

Hot Amazon Data-Engineer-Associate Questions - Data-Engineer-Associate Free Updates



P.S. Free & New Data-Engineer-Associate dumps are available on Google Drive shared by TestKingFree:
<https://drive.google.com/open?id=11G3MACAhNMyAfjH6H3RPeAE-DGc0HIzl>

Under the tremendous stress of fast pace in modern life, this version of our Data-Engineer-Associate test prep suits office workers perfectly. It can match your office software and as well as help you spare time practicing the Data-Engineer-Associate exam. As for its shining points, the PDF version can be readily downloaded and printed out so as to be read by you. It's really a convenient way for those who are fond of paper learning. With this kind of version, you can flip through the pages at liberty and quickly finish the check-up Data-Engineer-Associate Test Prep. And you can take notes on this version of our Data-Engineer-Associate exam questions.

Amazon Data-Engineer-Associate Exam Syllabus Topics:

Section	Weight	Objectives

Data Ingestion and Transformation	34%	- Transform and process data • 1. ETL/ELT patterns • 2. Data partitioning and compression • 3. Data transformation services (AWS Glue, Amazon EMR, AWS Lambda) • 4. Batch and stream processing - Apply programming concepts • 1. Version control • 2. Infrastructure as Code (IaC) • 3. SQL, Python, Scala - Orchestrate data pipelines • 1. Event-driven architectures • 2. Amazon Managed Workflows for Apache Airflow (MWAA) • 3. AWS Step Functions • 4. AWS Glue Workflows - Perform data ingestion • 1. Throughput and latency characteristics for AWS services • 2. Replayability of data • 3. Batch data ingestion (scheduled ingestion, event-driven ingestion) • 4. Streaming data ingestion • 5. Data ingestion patterns (frequency and data history)
Data Security and Governance	18%	- Implement data quality checks • 1. Data validation • 2. AWS Glue DataBrew - Ensure data encryption • 1. AWS KMS • 2. Encryption at rest and in transit - Apply authentication and authorization • 1. Service control policies (SCPs) • 2. AWS IAM policies and roles • 3. Amazon S3 bucket policies - Manage data privacy and compliance • 1. AWS Lake Formation permissions • 2. Data masking and tokenization • 3. PII data handling

Data Operations and Support	22%	- Automate data pipelines • 1. Scheduling jobs • 2. Event-driven triggers • 3. AWS Lambda triggers - Monitor data pipelines • 1. Logging and metrics • 2. AWS CloudTrail • 3. Amazon CloudWatch - Manage and troubleshoot data processes • 1. Performance tuning • 2. Cost optimization • 3. Debugging failed jobs
Data Store Management	26%	- Understand data cataloging • 1. Data discovery and classification • 2. Schema evolution • 3. AWS Glue Data Catalog - Choose a data store • 1. Access and storage patterns • 2. Data characteristics (structured, semi-structured, unstructured) • 3. Data lakes vs. data warehouses • 4. Amazon S3, Amazon RDS, Amazon DynamoDB, Amazon Redshift - Manage data lifecycle • 1. Data archiving • 2. Data retention policies • 3. Amazon S3 storage classes - Design data models • 1. Schema design • 2. Normalization and denormalization • 3. Partitioning and indexing strategies

>> Hot Amazon Data-Engineer-Associate Questions <<

Data-Engineer-Associate Free Updates & Data-Engineer-Associate Exam Duration

If you are looking to be Amazon Data-Engineer-Associate certified, TestKingFree is here to provide you with the best AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) exam dumps through which you can clear your AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) certification exam. We are providing practice exams in three formats including PDF which is the downloadable file from which you can study for your AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) exam questions and our Web-based application provides you the facility to assess yourself without installing any software on your device to prepare you for AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) exam dumps.

Amazon AWS Certified Data Engineer - Associate (DEA-C01) Sample Questions (Q208-Q213):

NEW QUESTION # 208

A company runs an extract, transform, and load (ETL) job in AWS Glue. The job processes personally identifiable information (PII) data and writes logs to an Amazon CloudWatch Logs log group. A data engineer needs to mask PII data in the CloudWatch Logs log group.

Which solution will meet these requirements?

- **A. Configure a data protection policy. Attach the policy to the CloudWatch log group.**
- B. Run an Amazon Macie sensitive data discovery job.
- C. Call AWS Glue sensitive data detection APIs in the ETL job.
- D. Attach an AWS Glue security configuration to the ETL job.

Answer: A

Explanation:

Option B is the right approach because the requirement is to mask PII in the log destination (the CloudWatch Logs log group). The exam guide explicitly calls out security responsibilities that include "data encryption and masking" and also emphasizes enabling and preparing logs for audit and governance needs. A log-group-level masking mechanism is therefore the most direct control point to prevent sensitive values from being exposed to anyone who can view logs.

Option A (Glue security configuration) is primarily used to apply protections such as encryption settings for Glue jobs and related outputs; it does not inherently solve the problem of masking PII that has already been emitted into application logs. Option C (Macie) is intended to discover and classify sensitive data—most commonly in Amazon S3—and to produce findings; it is not a log-masking control for CloudWatch Logs.

The material reinforces Macie as a discovery/classification service for PII rather than a masking mechanism.

Option D adds custom code paths in the ETL job and still may miss PII that appears in framework/system logs; it also violates the "least operational effort" spirit compared to applying a centralized log policy.

Therefore, attaching a data protection policy directly to the CloudWatch Logs log group best meets the masking requirement.

NEW QUESTION # 209

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

Answer: C

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.

Reference:

Amazon S3 Lifecycle Policies

S3 Storage Classes

Cost Management and Data Optimization Using Lifecycle Policies

AWS Data Engineering Documentation

NEW QUESTION # 210

A company receives test results from testing facilities that are located around the world. The company stores the test results in millions of 1 KB JSON files in an Amazon S3 bucket. A data engineer needs to process the files, convert them into Apache Parquet format, and load them into Amazon Redshift tables. The data engineer uses AWS Glue to process the files, AWS Step Functions to orchestrate the processes, and Amazon EventBridge to schedule jobs.

The company recently added more testing facilities. The time required to process files is increasing. The data engineer must reduce the data processing time.

Which solution will MOST reduce the data processing time?

- A. Use Amazon EMR instead of AWS Glue to group the raw input files. Process the files in Amazon EMR. Load the files into the Amazon Redshift tables.
- B. Use AWS Lambda to group the raw input files into larger files. Write the larger files back to Amazon S3. Use AWS Glue to process the files. Load the files into the Amazon Redshift tables.
- **C. Use the AWS Glue dynamic frame file-grouping option to ingest the raw input files. Process the files. Load the files into the Amazon Redshift tables.**
- D. Use the Amazon Redshift COPY command to move the raw input files from Amazon S3 directly into the Amazon Redshift tables. Process the files in Amazon Redshift.

Answer: C

Explanation:

Problem Analysis:

Millions of 1 KB JSON files in S3 are being processed and converted to Apache Parquet format using AWS Glue.

Processing time is increasing due to the additional testing facilities.

The goal is to reduce processing time while using the existing AWS Glue framework.

Key Considerations:

AWS Glue offers the dynamic frame file-grouping feature, which consolidates small files into larger, more efficient datasets during processing.

Grouping smaller files reduces overhead and speeds up processing.

Solution Analysis:

Option A: Lambda for File Grouping

Using Lambda to group files would add complexity and operational overhead. Glue already offers built-in grouping functionality.

Option B: AWS Glue Dynamic Frame File-Grouping

This option directly addresses the issue by grouping small files during Glue job execution.

Minimizes data processing time with no extra overhead.

Option C: Redshift COPY Command

COPY directly loads raw files but is not designed for pre-processing (conversion to Parquet).

Option D: Amazon EMR

While EMR is powerful, replacing Glue with EMR increases operational complexity.

Final Recommendation:

Use AWS Glue dynamic frame file-grouping for optimized data ingestion and processing.

AWS Glue Dynamic Frames

Optimizing Glue Performance

NEW QUESTION # 211

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. 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.
- B. 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.
- C. 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.
- D. 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.

Answer: D

NEW QUESTION # 212

A company has three subsidiaries. Each subsidiary uses a different data warehousing solution. The first subsidiary hosts its data warehouse in Amazon Redshift. The second subsidiary uses Teradata Vantage on AWS. The third subsidiary uses Google BigQuery.

The company wants to aggregate all the data into a central Amazon S3 data lake. The company wants to use Apache Iceberg as the table format.

A data engineer needs to build a new pipeline to connect to all the data sources, run transformations by using each source engine, join the data, and write the data to Iceberg.

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

- A. Use the Amazon Athena federated query connectors for Amazon Redshift, Teradata, and BigQuery to build the pipeline in Athena. Write a SQL query to read from all the data sources, join the data, and run a Merge operation on the data lake Iceberg table.
- B. Use the native Amazon Redshift connector, the Java Database Connectivity (JDBC) connector for Teradata, and the open source Apache Spark BigQuery connector to build the pipeline in Amazon EMR. Write code in PySpark to join the data. Run a Merge operation on the data lake Iceberg table.
- C. Use native Amazon Redshift, Teradata, and BigQuery connectors to build the pipeline in AWS Glue. Use native AWS Glue transforms to join the data. Run a Merge operation on the data lake Iceberg table.
- D. Use the native Amazon Redshift, Teradata, and BigQuery connectors in Amazon Appflow to write data to Amazon S3 and AWS Glue Data Catalog. Use Amazon Athena to join the data. Run a Merge operation on the data lake Iceberg table.

Answer: A

Explanation:

Amazon Athena provides federated query connectors that allow querying multiple data sources, such as Amazon Redshift, Teradata, and Google BigQuery, without needing to extract the data from the original source. This solution is optimal because it offers the least operational effort by avoiding complex data movement and transformation processes.

* Amazon Athena Federated Queries:

* Athena's federated queries allow direct querying of data stored across multiple sources, including Amazon Redshift, Teradata, and BigQuery. With Athena's support for Apache Iceberg, the company can easily run a Merge operation on the Iceberg table.

* The solution reduces complexity by centralizing the query execution and transformation process in Athena using SQL queries.

Reference: Amazon Athena Federated Query

Alternatives Considered:

A (AWS Glue pipeline): This would work but requires more operational effort to manage and transform the data in AWS Glue.

C (Amazon EMR): Using EMR and writing PySpark code introduces more operational overhead and complexity compared to a SQL-based solution in Athena.

D (Amazon AppFlow): AppFlow is more suitable for transferring data between services but is not as efficient for transformations and joins as Athena federated queries.

References:

Amazon Athena Documentation

Federated Queries in Amazon Athena

NEW QUESTION # 213

.....

if you choose to use the software version of our Data-Engineer-Associate study guide, you will find that you can download our

Data-Engineer-Associate exam prep on more than one computer and you can practice our Data-Engineer-Associate exam questions offline as well. We strongly believe that the software version of our Data-Engineer-Associate Study Materials will be of great importance for you to prepare for the exam and all of the employees in our company wish you early success!

Data-Engineer-Associate Free Updates: <https://www.testkingfree.com/Amazon/Data-Engineer-Associate-practice-exam-dumps.html>

- Reliable Data-Engineer-Associate Exam Blueprint □ Reliable Data-Engineer-Associate Exam Blueprint □ Data-Engineer-Associate Exam Introduction □ Go to website (www.pdf.dumps.com) open and search for □ Data-Engineer-Associate □ to download for free □ Data-Engineer-Associate Lead2pass Review
- 2026 Hot Data-Engineer-Associate Questions | Trustable 100% Free AWS Certified Data Engineer - Associate (DEA-C01) Free Updates □ Search for ➡ Data-Engineer-Associate □ on { www.pdfvce.com } immediately to obtain a free download □ Reliable Data-Engineer-Associate Exam Pdf
- Latest Real Data-Engineer-Associate Exam □ Data-Engineer-Associate Valid Exam Pdf ~ Data-Engineer-Associate Latest Demo ~ Search for ▷ Data-Engineer-Associate ◁ and easily obtain a free download on ✓ www.vceengine.com □ ✓ □ □ Data-Engineer-Associate Dump Collection
- Data-Engineer-Associate Exam Reviews □ Data-Engineer-Associate Lead2pass Review □ Reliable Data-Engineer-Associate Exam Pdf □ Easily obtain free download of ➡ Data-Engineer-Associate □ □ □ by searching on ✓ www.pdfvce.com □ ✓ □ □ New Data-Engineer-Associate Study Guide
- Pass Guaranteed 2026 High Pass-Rate Amazon Hot Data-Engineer-Associate Questions □ Open website ▶ www.examdiscuss.com ◀ and search for “Data-Engineer-Associate” for free download □ Guide Data-Engineer-Associate Torrent
- Free PDF 2026 Amazon Latest Data-Engineer-Associate: Hot AWS Certified Data Engineer - Associate (DEA-C01) Questions □ Search for 「 Data-Engineer-Associate 」 and download it for free immediately on ✓ www.pdfvce.com □ ✓ □ □ Data-Engineer-Associate Exam Introduction
- Money-Back Guarantee for Amazon Data-Engineer-Associate Exam Questions □ Enter 「 www.prepawaypdf.com 」 and search for □ Data-Engineer-Associate □ to download for free □ Reliable Data-Engineer-Associate Exam Pdf
- 100% Pass Quiz 2026 Amazon Data-Engineer-Associate: High-quality Hot AWS Certified Data Engineer - Associate (DEA-C01) Questions □ Download □ Data-Engineer-Associate □ for free by simply searching on ➡ www.pdfvce.com □ □ Certification Data-Engineer-Associate Cost
- 2026 Hot Data-Engineer-Associate Questions | Trustable 100% Free AWS Certified Data Engineer - Associate (DEA-C01) Free Updates □ Copy URL ➡ www.prepawaypdf.com □ □ □ open and search for ▶ Data-Engineer-Associate ◀ to download for free □ Data-Engineer-Associate Dump Collection
- Reliable Data-Engineer-Associate Exam Pdf □ Data-Engineer-Associate Dump Collection □ Data-Engineer-Associate Exam Introduction □ Simply search for □ Data-Engineer-Associate □ for free download on ✓ www.pdfvce.com □ ✓ □ □ Data-Engineer-Associate Valid Exam Pdf
- Data-Engineer-Associate Dumps Guide □ Data-Engineer-Associate New Study Guide □ Certification Data-Engineer-Associate Cost □ Easily obtain free download of ➤ Data-Engineer-Associate □ by searching on ⇒ www.exam4labs.com ⇐ □ Data-Engineer-Associate Valid Exam Pdf
- 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, app.parler.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, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, Disposable vapes

BONUS!!! Download part of TestKingFree Data-Engineer-Associate dumps for free: <https://drive.google.com/open?id=11G3MACAhNMfAfJH6H3RPeAE-DGc0HIzI>