Updated Data-Engineer-Associate Testkings, Data-Engineer-Associate Exam Experience

P.S. Free & New Data-Engineer-Associate dumps are available on Google Drive shared by TrainingDump: https://drive.google.com/open?id=1mOufUWz8j_6OkuprcOgvIZaxo1btVuQ3

Our Data-Engineer-Associate study materials are compiled and verified by the first-rate experts in the industry domestically and they are linked closely with the real exam. Our products' contents cover the entire syllabus of the exam and refer to the past years' exam papers. Our test bank provides all the questions which may appear in the real exam and all the important information about the exam. You can use the practice test software to test whether you have mastered the Data-Engineer-Associate Study Materials and the function of stimulating the exam to be familiar with the real exam's pace, atmosphere and environment. So our Data-Engineer-Associate study materials are real-exam-based and convenient for the clients to prepare for the exam.

In the era of informational globalization, the world has witnessed climax of science and technology development, and has enjoyed the prosperity of various scientific blooms. In 21st century, every country had entered the period of talent competition, therefore, we must begin to extend our Data-Engineer-Associate personal skills, only by this can we become the pioneer among our competitors. At the same time, our competitors are trying to capture every opportunity and get a satisfying job. In this case, we need a professional Data-Engineer-Associate Certification, which will help us stand out of the crowd and knock out the door of great company.

>> Updated Data-Engineer-Associate Testkings <<

Amazon Data-Engineer-Associate Practice Exams for Thorough Preparation (Desktop/Online/PDF)

The former exam candidates get the passing rate over 98 percent in recent years by choosing our Data-Engineer-Associate practice materials. You must be curious about the advantages of them. These traits briefly sum up our Data-Engineer-Associate study questions. So we take liberty of introducing our Data-Engineer-Associate learning guide for you, hoping you can find the best way to pass the exam. With our Data-Engineer-Associate exam prep, you will pass the exam with ease.

Amazon AWS Certified Data Engineer - Associate (DEA-C01) Sample Questions (Q79-Q84):

NEW QUESTION # 79
A company receives a data file from a partner each day in an Amazon S3 bucket. The company uses a daily AW5 Glue extract, transform, and load (ETL) pipeline to clean and transform each data file. The output of the ETL pipeline is written to a CSV file named Dairy.csv in a second 53 bucket.
Occasionally, the daily data file is empty or is missing values for required fields. When the file is missing data, the company can use the previous day's CSV file.
A data engineer needs to ensure that the previous day's data file is overwritten only if the new daily file is complete and valid.
Which solution will meet these requirements with the LEAST effort?

- A. Invoke an AWS Lambda function to check the file for missing data and to fill in missing values in required fields.

- B. Configure the AWS Glue ETL pipeline to use AWS Glue Data Quality rules. Develop rules in Data Quality Definition Language (DQDL) to check for missing values in required files and empty files.
- C. Run a SQL query in Amazon Athena to read the CSV file and drop missing rows. Copy the corrected CSV file to the second S3 bucket.
- D. Use AWS Glue Studio to change the code in the ETL pipeline to fill in any missing values in the required fields with the most common values for each field.

Answer: B

Explanation:
Problem Analysis:
The company runs a daily AWS Glue ETL pipeline to clean and transform files received in an S3 bucket.
If a file is incomplete or empty, the previous day's file should be retained.
Need a solution to validate files before overwriting the existing file.
Key Considerations:
Automate data validation with minimal human intervention.
Use built-in AWS Glue capabilities for ease of integration.
Ensure robust validation for missing or incomplete data.
Solution Analysis:
Option A: Lambda Function for Validation
Lambda can validate files, but it would require custom code.
Does not leverage AWS Glue's built-in features, adding operational complexity.
Option B: AWS Glue Data Quality Rules
AWS Glue Data Quality allows defining Data Quality Definition Language (DQDL) rules.
Rules can validate if required fields are missing or if the file is empty.
Automatically integrates into the existing ETL pipeline.
If validation fails, retain the previous day's file.
Option C: AWS Glue Studio with Filling Missing Values
Modifying ETL code to fill missing values with most common values risks introducing inaccuracies.
Does not handle empty files effectively.
Option D: Athena Query for Validation
Athena can drop rows with missing values, but this is a post-hoc solution.
Requires manual intervention to copy the corrected file to S3, increasing complexity.
Final Recommendation:
Use AWS Glue Data Quality to define validation rules in DQDL for identifying missing or incomplete data.
This solution integrates seamlessly with the ETL pipeline and minimizes manual effort.
Implementation Steps:
Enable AWS Glue Data Quality in the existing ETL pipeline.
Define DQDL Rules, such as:
Check if a file is empty.
Verify required fields are present and non-null.
Configure the pipeline to proceed with overwriting only if the file passes validation.
In case of failure, retain the previous day's file.
AWS Glue Data Quality Overview
Defining DQDL Rules
AWS Glue Studio Documentation

NEW QUESTION # 80
A company wants to analyze sales records that the company stores in a MySQL database. The company wants to correlate the records with sales opportunities identified by Salesforce.
The company receives 2 GB erf sales records every day. The company has 100 GB of identified sales opportunities. A data engineer needs to develop a process that will analyze and correlate sales records and sales opportunities. The process must run once each night.
Which solution will meet these requirements with the LEAST operational overhead?

- A. Use Amazon Managed Workflows for Apache Airflow (Amazon MWAA) to fetch both datasets. Use AWS Lambda functions to correlate the datasets. Use AWS Step Functions to orchestrate the process.
- B. Use Amazon AppFlow to fetch sales opportunities from Salesforce. Use AWS Glue to fetch sales records from the MySQL database. Correlate the sales records with sales opportunities. Use AWS Step Functions to orchestrate the process.
- C. Use Amazon AppFlow to fetch sales opportunities from Salesforce. Use AWS Glue to fetch sales records from the

MySQL database. Correlate the sales records with the sales opportunities. Use Amazon Managed Workflows for Apache Airflow (Amazon MWAA) to orchestrate the process.

- D. Use Amazon AppFlow to fetch sales opportunities from Salesforce. Use Amazon Kinesis Data Streams to fetch sales records from the MySQL database. Use Amazon Managed Service for Apache Flink to correlate the datasets. Use AWS Step Functions to orchestrate the process.

Answer: B

Explanation:
Problem Analysis:
The company processes 2 GB of daily sales records and 100 GB of Salesforce sales opportunities.
The goal is to analyze and correlate the two datasets with low operational overhead.
The process must run once nightly.
Key Considerations:
Amazon AppFlow simplifies data integration with Salesforce.
AWS Glue can extract data from MySQL and perform ETL operations.
Step Functions can orchestrate workflows with minimal manual intervention.
Apache Airflow and Flink add complexity, which conflicts with the requirement for low operational overhead.
Solution Analysis:
Option A: MWAA + Lambda + Step Functions
Requires custom Lambda code for dataset correlation, increasing development and operational complexity.
Option B: AppFlow + Glue + MWAA
MWAA adds orchestration overhead compared to the simpler Step Functions.
Option C: AppFlow + Glue + Step Functions
AppFlow fetches Salesforce data, Glue extracts MySQL data, and Step Functions orchestrate the entire process.
Minimal setup and operational overhead, making it the best choice.
Option D: AppFlow + Kinesis + Flink + Step Functions
Using Kinesis and Flink for batch processing introduces unnecessary complexity.
Final Recommendation:
Use Amazon AppFlow to fetch Salesforce data, AWS Glue to process MySQL data, and Step Functions for orchestration.
Amazon AppFlow Overview
AWS Glue ETL Documentation
AWS Step Functions

NEW QUESTION # 81
A company uses Amazon S3 as a data lake. The company sets up a data warehouse by using a multi-node Amazon Redshift cluster.
The company organizes the data files in the data lake based on the data source of each data file.
The company loads all the data files into one table in the Redshift cluster by using a separate COPY command for each data file location. This approach takes a long time to load all the data files into the table. The company must increase the speed of the data ingestion. The company does not want to increase the cost of the process.
Which solution will meet these requirements?

- A. Use an AWS Glue job to copy all the data files into one folder. Use a COPY command to load the data into Amazon Redshift.
- B. Use a provisioned Amazon EMR cluster to copy all the data files into one folder. Use a COPY command to load the data into Amazon Redshift.
- C. Create a manifest file that contains the data file locations. Use a COPY command to load the data into Amazon Redshift.
- D. Load all the data files in parallel into Amazon Aurora. Run an AWS Glue job to load the data into Amazon Redshift.

Answer: C

Explanation:
The company is facing performance issues loading data into Amazon Redshift because it is issuing separate COPY commands for each data file location. The most efficient way to increase the speed of data ingestion into Redshift without increasing the cost is to use a manifest file.
Option D: Create a manifest file that contains the data file locations. Use a COPY command to load the data into Amazon Redshift.
A manifest file provides a list of all the data files, allowing the COPY command to load all files in parallel from different locations in Amazon S3. This significantly improves the loading speed without adding costs, as it optimizes the data loading process in a single COPY operation.
Other options (A, B, C) involve additional steps that would either increase the cost (provisioning clusters, using Glue, etc.) or do not

address the core issue of needing a unified and efficient COPY process.
Reference:
Amazon Redshift COPY Command
Redshift Manifest File Documentation

NEW QUESTION # 82
A company stores data in a data lake that is in Amazon S3. Some data that the company stores in the data lake contains personally identifiable information (PII). Multiple user groups need to access the raw data. The company must ensure that user groups can access only the PII that they require.
Which solution will meet these requirements with the LEAST effort?

- A. Use Amazon Athena to query the data. Set up AWS Lake Formation and create data filters to establish levels of access for the company's IAM roles. Assign each user to the IAM role that matches the user's PII access requirements.
- B. Build a custom query builder UI that will run Athena queries in the background to access the data.
  Create user groups in Amazon Cognito. Assign access levels to the user groups based on the PII access requirements of the users.
- C. Create IAM roles that have different levels of granular access. Assign the IAM roles to IAM user groups. Use an identity-based policy to assign access levels to user groups at the column level.
- D. Use Amazon QuickSight to access the data. Use column-level security features in QuickSight to limit the PII that users can retrieve from Amazon S3 by using Amazon Athena. Define QuickSight access levels based on the PII access requirements of the users.

Answer: A

Explanation:
Amazon Athena is a serverless, interactive query service that enables you to analyze data in Amazon S3 using standard SQL. AWS Lake Formation is a service that helps you build, secure, and manage data lakes on AWS.
You can use AWS Lake Formation to create data filters that define the level of access for different IAM roles based on the columns, rows, or tags of the data. By using Amazon Athena to query the data and AWS Lake Formation to create data filters, the company can meet the requirements of ensuring that user groups can access only the PII that they require with the least effort. The solution is to use Amazon Athena to query the data in the data lake that is in Amazon S3. Then, set up AWS Lake Formation and create data filters to establish levels of access for the company's IAM roles. For example, a data filter can allow a user group to access only the columns that contain the PII that they need, such as name and email address, and deny access to the columns that contain the PII that they do not need, such as phone number and social security number.
Finally, assign each user to the IAM role that matches the user's PII access requirements. This way, the user groups can access the data in the data lake securely and efficiently. The other options are either not feasible or not optimal. Using Amazon QuickSight to access the data (option B) would require the company to pay for the QuickSight service and to configure the column-level security features for each user. Building a custom query builder UI that will run Athena queries in the background to access the data (option C) would require the company to develop and maintain the UI and to integrate it with Amazon Cognito. Creating IAM roles that have different levels of granular access (option D) would require the company to manage multiple IAM roles and policies and to ensure that they are aligned with the data schema. References:
Amazon Athena
AWS Lake Formation
AWS Certified Data Engineer - Associate DEA-C01 Complete Study Guide, Chapter 4: Data Analysis and Visualization, Section 4.3: Amazon Athena

NEW QUESTION # 83
An airline company is collecting metrics about flight activities for analytics. The company is conducting a proof of concept (POC) test to show how analytics can provide insights that the company can use to increase on-time departures.
The POC test uses objects in Amazon S3 that contain the metrics in .csv format. The POC test uses Amazon Athena to query the data. The data is partitioned in the S3 bucket by date.
As the amount of data increases, the company wants to optimize the storage solution to improve query performance.
Which combination of solutions will meet these requirements? (Choose two.)

- A. Preprocess the .csvdata to JSON format by fetchingonly the document keys that the query requires.
- B. Use an S3 bucket that is in the same account that uses Athena to query the data.
- C. Add a randomized string to the beginning of the keys in Amazon S3 to get more throughput across partitions.
- D. Preprocess the .csv data to Apache Parquet format by fetching only the data blocks that are needed for predicates.

- E. Use an S3 bucket that is in the same AWS Region where the company runs Athena queries.

Answer: D,E

Explanation:
Using an S3 bucket that is in the same AWS Region where the company runs Athena queries can improve query performance by reducing data transfer latency and costs. Preprocessing the .csv data to Apache Parquet format can also improve query performance by enabling columnar storage, compression, and partitioning, which can reduce the amount of data scanned and fetched by the query. These solutions can optimize the storage solution for the POC test without requiring much effort or changes to the existing data pipeline. The other solutions are not optimal or relevant for this requirement. Adding a randomized string to the beginning of the keys in Amazon S3 can improve the throughput across partitions, but it can also make the data harder to query and manage. Using an S3 bucket that is in the same account that uses Athena to query the data does not have any significant impact on query performance, as long as the proper permissions are granted.
Preprocessing the .csv data to JSON format does not offer any benefits over the .csv format, as both are row-based and verbose formats that require more data scanning and fetching than columnar formats like Parquet. References:
Best Practices When Using Athena with AWS Glue
Optimizing Amazon S3 Performance
AWS Certified Data Engineer - Associate DEA-C01 Complete Study Guide

NEW QUESTION # 84
.....

In order to serve you better, we have a complete system for you if you choose us. We have free demo for Data-Engineer-Associate training materials for you to have a try. If you have decided to buy Data-Engineer-Associate exam dumps of us, just add them to your cart, and pay for it, our system will send the downloading link and password to you within ten minutes, and if you don't receive, just contact us, we will solve this problem for you as quickly as possible. For Data-Engineer-Associate Training Materials, we also have after-service, if you have questions about the exam dumps, you can contact us by email.

Data-Engineer-Associate Exam Experience: https://www.trainingdump.com/Amazon/Data-Engineer-Associate-practice-exam-dumps.html

Amazon Updated Data-Engineer-Associate Testkings It support any electronics, IPhone, Android or Windows, Our product Data-Engineer-Associate test guide delivers more important information with fewer questions and answers, That's also the reason we strongly recommend that you choose our Amazon Data-Engineer-Associate exam engine, Amazon Updated Data-Engineer-Associate Testkings Life needs to be colorful and meaningful, Web-based Data-Engineer-Associate practice exam is customizable and you can adjust its time and type of Data-Engineer-Associate AWS Certified Data Engineer - Associate (DEA-C01) questions.

Because there are some rather significant additions to what Data-Engineer-Associate you now need to know to pass the A+ exams and become certified that you did not need to know just a short while ago.

Try Our Amazon Data-Engineer-Associate Latest Exam Questions & Answers, It support any electronics, IPhone, Android or Windows, Our product Data-Engineer-Associate test guide delivers more important information with fewer questions and answers.

Free PDF Quiz 2026 Amazon Data-Engineer-Associate – Efficient Updated Testkings

That's also the reason we strongly recommend that you choose our Amazon Data-Engineer-Associate exam engine, Life needs to be colorful and meaningful, Web-based Data-Engineer-Associate practice exam is customizable and you can adjust its time and type of Data-Engineer-Associate AWS Certified Data Engineer - Associate (DEA-C01) questions.

- Download Data-Engineer-Associate Free Dumps □ Data-Engineer-Associate Training Materials □ Download Data-Engineer-Associate Free Dumps □ The page for free download of （ Data-Engineer-Associate ） on □ www.practicevce.com □ will open immediately ☑Data-Engineer-Associate Valid Braindumps
- Latest Data-Engineer-Associate Dumps Free □ Data-Engineer-Associate Reliable Test Pattern □ Latest Data-Engineer-Associate Dumps Free □ Search for { Data-Engineer-Associate } and easily obtain a free download on 【 www.pdfvce.com 】 □Valid Dumps Data-Engineer-Associate Free
- 2026 Amazon Updated Data-Engineer-Associate Testkings - AWS Certified Data Engineer - Associate (DEA-C01) Realistic Exam Experience 100% Pass □ Easily obtain free download of □ Data-Engineer-Associate □ by searching on 「 www.prepawayete.com 」 □Valid Data-Engineer-Associate Exam Bootcamp
- Updated Data-Engineer-Associate Testkings and Amazon Data-Engineer-Associate Exam Experience: AWS Certified Data

Engineer - Associate (DEA-C01) Pass Certify ☣ Open ➡ www.pdfvce.com □□□ and search for ☀ Data-Engineer-Associate □☀□ to download exam materials for free □Data-Engineer-Associate High Quality

- Ace Your Amazon Data-Engineer-Associate Exam with www.testkingpass.com: Comprehensive Study Material and Real Exam Questions □ Copy URL 《 www.testkingpass.com 》 open and search for ➥ Data-Engineer-Associate □ to download for free □Exam Data-Engineer-Associate Lab Questions
- Data-Engineer-Associate Valid Braindumps □ Data-Engineer-Associate Valid Braindumps □ New Data-Engineer-Associate Test Blueprint □ Search for ▹ Data-Engineer-Associate ◃ and download exam materials for free through 「 www.pdfvce.com 」 □Data-Engineer-Associate Real Braindumps
- 100% Pass Newest Amazon - Updated Data-Engineer-Associate Testkings □ Simply search for □ Data-Engineer-Associate □ for free download on { www.prep4sures.top } □Data-Engineer-Associate Boot Camp
- Access Real Pdfvce Amazon Data-Engineer-Associate Exam Questions Easily in dumps PDF Form □ Search for ➡ Data-Engineer-Associate □□□ and download it for free immediately on " www.pdfvce.com " □New Data-Engineer-Associate Test Blueprint
- Updated Data-Engineer-Associate Testkings - The Best Amazon Data-Engineer-Associate Exam Experience: AWS Certified Data Engineer - Associate (DEA-C01) □ Search for 「 Data-Engineer-Associate 」 on ▸ www.vceengine.com ◂ immediately to obtain a free download □Data-Engineer-Associate Reliable Study Guide
- Ace Your Amazon Data-Engineer-Associate Exam with Pdfvce: Comprehensive Study Material and Real Exam Questions □ □ Search on （ www.pdfvce.com ） for " Data-Engineer-Associate " to obtain exam materials for free download □Data-Engineer-Associate Boot Camp
- Pass Guaranteed Quiz Data-Engineer-Associate - Professional Updated AWS Certified Data Engineer - Associate (DEA-C01) Testkings □ Go to website 《 www.verifieddumps.com 》 open and search for □ Data-Engineer-Associate □ to download for free □Data-Engineer-Associate Reliable Study Guide
- ladsom.acts2.courses, www.stes.tyc.edu.tw, 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, www.stes.tyc.edu.tw, 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, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, writeablog.net, 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, aseducativa.com, Disposable vapes

P.S. Free & New Data-Engineer-Associate dumps are available on Google Drive shared by TrainingDump: https://drive.google.com/open?id=1mOufUWz8j_6OkuprcOgvIZaxo1btVuQ3