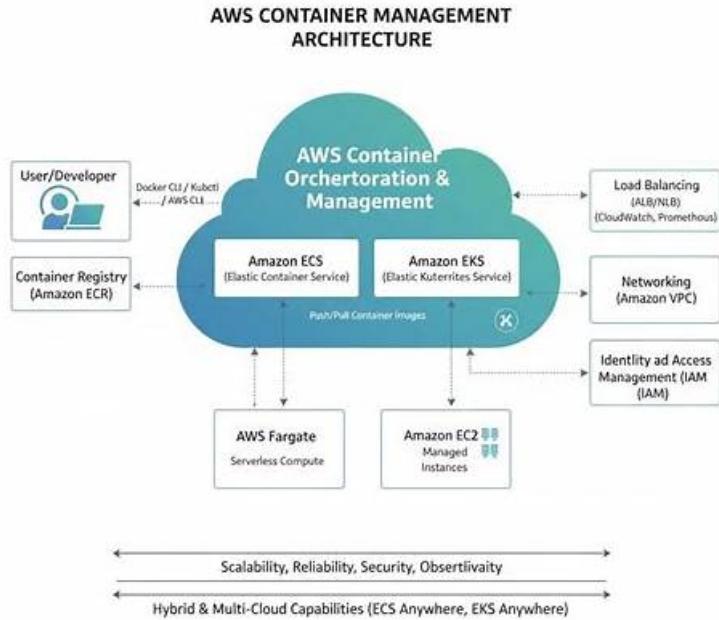


Amazon Data-Engineer-Associate Exam Tips | Data-Engineer-Associate Download Free Dumps



2026 Latest Exams4Collection Data-Engineer-Associate PDF Dumps and Data-Engineer-Associate Exam Engine Free Share:
<https://drive.google.com/open?id=1pSpbNU0BnQ0DhqpPp2qImEYXC6FtuGCI>

All kinds of exams are changing with dynamic society because the requirements are changing all the time. To keep up with the newest regulations of the Data-Engineer-Associate exam, our experts keep their eyes focusing on it. Our Data-Engineer-Associate practice materials are updating according to the precise of the real exam. Our test prep can help you to conquer all difficulties you may encounter. In other words, we will be your best helper.

Although the passing rate of our Data-Engineer-Associate simulating exam is nearly 100%, we can refund money in full if you are still worried that you may not pass. You don't need to worry about the complexity of the refund process at all, we've made it quite simple. As long as you provide us with proof that you failed the exam after using our Data-Engineer-Associate, we can refund immediately. If you encounter any problems during the refund process, you can also contact our customer service staff at any time. They will help you solve the problem as quickly as possible. That is to say, our Data-Engineer-Associate Exam Questions almost guarantee that you pass the exam. Even if you don't pass, you don't have to pay any price for our Data-Engineer-Associate simulating exam. I hope we have enough sincerity to impress you.

>> Amazon Data-Engineer-Associate Exam Tips <<

Data-Engineer-Associate Download Free Dumps | Exam Data-Engineer-Associate Material

As we all know, HR from many companies hold the view that candidates who own a Data-Engineer-Associate professional certification are preferred, because they are more likely to solve potential problems during work. And the Data-Engineer-Associate certification vividly demonstrates the fact that they are better learners. As for candidates who possessed with a Data-Engineer-Associate professional certification are more competitive. The current world 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.

Amazon AWS Certified Data Engineer - Associate (DEA-C01) Sample Questions (Q193-Q198):

NEW QUESTION # 193

A data engineer needs to create an Amazon Athena table based on a subset of data from an existing Athena table named cities_world. The cities_world table contains cities that are located around the world. The data engineer must create a new table named cities_us to contain only the cities from cities_world that are located in the US.

Which SQL statement should the data engineer use to meet this requirement?

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

Answer: A

Explanation:

To create a new table named cities_usa in Amazon Athena based on a subset of data from the existing cities_world table, you should use an INSERT INTO statement combined with a SELECT statement to filter only the records where the country is 'usa'. The correct SQL syntax would be:

Option A: `INSERT INTO cities_usa (city, state) SELECT city, state FROM cities_world WHERE country='usa';` This statement inserts only the cities and states where the country column has a value of 'usa' from the cities_world table into the cities_usa table. This is a correct approach to create a new table with data filtered from an existing table in Athena.

Options B, C, and D are incorrect due to syntax errors or incorrect SQL usage (e.g., the MOVE command or the use of UPDATE in a non-relevant context).

References:

Amazon Athena SQL Reference

Creating Tables in Athena

NEW QUESTION # 194

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

Answer: A

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 overhead¹³.

* 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.

References:

* 1: AWS Data Exchange User Guide

* 2: AWS Data Exchange FAQs

* 3: AWS Glue Developer Guide

* : AWS CodeCommit User Guide

* : Amazon Kinesis Data Streams Developer Guide

* : Amazon Elastic Container Registry User Guide

* : Build a Continuous Delivery Pipeline for Your Container Images with Amazon ECR as Source

NEW QUESTION # 195

A company is building an analytics solution. The solution uses Amazon S3 for data lake storage and Amazon Redshift for a data warehouse. The company wants to use Amazon Redshift Spectrum to query the data that is in Amazon S3.

Which actions will provide the FASTEST queries? (Choose two.)

- A. Split the data into files that are less than 10 KB.
- B. Partition the data based on the most common query predicates.
- C. Use file formats that are not
- D. Use a columnar storage file format.
- E. Use gzip compression to compress individual files to sizes that are between 1 GB and 5 GB.

Answer: B,D

NEW QUESTION # 196

A company extracts approximately 1 TB of data every day from data sources such as SAP HANA, Microsoft SQL Server, MongoDB, Apache Kafka, and Amazon DynamoDB. Some of the data sources have undefined data schemas or data schemas that change.

A data engineer must implement a solution that can detect the schema for these data sources. The solution must extract, transform, and load the data to an Amazon S3 bucket. The company has a service level agreement (SLA) to load the data into the S3 bucket within 15 minutes of data creation.

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

- A. Use Amazon EMR to detect the schema and to extract, transform, and load the data into the S3 bucket.
Create a pipeline in Apache Spark.
- B. Use AWS Glue to detect the schema and to extract, transform, and load the data into the S3 bucket.
Create a pipeline in Apache Spark.
- C. Create a stored procedure in Amazon Redshift to detect the schema and to extract, transform, and load the data into a Redshift Spectrum table. Access the table from Amazon S3.
- D. Create a PySpark program in AWS Lambda to extract, transform, and load the data into the S3 bucket.

Answer: B

Explanation:

AWS Glue is a fully managed service that provides a serverless data integration platform. It can automatically discover and categorize data from various sources, including SAP HANA, Microsoft SQL Server, MongoDB, Apache Kafka, and Amazon DynamoDB. It can also infer the schema of the data and store it in the AWS Glue Data Catalog, which is a central metadata repository. AWS Glue can then use the schema information to generate and run Apache Spark code to extract, transform, and load the data into an Amazon S3 bucket. AWS Glue can also monitor and optimize the performance and cost of the data pipeline, and handle any schema changes that may occur in the source data. AWS Glue can meet the SLA of loading the data into the S3 bucket within 15 minutes of data creation, as it can trigger the data pipeline based on events, schedules, or on-demand. AWS Glue has the least operational overhead among the options, as it does not require provisioning, configuring, or managing any servers or clusters. It

also handles scaling, patching, and security automatically. References:

AWS Glue

[AWS Glue Data Catalog]

[AWS Glue Developer Guide]

AWS Certified Data Engineer - Associate DEA-C01 Complete Study Guide

NEW QUESTION # 197

A company maintains a data warehouse in an on-premises Oracle database. The company wants to build a data lake on AWS. The company wants to load data warehouse tables into Amazon S3 and synchronize the tables with incremental data that arrives from the data warehouse every day.

Each table has a column that contains monotonically increasing values. The size of each table is less than 50 GB. The data warehouse tables are refreshed every night between 1 AM and 2 AM. A business intelligence team queries the tables between 10 AM and 8 PM every day.

Which solution will meet these requirements in the MOST operationally efficient way?

- A. Use AWS Glue to load a full copy of the data warehouse tables into Amazon S3 every day. Overwrite the previous day's full-load copy every day.
- B. Use an AWS Glue Java Database Connectivity (JDBC) connection. Configure a job bookmark for a column that contains monotonically increasing values. Write custom logic to append the daily incremental data to a full-load copy that is in Amazon S3.
- C. Use an AWS Database Migration Service (AWS DMS) full load plus CDC job to load tables that contain monotonically increasing data columns from the on-premises data warehouse to Amazon S3.
Use custom logic in AWS Glue to append the daily incremental data to a full-load copy that is in Amazon S3.
- D. Use an AWS Database Migration Service (AWS DMS) full load migration to load the data warehouse tables into Amazon S3 every day Overwrite the previous day's full-load copy every day.

Answer: C

Explanation:

The company needs to load data warehouse tables into Amazon S3 and perform incremental synchronization with daily updates. The most efficient solution is to use AWS Database Migration Service (AWS DMS) with a combination of full load and change data capture (CDC) to handle the initial load and daily incremental updates.

Option A: Use an AWS Database Migration Service (AWS DMS) full load plus CDC job to load tables that contain monotonically increasing data columns from the on-premises data warehouse to Amazon S3. Use custom logic in AWS Glue to append the daily incremental data to a full-load copy that is in Amazon S3.

DMS is designed to migrate databases to AWS, and the combination of full load plus CDC is ideal for handling incremental data changes efficiently. AWS Glue can then be used to append the incremental data to the full data set in S3. This solution is highly operationally efficient because it automates both the full load and incremental updates.

Options B, C, and D are less operationally efficient because they either require writing custom logic to handle bookmarks manually or involve unnecessary daily full loads.

References:

AWS Database Migration Service Documentation

AWS Glue Documentation

NEW QUESTION # 198

.....

We provide free update to the client and the discounts to the old client. We provide free update of our Data-Engineer-Associate exam materials within one year and after one year the client can enjoy the 50% discounts. The old clients enjoy some certain discounts when they buy our Data-Engineer-Associate exam torrent. Our experts check whether there is the update of the test bank every day and if there is the system will send to the client automatically. We choose the most typical questions and answers which seize the focus and important information and the questions and answers are based on the real exam. So you can master the most important Data-Engineer-Associate Exam Torrent in the shortest time and finally pass the exam successfully.

Data-Engineer-Associate Download Free Dumps: <https://www.exams4collection.com/Data-Engineer-Associate-latest-braindumps.html>

That means if you study our Data-Engineer-Associate training prep, your passing rate is much higher than other candidates. Our company are established more than 10 years, our quality of Data-Engineer-Associate valid practice test questions are the leading

position in this field. In addition, Data-Engineer-Associate exam materials are high-quality, and you can improve your efficiency by using them. You can pass the real exam easily with our latest Data-Engineer-Associate vce dumps and this is the only smartest way to get success.

Every time they drop a pencil, the analysts jump on it, Background color of this cell, That means if you study our Data-Engineer-Associate training prep, your passing rate is much higher than other candidates.

High Pass-Rate Amazon Data-Engineer-Associate Exam Tips & The Best Exams4Collection - Leading Provider in Qualification Exams

Our company are established more than 10 years, our quality of Data-Engineer-Associate valid practice test questions are the leading position in this filed, In addition, Data-Engineer-Associate exam materials are high-quality, and you can improve your efficiency by using them.

You can pass the real exam easily with our latest Data-Engineer-Associate vce dumps and this is the only smartest way to get success, Then I started using the Test King website and was amazed by the way they have made things easier.

2026 Latest Exams4Collection Data-Engineer-Associate PDF Dumps and Data-Engineer-Associate Exam Engine Free Share:
<https://drive.google.com/open?id=1pSpbNU0BnQ0DhqpPp2qImEYXC6FtuGCI>