPC. The flow logs can provide information about the network traffic to and from the EMR cluster that is launched in the VPC. The company can use the flow logs to determine whether the traffic that originates from the EMR cluster can successfully reach the data providers, such as Amazon RDS, Amazon Redshift, or other external sources. The company can also determine whether any security group that might be attached to the EMR cluster allows connections to the data source servers on the informed ports. The company can use Amazon VPC Flow Logs or Amazon CloudWatch Logs Insights to query the flow logs3 .
Therefore, option D is correct.
Option A is incorrect because it suggests using AWS CloudFormation to automate the Step Functions state machine deployment. While this is a good practice to ensure consistency and repeatability of the deployment, it does not help to identify the reason why the state machine is not able to run the EMR jobs. Moreover, creating a step to pause the state machine during the EMR jobs that fail and wait for a human user to send approval through an email message is not a reliable way to troubleshoot the issue. The company should use the Step Functions console or API to monitor the execution history and status of the state machine, and use Amazon CloudWatch to view the logs and metrics of the EMR jobs .
Option C is incorrect because it suggests changing the AWS Step Functions state machine code to use Amazon EMR on EKS. Amazon EMR on EKS is a service that allows you to run EMR jobs on Amazon Elastic Kubernetes Service (Amazon EKS) clusters. While this service has some benefits, such as lower cost and faster execution time, it does not support all the features and integrations that EMR on EC2 does, such as EMR Notebooks, EMR Studio, and EMRFS. Therefore, changing the state machine code to use EMR on EKS may not be compatible with the existing data pipeline and may introduce new issues.
Option E is incorrect because it suggests checking the retry scenarios that the company configured for the EMR jobs. While this is a good practice to handle transient failures and errors, it does not help to identify the root cause of why the state machine is not able to run the EMR jobs. Moreover, increasing the number of seconds in the interval between each EMR task may not improve the success rate of the jobs, and may increase the execution time and cost of the state machine. Configuring an Amazon SNS topic to store the error messages may help to notify the company of any failures, but it does not provide enough information to troubleshoot the issue.
1: Manage an Amazon EMR Job - AWS Step Functions
2: Access Analyzer for S3 - Amazon Simple Storage Service
3: Working with Amazon EMR and VPC Flow Logs - Amazon EMR
[4]: Analyzing VPC Flow Logs with Amazon CloudWatch Logs Insights - Amazon Virtual Private Cloud
[5]: Monitor AWS Step Functions - AWS Step Functions
[6]: Monitor Amazon EMR clusters - Amazon EMR
[7]: Amazon EMR on Amazon EKS - Amazon EMR

NEW QUESTION # 32
.....

Discount Data-Engineer-Associate Code: https://www.real4prep.com/Data-Engineer-Associate-exam.html

Amazon Data-Engineer-Associate Reliable Exam Braindumps there was so much to study, Amazon Data-Engineer-Associate Reliable Exam Braindumps Products can be accessed instantly after the confirmation of payment is received, We are legal authorized company devoting to researching and selling professional Amazon Data-Engineer-Associate examcollection many years, Amazon Data-Engineer-Associate Reliable Exam Braindumps All of the content are the absolute same, just in different ways to use.

Currently, my company has introduced a variety Data-Engineer-Associate Study Plan of learning materials, covering almost all the official certification of qualification exams, and each Data-Engineer-Associate practice dump in our online store before the listing, are subject to stringent quality checks within the company.

Top Data-Engineer-Associate Reliable Exam Braindumps | Easy To Study and Pass Exam at first attempt & Latest updated Data-Engineer-Associate: AWS Certified Data Engineer - Associate (DEA-C01)

Creating a Soundtrack, there was so much to Data-Engineer-Associate Hottest Certification study, Products can be accessed instantly after the confirmation of payment is received, We are legal authorized company devoting to researching and selling professional Amazon Data-Engineer-Associate Examcollection many years.

All of the content are the absolute same, just in different Data-Engineer-Associate ways to use, That means if you study our study guide, your passing rate is much higher than other candidates.

- Valid Data-Engineer-Associate Reliable Exam Braindumps - Free Download Discount Data-Engineer-Associate Code: AWS Certified Data Engineer - Associate (DEA-C01) □ Open website ➠ www.prepawaypdf.com □ and search for { Data-Engineer-Associate } for free download □Online Data-Engineer-Associate Training
- Data-Engineer-Associate Reliable Exam Braindumps: 2026 Realistic Amazon Discount AWS Certified Data Engineer - Associate (DEA-C01) Code Pass Guaranteed □ Enter 《 www.pdfvce.com 》 and search for （ Data-Engineer-Associate ） to download for free □Online Data-Engineer-Associate Training
- Test Data-Engineer-Associate Valid □ New Data-Engineer-Associate Mock Exam □ Valid Data-Engineer-Associate Test Cram □ Search for ➠ Data-Engineer-Associate □ and obtain a free download on ➤ www.prepawayete.com □ □ □Data-Engineer-Associate Latest Cram Materials
- Valid Data-Engineer-Associate Test Cram □ Pdf Data-Engineer-Associate Format □ Latest Braindumps Data-Engineer-Associate Ppt □ Search for ➥ Data-Engineer-Associate □ on ✔ www.pdfvce.com □✔□ immediately to obtain a free download □Certification Data-Engineer-Associate Cost
- Data-Engineer-Associate Reliable Exam Braindumps: 2026 Realistic Amazon Discount AWS Certified Data Engineer - Associate (DEA-C01) Code Pass Guaranteed □ Go to website □ www.prep4away.com □ open and search for " Data-Engineer-Associate " to download for free □Questions Data-Engineer-Associate Exam
- 2026 Amazon Data-Engineer-Associate Pass-Sure Reliable Exam Braindumps □ Open website ➤ www.pdfvce.com □ and search for □ Data-Engineer-Associate □ for free download □Data-Engineer-Associate Valid Test Forum
- Quiz 2026 Amazon Data-Engineer-Associate: AWS Certified Data Engineer - Associate (DEA-C01) – High Pass-Rate Reliable Exam Braindumps □ Enter ✔ www.prep4sures.top □✔□ and search for 【 Data-Engineer-Associate 】 to download for free □Data-Engineer-Associate New Soft Simulations
- Online Data-Engineer-Associate Training ☀ Data-Engineer-Associate Valid Test Forum □ Data-Engineer-Associate Valid Study Questions □ Open （ www.pdfvce.com ） enter 「 Data-Engineer-Associate 」 and obtain a free download □ □Data-Engineer-Associate Valid Study Questions
- Data-Engineer-Associate Reliable Exam Braindumps: 2026 Realistic Amazon Discount AWS Certified Data Engineer - Associate (DEA-C01) Code Pass Guaranteed □ Go to website ✔ www.troytecdumps.com □✔□ open and search for 「 Data-Engineer-Associate 」 to download for free □Latest Braindumps Data-Engineer-Associate Ppt
- Free PDF Quiz Data-Engineer-Associate - AWS Certified Data Engineer - Associate (DEA-C01) High Hit-Rate Reliable Exam Braindumps □ Easily obtain free download of ⇛ Data-Engineer-Associate ⇚ by searching on ➡ www.pdfvce.com □□□ □Data-Engineer-Associate Latest Cram Materials
- Data-Engineer-Associate Latest Cram Materials □ Data-Engineer-Associate Actual Test Answers □ Valid Data-Engineer-Associate Test Cram □ Open ▶ www.prepawayexam.com ◀ enter " Data-Engineer-Associate " and obtain a free download □Data-Engineer-Associate Valid Study Questions
- shortcourses.russellcollege.edu.au, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, courses.learnwells.com, ncon.edu.sa, rcmspace.com, www.stes.tyc.edu.tw, Disposable vapes

DOWNLOAD the newest Real4Prep Data-Engineer-Associate PDF dumps from Cloud Storage for free: https://drive.google.com/open?id=1oY_VTKUWYrHZG8COhtvO9A5G8noDShRy