

Quiz 2026 Amazon Accurate Data-Engineer-Associate Valid Test Sample

Question 3 of 10

A data engineer wants to create a new table containing the names of customers that live in France. They have written the following command:

```
CREATE TABLE customersInFrance
AS
SELECT id,
       firstName,
       lastName,
FROM customersLocations
WHERE country = 'FRANCE';
```

A senior data engineer mentions that it is organization policy to include a table property indicating that the new table includes personally identifiable information (PII).

Which of the following lines of code fills in the above blank to successfully complete the task?

- ☐ A. There is no way to indicate whether a table contains PII.
- ☐ B. "COMMENT PII"
- ☒ C. TBLPARTIAL PII
- ☐ D. COMMENT "Contains PII"
- ☐ E. PII

Navigation buttons: Previous, Next, Show Answer, Review, Show List, Save Session, End Exam.

P.S. Free & New Data-Engineer-Associate dumps are available on Google Drive shared by TestkingPass:
<https://drive.google.com/open?id=1BMsZJgh8hqUonfApu0Ka0nbCKNSdcyrV>

As for candidates who possessed with a Data-Engineer-Associate professional certification are more competitive. The current word is a stage of science and technology, social media and social networking has already become a popular means of Data-Engineer-Associate exam materials. As a result, more and more people study or prepare for exam through social networking. By this way, our Data-Engineer-Associate learning guide can be your best learn partner. The pass rate of our Data-Engineer-Associate exam questions is high as 99% to 100%, and it is a wise choice to have our Data-Engineer-Associate training guide.

The TestkingPass product here is better, cheaper, higher quality and unlimited for all time; kiss the days of purchasing multiple Amazon braindumps repeatedly, or renewing Data-Engineer-Associate training courses because you ran out of time. Now you can learn Data-Engineer-Associate skills and theory at your own pace and anywhere you want with top of the Data-Engineer-Associate braindumps, you will find it's just like a piece of cake to pass Data-Engineer-Associate exam.

>> Data-Engineer-Associate Valid Test Sample <<

Test Data-Engineer-Associate Answers, Data-Engineer-Associate Latest Braindumps Free

In order to pass Amazon certification Data-Engineer-Associate exam, selecting the appropriate training tools is very necessary. And professional study materials about Amazon certification Data-Engineer-Associate exam is a very important part. Our TestkingPass can have a good and quick provide of professional study materials about Amazon Certification Data-Engineer-Associate Exam. Our TestkingPass IT experts are very experienced and their study materials are very close to the actual exam questions, almost the same. TestkingPass is a convenient website specifically for people who want to take the certification exams, which can effectively help the candidates to pass the exam.

Amazon AWS Certified Data Engineer - Associate (DEA-C01) Sample Questions (Q24-Q29):

NEW QUESTION # 24

A company uses an Amazon QuickSight dashboard to monitor usage of one of the company's applications.

The company uses AWS Glue jobs to process data for the dashboard. The company stores the data in a single Amazon S3 bucket. The company adds new data every day.

A data engineer discovers that dashboard queries are becoming slower over time. The data engineer determines that the root cause of the slowing queries is long-running AWS Glue jobs.

Which actions should the data engineer take to improve the performance of the AWS Glue jobs? (Choose two.)

- A. Convert the AWS Glue schema to the DynamicFrame schema class.
- B. Modify the IAM role that grants access to AWS glue to grant access to all S3 features.
- C. Adjust AWS Glue job scheduling frequency so the jobs run half as many times each day.
- **D. Increase the AWS Glue instance size by scaling up the worker type.**
- **E. Partition the data that is in the S3 bucket. Organize the data by year, month, and day.**

Answer: D,E

Explanation:

Partitioning the data in the S3 bucket can improve the performance of AWS Glue jobs by reducing the amount of data that needs to be scanned and processed. By organizing the data by year, month, and day, the AWS Glue job can use partition pruning to filter out irrelevant data and only read the data that matches the query criteria. This can speed up the data processing and reduce the cost of running the AWS Glue job.

Increasing the AWS Glue instance size by scaling up the worker type can also improve the performance of AWS Glue jobs by providing more memory and CPU resources for the Spark execution engine. This can help the AWS Glue job handle larger data sets and complex transformations more efficiently. The other options are either incorrect or irrelevant, as they do not affect the performance of the AWS Glue jobs. Converting the AWS Glue schema to the DynamicFrame schema class does not improve the performance, but rather provides additional functionality and flexibility for data manipulation. Adjusting the AWS Glue job scheduling frequency does not improve the performance, but rather reduces the frequency of data updates. Modifying the IAM role that grants access to AWS Glue does not improve the performance, but rather affects the security and permissions of the AWS Glue service.

References:

* Optimising Glue Scripts for Efficient Data Processing: Part 1 (Section: Partitioning Data in S3)

* Best practices to optimize cost and performance for AWS Glue streaming ETL jobs (Section:

Development tools)

* Monitoring with AWS Glue job run insights (Section: Requirements)

* AWS Certified Data Engineer - Associate DEA-C01 Complete Study Guide (Chapter 5, page 133)

NEW QUESTION # 25

A car sales company maintains data about cars that are listed for sale in an area. The company receives data about new car listings from vendors who upload the data daily as compressed files into Amazon S3. The compressed files are up to 5 KB in size. The company wants to see the most up-to-date listings as soon as the data is uploaded to Amazon S3.

A data engineer must automate and orchestrate the data processing workflow of the listings to feed a dashboard. The data engineer must also provide the ability to perform one-time queries and analytical reporting. The query solution must be scalable.

Which solution will meet these requirements MOST cost-effectively?

- A. Use AWS Glue to process incoming data. Use AWS Step Functions to orchestrate workflows. Use Amazon Redshift Spectrum for one-time queries and analytical reporting. Use OpenSearch Dashboards in Amazon OpenSearch Service for the dashboard.
- **B. Use AWS Glue to process incoming data. Use AWS Lambda and S3 Event Notifications to orchestrate workflows. Use Amazon Athena for one-time queries and analytical reporting. Use Amazon QuickSight for the dashboard.**
- C. Use a provisioned Amazon EMR cluster to process incoming data. Use AWS Step Functions to orchestrate workflows. Use Amazon Athena for one-time queries and analytical reporting. Use Amazon QuickSight for the dashboard.
- D. Use an Amazon EMR cluster to process incoming data. Use AWS Step Functions to orchestrate workflows. Use Apache Hive for one-time queries and analytical reporting. Use Amazon OpenSearch Service to bulk ingest the data into compute optimized instances. Use OpenSearch Dashboards in OpenSearch Service for the dashboard.

Answer: B

Explanation:

For processing the incoming car listings in a cost-effective, scalable, and automated way, the ideal approach involves using AWS Glue for data processing, AWS Lambda with S3 Event Notifications for orchestration, Amazon Athena for one-time queries and analytical reporting, and Amazon QuickSight for visualization on the dashboard. Let's break this down:

* AWS Glue: This is a fully managed ETL (Extract, Transform, Load) service that automatically processes the incoming data files. Glue is serverless and supports diverse data sources, including Amazon S3 and Redshift.

* AWS Lambda and S3 Event Notifications: Using Lambda and S3 Event Notifications allows near real-time triggering of processing workflows as soon as new data is uploaded into S3. This approach is event-driven, ensuring that the listings are processed as soon as they are uploaded, reducing the latency for data processing.

* Amazon Athena: A serverless, pay-per-query service that allows interactive queries directly against data in S3 using standard SQL. It is ideal for the requirement of one-time queries and analytical reporting without the need for provisioning or managing servers.

* Amazon QuickSight: A business intelligence tool that integrates with a wide range of AWS data sources, including Athena, and is used for creating interactive dashboards. It scales well and provides real-time insights for the car listings.

This solution (Option D) is the most cost-effective, because both Glue and Athena are serverless and priced based on usage, reducing costs when compared to provisioning EMR clusters in the other options. Moreover, using Lambda for orchestration is more cost-effective than AWS Step Functions due to its lightweight nature.

References:

- * AWS Glue Documentation
- * Amazon Athena Documentation
- * Amazon QuickSight Documentation
- * S3 Event Notifications and Lambda

NEW QUESTION # 26

A company has five offices in different AWS Regions. Each office has its own human resources (HR) department that uses a unique IAM role. The company stores employee records in a data lake that is based on Amazon S3 storage.

A data engineering team needs to limit access to the records. Each HR department should be able to access records for only employees who are within the HR department's Region.

Which combination of steps should the data engineering team take to meet this requirement with the LEAST operational overhead? (Choose two.)

- A. Modify the IAM roles of the HR departments to add a data filter for each department's Region.
- **B. Enable fine-grained access control in AWS Lake Formation. Add a data filter for each Region.**
- C. Create a separate S3 bucket for each Region. Configure an IAM policy to allow S3 access. Restrict access based on Region.
- D. Use data filters for each Region to register the S3 paths as data locations.
- **E. Register the S3 path as an AWS Lake Formation location.**

Answer: B,E

Explanation:

AWS Lake Formation is a service that helps you build, secure, and manage data lakes on Amazon S3. You can use AWS Lake Formation to register the S3 path as a data lake location, and enable fine-grained access control to limit access to the records based on the HR department's Region. You can use data filters to specify which S3 prefixes or partitions each HR department can access, and grant permissions to the IAM roles of the HR departments accordingly. This solution will meet the requirement with the least operational overhead, as it simplifies the data lake management and security, and leverages the existing IAM roles of the HR departments.

The other options are not optimal for the following reasons:

A: Use data filters for each Region to register the S3 paths as data locations. This option is not possible, as data filters are not used to register S3 paths as data locations, but to grant permissions to access specific S3 prefixes or partitions within a data location. Moreover, this option does not specify how to limit access to the records based on the HR department's Region.

C: Modify the IAM roles of the HR departments to add a data filter for each department's Region. This option is not possible, as data filters are not added to IAM roles, but to permissions granted by AWS Lake Formation. Moreover, this option does not specify how to register the S3 path as a data lake location, or how to enable fine-grained access control in AWS Lake Formation.

E: Create a separate S3 bucket for each Region. Configure an IAM policy to allow S3 access. Restrict access based on Region. This option is not recommended, as it would require more operational overhead to create and manage multiple S3 buckets, and to configure and maintain IAM policies for each HR department. Moreover, this option does not leverage the benefits of AWS Lake Formation, such as data cataloging, data transformation, and data governance.

References:

- 1: AWS Lake Formation
- 2: AWS Lake Formation Permissions
- 3: AWS Identity and Access Management
- 4: Amazon S3

NEW QUESTION # 27

A company is planning to use a provisioned Amazon EMR cluster that runs Apache Spark jobs to perform big data analysis. The company requires high reliability. A big data team must follow best practices for running cost-optimized and long-running workloads on Amazon EMR. The team must find a solution that will maintain the company's current level of performance.

Which combination of resources will meet these requirements MOST cost-effectively? (Choose two.)

- A. Use Hadoop Distributed File System (HDFS) as a persistent data store.
- B. Use Spot Instances for all primary nodes.
- **C. Use Graviton instances for core nodes and task nodes.**
- **D. Use Amazon S3 as a persistent data store.**
- E. Use x86-based instances for core nodes and task nodes.

Answer: C,D

Explanation:

The best combination of resources to meet the requirements of high reliability, cost-optimization, and performance for running Apache Spark jobs on Amazon EMR is to use Amazon S3 as a persistent data store and Graviton instances for core nodes and task nodes.

Amazon S3 is a highly durable, scalable, and secure object storage service that can store any amount of data for a variety of use cases, including big data analytics¹. Amazon S3 is a better choice than HDFS as a persistent data store for Amazon EMR, as it decouples the storage from the compute layer, allowing for more flexibility and cost-efficiency. Amazon S3 also supports data encryption, versioning, lifecycle management, and cross-region replication¹. Amazon EMR integrates seamlessly with Amazon S3, using EMR File System (EMRFS) to access data stored in Amazon S3 buckets². EMRFS also supports consistent view, which enables Amazon EMR to provide read-after-write consistency for Amazon S3 objects that are accessed through EMRFS².

Graviton instances are powered by Arm-based AWS Graviton² processors that deliver up to 40% better price performance over comparable current generation x86-based instances³. Graviton instances are ideal for running workloads that are CPU-bound, memory-bound, or network-bound, such as big data analytics, web servers, and open-source databases³. Graviton instances are compatible with Amazon EMR, and can be used for both core nodes and task nodes. Core nodes are responsible for running the data processing frameworks, such as Apache Spark, and storing data in HDFS or the local file system. Task nodes are optional nodes that can be added to a cluster to increase the processing power and throughput. By using Graviton instances for both core nodes and task nodes, you can achieve higher performance and lower cost than using x86-based instances.

Using Spot Instances for all primary nodes is not a good option, as it can compromise the reliability and availability of the cluster. Spot Instances are spare EC2 instances that are available at up to 90% discount compared to On-Demand prices, but they can be interrupted by EC2 with a two-minute notice when EC2 needs the capacity back. Primary nodes are the nodes that run the cluster software, such as Hadoop, Spark, Hive, and Hue, and are essential for the cluster operation. If a primary node is interrupted by EC2, the cluster will fail or become unstable. Therefore, it is recommended to use On-Demand Instances or Reserved Instances for primary nodes, and use Spot Instances only for task nodes that can tolerate interruptions. References:

Amazon S3 - Cloud Object Storage

EMR File System (EMRFS)

AWS Graviton² Processor-Powered Amazon EC2 Instances

[Plan and Configure EC2 Instances]

[Amazon EC2 Spot Instances]

[Best Practices for Amazon EMR]

NEW QUESTION # 28

A company uses AWS Glue jobs to implement several data pipelines. The pipelines are critical to the company.

The company needs to implement a monitoring mechanism that will alert stakeholders if the pipelines fail.

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

- A. Create an Amazon EventBridge rule to match AWS Glue job failure events. Define an Amazon CloudWatch metric based on the EventBridge rule. Set up a CloudWatch alarm based on the metric to send notifications to an Amazon Simple Notification Service (Amazon SNS) topic.
- B. Configure an Amazon CloudWatch Logs log group for the AWS Glue jobs. Create an Amazon EventBridge rule to match new log creation events in the log group. Configure the rule to send notifications to an Amazon Simple Notification Service (Amazon SNS) topic.
- **C. Create an Amazon EventBridge rule to match AWS Glue job failure events. Configure the rule to target an AWS Lambda function to process events. Configure the function to send notifications to an Amazon Simple Notification Service (Amazon SNS) topic.**
- D. Configure an Amazon CloudWatch Logs log group for the AWS Glue jobs. Create an Amazon EventBridge rule to match new log creation events in the log group. Configure the rule to target an AWS Lambda function that reads the logs and sends notifications to an Amazon Simple Notification Service (Amazon SNS) topic if AWS Glue job failure logs are present.

Answer: C

Explanation:

Creating an EventBridge rule that triggers a Lambda function on AWS Glue job failure events and then sends notifications via

Amazon SNS is the most direct and operationally efficient method:

"Practice Quiz 10: A data engineer must monitor the data pipeline... Which solution will meet these requirements?

A: Inspect the job run monitoring section of the AWS Glue console.

-Ace the AWS Certified Data Engineer - Associate Certification - version 2 - apple.pdf Although this reference directly supports using AWS Glue's monitoring features via EventBridge, it implies that solutions like A-which directly use EventBridge failure events for automation-are more optimal and less complex than constructing custom logs and metrics.

NEW QUESTION # 29

.....

Forget complaining for your failure. Please think about why there are candidates to pass exam every day. Option is more important than effort sometimes. Amazon Data-Engineer-Associate reliable exam collection pdf are being searched about 100,000 in the website every day. There are more than 600 candidates choosing valid Amazon Data-Engineer-Associate reliable exam collection pdf every day. We help thousands of people clear exams every year. The success is close at hand, why do you grab it?

Test Data-Engineer-Associate Answers: <https://www.testkingpass.com/Data-Engineer-Associate-testking-dumps.html>

Amazon Data-Engineer-Associate Valid Test Sample So passing the exam is precondition of holding the important certificate, Amazon Data-Engineer-Associate Valid Test Sample How to solve this problem is your new trouble, The Amazon introduces changes in the Data-Engineer-Associate format and topics, which are reported to our valued customers, Besides, we promise you full refund if you lose exam with our Data-Engineer-Associate test questions.

Windows offers an auto hide" mode for the taskbar, The problem Data-Engineer-Associate is the calculation of passer ratings, So passing the exam is precondition of holding the important certificate.

How to solve this problem is your new trouble, The Amazon introduces changes in the Data-Engineer-Associate format and topics, which are reported to our valued customers.

Excellent Data-Engineer-Associate Valid Test Sample - Win Your Amazon Certificate with Top Score

Besides, we promise you full refund if you lose exam with our Data-Engineer-Associate test questions, Align ourselves with most adroit experts in the market, and our Data-Engineer-Associate practice materials have gained great reputation among the market for over ten years.

- 2026 Data-Engineer-Associate Valid Test Sample Free PDF | Professional Test Data-Engineer-Associate Answers: AWS Certified Data Engineer - Associate (DEA-C01) x Enter ➡ www.prepawayete.com □ and search for ☀ Data-Engineer-Associate □☀□ to download for free □Data-Engineer-Associate Test Question
- Latest Data-Engineer-Associate Study Guide □ Data-Engineer-Associate Test Question □ Latest Data-Engineer-Associate Test Cost □ Enter ➤ www.pdfvce.com □ and search for 【 Data-Engineer-Associate 】 to download for free □Valid Data-Engineer-Associate Exam Camp
- Latest Data-Engineer-Associate Study Guide □ Exam Data-Engineer-Associate Papers □ Data-Engineer-Associate Reliable Test Cost □ Download □ Data-Engineer-Associate □ for free by simply searching on (www.troytecdumps.com) □Exam Data-Engineer-Associate Guide
- Data-Engineer-Associate Certification Dumps - Data-Engineer-Associate Study Guide Files - Data-Engineer-Associate Practice Test Questions □ Open ➡ www.pdfvce.com □ and search for ➡ Data-Engineer-Associate □ to download exam materials for free □Latest Data-Engineer-Associate Test Cost
- 2026 Data-Engineer-Associate Valid Test Sample | Latest Amazon Test Data-Engineer-Associate Answers: AWS Certified Data Engineer - Associate (DEA-C01) □ Immediately open □ www.vce4dumps.com □ and search for ⇒ Data-Engineer-Associate ⇐ to obtain a free download □Latest Data-Engineer-Associate Study Guide
- Valid Data-Engineer-Associate Exam Camp □ Data-Engineer-Associate Reliable Test Cost ☒ Exam Dumps Data-Engineer-Associate Free ☒ Open (www.pdfvce.com) and search for ☀ Data-Engineer-Associate □☀□ to download exam materials for free □Exam Data-Engineer-Associate Papers
- Data-Engineer-Associate Certification Dumps - Data-Engineer-Associate Study Guide Files - Data-Engineer-Associate Practice Test Questions □ Easily obtain free download of ➡ Data-Engineer-Associate □□□ by searching on 「 www.dumpsmaterials.com 」 □Latest Data-Engineer-Associate Test Cost
- Pass Guaranteed Data-Engineer-Associate - AWS Certified Data Engineer - Associate (DEA-C01) Useful Valid Test Sample □ Immediately open ☀ www.pdfvce.com □☀□ and search for □ Data-Engineer-Associate □ to obtain a free download □New Data-Engineer-Associate Test Guide
- Pass Guaranteed Data-Engineer-Associate - AWS Certified Data Engineer - Associate (DEA-C01) Useful Valid Test

Data-Engineer-Associate Labs □ Data-Engineer-Associate Reliable Test Cost □ Training Data-Engineer-Associate Solutions □ The page for free download of 「 Data-Engineer-Associate 」 on (www.pdfvce.com) will open immediately □ Data-Engineer-Associate Latest Exam Pdf

- P.S. Free & New Data-Engineer-Associate dumps are available on Google Drive shared by TestkingPass:
<https://drive.google.com/open?id=1BMszJgh8hqUonfApu0Ka0nbCKNSdcyrV>

P.S. Free & New Data-Engineer-Associate dumps are available on Google Drive shared by TestkingPass: <https://drive.google.com/open?id=1BMszJgh8hqUonfApu0Ka0nbCKNSdcyrV>