Valid Data-Engineer-Associate Vce Dumps - Data-Engineer-Associate Test Pdf

Actually, most people do not like learning the boring knowledge. It is hard to understand if our brain rejects taking the initiative. Now, our company has researched the Data-Engineer-Associate study materials, a kind of high efficient learning tool. Firstly, we have deleted all irrelevant knowledge, which decreases your learning pressure. Then, the difficult questions of the Data-Engineer-Associate Study Materials will have vivid explanations. So you will have a better understanding after you carefully see the explanations.

By using the Pass4Test Data-Engineer-Associate valid exam lab questions, there is no need to purchase anything else or attend expensive training. We promise that you can pass the Data-Engineer-Associate certification at the first try. We will update our questions and answers in time after thoroughly analysis of latest real exams, so the Data-Engineer-Associate Exam Dumps shown front of you is the latest and valid. Besides, we offer you one year free update Amazon study dumps after your purchase.

>> Valid Data-Engineer-Associate Vce Dumps <<

Data-Engineer-Associate Training Materials are Your Excellent Chance to Master More Useful Knowledge - Pass4Test

Our Data-Engineer-Associate training materials are the latest, valid and accurate study material for candidates who are eager to clear Data-Engineer-Associate exams. You can actually grasp the shortest time to do as much interesting and effective things you like as possible. Data-Engineer-Associate real questions are high value & high pass rate with competitive price products. And our pass rate of Data-Engineer-Associate Study Guide is as high as 99% to 100%. As long as you study with our Data-Engineer-Associate exam questions, you will pass the Data-Engineer-Associate exam easily.

Amazon AWS Certified Data Engineer - Associate (DEA-C01) Sample Questions (Q133-Q138):

NEW QUESTION # 133

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 # 134

A company is developing machine learning (ML) models. A data engineer needs to apply data quality rules to training data. The company stores the training data in an Amazon S3 bucket.

- A. Create an AWS Lambda function to check data quality and to raise exceptions in the code.
- B. Create AWS Lambda functions to evaluate data quality rules and orchestrate with AWS Step Functions.
- C. Create an Amazon EMR provisioned cluster. Add a Python data quality package.
- D. Create an AWS Glue DataBrew project for the data in the S3 bucket. Create a ruleset for the data quality rules. Create a profile job to run the data quality rules. Use Amazon EventBridge to run the profile job when data is added to the S3 bucket.

Answer: D

Explanation:
AWS Glue DataBrew provides a no-code way to define and run data quality rulesets for data stored in S3.
You can trigger profiling jobs via Amazon EventBridge on new uploads for automated checks.
"Use AWS Glue DataBrew to define and run data quality rules on S3 datasets with minimal coding effort.
Automate validation by triggering jobs through EventBridge."
- Ace the AWS Certified Data Engineer - Associate Certification - version 2 - apple.pdf

NEW QUESTION # 135

A company needs to partition the Amazon S3 storage that the company uses for a data lake. The partitioning will use a path of the S3 object keys in the following format: s3://bucket/prefix/year=2023/month=01/day=01.
A data engineer must ensure that the AWS Glue Data Catalog synchronizes with the S3 storage when the company adds new partitions to the bucket.
Which solution will meet these requirements with the LEAST latency?

- A. Run the MSCK REPAIR TABLE command from the AWS Glue console.
- B. Schedule an AWS Glue crawler to run every morning.
- C. Use code that writes data to Amazon S3 to invoke the Boto3 AWS Glue create partition API call.

- D. Manually run the AWS Glue CreatePartition API twice each day.

Answer: B

Explanation:
The best solution to ensure that the AWS Glue Data Catalog synchronizes with the S3 storage when the company adds new partitions to the bucket with the least latency is to use code that writes data to Amazon S3 to invoke the Boto3 AWS Glue create partition API call. This way, the Data Catalog is updated as soon as new data is written to S3, and the partition information is immediately available for querying by other services. The Boto3 AWS Glue create partition API call allows you to create a new partition in the Data Catalog by specifying the table name, the database name, and the partition values1. You can use this API call in your code that writes data to S3, such as a Python script or an AWS Glue ETL job, to create a partition for each new S3 object key that matches the partitioning scheme.
Option A is not the best solution, as scheduling an AWS Glue crawler to run every morning would introduce a significant latency between the time new data is written to S3 and the time the Data Catalog is updated. AWS Glue crawlers are processes that connect to a data store, progress through a prioritized list of classifiers to determine the schema for your data, and then create metadata tables in the Data Catalog2. Crawlers can be scheduled to run periodically, such as daily or hourly, but they cannot run continuously or in real-time.
Therefore, using a crawler to synchronize the Data Catalog with the S3 storage would not meet the requirement of the least latency.
Option B is not the best solution, as manually running the AWS Glue CreatePartition API twice each day would also introduce a significant latency between the time new data is written to S3 and the time the Data Catalog is updated. Moreover, manually running the API would require more operational overhead and human intervention than using code that writes data to S3 to invoke the API automatically.
Option D is not the best solution, as running the MSCK REPAIR TABLE command from the AWS Glue console would also introduce a significant latency between the time new data is written to S3 and the time the Data Catalog is updated. The MSCK REPAIR TABLE command is a SQL command that you can run in the AWS Glue console to add partitions to the Data Catalog based on the S3 object keys that match the partitioning scheme3. However, this command is not meant to be run frequently or in real-time, as it can take a long time to scan the entire S3 bucket and add the partitions. Therefore, using this command to synchronize the Data Catalog with the S3 storage would not meet the requirement of the least latency. References:
* AWS Glue CreatePartition API
* Populating the AWS Glue Data Catalog
* MSCK REPAIR TABLE Command
* AWS Certified Data Engineer - Associate DEA-C01 Complete Study Guide

NEW QUESTION # 136
A data engineer uses Amazon Kinesis Data Streams to ingest and process records that contain user behavior data from an application every day.
The data engineer notices that the data stream is experiencing throttling because hot shards receive much more data than other shards in the data stream.
How should the data engineer resolve the throttling issue?

- A. Decrease the size of the records that the producer sends to match the capacity of the stream.
- B. Increase the number of shards in the data stream. Distribute the records across the shards.
- C. Limit the number of records that are sent each second by the producer to match the capacity of the stream.
- D. Use a random partition key to distribute the ingested records.

Answer: C

NEW QUESTION # 137
A company stores data in a data lake that is in Amazon S3. Some data that the company stores in the data lake contains personally identifiable information (PII). Multiple user groups need to access the raw dat a. The company must ensure that user groups can access only the PII that they require.
Which solution will meet these requirements with the LEAST effort?

- A. Use Amazon QuickSight to access the data. Use column-level security features in QuickSight to limit the PII that users can retrieve from Amazon S3 by using Amazon Athena. Define QuickSight access levels based on the PII access requirements of the users.
- B. Create IAM roles that have different levels of granular access. Assign the IAM roles to IAM user groups. Use an identity-based policy to assign access levels to user groups at the column level.
- C. Build a custom query builder UI that will run Athena queries in the background to access the data. Create user groups in

Amazon Cognito. Assign access levels to the user groups based on the PII access requirements of the users.
- D. Use Amazon Athena to query the data. Set up AWS Lake Formation and create data filters to establish levels of access for the company's IAM roles. Assign each user to the IAM role that matches the user's PII access requirements.

Answer: D

Explanation:
Amazon Athena is a serverless, interactive query service that enables you to analyze data in Amazon S3 using standard SQL. AWS Lake Formation is a service that helps you build, secure, and manage data lakes on AWS. You can use AWS Lake Formation to create data filters that define the level of access for different IAM roles based on the columns, rows, or tags of the data. By using Amazon Athena to query the data and AWS Lake Formation to create data filters, the company can meet the requirements of ensuring that user groups can access only the PII that they require with the least effort. The solution is to use Amazon Athena to query the data in the data lake that is in Amazon S3. Then, set up AWS Lake Formation and create data filters to establish levels of access for the company's IAM roles. For example, a data filter can allow a user group to access only the columns that contain the PII that they need, such as name and email address, and deny access to the columns that contain the PII that they do not need, such as phone number and social security number. Finally, assign each user to the IAM role that matches the user's PII access requirements. This way, the user groups can access the data in the data lake securely and efficiently. The other options are either not feasible or not optimal. Using Amazon QuickSight to access the data (option B) would require the company to pay for the QuickSight service and to configure the column-level security features for each user. Building a custom query builder UI that will run Athena queries in the background to access the data (option C) would require the company to develop and maintain the UI and to integrate it with Amazon Cognito. Creating IAM roles that have different levels of granular access (option D) would require the company to manage multiple IAM roles and policies and to ensure that they are aligned with the data schema. Reference:
Amazon Athena
AWS Lake Formation
AWS Certified Data Engineer - Associate DEA-C01 Complete Study Guide, Chapter 4: Data Analysis and Visualization, Section 4.3: Amazon Athena

NEW QUESTION # 138
......

However, when asked whether the Data-Engineer-Associate latest dumps are reliable, costumers may be confused. For us, we strongly recommend the Data-Engineer-Associate exam questions compiled by our company, here goes the reason. On one hand, our Data-Engineer-Associate test material owns the best quality. When it comes to the study materials selling in the market, qualities are patchy. But our Amazon test material has been recognized by multitude of customers, which possess of the top-class quality, can help you pass exam successfully. On the other hand, our Data-Engineer-Associate Latest Dumps are designed by the most experienced experts, thus it can not only teach you knowledge, but also show you the method of learning in the most brief and efficient ways.

Data-Engineer-Associate Test Pdf: https://www.pass4test.com/Data-Engineer-Associate.html

Before you buy our product, you can download and try out it freely so you can have a good understanding of our Data-Engineer-Associate quiz prep, You only need little time and energy to review and prepare for the exam if you use our Data-Engineer-Associate prep torrent as the studying materials, Already a Member, As more people realize the importance of Amazon Data-Engineer-Associate Test Pdf certificate, many companies raise their prices, No matter when you have questions or problem about our Data-Engineer-Associate test questions and dumps, we will be pleased to reply and solve with you in three hours.

Selective Sharpening , Working with the Filmstrip, Before you buy our product, you can download and try out it freely so you can have a good understanding of our Data-Engineer-Associate Quiz prep.

You only need little time and energy to review and prepare for the exam if you use our Data-Engineer-Associate prep torrent as the studying materials, Already a Member, As more people Data-Engineer-Associate realize the importance of Amazon certificate, many companies raise their prices.

Free PDF Data-Engineer-Associate - High Hit-Rate Valid AWS Certified Data Engineer - Associate (DEA-C01) Vce Dumps

No matter when you have questions or problem about our Data-Engineer-Associate test questions and dumps, we will be pleased to reply and solve with you in three hours.

- Latest Valid Data-Engineer-Associate Vce Dumps - How to Download for PDF Free Data-Engineer-Associate Test Pdf 🔲

- ▢ Search for [Data-Engineer-Associate] and download it for free on ➤ www.torrentvce.com ▢ website ▢Latest Data-Engineer-Associate Practice Materials
- Reliable Data-Engineer-Associate Test Question ▢ Data-Engineer-Associate Exam Objectives ▢ Free Data-Engineer-Associate Exam ▢ Search for ➠ Data-Engineer-Associate ▢ and easily obtain a free download on ➠ www.pdfvce.com ▢ ▢Data-Engineer-Associate Reliable Dumps Ppt
- Data-Engineer-Associate Latest Exam Tips ▢ Data-Engineer-Associate Dumps PDF ▢ Data-Engineer-Associate Latest Exam Book ▢ ✔ www.examdiscuss.com ▢✔ ▢ is best website to obtain ➠ Data-Engineer-Associate ▢ for free download ▢Data-Engineer-Associate Dumps PDF
- Pdfvce Amazon Data-Engineer-Associate Dumps - Improve Your Exam Preparation Quickly ▢ Search for ➠ Data-Engineer-Associate ▢ and download it for free on ➠ www.pdfvce.com ▢▢▢ website ▢Data-Engineer-Associate Latest Exam Discount
- New Data-Engineer-Associate Exam Vce ▢ Reliable Data-Engineer-Associate Study Plan ▢ Data-Engineer-Associate Relevant Exam Dumps ▢ Search for ☀ Data-Engineer-Associate ▢☀ and easily obtain a free download on 【 www.validtorrent.com 】 ▢Latest Real Data-Engineer-Associate Exam
- 100% Pass 2026 Amazon Data-Engineer-Associate: AWS Certified Data Engineer - Associate (DEA-C01) –Efficient Valid Vce Dumps ▢ Open ➠ www.pdfvce.com ▢ and search for [Data-Engineer-Associate] to download exam materials for free ▢Latest Real Data-Engineer-Associate Exam
- Amazon Valid Data-Engineer-Associate Vce Dumps Spend Your Little Time and Energy to Pass Data-Engineer-Associate exam ▢ Download ▢ Data-Engineer-Associate ▢ for free by simply searching on ➠ www.vce4dumps.com ▢ ▢Free Data-Engineer-Associate Exam
- Download Amazon Data-Engineer-Associate Exam Dumps Instantly ▢ Search on （ www.pdfvce.com ） for ➠ Data-Engineer-Associate ▢ to obtain exam materials for free download ▢Reliable Data-Engineer-Associate Study Plan
- Download Amazon Data-Engineer-Associate Exam Dumps Instantly ▢ Go to website ➠ www.exam4labs.com ▢ open and search for ▸ Data-Engineer-Associate ◂ to download for free ▢Data-Engineer-Associate Latest Exam Discount
- High Pass Rate Data-Engineer-Associate Exam Questions Convey All Important Information of Data-Engineer-Associate Exam ▢ Download ▷ Data-Engineer-Associate ◁ for free by simply searching on ▢ www.pdfvce.com ▢ ▢Data-Engineer-Associate Exam Objectives
- New Data-Engineer-Associate Exam Vce ▢ Latest Real Data-Engineer-Associate Exam ▢ New Data-Engineer-Associate Exam Questions ▢ Open website ➤ www.dumpsquestion.com ▢ and search for ➠ Data-Engineer-Associate ▢ for free download ▢Data-Engineer-Associate Valid Study Questions
- 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, 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, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, myportal.utt.edu.tt, Disposable vapes

P.S. Free & New Data-Engineer-Associate dumps are available on Google Drive shared by Pass4Test:
https://drive.google.com/open?id=1XvSeO--0pI7Pd9xykIdClEgPI8Qn32RN