Data-Engineer-Associate Online Training Materials | Data-Engineer-Associate Test Answers

What's more, part of that ITExamSimulator Data-Engineer-Associate dumps now are free: https://drive.google.com/open?id=1XP9doBLOFulJ4V7Ex_LjH_SkSbooL9iz

Before you place orders, you can download the free demos of Data-Engineer-Associate practice test as experimental acquaintance. Once you decide to buy, you will have many benefits like free update lasting one-year and convenient payment mode. We will inform you immediately once there are latest versions of Data-Engineer-Associate Test Question released. And if you get any questions, please get contact with us, our staff will be online 24/7 to solve your problems all the way.

Our Data-Engineer-Associate exam Braindumps are available in PDF, software, and online three modes, which allowing you to switch learning materials on paper, on your phone or on your computer, and to study anywhere and anytime. And in any version of Data-Engineer-Associate practice materials, the number of downloads and the number of people used at the same time are not

limited. You can practice repeatedly for the same set of Data-Engineer-Associate Questions and continue to consolidate important knowledge points.

>> Data-Engineer-Associate Online Training Materials <<

Data-Engineer-Associate Test Answers, Data-Engineer-Associate Latest Dumps Questions

There are three effect versions of the date available for candidates who want to pass the Data-Engineer-Associate exam. PDF, APP and Software, each version has its advantage, and each version is the most effect way. You can choose the most suitable version to learn. Of course, if you want to, you can choose more than one version to prepare your Data-Engineer-Associate exam. Our exam materials allow you to prepare for the Real Data-Engineer-Associate Exam and will help you with the self-assessment. If you like use paper to learn, you can print in PDF; if you like learn with electronic equipment, you can use our APP online version offline. Our Amazon practice test software will give you a real exam environment with multiple learning tools that allow you to do a selective study and will help you to get the job that you are looking for.

Amazon AWS Certified Data Engineer - Associate (DEA-C01) Sample Questions (Q62-Q67):

NEW QUESTION # 62
A transportation company wants to track vehicle movements by capturing geolocation records. The records are 10 bytes in size. The company receives up to 10,000 records every second. Data transmission delays of a few minutes are acceptable because of unreliable network conditions.
The transportation company wants to use Amazon Kinesis Data Streams to ingest the geolocation data. The company needs a reliable mechanism to send data to Kinesis Data Streams. The company needs to maximize the throughput efficiency of the Kinesis shards.
Which solution will meet these requirements in the MOST operationally efficient way?

- A. Kinesis Producer Library (KPL)
- B. Kinesis Agent
- C. Amazon Data Firehose
- D. Kinesis SDK

Answer: A

Explanation:
Problem Analysis:
The company ingests geolocation records (10 bytes each) at 10,000 records per second into Kinesis Data Streams.
Data transmission delays are acceptable, but the solution must maximize throughput efficiency.
Key Considerations:
The Kinesis Producer Library (KPL) batches records and uses aggregation to optimize shard throughput.
Efficiently handles high-throughput scenarios with minimal operational overhead.
Solution Analysis:
Option A: Kinesis Agent
Designed for file-based ingestion; not optimized for geolocation records.
Option B: KPL
Aggregates records into larger payloads, significantly improving shard throughput.
Suitable for applications generating small, high-frequency records.
Option C: Kinesis Firehose
Firehose is for delivery to destinations like S3 or Redshift and is not optimized for direct ingestion to Kinesis Data Streams.
Option D: Kinesis SDK
The SDK lacks advanced features like aggregation, resulting in lower throughput efficiency.
Final Recommendation:
Use Kinesis Producer Library (KPL) for its built-in aggregation and batching capabilities.
Reference:
Kinesis Producer Library (KPL) Overview

NEW QUESTION # 63
A company stores CSV files in an Amazon S3 bucket. A data engineer needs to process the data in the CSV files and store the processed data in a new S3 bucket.
The process needs to rename a column, remove specific columns, ignore the second row of each file, create a new column based on the values of the first row of the data, and filter the results by a numeric value of a column.
Which solution will meet these requirements with the LEAST development effort?

- A. Use an AWS Glue workflow to build a set of jobs to crawl and transform the CSV files.
- B. Use AWS Glue Python jobs to read and transform the CSV files.
- C. Use an AWS Glue custom crawler to read and transform the CSV files.
- D. Use AWS Glue DataBrew recipes to read and transform the CSV files.

Answer: D

NEW QUESTION # 64
A company is planning to migrate on-premises Apache Hadoop clusters to Amazon EMR. The company also needs to migrate a data catalog into a persistent storage solution.
The company currently stores the data catalog in an on-premises Apache Hive metastore on the Hadoop clusters. The company requires a serverless solution to migrate the data catalog.
Which solution will meet these requirements MOST cost-effectively?

- A. Use AWS Database Migration Service (AWS DMS) to migrate the Hive metastore into Amazon S3. Configure AWS Glue Data Catalog to scan Amazon S3 to produce the data catalog.
- B. Configure a Hive metastore in Amazon EMR. Migrate the existing on-premises Hive metastore into Amazon EMR. Use AWS Glue Data Catalog to store the company's data catalog as an external data catalog.
- C. Configure an external Hive metastore in Amazon EMR. Migrate the existing on-premises Hive metastore into Amazon EMR. Use Amazon Aurora MySQL to store the company's data catalog.
- D. Configure a new Hive metastore in Amazon EMR. Migrate the existing on-premises Hive metastore into Amazon EMR. Use the new metastore as the company's data catalog.

Answer: A

Explanation:
AWS Database Migration Service (AWS DMS) is a service that helps you migrate databases to AWS quickly and securely. You can use AWS DMS to migrate the Hive metastore from the on-premises Hadoop clusters into Amazon S3, which is a highly scalable, durable, and cost-effective object storage service. AWS Glue Data Catalog is a serverless, managed service that acts as a central metadata repository for your data assets. You can use AWS Glue Data Catalog to scan the Amazon S3 bucket that contains the migrated Hive metastore and create a data catalog that is compatible with Apache Hive and other AWS services. This solution meets the requirements of migrating the data catalog into a persistent storage solution and using a serverless solution. This solution is also the most cost-effective, as it does not incur any additional charges for running Amazon EMR or Amazon Aurora MySQL clusters. The other options are either not feasible or not optimal. Configuring a Hive metastore in Amazon EMR (option B) or an external Hive metastore in Amazon EMR (option C) would require running and maintaining Amazon EMR clusters, which would incur additional costs and complexity. Using Amazon Aurora MySQL to store the company's data catalog (option C) would also incur additional costs and complexity, as well as introduce compatibility issues with Apache Hive. Configuring a new Hive metastore in Amazon EMR (option D) would not migrate the existing data catalog, but create a new one, which would result in data loss and inconsistency. Reference:
Using AWS Database Migration Service
Populating the AWS Glue Data Catalog
AWS Certified Data Engineer - Associate DEA-C01 Complete Study Guide, Chapter 4: Data Analysis and Visualization, Section 4.2: AWS Glue Data Catalog

NEW QUESTION # 65
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. 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.
- B. 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).
- C. 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.
- D. 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.
- E. 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.

Answer: A,C

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 # 66
A media company wants to improve a system that recommends media content to customer based on user behavior and preferences. To improve the recommendation system, the company needs to incorporate insights from third-party datasets into the company's existing analytics platform.
The company wants to minimize the effort and time required to incorporate third-party datasets.
Which solution will meet these requirements with the LEAST operational overhead?

- A. Use Amazon Kinesis Data Streams to access and integrate third-party datasets from Amazon Elastic Container Registry (Amazon ECR).
- B. Use API calls to access and integrate third-party datasets from AWS Data Exchange.
- C. Use API calls to access and integrate third-party datasets from AWS
- D. Use Amazon Kinesis Data Streams to access and integrate third-party datasets from AWS CodeCommit repositories.

Answer: B

Explanation:
AWS Data Exchange is a service that makes it easy to find, subscribe to, and use third-party data in the cloud.
It provides a secure and reliable way to access and integrate data from various sources, such as data providers, public datasets, or AWS services. Using AWS Data Exchange, you can browse and subscribe to data products that suit your needs, and then use API calls or the AWS Management Console to export the data to Amazon S3, where you can use it with your existing analytics platform. This solution minimizes the effort and time required to incorporate third-party datasets, as you do not need to set up and manage data pipelines, storage, or access controls. You also benefit from the data quality and freshness provided by the data providers, who can update their data products as frequently as needed12.
The other options are not optimal for the following reasons:
* B. Use API calls to access and integrate third-party datasets from AWS. This option is vague and does not specify which AWS service or feature is used to access and integrate third-party datasets. AWS offers a variety of services and features that can help with data ingestion, processing, and analysis, but not all of them are suitable for the given scenario. For example, AWS Glue is a serverless data integration service that can help you discover, prepare, and combine data from various sources, but it requires you to create and run data extraction, transformation, and loading (ETL) jobs, which can add operational overhead3.
* C. Use Amazon Kinesis Data Streams to access and integrate third-party datasets from AWS CodeCommit repositories. This option is not feasible, as AWS CodeCommit is a source control service that hosts secure Git-based repositories, not a data source that can be accessed by Amazon Kinesis Data Streams. Amazon Kinesis Data Streams is a service that enables you to capture, process, and analyze data streams in real time, such as clickstream data, application logs, or IoT telemetry. It does not support accessing and integrating data from AWS CodeCommit repositories, which are meant for storing and managing code, not data .
* D. Use Amazon Kinesis Data Streams to access and integrate third-party datasets from Amazon Elastic Container Registry (Amazon ECR). This option is also not feasible, as Amazon ECR is a fully managed container registry service that stores, manages, and deploys container images, not a data source that can be accessed by Amazon Kinesis Data Streams. Amazon Kinesis Data Streams does not support accessing and integrating data from Amazon ECR, which is meant for storing and managing container images, not data .
:
1: AWS Data Exchange User Guide
2: AWS Data Exchange FAQs
3: AWS Glue Developer Guide
4: AWS CodeCommit User Guide
5: Amazon Kinesis Data Streams Developer Guide
6: Amazon Elastic Container Registry User Guide
7: Build a Continuous Delivery Pipeline for Your Container Images with Amazon ECR as Source

NEW QUESTION # 67
.....

Training Materials It is essential to equip yourself with international admitted certifications.

Your desktop standard is Windows XP Professional, Cisco Modular Network Architecture, Our supporter of Data-Engineer-Associate ITExamSimulator Pass Guide study guide has exceeded Data-Engineer-Associate tens of thousands around the world, which directly reflects the quality of them.

Quiz 2026 Amazon Data-Engineer-Associate: Reliable AWS Certified Data Engineer - Associate (DEA-C01) Online Training Materials

For one thing, the most advanced operation system in our company which can assure you the fastest delivery speed on our Data-Engineer-Associate Exam Questions, It is essential to equip yourself with international admitted certifications.

High-quality contents and flexible choices of learning mode would bring about the convenience and easiness for you, This is our target that helps you to make it easier to get Data-Engineer-Associate certification and you can find job more easily.

- Newest Data-Engineer-Associate Online Training Materials | 100% Free Data-Engineer-Associate Test Answers ✔□ Open ➤ www.practicevce.com □ enter ➤ Data-Engineer-Associate □ and obtain a free download □Practice Data-Engineer-Associate Engine
- Practice Data-Engineer-Associate Engine □ Test Data-Engineer-Associate Questions Answers □ Data-Engineer-Associate Latest Exam Question □ The page for free download of ➠ Data-Engineer-Associate □ on 「www.pdfvce.com」 will open immediately □Practice Data-Engineer-Associate Engine
- Updated Data-Engineer-Associate Online Training Materials – Pass Data-Engineer-Associate First Attempt □ Search on 「www.troytecdumps.com」 for 「Data-Engineer-Associate」 to obtain exam materials for free download □Data-Engineer-Associate Test Assessment
- Data-Engineer-Associate Test Assessment □ Latest Data-Engineer-Associate Examprep □ Practice Data-Engineer-Associate Engine □ Search for ▶ Data-Engineer-Associate ◀ and download it for free on □ www.pdfvce.com □ website □Test Data-Engineer-Associate Dumps
- Practice Data-Engineer-Associate Engine □ Dumps Data-Engineer-Associate Collection □ Passing Data-Engineer-Associate Score □ Search for ☀ Data-Engineer-Associate □☀□ and download it for free on ✔ www.pdfdumps.com □✔□ website □Data-Engineer-Associate New Exam Camp
- Buy Amazon Data-Engineer-Associate Real Exam Dumps Today and Get Massive Benefits □ Easily obtain free download of ▶ Data-Engineer-Associate ◀ by searching on ➡ www.pdfvce.com □ □Exam Data-Engineer-Associate Blueprint
- Amazon Data-Engineer-Associate Exam | Data-Engineer-Associate Online Training Materials - Full Refund if Failing Data-Engineer-Associate: AWS Certified Data Engineer - Associate (DEA-C01) Exam □ Copy URL （www.prep4away.com） open and search for { Data-Engineer-Associate } to download for free □Latest Data-Engineer-Associate Examprep
- Data-Engineer-Associate Reliable Test Question □ Practice Data-Engineer-Associate Engine □ Practice Data-Engineer-Associate Engine □ Open website □ www.pdfvce.com □ and search for □ Data-Engineer-Associate □ for free download □Real Data-Engineer-Associate Exam Answers
- Passing Data-Engineer-Associate Score □ Real Data-Engineer-Associate Exam Answers □ Data-Engineer-Associate Quiz □ Open { www.prepawaypdf.com } enter ➠ Data-Engineer-Associate □ and obtain a free download □Test Data-Engineer-Associate Assessment
- Latest AWS Certified Data Engineer - Associate (DEA-C01) exam pdf, Data-Engineer-Associate practice exam □ Search for { Data-Engineer-Associate } on ▶ www.pdfvce.com ◀ immediately to obtain a free download □Test Data-Engineer-Associate Dumps
- 100% Pass Quiz 2026 The Best Amazon Data-Engineer-Associate Online Training Materials □ Search for □ Data-Engineer-Associate □ and obtain a free download on （www.examcollectionpass.com） □Data-Engineer-Associate Quiz
- myportal.utt.edu.tt, www.stes.tyc.edu.tw, myportal.utt.edu.tt, bbs.t-firefly.com, 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, Disposable vapes

What's more, part of that ITExamSimulator Data-Engineer-Associate dumps now are free: https://drive.google.com/open?id=1XP9doBLOFulJ4V7Ex_LjH_SkSbooL9iz