

Data-Engineer-Associate Latest Exam Papers | Data-Engineer-Associate Valid Exam Tips



P.S. Free 2026 Amazon Data-Engineer-Associate dumps are available on Google Drive shared by ActualTestsIT:
<https://drive.google.com/open?id=1TtkxKjvL9IF9svMJQPvzpzsVHFoWYer>

When we are in some kind of learning web site, often feel dazzling, because web page appear too desultory. Absorbing the lessons of the Data-Engineer-Associate test prep, will be all kinds of qualification examination classify layout, at the same time on the front page of the Data-Engineer-Associate test materials have clear test module classification, so clear page design greatly convenient for the users, can let users in a very short period of time to find what they want to study, and then targeted to study. Saving the precious time of users, also makes the Data-Engineer-Associate Quiz torrent look more rich.

The ActualTestsIT Data-Engineer-Associate exam practice test questions will provide you with everything that you need to learn, prepare and pass the AWS Certified Data Engineer - Associate (DEA-C01) Data-Engineer-Associate exam. The ActualTestsIT Data-Engineer-Associate exam questions are the real PSE questions that will help you to understand the real AWS Certified Data Engineer - Associate (DEA-C01) Data-Engineer-Associate Exam Pattern and answers and you can easily pass the final AWS Certified Data Engineer - Associate (DEA-C01) Data-Engineer-Associate exam.

>> Data-Engineer-Associate Latest Exam Papers <<

Free PDF 2026 Amazon Updated Data-Engineer-Associate: AWS Certified Data Engineer - Associate (DEA-C01) Latest Exam Papers

Once you submit your practice, the system of our Data-Engineer-Associate exam quiz will automatically generate a report. The system is highly flexible, which has short reaction time. So you will quickly get a feedback about your exercises of the Data-Engineer-Associate preparation questions. For example, it will note that how much time you have used to finish the Data-Engineer-Associate Study Guide, and how much marks you got for your practice as well as what kind of the questions and answers you are wrong with.

Amazon AWS Certified Data Engineer - Associate (DEA-C01) Sample Questions (Q39-Q44):

NEW QUESTION # 39

A company maintains multiple extract, transform, and load (ETL) workflows that ingest data from the company's operational databases into an Amazon S3 based data lake. The ETL workflows use AWS Glue and Amazon EMR to process data. The company wants to improve the existing architecture to provide automated orchestration and to require minimal manual effort. Which solution will meet these requirements with the LEAST operational overhead?

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

Answer: D

Explanation:

AWS Glue workflows are a feature of AWS Glue that enable you to create and visualize complex ETL pipelines using AWS Glue components, such as crawlers, jobs, triggers, and development endpoints. AWS Glue workflows provide automated orchestration and require minimal manual effort, as they handle dependency resolution, error handling, state management, and resource allocation for your ETL workflows.

You can use AWS Glue workflows to ingest data from your operational databases into your Amazon S3 based data lake, and then use AWS Glue and Amazon EMR to process the data in the data lake. This solution will meet the requirements with the least operational overhead, as it leverages the serverless and fully managed nature of AWS Glue, and the scalability and flexibility of Amazon EMR12.

The other options are not optimal for the following reasons:

B. AWS Step Functions tasks. AWS Step Functions is a service that lets you coordinate multiple AWS services into serverless workflows. You can use AWS Step Functions tasks to invoke AWS Glue and Amazon EMR jobs as part of your ETL workflows, and use AWS Step Functions state machines to define the logic and flow of your workflows. However, this option would require more manual effort than AWS Glue workflows, as you would need to write JSON code to define your state machines, handle errors and retries, and monitor the execution history and status of your workflows3.

C. AWS Lambda functions. AWS Lambda is a service that lets you run code without provisioning or managing servers. You can use AWS Lambda functions to trigger AWS Glue and Amazon EMR jobs as part of your ETL workflows, and use AWS Lambda event sources and destinations to orchestrate the flow of your workflows. However, this option would also require more manual effort than AWS Glue workflows, as you would need to write code to implement your business logic, handle errors and retries, and monitor the invocation and execution of your Lambda functions. Moreover, AWS Lambda functions have limitations on the execution time, memory, and concurrency, which may affect the performance and scalability of your ETL workflows.

D. Amazon Managed Workflows for Apache Airflow (Amazon MWAA) workflows. Amazon MWAA is a managed service that makes it easy to run open source Apache Airflow on AWS. Apache Airflow is a popular tool for creating and managing complex ETL pipelines using directed acyclic graphs (DAGs). You can use Amazon MWAA workflows to orchestrate AWS Glue and Amazon EMR jobs as part of your ETL workflows, and use the Airflow web interface to visualize and monitor your workflows. However, this option would have more operational overhead than AWS Glue workflows, as you would need to set up and configure your Amazon MWAA environment, write Python code to define your DAGs, and manage the dependencies and versions of your Airflow plugins and operators.

1: AWS Glue Workflows

2: AWS Glue and Amazon EMR

3: AWS Step Functions

AWS Lambda

Amazon Managed Workflows for Apache Airflow

NEW QUESTION # 40

A company stores customer records in Amazon S3. The company must not delete or modify the customer record data for 7 years after each record is created. The root user also must not have the ability to delete or modify the data.

A data engineer wants to use S3 Object Lock to secure the data.

Which solution will meet these requirements?

- A. Enable governance mode on the S3 bucket. Use a default retention period of 7 years.
- B. Set the retention period for individual objects in the S3 bucket to 7 years.
- C. Place a legal hold on individual objects in the S3 bucket. Set the retention period to 7 years.
- **D. Enable compliance mode on the S3 bucket. Use a default retention period of 7 years.**

Answer: D

Explanation:

The company wants to ensure that no customer records are deleted or modified for 7 years, and even the root user should not have the ability to change the data. S3 Object Lock in Compliance Mode is the correct solution for this scenario.

* Option B: Enable compliance mode on the S3 bucket. Use a default retention period of 7 years. In Compliance Mode, even the root user cannot delete or modify locked objects during the retention period. This ensures that the data is protected for the entire 7-year duration as required. Compliance mode is stricter than governance mode and prevents all forms of alteration, even by privileged users.

Option A (Governance Mode) still allows certain privileged users (like the root user) to bypass the lock, which does not meet the company's requirement. Option C (legal hold) and Option D (setting retention per object) do not fully address the requirement to block root user modifications.

References:

* Amazon S3 Object Lock Documentation

NEW QUESTION # 41

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

Answer: D

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. References:

* AWS Glue Workflows

* AWS Step Functions

* AWS Glue Studio

* Amazon MWAA

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

NEW QUESTION # 42

A company uses a variety of AWS and third-party data stores. The company wants to consolidate all the data into a central data warehouse to perform analytics. Users need fast response times for analytics queries.

The company uses Amazon QuickSight in direct query mode to visualize the data. Users normally run queries during a few hours

each day with unpredictable spikes.

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

- A. Use Amazon Redshift provisioned clusters to load all the data into Amazon Redshift managed storage (RMS).
- B. Use Amazon Aurora PostgreSQL to load all the data into Aurora.
- C. Use Amazon Athena to load all the data into Amazon S3 in Apache Parquet format.
- **D. Use Amazon Redshift Serverless to load all the data into Amazon Redshift managed storage (RMS).**

Answer: D

Explanation:

Problem Analysis:

The company requires a centralized data warehouse for consolidating data from various sources.

They use Amazon QuickSight in direct query mode, necessitating fast response times for analytical queries.

Users query the data intermittently, with unpredictable spikes during the day.

Operational overhead should be minimal.

Key Considerations:

The solution must support fast, SQL-based analytics.

It must handle unpredictable spikes efficiently.

Must integrate seamlessly with QuickSight for direct querying.

Minimize operational complexity and scaling concerns.

Solution Analysis:

Option A: Amazon Redshift Serverless

Redshift Serverless eliminates the need for provisioning and managing clusters.

Automatically scales compute capacity up or down based on query demand.

Reduces operational overhead by handling performance optimization.

Fully integrates with Amazon QuickSight, ensuring low-latency analytics.

Reduces costs as it charges only for usage, making it ideal for workloads with intermittent spikes.

Option B: Amazon Athena with S3 (Apache Parquet)

Athena supports querying data directly from S3 in Parquet format.

While it's cost-effective, performance depends on the size and complexity of the data.

It is not optimized for high-speed analytics needed by QuickSight in direct query mode.

Option C: Amazon Redshift Provisioned Clusters

Requires manual cluster provisioning, scaling, and maintenance.

Higher operational overhead compared to Redshift Serverless.

Option D: Amazon Aurora PostgreSQL

Aurora is optimized for transactional databases, not data warehousing or analytics.

Does not meet the requirement for fast analytics queries.

Final Recommendation:

Amazon Redshift Serverless is the best choice for this use case because it provides fast analytics, integrates natively with QuickSight, and minimizes operational complexity while efficiently handling unpredictable spikes.

Reference:

Amazon Redshift Serverless Overview

Amazon QuickSight and Redshift Integration

NEW QUESTION # 43

The company stores a large volume of customer records in Amazon S3. To comply with regulations, the company must be able to access new customer records immediately for the first 30 days after the records are created. The company accesses records that are older than 30 days infrequently.

The company needs to cost-optimize its Amazon S3 storage.

Which solution will meet these requirements MOST cost-effectively?

- **A. Apply a lifecycle policy to transition records to S3 Standard Infrequent-Access (S3 Standard-IA) storage after 30 days.**
- B. Use S3 Standard-Infrequent Access (S3 Standard-IA) storage for all customer records.
- C. Transition records to S3 Glacier Deep Archive storage after 30 days.
- D. Use S3 Intelligent-Tiering storage.

Answer: A

Explanation:

The most cost-effective solution in this case is to apply a lifecycle policy to transition records to Amazon S3 Standard-IA storage after 30 days. Here's why:

- * **Amazon S3 Lifecycle Policies:** Amazon S3 offers lifecycle policies that allow you to automatically transition objects between different storage classes to optimize costs. For data that is frequently accessed in the first 30 days and infrequently accessed after that, transitioning from the S3 Standard storage class to S3 Standard-Infrequent Access (S3 Standard-IA) after 30 days makes the most sense. S3 Standard-IA is designed for data that is accessed less frequently but still needs to be retained, offering lower storage costs than S3 Standard with a retrieval cost for access.

- * **Cost Optimization:** S3 Standard-IA offers a lower price per GB than S3 Standard. Since the data will be accessed infrequently after 30 days, using S3 Standard-IA will lower storage costs while still allowing for immediate retrieval when necessary.

- * **Compliance with Regulations:** Since the records need to be immediately accessible for the first 30 days, the use of S3 Standard for that period ensures compliance with regulatory requirements. After 30 days, transitioning to S3 Standard-IA continues to meet access requirements for infrequent access while reducing storage costs.

- * **Alternatives Considered:**

- * **Option B (S3 Intelligent-Tiering):** While S3 Intelligent-Tiering automatically moves data between access tiers based on access patterns, it incurs a small monthly monitoring and automation charge per object. It could be a viable option, but transitioning data to S3 Standard-IA directly would be more cost-effective since the pattern of access is well-known (frequent for 30 days, infrequent thereafter).

- * **Option C (S3 Glacier Deep Archive):** Glacier Deep Archive is the lowest-cost storage class, but it is not suitable in this case because the data needs to be accessed immediately within 30 days and on an infrequent basis thereafter. Glacier Deep Archive requires hours for data retrieval, which is not acceptable for infrequent access needs.

- * **Option D (S3 Standard-IA for all records):** Using S3 Standard-IA for all records would result in higher costs for the first 30 days, as the data is frequently accessed. S3 Standard-IA incurs retrieval charges, making it less suitable for frequently accessed data.

:

Amazon S3 Lifecycle Policies

S3 Storage Classes

Cost Management and Data Optimization Using Lifecycle Policies

AWS Data Engineering Documentation

NEW QUESTION # 44

.....

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 exam torrent are updating according to the precise of the real exam. Our Data-Engineer-Associate test prep to help you to conquer all difficulties you may encounter. Once you choose our Data-Engineer-Associate Quiz torrent, we will send the new updates for one year long, which is new enough to deal with the exam for you and guide you through difficulties in your exam preparation.

Data-Engineer-Associate Valid Exam Tips: <https://www.actualtestsit.com/Amazon/Data-Engineer-Associate-exam-prep-dumps.html>

It depends on the client to choose the version they favor to learn our Data-Engineer-Associate study materials, Amazon Data-Engineer-Associate Latest Exam Papers And besides, you can take it with you wherever you go for it is portable and takes no place, These special offers help you save huge money that you spend on buying individual Data-Engineer-Associate ActualTestsIT exam files, According to the market research, we have found that a lot of people preparing for the Data-Engineer-Associate exam want to gain the newest information about the exam.

To illustrate one useful arrangement, you won't go wrong if you put different variables Data-Engineer-Associate in different columns and different records in different rows, The use of sequence numbers offers flexibility when modifying a portion of a prefix list.

Free PDF Quiz 2026 Accurate Amazon Data-Engineer-Associate Latest Exam Papers

It depends on the client to choose the version they favor to learn our Data-Engineer-Associate Study Materials, And besides, you can take it with you wherever you go for it is portable and takes no place.

These special offers help you save huge money that you spend on buying individual Data-Engineer-Associate ActualTestsIT exam files, According to the market research, we have found that a lot of people preparing for the Data-Engineer-Associate exam want to gain the newest information about the exam.

So our simulating exam is definitely making your review more durable.

- P.S. Free 2026 Amazon Data-Engineer-Associate dumps are available on Google Drive shared by ActualTestsIT: <https://drive.google.com/open?id=1TtkXKjvL9IF9svMJOPVpzssVHFoWYer>

P.S. Free 2026 Amazon Data-Engineer-Associate dumps are available on Google Drive shared by ActualTestsIT: <https://drive.google.com/open?id=1TtkXKjvL9IF9svMJOPVpzssVHFoWYer>