

Sample Amazon Data-Engineer-Associate Questions & New Data-Engineer-Associate Exam Questions



DOWNLOAD the newest Pass4sures Data-Engineer-Associate PDF dumps from Cloud Storage for free:
https://drive.google.com/open?id=16LUslRF336O3UApq_5sx9XOjyQmInpA

The pass rate for Data-Engineer-Associate training materials is 98.95%, and you can pass and get the certificate successfully if you buy Data-Engineer-Associate training materials from us. Besides, we have experienced experts to compile and verify Data-Engineer-Associate training materials, therefore quality and accuracy can be guaranteed. We are pass guarantee and money back guarantee if you buy Data-Engineer-Associate Exam Dumps from us. We provide you with free update for one year for the Data-Engineer-Associate training materials, so that you can know the latest information about the exam.

There are multiple choices on the versions of our Data-Engineer-Associate learning guide to select according to our interests and habits since we have three different versions of our Data-Engineer-Associate exam questions: the PDF, the Software and the APP online. The Software and APP online versions of our Data-Engineer-Associate preparation materials can be practiced on computers or phones. They are new developed for the reason that electronics products have been widely applied to our life and work style. The PDF version of our Data-Engineer-Associate Actual Exam supports printing, and you can practice with papers and take notes on it.

>> **Sample Amazon Data-Engineer-Associate Questions** <<

New Data-Engineer-Associate Exam Questions & Data-Engineer-Associate Valid Exam Guide

These are expertly designed Amazon Data-Engineer-Associate mock tests, under the supervision of thousands of professionals. A 24/7 customer service is available for assistance in case of any sort of pinch. It shows results at the end of every Data-Engineer-Associate mock test attempt so you don't repeat mistakes in the next try. To confirm the license of the product, you need an active internet connection. Pass4sures desktop AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) practice test is compatible with every Windows-based computer. You can use this software without an active internet connection.

Amazon AWS Certified Data Engineer - Associate (DEA-C01) Sample

Questions (Q84-Q89):

NEW QUESTION # 84

A company uses AWS Step Functions to orchestrate a data pipeline. The pipeline consists of Amazon EMR jobs that ingest data from data sources and store the data in an Amazon S3 bucket. The pipeline also includes EMR jobs that load the data to Amazon Redshift.

The company's cloud infrastructure team manually built a Step Functions state machine. The cloud infrastructure team launched an EMR cluster into a VPC to support the EMR jobs. However, the deployed Step Functions state machine is not able to run the EMR jobs.

Which combination of steps should the company take to identify the reason the Step Functions state machine is not able to run the EMR jobs? (Choose two.)

- A. Query the flow logs for the VPC. Determine whether the traffic that originates from the EMR cluster can successfully reach the data providers. Determine whether any security group that might be attached to the Amazon EMR cluster allows connections to the data source servers on the informed ports.
- B. Use AWS CloudFormation to automate the Step Functions state machine deployment. Create a step to pause the state machine during the EMR jobs that fail. Configure the step to wait for a human user to send approval through an email message. Include details of the EMR task in the email message for further analysis.
- C. Verify that the Step Functions state machine code has all IAM permissions that are necessary to create and run the EMR jobs. Verify that the Step Functions state machine code also includes IAM permissions to access the Amazon S3 buckets that the EMR jobs use. Use Access Analyzer for S3 to check the S3 access properties.
- D. Check the retry scenarios that the company configured for the EMR jobs. Increase the number of seconds in the interval between each EMR task. Validate that each fallback state has the appropriate catch for each decision state. Configure an Amazon Simple Notification Service (Amazon SNS) topic to store the error messages.
- E. Check for entries in Amazon CloudWatch for the newly created EMR cluster. Change the AWS Step Functions state machine code to use Amazon EMR on EKS. Change the IAM access policies and the security group configuration for the Step Functions state machine code to reflect inclusion of Amazon Elastic Kubernetes Service (Amazon EKS).

Answer: A,C

Explanation:

To identify the reason why the Step Functions state machine is not able to run the EMR jobs, the company should take the following steps:

Verify that the Step Functions state machine code has all IAM permissions that are necessary to create and run the EMR jobs. The state machine code should have an IAM role that allows it to invoke the EMR APIs, such as RunJobFlow, AddJobFlowSteps, and DescribeStep. The state machine code should also have IAM permissions to access the Amazon S3 buckets that the EMR jobs use as input and output locations. The company can use Access Analyzer for S3 to check the access policies and permissions of the S3 buckets.

Therefore, option B is correct.

Query the flow logs for the VPC. The flow logs can provide information about the network traffic to and from the EMR cluster that is launched in the VPC. The company can use the flow logs to determine whether the traffic that originates from the EMR cluster can successfully reach the data providers, such as Amazon RDS, Amazon Redshift, or other external sources. The company can also determine whether any security group that might be attached to the EMR cluster allows connections to the data source servers on the informed ports. The company can use Amazon VPC Flow Logs or Amazon CloudWatch Logs Insights to query the flow logs.

Therefore, option D is correct.

Option A is incorrect because it suggests using AWS CloudFormation to automate the Step Functions state machine deployment. While this is a good practice to ensure consistency and repeatability of the deployment, it does not help to identify the reason why the state machine is not able to run the EMR jobs. Moreover, creating a step to pause the state machine during the EMR jobs that fail and wait for a human user to send approval through an email message is not a reliable way to troubleshoot the issue. The company should use the Step Functions console or API to monitor the execution history and status of the state machine, and use Amazon CloudWatch to view the logs and metrics of the EMR jobs.

Option C is incorrect because it suggests changing the AWS Step Functions state machine code to use Amazon EMR on EKS. Amazon EMR on EKS is a service that allows you to run EMR jobs on Amazon Elastic Kubernetes Service (Amazon EKS) clusters. While this service has some benefits, such as lower cost and faster execution time, it does not support all the features and integrations that EMR on EC2 does, such as EMR Notebooks, EMR Studio, and EMRFS. Therefore, changing the state machine code to use EMR on EKS may not be compatible with the existing data pipeline and may introduce new issues.

Option E is incorrect because it suggests checking the retry scenarios that the company configured for the EMR jobs. While this is a good practice to handle transient failures and errors, it does not help to identify the root cause of why the state machine is not able to run the EMR jobs. Moreover, increasing the number of seconds in the interval between each EMR task may not improve the success rate of the jobs, and may increase the execution time and cost of the state machine. Configuring an Amazon SNS topic to store the error messages may help to notify the company of any failures, but it does not provide enough information to troubleshoot

the issue.

1: Manage an Amazon EMR Job - AWS Step Functions

2: Access Analyzer for S3 - Amazon Simple Storage Service

3: Working with Amazon EMR and VPC Flow Logs - Amazon EMR

[4]: Analyzing VPC Flow Logs with Amazon CloudWatch Logs Insights - Amazon Virtual Private Cloud

[5]: Monitor AWS Step Functions - AWS Step Functions

[6]: Monitor Amazon EMR clusters - Amazon EMR

[7]: Amazon EMR on Amazon EKS - Amazon EMR

NEW QUESTION # 85

A company runs an AWS Glue workflow every day to process time series data from an Amazon S3 bucket.

The workflow loads the data into an Amazon Redshift Serverless table. The company observes that some of the jobs in the workflow occasionally fail.

A data engineer must receive a notification when the Redshift table does not contain the most recent data.

Which solution will meet this requirement in the MOST operationally efficient way?

- A. Schedule an AWS Glue Data Quality job to check the freshness of the data. Create an Amazon EventBridge rule to notify an Amazon Simple Notification Service (Amazon SNS) topic when a data quality rule fails.
- B. Configure an Amazon EventBridge Scheduler to run an Amazon Macie job to scan the Redshift table for data freshness. Configure Macie to notify an Amazon Simple Notification Service (Amazon SNS) topic when an AWS Glue job fails.
- C. Create an Amazon CloudWatch dashboard that displays a metric named Failed AWS Glue Jobs that counts AWS Glue job failures during the previous day. Set a CloudWatch alarm to send a notification when the metric value exceeds zero.
- D. Load AWS Glue job logs to an Amazon S3 bucket. Configure an Amazon CloudWatch alarm to send a notification when the job logs in the S3 bucket contain Job.State=FAILED.

Answer: A

Explanation:

Option B is the most operationally efficient because it checks the business requirement directly: whether the target table contains the most recent data, not merely whether a job failed. Monitoring only failures (Options C and D) can produce false positives (a job failure might not impact freshness) and false negatives (a job can succeed but still load stale or incomplete data). The study material emphasizes implementing data quality validation as part of the ETL process so data can be verified before or as it is stored, rather than relying only on pipeline execution status.

Using a data quality rule focused on freshness (for example, validating that a "max event timestamp" or "latest partition date" meets today's expected value) lets the pipeline detect stale loads even when the workflow runs. Then, an EventBridge rule can route failures of that data quality check to SNS for immediate notification, keeping operations serverless and centralized. Macie (Option A) is designed for sensitive-data discovery/classification, not operational "freshness" checks on Redshift tables, so it adds unnecessary services and effort compared to a Glue-native data quality validation approach.

NEW QUESTION # 86

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 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.
- B. 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.
- 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. Allocate data processing units (DPUs) to run the crawler every day. Specify a database name for the output.

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.

References:

- * 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 # 87

A company has several new datasets in CSV and JSON formats. A data engineer needs to make the data available to a team of data analysts who will analyze the data by using SQL queries.

Which solution will meet these requirements in the MOST cost-effective way?

- A. Create an AWS Glue DataBrew project that contains the new data. Make the DataBrew project available to the data analysts.
- B. Load the data into SPICE (Super-fast, Parallel, In-memory Calculation Engine) in Amazon QuickSight. Allow the data analysts to create analyses and dashboards in QuickSight.
- C. Create an Amazon RDS for MySQL DB cluster. Use AWS Glue to transform and load the CSV and JSON files into database tables. Provide the data analysts access to the DB cluster.
- **D. Store the data in an Amazon S3 bucket. Use an AWS Glue crawler to catalog the S3 data as tables. Create an Amazon Athena workgroup that has a data usage threshold. Grant the data analysts access to the Athena workgroup.**

Answer: D

Explanation:

Option C is the most cost-effective because it keeps the datasets in Amazon S3 and uses Amazon Athena to query them with SQL only when needed, avoiding the cost of running always-on database infrastructure. The study material states that "Amazon Athena is a serverless service that allows you to query data stored in Amazon S3 using standard SQL". This directly matches the requirement that analysts will "analyze the data by using SQL queries," while remaining cost-efficient due to serverless, on-demand querying.

To make CSV and JSON in S3 easily queryable, metadata must be discoverable and managed. The material also highlights that AWS Glue automates cataloging data in S3 through the AWS Glue Data Catalog, which

"helps discover and manage metadata for data stored in AWS," enabling query engines to treat file data as tables .

Other options are less aligned: RDS MySQL adds continuous capacity and administration costs; DataBrew is primarily for visual preparation (clean/normalize) rather than serving as the SQL query layer ; and QuickSight /SPICE targets BI dashboards, not general-purpose ad hoc SQL over raw files.

NEW QUESTION # 88

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

Answer: C

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).

Reference:

Amazon Athena SQL Reference

Creating Tables in Athena

NEW QUESTION # 89

.....

Cease to struggle and you cease to live. Only by continuous learning can we not be surpassed by others. Many people do not like to study and think that learning is a very vexing thing. This kind of cognition makes their careers stagnate. Data-Engineer-Associate test question will change your perception. Data-Engineer-Associate learning dumps aim to help students learn easily and effectively that has been developed over many years by many industry experts. With Data-Engineer-Associate study tool, you no longer need to look at a drowsy textbook. You do not need to study day and night. With Data-Engineer-Associate learning dumps, you only need to spend 20-30 hours on studying, and then you can easily pass the exam. At the same time, the language in Data-Engineer-Associate test question is very simple and easy to understand. Even if you are a newcomer who has just entered the industry, you can learn all the knowledge points without any obstacles. We believe that Data-Engineer-Associate study tool will make you fall in love with learning. Come and buy it now.

New Data-Engineer-Associate Exam Questions: <https://www.pass4sures.top/AWS-Certified-Data-Engineer/Data-Engineer-Associate-testking-braindumps.html>

Amazon Sample Data-Engineer-Associate Questions Besides, printed material would be suitable for some candidates who are not convenient to use electronic products, Pass4sures Data-Engineer-Associate test questions will be your best choice, Amazon Sample Data-Engineer-Associate Questions So you must act from now, By engaging in learning modes and Data-Engineer-Associate test modes, you can effectively enhance your understanding of the Data-Engineer-Associate exam and build the confidence needed to succeed, Amazon Sample Data-Engineer-Associate Questions With each learning option, you get quality material that covers all exam objectives to boost your performance.

Organize and rate images and assigning tags in the field, and Valid Exam Data-Engineer-Associate Blueprint integrating them later with Adobe Photoshop Lightroom on your computer, And Now for Something Completely Different!

Besides, printed material would be suitable for some candidates who are not convenient to use electronic products, Pass4sures Data-Engineer-Associate Test Questions will be your best choice.

Passing Data-Engineer-Associate Exam Prep Materials - Data-Engineer-Associate Valid Braindumps - Pass4sures

So you must act from now, By engaging in learning modes and Data-Engineer-Associate test modes, you can effectively enhance your understanding of the Data-Engineer-Associate exam and build the confidence needed to succeed.

With each learning option, you get quality Data-Engineer-Associate material that covers all exam objectives to boost your performance.

- Pass Guaranteed 2026 Amazon Data-Engineer-Associate: High Hit-Rate Sample AWS Certified Data Engineer - Associate (DEA-C01) Questions Search on 《 www.verifieddumps.com 》 for 《 Data-Engineer-Associate 》 to obtain exam materials for free download Data-Engineer-Associate Popular Exams
- Training Data-Engineer-Associate Kit Reliable Data-Engineer-Associate Test Book Data-Engineer-Associate Exam Forum Download Data-Engineer-Associate for free by simply entering www.pdfvce.com website Latest Data-Engineer-Associate Exam Notes
- Pass Guaranteed 2026 Amazon Data-Engineer-Associate: High Hit-Rate Sample AWS Certified Data Engineer - Associate (DEA-C01) Questions Immediately open www.vce4dumps.com and search for Data-Engineer-Associate to obtain a free download Latest Data-Engineer-Associate Exam Notes
- Data-Engineer-Associate Trustworthy Exam Content Valid Test Data-Engineer-Associate Vce Free Data-Engineer-Associate Trustworthy Exam Content Download Data-Engineer-Associate for free by simply searching on www.pdfvce.com Data-Engineer-Associate Popular Exams
- 100% Pass Quiz 2026 Amazon Data-Engineer-Associate: AWS Certified Data Engineer - Associate (DEA-C01) – Trustable Sample Questions www.prepawaypdf.com is best website to obtain [Data-Engineer-Associate] for free download Latest Data-Engineer-Associate Test Vce
- Examcollection Data-Engineer-Associate Dumps Latest Data-Engineer-Associate Braindumps Pdf Latest Data-Engineer-Associate Braindumps Pdf Easily obtain Data-Engineer-Associate for free download through 《 www.pdfvce.com 》 Latest Data-Engineer-Associate Mock Exam
- Pass Guaranteed 2026 Amazon Data-Engineer-Associate: High Hit-Rate Sample AWS Certified Data Engineer - Associate (DEA-C01) Questions Search for Data-Engineer-Associate and download it for free immediately on [www.easy4engine.com] Latest Braindumps Data-Engineer-Associate Book
- 100% Pass Quiz Amazon - Data-Engineer-Associate - Authoritative Sample AWS Certified Data Engineer - Associate (DEA-C01) Questions Simply search for 《 Data-Engineer-Associate 》 for free download on www.pdfvce.com Data-Engineer-Associate Certification Test Answers
- Latest Braindumps Data-Engineer-Associate Book Data-Engineer-Associate Reliable Exam Voucher Data-Engineer-Associate Trustworthy Exam Content Search for 《 Data-Engineer-Associate 》 on (www.verifieddumps.com) immediately to obtain a free download Data-Engineer-Associate Valid Exam Cost
- Data-Engineer-Associate Exam Questions without being overloaded with unnecessary details Copy URL [www.pdfvce.com] open and search for Data-Engineer-Associate to download for free Latest Braindumps Data-Engineer-Associate Book
- Data-Engineer-Associate Reliable Exam Voucher Data-Engineer-Associate Certification Test Answers Data-Engineer-Associate Reliable Exam Voucher Open website www.prepawayete.com and search for Data-Engineer-Associate for free download Data-Engineer-Associate Free Exam Questions
- www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, bbs.t-firefly.com, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, 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, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, Disposable vapes

BTW, DOWNLOAD part of Pass4sures Data-Engineer-Associate dumps from Cloud Storage: https://drive.google.com/open?id=16LUsIRF336O3UApq_5sxf9XOjyQmJnpA