Free Amazon Data-Engineer-Associate Exam Questions Updates and Demos

Databricks Certified Data Engineer Associate Exam Dumps 2024

Databricks Certified Data Engineer Associate Practice Tests 2024. Contains 760+ exam questions to pass the exam in first attempt.

SkillCertPro offers real exam questions for practice for all major IT certifications.

- For a full set of 760+ questions. Go to
  https://skillcertpro.com/product/databricks-certified-data-engineer-associate-exam-questions/
- SkillCertPro offers detailed explanations to each question which helps to understand the concepts better.
- It is recommended to score above 85% in SkillCertPro exams before attempting a real exam.
- SkillCertPro updates exam questions every 2 weeks.
- You will get life time access and life time free updates
- SkillCertPro assures 100% pass guarantee in first attempt.

Below are the free 10 sample questions.

Question 1:

How does a Delta Lake differ from a traditional data lake?

A. Delta lake is Datawarehouse service on top of data lake that can provide reliability, security, and performance
B. Delta lake is a caching layer on top of data lake that can provide reliability, security, and performance
C. Delta lake is an open storage format like parquet with additional capabilities that can provide reliability, security, and performance
D. Delta lake is an open storage format designed to replace flat files with additional capabilities that can provide reliability, security, and performance
E. Delta lake is proprietary software designed by Databricks that can provide reliability, security, and performance

P.S. Free & New Data-Engineer-Associate dumps are available on Google Drive shared by Actual4Exams:
https://drive.google.com/open?id=1Jyg0z3QcoygVJpqIPbB__7d7lbC6jtr2

If you don't prepare with real Data-Engineer-Associate questions, you fail, lose time and money. Actual4Exams product is specially designed to help you pass the exam on the first try. The study material is easy to use. You can choose from 3 different formats available according to your needs. The 3 formats are Amazon Data-Engineer-Associate desktop practice test software, browser based practice exam, and PDF.

The AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) practice questions are designed by experienced and qualified AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) exam trainers. They have the expertise, knowledge, and experience to design and maintain the top standard of AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) exam dumps. So rest assured that with the AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) exam real questions you can not only ace your AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) exam dumps preparation but also get deep insight knowledge about Amazon Data-Engineer-Associate exam topics. So download AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) exam questions now and start this journey.

>> Valid Real Data-Engineer-Associate Exam <<

Valid Real Data-Engineer-Associate Exam | Pass-Sure Data-Engineer-Associate Valid Exam Vce Free: AWS Certified Data Engineer - Associate

(DEA-C01) 100% Pass

After you really improve your strength, you will find that your strength can bring you many benefits. Users of our Data-Engineer-Associate practice prep can prove this to you. You have to believe that your strength matches the opportunities you have gained. And the opportunities you get are the basic prerequisite for your promotion and salary increase. After you use our Data-Engineer-Associate Exam Materials, you will more agree with this. With the help of our Data-Engineer-Associate study guide, nothing is impossible to you.

Amazon AWS Certified Data Engineer - Associate (DEA-C01) Sample Questions (Q82-Q87):

NEW QUESTION # 82

A company uses an Amazon Redshift provisioned cluster as its database. The Redshift cluster has five reserved ra3.4xlarge nodes and uses key distribution.

A data engineer notices that one of the nodes frequently has a CPU load over 90%. SQL Queries that run on the node are queued. The other four nodes usually have a CPU load under 15% during daily operations.

The data engineer wants to maintain the current number of compute nodes. The data engineer also wants to balance the load more evenly across all five compute nodes.

Which solution will meet these requirements?

- A. Change the primary key to be the data column that is most often used in a WHERE clause of the SQL SELECT statement.
- B. Change the sort key to be the data column that is most often used in a WHERE clause of the SQL SELECT statement.
- C. Upgrade the reserved node from ra3.4xlarqe to ra3.16xlarqe.
- D. Change the distribution key to the table column that has the largest dimension.

Answer: D

Explanation:
Changing the distribution key to the table column that has the largest dimension will help to balance the load more evenly across all five compute nodes. The distribution key determines how the rows of a table are distributed among the slices of the cluster. If the distribution key is not chosen wisely, it can cause data skew, meaning some slices will have more data than others, resulting in uneven CPU load and query performance.
By choosing the table column that has the largest dimension, meaning the column that has the most distinct values, as the distribution key, the data engineer can ensure that the rows are distributed more uniformly across the slices, reducing data skew and improving query performance.
The other options are not solutions that will meet the requirements. Option A, changing the sort key to be the data column that is most often used in a WHERE clause of the SQL SELECT statement, will not affect the data distribution or the CPU load. The sort key determines the order in which the rows of a table are stored on disk, which can improve the performance of range-restricted queries, but not the load balancing. Option C, upgrading the reserved node from ra3.4xlarge to ra3.16xlarge, will not maintain the current number of compute nodes, as it will increase the cost and the capacity of the cluster. Option D, changing the primary key to be the data column that is most often used in a WHERE clause of the SQL SELECT statement, will not affect the data distribution or the CPU load either. The primary key is a constraint that enforces the uniqueness of the rows in a table, but it does not influence the data layout or the query optimization. References:
Choosing a data distribution style
Choosing a data sort key
Working with primary keys

NEW QUESTION # 83

A company's data engineer needs to optimize the performance of table SQL queries. The company stores data in an Amazon Redshift cluster. The data engineer cannot increase the size of the cluster because of budget constraints.
The company stores the data in multiple tables and loads the data by using the EVEN distribution style. Some tables are hundreds of gigabytes in size. Other tables are less than 10 MB in size.
Which solution will meet these requirements?

- A. Use the ALL distribution style for large tables. Specify primary and foreign keys for all tables.
- B. Keep using the EVEN distribution style for all tables. Specify primary and foreign keys for all tables.
- C. Specify a combination of distribution, sort, and partition keys for all tables.
- D. Use the ALL distribution style for rarely updated small tables. Specify primary and foreign keys for all tables.

Answer: C

Explanation:
This solution meets the requirements of optimizing the performance of table SQL queries without increasing the size of the cluster. By using the ALL distribution style for rarely updated small tables, you can ensure that the entire table is copied to every node in the cluster, which eliminates the need for data redistribution during joins. This can improve query performance significantly, especially for frequently joined dimension tables.
However, using the ALL distribution style also increases the storage space and the load time, so it is only suitable for small tables that are not updated frequently or extensively. By specifying primary and foreign keys for all tables, you can help the query optimizer to generate better query plans and avoid unnecessary scans or joins. You can also use the AUTO distribution style to let Amazon Redshift choose the optimal distribution style based on the table size and the query patterns. References:
* Choose the best distribution style
* Distribution styles
* Working with data distribution styles

NEW QUESTION # 84
A company is migrating on-premises workloads to AWS. The company wants to reduce overall operational overhead. The company also wants to explore serverless options.
The company's current workloads use Apache Pig, Apache Oozie, Apache Spark, Apache Hbase, and Apache Flink. The on-premises workloads process petabytes of data in seconds. The company must maintain similar or better performance after the migration to AWS.
Which extract, transform, and load (ETL) service will meet these requirements?

- A. Amazon Redshift
- B. AWS Glue
- C. Amazon EMR
- D. AWS Lambda

Answer: B

Explanation:
AWS Glue is a fully managed serverless ETL service that can handle petabytes of data in seconds. AWS Glue can run Apache Spark and Apache Flink jobs without requiring any infrastructure provisioning or management. AWS Glue can also integrate with Apache Pig, Apache Oozie, and Apache Hbase using AWS Glue Data Catalog and AWS Glue workflows. AWS Glue can reduce the overall operational overhead by automating the data discovery, data preparation, and data loading processes. AWS Glue can also optimize the cost and performance of ETL jobs by using AWS Glue Job Bookmarking, AWS Glue Crawlers, and AWS Glue Schema Registry. References:
AWS Glue
AWS Glue Data Catalog
AWS Glue Workflows
[AWS Glue Job Bookmarking]
[AWS Glue Crawlers]
[AWS Glue Schema Registry]
[AWS Certified Data Engineer - Associate DEA-C01 Complete Study Guide]

NEW QUESTION # 85
A company has an application that uses a microservice architecture. The company hosts the application on an Amazon Elastic Kubernetes Services (Amazon EKS) cluster.
The company wants to set up a robust monitoring system for the application. The company needs to analyze the logs from the EKS cluster and the application. The company needs to correlate the cluster's logs with the application's traces to identify points of failure in the whole application request flow.
Which combination of steps will meet these requirements with the LEAST development effort? (Select TWO.)

- A. Use AWS Glue to correlate the logs and traces.
- B. Use Amazon CloudWatch to collect logs. Use Amazon Kinesis to collect traces.
- C. Use Amazon CloudWatch to collect logs. Use Amazon Managed Streaming for Apache Kafka (Amazon MSK) to collect traces.
- D. Use Amazon OpenSearch to correlate the logs and traces.
- E. Use FluentBit to collect logs. Use OpenTelemetry to collect traces.

Answer: D,E

Explanation:
Step 1: Log Collection (FluentBit and CloudWatch)
Option A suggests using FluentBit to collect logs and OpenTelemetry to collect traces.
FluentBit is a lightweight log processor that integrates with Amazon EKS to collect and forward logs from Kubernetes clusters. It is widely used with minimal overhead, making it an ideal choice for log collection in this scenario. FluentBit is also natively compatible with AWS services.
OpenTelemetry is a popular framework to collect traces from distributed applications. It provides observability, making it easier to monitor microservices.
This combination allows you to effectively gather both logs and traces with minimal setup and configuration, aligning with the goal of least development effort.
CloudWatch can be used to monitor logs (Option B and C). However, for applications that need more custom and fine-grained control over logging mechanisms, FluentBit and OpenTelemetry are the preferred choice in microservice environments.
Step 2: Log and Trace Correlation (Amazon OpenSearch)
Option D (Amazon OpenSearch) is specifically designed to search, analyze, and visualize logs, metrics, and traces in real-time. OpenSearch allows you to correlate logs and traces effectively.
With Amazon OpenSearch, you can set up dashboards that help in visualizing both logs and traces together, which assists in identifying any failure points across the entire request flow.
It offers integrations with FluentBit and OpenTelemetry, ensuring that both logs from the EKS cluster and application traces are centrally collected, stored, and correlated without additional heavy development.
Step 3: Why Other Options Are Not Suitable
Option B (Amazon Kinesis) is designed for real-time data streaming and analytics but is not as well-suited for tracing microservice requests and logs correlation compared to OpenSearch.
Option C (Amazon MSK) provides a managed Kafka streaming service, but this adds complexity when trying to integrate and correlate logs and traces from a microservice environment. Setting up Kafka requires more development effort compared to using FluentBit and OpenTelemetry.
Option E (AWS Glue) is primarily an ETL (Extract, Transform, Load) service. While Glue is powerful for data processing, it is not a native tool for log and trace correlation, and using it would add unnecessary complexity for this use case.
Conclusion:
To meet the requirements with the least development effort:
Use FluentBit for log collection and OpenTelemetry for tracing (Option A).
Correlate logs and traces using Amazon OpenSearch (Option D).
This approach leverages AWS-native services designed for seamless integration with microservices hosted on Amazon EKS and ensures effective monitoring with minimal overhead.

NEW QUESTION # 86

A company saves customer data to an Amazon S3 bucket. The company uses server-side encryption with AWS KMS keys (SSE-KMS) to encrypt the bucket. The dataset includes personally identifiable information (PII) such as social security numbers and account details.
Data that is tagged as PII must be masked before the company uses customer data for analysis. Some users must have secure access to the PII data during the preprocessing phase. The company needs a low-maintenance solution to mask and secure the PII data throughout the entire engineering pipeline.
Which combination of solutions will meet these requirements? (Select TWO.)

- A. Configure an Amazon Made discovery job for the S3 bucket.
- B. Use AWS Identity and Access Management (IAM) to manage permissions and to control access to the PII data.
- C. Use AWS Glue DataBrew to perform extract, transform, and load (ETL) tasks that mask the PII data before analysis.
- D. Use Amazon GuardDuty to monitor access patterns for the PII data that is used in the engineering pipeline.
- E. Write custom scripts in an application to mask the PII data and to control access.

Answer: B,C

Explanation:
To address the requirement of masking PII data and ensuring secure access throughout the data pipeline, the combination of AWS Glue DataBrew and IAM provides a low-maintenance solution.
A . AWS Glue DataBrew for Masking:
AWS Glue DataBrew provides a visual tool to perform data transformations, including masking PII data. It allows for easy configuration of data transformation tasks without requiring manual coding, making it ideal for this use case.

Reference:

D . AWS Identity and Access Management (IAM):

Using IAM policies allows fine-grained control over access to PII data, ensuring that only authorized users can view or process sensitive data during the pipeline stages.

Alternatives Considered:

B (Amazon GuardDuty): GuardDuty is for threat detection and does not handle data masking or access control for PII.

C (Amazon Macie): Macie can help discover sensitive data but does not handle the masking of PII or access control.

E (Custom scripts): Custom scripting increases the operational burden compared to a built-in solution like DataBrew.

AWS Glue DataBrew for Data Masking

IAM Policies for PII Access Control

NEW QUESTION # 87

......

If you are still hesitate to choose our Actual4Exams, you can try to free download part of Amazon Data-Engineer-Associate exam certification exam questions and answers provided in our Actual4Exams. So that you can know the high reliability of our Actual4Exams. Our Actual4Exams will be your best selection and guarantee to pass Amazon Data-Engineer-Associate Exam Certification. Your choose of our Actual4Exams is equal to choose success.

Data-Engineer-Associate Valid Exam Vce Free: https://www.actual4exams.com/Data-Engineer-Associate-valid-dump.html

Amazon Valid Real Data-Engineer-Associate Exam You can choose learning tools to pass the exam, Secondly, since our Data-Engineer-Associate training quiz appeared on the market, seldom do we have the cases of customer information disclosure, Actual4Exams is the right place where you can go and find the best products for your online Data-Engineer-Associate Amazon audio lectures preparation, To contribute the long-term of cooperation with our customers, we offer great discount for purchasing our Data-Engineer-Associate exam pdf.

By contrast, I know customers who have picked Data-Engineer-Associate the Struts framework and did their own modifications, An array is a group of items of the same type either value or reference types) For Data-Engineer-Associate Valid Exam Vce Free instance, you might create an array that contains all integer values or all string values.

Amazon Valid Real Data-Engineer-Associate Exam: AWS Certified Data Engineer - Associate (DEA-C01) - Actual4Exams Authoritative Provider

You can choose learning tools to pass the exam, Secondly, since our Data-Engineer-Associate training quiz appeared on the market, seldom do we have the cases of customer information disclosure.

Actual4Exams is the right place where you can go and find the best products for your online Data-Engineer-Associate Amazon audio lectures preparation, To contribute the long-term of cooperation with our customers, we offer great discount for purchasing our Data-Engineer-Associate exam pdf.

Download Data-Engineer-Associate Testing Engine File.

- Data-Engineer-Associate Reliable Test Dumps □ Data-Engineer-Associate Latest Exam Discount □ Data-Engineer-Associate Book Pdf □ Copy URL ✔ www.prep4away.com □✔ □ open and search for ▷ Data-Engineer-Associate ◁ to download for free □Authentic Data-Engineer-Associate Exam Hub
- Data-Engineer-Associate New Braindumps □ Data-Engineer-Associate Knowledge Points □ Data-Engineer-Associate Pdf Version □ Easily obtain free download of □ Data-Engineer-Associate □ by searching on ⇒ www.pdfvce.com ⇐ □ □Certification Data-Engineer-Associate Exam Infor
- Realistic Valid Real Data-Engineer-Associate Exam| Easy To Study and Pass Exam at first attempt - Authoritative Amazon AWS Certified Data Engineer - Associate (DEA-C01) □ Easily obtain free download of ➡ Data-Engineer-Associate □ by searching on 《 www.passtestking.com 》 □Data-Engineer-Associate Test Simulator Fee
- Valid Real Data-Engineer-Associate Exam - 2025 First-grade Amazon Data-Engineer-Associate Valid Exam Vce Free 100% Pass □ Enter □ www.pdfvce.com □ and search for ▸ Data-Engineer-Associate ◂ to download for free □Data-Engineer-Associate Free Practice
- Data-Engineer-Associate Pdf Version □ New Data-Engineer-Associate Exam Question □ Data-Engineer-Associate Test Dumps.zip □ Open website □ www.prep4away.com □ and search for ➤ Data-Engineer-Associate □ for free download □Data-Engineer-Associate Free Practice
- The best way to Prepare Exam With Amazon Data-Engineer-Associate Exam Dumps □ Easily obtain free download of ✔ Data-Engineer-Associate □✔ □ by searching on { www.pdfvce.com } □Valid Data-Engineer-Associate Exam Question

- Amazon - Data-Engineer-Associate –The Best Valid Real Exam 🡒 The page for free download of （Data-Engineer-Associate） on 【 www.exams4collection.com 】 will open immediately ✿ Vce Data-Engineer-Associate Test Simulator
- Data-Engineer-Associate Test Simulator Fee 🡒 Data-Engineer-Associate Test Simulator Fee 🡒 Valid Exam Data-Engineer-Associate Practice 🡒 Download 《 Data-Engineer-Associate 》 for free by simply searching on " www.pdfvce.com " 🡒Data-Engineer-Associate New Braindumps
- Data-Engineer-Associate Reliable Test Dumps 🡒 Valid Data-Engineer-Associate Exam Question 🡒 Data-Engineer-Associate Reliable Exam Book 🡒 Search on ➡ www.real4dumps.com 🡐 for 🡒 Data-Engineer-Associate 🡐 to obtain exam materials for free download 🡒Exam Data-Engineer-Associate Sample
- Amazon - Data-Engineer-Associate –The Best Valid Real Exam 🡒 Copy URL ➡ www.pdfvce.com 🡐🡐🡐 open and search for ➡ Data-Engineer-Associate 🡐🡐🡐 to download for free 🡒Data-Engineer-Associate Actual Dump
- Valid Real Data-Engineer-Associate Exam - 2025 First-grade Amazon Data-Engineer-Associate Valid Exam Vce Free 100% Pass 🡒 Open website 「 www.torrentvce.com 」 and search for { Data-Engineer-Associate } for free download 🡒 🡒Data-Engineer-Associate New Braindumps
- stancoo822.designertoblog.com, teteclass.com, www.stes.tyc.edu.tw, learn.anantnaad.in, learnvernac.co.za, www.stes.tyc.edu.tw, gradenet.ng, www.stes.tyc.edu.tw, 0854422957s.blogspot.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, Disposable vapes

DOWNLOAD the newest Actual4Exams Data-Engineer-Associate PDF dumps from Cloud Storage for free:
https://drive.google.com/open?id=1Jyg0z3QcoygVJpqIPbB__7d7lbC6jtr2