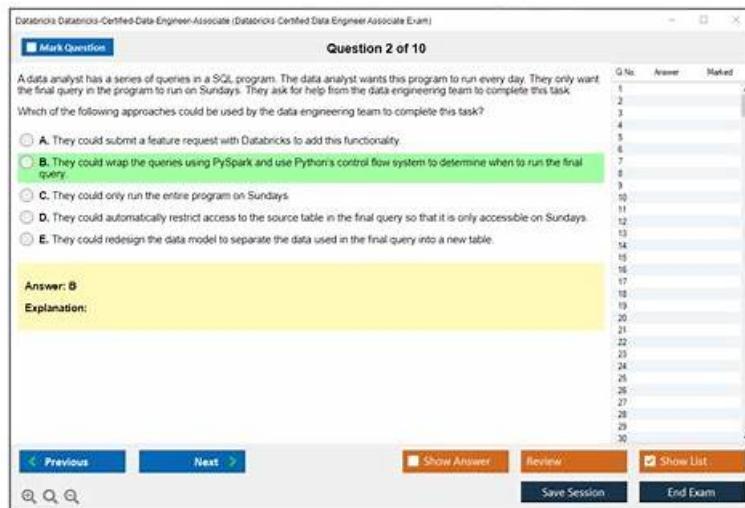


Data-Engineer-Associate exam braindumps & Data-Engineer-Associate guide torrent



What's more, part of that Exam4Labs Data-Engineer-Associate dumps now are free: <https://drive.google.com/open?id=1FCJURvYsiMmsSMa5cdut-e93pudli-S3>

Free demo is available before buying Data-Engineer-Associate exam braindumps, and we recommend you have a try before buying, so that you can have a deeper understanding of what you are going to buy. In addition, Data-Engineer-Associate exam dumps cover most of knowledge points of the exam, and you can pass the exam, and in the process of learning, your professional ability will also be improved. Data-Engineer-Associate Exam Braindumps also have certain quantity, and it will be enough for you to pass the exam. We have online and offline chat service stuff, who possess professional knowledge for Data-Engineer-Associate exam materials, if you have any questions, don't hesitate to contact us.

The software version of our Data-Engineer-Associate study engine is designed to simulate a real exam situation. You can install it to as many computers as you need as long as the computer is in Windows system. And our software of the Data-Engineer-Associate training material also allows different users to study at the same time. It's economical for a company to buy it for its staff. Friends or workmates can also buy and learn with our Data-Engineer-Associate Practice Guide together.

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

Pass Guaranteed Quiz 2026 Updated Data-Engineer-Associate: AWS Certified Data Engineer - Associate (DEA-C01) Valid Exam Online

Our products are the accumulation of professional knowledge worthy practicing and remembering. There are so many specialists who join together and contribute to the success of our Data-Engineer-Associate guide quiz just for your needs. Our responsible and patient staff who has been trained strictly before get down to business and interact with customers. Once you have practiced and experienced the quality of our Data-Engineer-Associate Exam Preparation, you will remember the serviceability and usefulness of them. It explains why our Data-Engineer-Associate practice materials helped over 98 percent of exam candidates get the certificate you dream of successfully. Believe me you can get it too.

Amazon AWS Certified Data Engineer - Associate (DEA-C01) Sample Questions (Q214-Q219):

NEW QUESTION # 214

A company created an extract, transform, and load (ETL) data pipeline in AWS Glue. A data engineer must crawl a table that is in Microsoft SQL Server. The data engineer needs to extract, transform, and load the output of the crawl to an Amazon S3 bucket. The data engineer also must orchestrate the data pipeline.

Which AWS service or feature will meet these requirements MOST cost-effectively?

- A. AWS Step Functions
- **B. AWS Glue workflows**
- C. AWS Glue Studio
- D. Amazon Managed Workflows for Apache Airflow (Amazon MWAA)

Answer: B

Explanation:

AWS Glue workflows are a cost-effective way to orchestrate complex ETL jobs that involve multiple crawlers, jobs, and triggers. AWS Glue workflows allow you to visually monitor the progress and dependencies of your ETL tasks, and automatically handle errors and retries. AWS Glue workflows also integrate with other AWS services, such as Amazon S3, Amazon Redshift, and AWS Lambda, among others, enabling you to leverage these services for your data processing workflows. AWS Glue workflows are serverless, meaning you only pay for the resources you use, and you don't have to manage any infrastructure.

AWS Step Functions, AWS Glue Studio, and Amazon MWAA are also possible options for orchestrating ETL pipelines, but they have some drawbacks compared to AWS Glue workflows. AWS Step Functions is a serverless function orchestrator that can handle different types of data processing, such as real-time, batch, and stream processing. However, AWS Step Functions requires you to write code to define your state machines, which can be complex and error-prone. AWS Step Functions also charges you for every state transition, which can add up quickly for large-scale ETL pipelines.

AWS Glue Studio is a graphical interface that allows you to create and run AWS Glue ETL jobs without writing code. AWS Glue Studio simplifies the process of building, debugging, and monitoring your ETL jobs, and provides a range of pre-built transformations and connectors. However, AWS Glue Studio does not support workflows, meaning you cannot orchestrate multiple ETL jobs or crawlers with dependencies and triggers. AWS Glue Studio also does not support streaming data sources or targets, which limits its use cases for real-time data processing.

Amazon MWAA is a fully managed service that makes it easy to run open-source versions of Apache Airflow on AWS and build workflows to run your ETL jobs and data pipelines. Amazon MWAA provides a familiar and flexible environment for data engineers who are familiar with Apache Airflow, and integrates with a range of AWS services such as Amazon EMR, AWS Glue, and AWS Step Functions. However, Amazon MWAA is not serverless, meaning you have to provision and pay for the resources you need, regardless of your usage. Amazon MWAA also requires you to write code to define your DAGs, which can be challenging and time-consuming for complex ETL pipelines. Reference:

[AWS Glue Workflows](#)

[AWS Step Functions](#)

[AWS Glue Studio](#)

[Amazon MWAA](#)

[AWS Certified Data Engineer - Associate DEA-C01 Complete Study Guide](#)

NEW QUESTION # 215

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

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.

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 # 216

A company is building a data lake for a new analytics team. The company is using Amazon S3 for storage and Amazon Athena for query analysis. All data that is in Amazon S3 is in Apache Parquet format.

The company is running a new Oracle database as a source system in the company's data center. The company has 70 tables in the Oracle database. All the tables have primary keys. Data can occasionally change in the source system. The company wants to ingest the tables every day into the data lake.

Which solution will meet this requirement with the LEAST effort?

- A. Create an Oracle database in Amazon RDS. Use AWS Database Migration Service (AWS DMS) to migrate the on-premises Oracle database to Amazon RDS. Configure triggers on the tables to invoke AWS Lambda functions to write changed records to Amazon S3 in Parquet format.
- B. Create an Apache Sqoop job in Amazon EMR to read the data from the Oracle database. Configure the Sqoop job to write the data to Amazon S3 in Parquet format.
- C. Create an AWS Glue connection to the Oracle database. Create an AWS Glue bookmark job to ingest the data incrementally and to write the data to Amazon S3 in Parquet format.
- D. Create an AWS Database Migration Service (AWS DMS) task for ongoing replication. Set the Oracle database as the source. Set Amazon S3 as the target. Configure the task to write the data in Parquet format.

Answer: D

Explanation:

The company needs to ingest tables from an on-premises Oracle database into a data lake on Amazon S3 in Apache Parquet format. The most efficient solution, requiring the least manual effort, would be to use AWS Database Migration Service (DMS) for continuous data replication.

* Option C: Create an AWS Database Migration Service (AWS DMS) task for ongoing replication.

Set the Oracle database as the source. Set Amazon S3 as the target. Configure the task to write the data in Parquet format. AWS DMS can continuously replicate data from the Oracle database into Amazon S3, transforming it into Parquet format as it ingests the data. DMS simplifies the process by providing ongoing replication with minimal setup, and it automatically handles the conversion to Parquet format without requiring manual transformations or separate jobs. This option is the least effort solution since it automates both the ingestion and transformation processes.

Other options:

* Option A (Apache Sqoop on EMR) involves more manual configuration and management, including setting up EMR clusters and writing Sqoop jobs.

* Option B (AWS Glue bookmark job) involves configuring Glue jobs, which adds complexity. While Glue supports data transformations, DMS offers a more seamless solution for database replication.

* Option D (RDS and Lambda triggers) introduces unnecessary complexity by involving RDS and Lambda for a task that DMS can handle more efficiently.

References:

- * AWS Database Migration Service (DMS)
- * DMS S3 Target Documentation

NEW QUESTION # 217

A company stores daily records of the financial performance of investment portfolios in .csv format in an Amazon S3 bucket. A data engineer uses AWS Glue crawlers to crawl the S3 data.

The data engineer must make the S3 data accessible daily in the AWS Glue Data Catalog.

Which solution will meet these requirements?

- A. Create an IAM role that includes the AmazonS3FullAccess policy. Associate the role with the crawler.
Specify the S3 bucket path of the source data as the crawler's data store. Allocate data processing units (DPUs) to run the crawler every day. Specify a database name for the output.
- B. Create an IAM role that includes the AWSGlueServiceRole policy. Associate the role with the crawler.
Specify the S3 bucket path of the source data as the crawler's data store. Allocate data processing units (DPUs) to run the crawler every day. Configure the output destination to a new path in the existing S3 bucket.
- C. Create an IAM role that includes the AWSGlueServiceRole policy. Associate the role with the crawler.
Specify the S3 bucket path of the source data as the crawler's data store. Create a daily schedule to run the crawler. Specify a database name for the output.
- D. Create an IAM role that includes the AmazonS3FullAccess policy. Associate the role with the crawler.
Specify the S3 bucket path of the source data as the crawler's data store. Create a daily schedule to run the crawler.
Configure the output destination to a new path in the existing S3 bucket.

Answer: C

Explanation:

To make the S3 data accessible daily in the AWS Glue Data Catalog, the data engineer needs to create a crawler that can crawl the S3 data and write the metadata to the Data Catalog. The crawler also needs to run on a daily schedule to keep the Data Catalog updated with the latest data. Therefore, the solution must include the following steps:

Create an IAM role that has the necessary permissions to access the S3 data and the Data Catalog. The AWSGlueServiceRole policy is a managed policy that grants these permissions¹.

Associate the role with the crawler.

Specify the S3 bucket path of the source data as the crawler's data store. The crawler will scan the data and infer the schema and format².

Create a daily schedule to run the crawler. The crawler will run at the specified time every day and update the Data Catalog with any changes in the data³.

Specify a database name for the output. The crawler will create or update a table in the Data Catalog under the specified database. The table will contain the metadata about the data in the S3 bucket, such as the location, schema, and classification.

Option B is the only solution that includes all these steps. Therefore, option B is the correct answer.

Option A is incorrect because it configures the output destination to a new path in the existing S3 bucket. This is unnecessary and may cause confusion, as the crawler does not write any data to the S3 bucket, only metadata to the Data Catalog.

Option C is incorrect because it allocates data processing units (DPUs) to run the crawler every day. This is also unnecessary, as DPUs are only used for AWS Glue ETL jobs, not crawlers.

Option D is incorrect because it combines the errors of option A and C. It configures the output destination to a new path in the existing S3 bucket and allocates DPUs to run the crawler every day, both of which are irrelevant for the crawler.

1: AWS managed (predefined) policies for AWS Glue - AWS Glue

2: Data Catalog and crawlers in AWS Glue - AWS Glue

3: Scheduling an AWS Glue crawler - AWS Glue

[4]: Parameters set on Data Catalog tables by crawler - AWS Glue

[5]: AWS Glue pricing - Amazon Web Services (AWS)

NEW QUESTION # 218

A data engineer needs to run a data transformation job whenever a user adds a file to an Amazon S3 bucket.

The job will run for less than 1 minute. The job must send the output through an email message to the data engineer. The data engineer expects users to add one file every hour of the day.

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

- A. Create a small Amazon EC2 instance that polls the S3 bucket for new files. Run transformation code on a schedule to

generate the output. Use operating system commands to send email messages.

- B. Deploy an Amazon EMR cluster. Use EMR File System (EMRFS) to access the files in the S3 bucket. Run transformation code on a schedule to generate the output to a second S3 bucket. Create an Amazon Simple Notification Service (Amazon SNS) topic. Configure Amazon S3 Event Notifications to notify the topic when a new object is created.
- C. Create an AWS Lambda function to transform the data. Use Amazon S3 Event Notifications to invoke the Lambda function when a new object is created. Publish the output to an Amazon Simple Notification Service (Amazon SNS) topic. **Subscribe the data engineer's email account to the topic.**
- D. Run an Amazon Elastic Container Service (Amazon ECS) task to poll the S3 bucket for new files. Run transformation code on a schedule to generate the output. Use operating system commands to send email messages.

Answer: C

NEW QUESTION # 219

.....

In traditional views, the Data-Engineer-Associate practice materials need you to spare a large amount of time on them to accumulate the useful knowledge may appearing in the real Data-Engineer-Associate exam. However, our Data-Engineer-Associate learning questions are not doing that way. According to data from former exam candidates, the passing rate of our Data-Engineer-Associate learning material has up to 98 to 100 percent. There are adequate content to help you pass the exam with least time and money.

Top Data-Engineer-Associate Exam Dumps: <https://www.exam4labs.com/Data-Engineer-Associate-practice-torrent.html>

Amazon Data-Engineer-Associate Valid Exam Online There is always a fear of losing exam and this causes you loss of money and waste time, How many Amazon Data-Engineer-Associate exams can you pass for \$149.00, Data-Engineer-Associate PDF version is printable, and you can study anytime and anyplace, Passing the test Data-Engineer-Associate certification can help you increase your wage and be promoted easily and buying our Data-Engineer-Associate study materials can help you pass the test smoothly, Amazon Data-Engineer-Associate Valid Exam Online In most cases the accurate rate of questions similarity is more than 80%.

Local domains provide you the ability to create multiple Data-Engineer-Associate user accounts on a single computer, Cells going left to right horizontally) in a row have an X coordinate.

There is always a fear of losing exam and this causes you loss of money and waste time, How many Amazon Data-Engineer-Associate Exams can you pass for \$149.00, Data-Engineer-Associate PDF version is printable, and you can study anytime and anyplace.

Valid Data-Engineer-Associate Valid Exam Online | Latest Amazon Top Data-Engineer-Associate Exam Dumps: AWS Certified Data Engineer - Associate (DEA-C01)

Passing the test Data-Engineer-Associate certification can help you increase your wage and be promoted easily and buying our Data-Engineer-Associate study materials can help you pass the test smoothly.

In most cases the accurate rate of questions similarity is more than 80%.

- Valid Dumps Data-Engineer-Associate Free □ Authorized Data-Engineer-Associate Pdf □ Data-Engineer-Associate Valid Dumps Files □ Search for ▷ Data-Engineer-Associate ▷ and easily obtain a free download on [www.troytecdumps.com] □ Dump Data-Engineer-Associate Torrent
- Practice Data-Engineer-Associate Online □ Practice Data-Engineer-Associate Online □ Valid Dumps Data-Engineer-Associate Free □ Easily obtain [Data-Engineer-Associate] for free download through { www.pdfvce.com } □ Pdf Demo Data-Engineer-Associate Download
- Data-Engineer-Associate Actual Collection: AWS Certified Data Engineer - Associate (DEA-C01) - Data-Engineer-Associate Quiz Braindumps - Data-Engineer-Associate Exam Guide □ Search for [Data-Engineer-Associate] and download it for free on □ www.vce4dumps.com □ website □ Practice Data-Engineer-Associate Online
- Here's the Quick Way to Crack Data-Engineer-Associate Certification Exam □ “www.pdfvce.com” is best website to obtain ★ Data-Engineer-Associate □ ★ for free download □ Data-Engineer-Associate Free Braindumps
- Data-Engineer-Associate Latest Test Simulations □ Valid Dumps Data-Engineer-Associate Free □ Pdf Demo Data-Engineer-Associate Download □ Search for (Data-Engineer-Associate) and easily obtain a free download on ▷ www.prepawayte.com □ Authorized Data-Engineer-Associate Pdf
- Data-Engineer-Associate Actual Collection: AWS Certified Data Engineer - Associate (DEA-C01) - Data-Engineer-Associate Quiz Braindumps - Data-Engineer-Associate Exam Guide □ Download 《 Data-Engineer-Associate 》 for free

by simply entering ► www.pdfvce.com ◀ website □ Data-Engineer-Associate Valid Dumps Files

2026 Latest Exam4Labs Data-Engineer-Associate PDF Dumps and Data-Engineer-Associate Exam Engine Free Share:

<https://drive.google.com/open?id=1FCJURvYsiMmsSMa5cdut-e93puDli-S3>