

Data-Engineer-Associate Exam Score, Data-Engineer-Associate Exam Preview

Exam:	Databricks Certified Data Engineer Associate (Version 2)
Date Completed:	29 October 2022
Time Taken:	51 minutes of 90 minutes total
Score:	80.00%
Result:	Pass
Topic	Percentage Correct
Databricks Lakehouse Platform	72.72%
ELT with Spark SQL and Python	92.30%
Incremental Data Processing	83.33%
Production Pipelines	57.14%
Data Governance	100.00%

Congratulations! You've passed the exam. You will receive the certification via Databricks Academy within 24 hours.

P.S. Free & New Data-Engineer-Associate dumps are available on Google Drive shared by ITExamDownload:
<https://drive.google.com/open?id=1skPLwEztrqK-pxRnztldwvfl7V2bSSy->

Make yourself more valuable in today's competitive computer industry ITExamDownload's preparation material includes the most excellent features, prepared by the same dedicated experts who have come together to offer an integrated solution. ITExamDownload's Data-Engineer-Associate preparation material includes the most excellent features, prepared by the same dedicated experts who have come together to offer an integrated solution. Data-Engineer-Associate Preparation material guarantee that you will get most excellent and simple method to pass your certification Data-Engineer-Associate exams on the first attempt.

Our Data-Engineer-Associate prep material target all users and any learners, regardless of their age, gender and education background. We provide 3 versions for the clients to choose based on the consideration that all the users can choose the most suitable version to learn. The 3 versions each support different using method and equipment and the client can use the Data-Engineer-Associate Exam Dump on the smart phones, laptops or the tablet computers. The clients can choose the version which supports their equipment on their hands to learn.

>> Data-Engineer-Associate Exam Score <<

Data-Engineer-Associate Exam Preview, Reliable Data-Engineer-Associate Exam Cram

The clients can download our products and use our Data-Engineer-Associate study materials immediately after they pay successfully with their credit cards. Our system will send our Data-Engineer-Associate learning prep in the form of mails to the client in 5-10 minutes after their successful payment. The mails provide the links and if only the clients click on the links they can log in our software immediately to learn our Data-Engineer-Associate Guide materials. If there are something they can't understand, they can contact with our service and we will solve them right away.

Amazon AWS Certified Data Engineer - Associate (DEA-C01) Sample Questions (Q55-Q60):

NEW QUESTION # 55

A data engineer needs to build an enterprise data catalog based on the company's Amazon S3 buckets and Amazon RDS databases. The data catalog must include storage format metadata for the data in the catalog. Which solution will meet these requirements with the LEAST effort?

- A. Use Amazon Macie to build a data catalog and to identify sensitive data elements. Collect the data format information from Macie.
- B. Use scripts to scan data elements and to assign data classifications based on the format of the data.
- C. Use an AWS Glue crawler to build a data catalog. Use AWS Glue crawler classifiers to recognize the format of data and store the format in the catalog.
- D. Use an AWS Glue crawler to scan the S3 buckets and RDS databases and build a data catalog. Use data stewards to inspect the data and update the data catalog with the data format.

Answer: C

Explanation:

To build an enterprise data catalog with metadata for storage formats, the easiest and most efficient solution is using an AWS Glue crawler. The Glue crawler can scan Amazon S3 buckets and Amazon RDS databases to automatically create a data catalog that includes metadata such as the schema and storage format (e.g., CSV, Parquet, etc.). By using AWS Glue crawler classifiers, you can configure the crawler to recognize the format of the data and store this information directly in the catalog.

Option B: Use an AWS Glue crawler to build a data catalog. Use AWS Glue crawler classifiers to recognize the format of data and store the format in the catalog.

This option meets the requirements with the least effort because Glue crawlers automate the discovery and cataloging of data from multiple sources, including S3 and RDS, while recognizing various file formats via classifiers.

Other options (A, C, D) involve additional manual steps, like having data stewards inspect the data, or using services like Amazon Macie that focus more on sensitive data detection rather than format cataloging.

Reference:

AWS Glue Crawler Documentation

AWS Glue Classifiers

NEW QUESTION # 56

A company is migrating its database servers from Amazon EC2 instances that run Microsoft SQL Server to Amazon RDS for Microsoft SQL Server DB instances. The company's analytics team must export large data elements every day until the migration is complete. The data elements are the result of SQL joins across multiple tables. The data must be in Apache Parquet format. The analytics team must store the data in Amazon S3.

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

- A. Use a SQL query to create a view in the EC2 instance-based SQL Server databases that contains the required data elements. Create and run an AWS Glue crawler to read the view. Create an AWS Glue job that retrieves the data and transfers the data in Parquet format to an S3 bucket. Schedule the AWS Glue job to run every day.
- B. Create an AWS Lambda function that queries the EC2 instance-based databases by using Java Database Connectivity (JDBC). Configure the Lambda function to retrieve the required data, transform the data into Parquet format, and transfer the data into an S3 bucket. Use Amazon EventBridge to schedule the Lambda function to run every day.
- **C. Create a view in the EC2 instance-based SQL Server databases that contains the required data elements. Create an AWS Glue job that selects the data directly from the view and transfers the data in Parquet format to an S3 bucket. Schedule the AWS Glue job to run every day.**
- D. Schedule SQL Server Agent to run a daily SQL query that selects the desired data elements from the EC2 instance-based SQL Server databases. Configure the query to direct the output .csv objects to an S3 bucket. Create an S3 event that invokes an AWS Lambda function to transform the output format from .csv to Parquet.

Answer: C

Explanation:

Option A is the most operationally efficient way to meet the requirements because it minimizes the number of steps and services involved in the data export process. AWS Glue is a fully managed service that can extract, transform, and load (ETL) data from various sources to various destinations, including Amazon S3. AWS Glue can also convert data to different formats, such as Parquet, which is a columnar storage format that is optimized for analytics. By creating a view in the SQL Server databases that contains the required data elements, the AWS Glue job can select the data directly from the view without having to perform any joins or transformations on the source data. The AWS Glue job can then transfer the data in Parquet format to an S3 bucket and run on a daily schedule.

Option B is not operationally efficient because it involves multiple steps and services to export the data. SQL Server Agent is a tool that can run scheduled tasks on SQL Server databases, such as executing SQL queries.

However, SQL Server Agent cannot directly export data to S3, so the query output must be saved as .csv objects on the EC2 instance. Then, an S3 event must be configured to trigger an AWS Lambda function that can transform the .csv objects to Parquet format and upload them to S3. This option adds complexity and latency to the data export process and requires additional resources and configuration.

Option C is not operationally efficient because it introduces an unnecessary step of running an AWS Glue crawler to read the view. An AWS Glue crawler is a service that can scan data sources and create metadata tables in the AWS Glue Data Catalog. The Data Catalog is a central repository that stores information about the data sources, such as schema, format, and location. However, in this scenario, the schema and format of the data elements are already known and fixed, so there is no need to run a crawler to discover them. The AWS Glue job can directly select the data from the view without using the Data Catalog. Running a crawler adds extra time and cost to the data export process.

Option D is not operationally efficient because it requires custom code and configuration to query the databases and transform the

data. An AWS Lambda function is a service that can run code in response to events or triggers, such as Amazon EventBridge. Amazon EventBridge is a service that can connect applications and services with event sources, such as schedules, and route them to targets, such as Lambda functions. However, in this scenario, using a Lambda function to query the databases and transform the data is not the best option because it requires writing and maintaining code that uses JDBC to connect to the SQL Server databases, retrieve the required data, convert the data to Parquet format, and transfer the data to S3.

This option also has limitations on the execution time, memory, and concurrency of the Lambda function, which may affect the performance and reliability of the data export process.

References:

- * AWS Certified Data Engineer - Associate DEA-C01 Complete Study Guide
- * AWS Glue Documentation
- * Working with Views in AWS Glue
- * Converting to Columnar Formats

NEW QUESTION # 57

A data engineer has a one-time task to read data from objects that are in Apache Parquet format in an Amazon S3 bucket. The data engineer needs to query only one column of the data.

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

- **A. Use S3 Select to write a SQL SELECT statement to retrieve the required column from the S3 objects.**
- B. Run an AWS Glue crawler on the S3 objects. Use a SQL SELECT statement in Amazon Athena to query the required column.
- C. Configure an AWS Lambda function to load data from the S3 bucket into a pandas dataframe- Write a SQL SELECT statement on the dataframe to query the required column.
- D. Prepare an AWS Glue DataBrew project to consume the S3 objects and to query the required column.

Answer: A

Explanation:

Option B is the best solution to meet the requirements with the least operational overhead because S3 Select is a feature that allows you to retrieve only a subset of data from an S3 object by using simple SQL expressions.

S3 Select works on objects stored in CSV, JSON, or Parquet format. By using S3 Select, you can avoid the need to download and process the entire S3 object, which reduces the amount of data transferred and the computation time. S3 Select is also easy to use and does not require any additional services or resources.

Option A is not a good solution because it involves writing custom code and configuring an AWS Lambda function to load data from the S3 bucket into a pandas dataframe and query the required column. This option adds complexity and latency to the data retrieval process and requires additional resources and configuration.

Moreover, AWS Lambda has limitations on the execution time, memory, and concurrency, which may affect the performance and reliability of the data retrieval process.

Option C is not a good solution because it involves creating and running an AWS Glue DataBrew project to consume the S3 objects and query the required column. AWS Glue DataBrew is a visual data preparation tool that allows you to clean, normalize, and transform data without writing code. However, in this scenario, the data is already in Parquet format, which is a columnar storage format that is optimized for analytics.

Therefore, there is no need to use AWS Glue DataBrew to prepare the data. Moreover, AWS Glue DataBrew adds extra time and cost to the data retrieval process and requires additional resources and configuration.

Option D is not a good solution because it involves running an AWS Glue crawler on the S3 objects and using a SQL SELECT statement in Amazon Athena to query the required column. An AWS Glue crawler is a service that can scan data sources and create metadata tables in the AWS Glue Data Catalog. The Data Catalog is a central repository that stores information about the data sources, such as schema, format, and location. Amazon Athena is a serverless interactive query service that allows you to analyze data in S3 using standard SQL. However, in this scenario, the schema and format of the data are already known and fixed, so there is no need to run a crawler to discover them. Moreover, running a crawler and using Amazon Athena adds extra time and cost to the data retrieval process and requires additional services and configuration.

:

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

S3 Select and Glacier Select - Amazon Simple Storage Service

AWS Lambda - FAQs

What Is AWS Glue DataBrew? - AWS Glue DataBrew

Populating the AWS Glue Data Catalog - AWS Glue

What is Amazon Athena? - Amazon Athena

NEW QUESTION # 58

A company is using Amazon Redshift to build a data warehouse solution. The company is loading hundreds of files into a table that is in a Redshift cluster.

The company wants the data warehouse solution to achieve the greatest possible throughput. The solution must use cluster resources optimally when the company loads data into the table.

Which solution will meet these requirements?

- A. Use S3DistCp to load multiple files into Hadoop Distributed File System (HDFS). Use an HDFS connector to ingest the data into the Redshift cluster.
- **B. Use a single COPY command to load the data into the Redshift cluster.**
- C. Use multiple COPY commands to load the data into the Redshift cluster.
- D. Use a number of INSERT statements equal to the number of Redshift cluster nodes. Load the data in parallel into each node.

Answer: B

Explanation:

To achieve the highest throughput and efficiently use cluster resources while loading data into an Amazon Redshift cluster, the optimal approach is to use a single COPY command that ingests data in parallel.

Option D: Use a single COPY command to load the data into the Redshift cluster.

The COPY command is designed to load data from multiple files in parallel into a Redshift table, using all the cluster nodes to optimize the load process. Redshift is optimized for parallel processing, and a single COPY command can load multiple files at once, maximizing throughput.

Options A, B, and C either involve unnecessary complexity or inefficient approaches, such as using multiple COPY commands or INSERT statements, which are not optimized for bulk loading.

Reference:

Amazon Redshift COPY Command Documentation

NEW QUESTION # 59

A company uses AWS Glue jobs to implement several data pipelines. The pipelines are critical to the company.

The company needs to implement a monitoring mechanism that will alert stakeholders if the pipelines fail.

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

- A. Configure an Amazon CloudWatch Logs log group for the AWS Glue jobs. Create an Amazon EventBridge rule to match new log creation events in the log group. Configure the rule to send notifications to an Amazon Simple Notification Service (Amazon SNS) topic.
- **B. Create an Amazon EventBridge rule to match AWS Glue job failure events. Configure the rule to target an AWS Lambda function to process events. Configure the function to send notifications to an Amazon Simple Notification Service (Amazon SNS) topic.**
- C. Create an Amazon EventBridge rule to match AWS Glue job failure events. Define an Amazon CloudWatch metric based on the EventBridge rule. Set up a CloudWatch alarm based on the metric to send notifications to an Amazon Simple Notification Service (Amazon SNS) topic.
- D. Configure an Amazon CloudWatch Logs log group for the AWS Glue jobs. Create an Amazon EventBridge rule to match new log creation events in the log group. Configure the rule to target an AWS Lambda function that reads the logs and sends notifications to an Amazon Simple Notification Service (Amazon SNS) topic if AWS Glue job failure logs are present.

Answer: B

Explanation:

Creating an EventBridge rule that triggers a Lambda function on AWS Glue job failure events and then sends notifications via Amazon SNS is the most direct and operationally efficient method:

"Practice Quiz 10: A data engineer must monitor the data pipeline... Which solution will meet these requirements?"

A: Inspect the job run monitoring section of the AWS Glue console.

-Ace the AWS Certified Data Engineer - Associate Certification - version 2 - apple.pdf Although this reference directly supports using AWS Glue's monitoring features via EventBridge, it implies that solutions like A-which directly use EventBridge failure events for automation-are more optimal and less complex than constructing custom logs and metrics.

NEW QUESTION # 60

.....

Our Data-Engineer-Associate study guide provides free trial services, so that you can learn about some of our topics and how to open the software before purchasing. During the trial period of our Data-Engineer-Associate study materials, the PDF versions of the sample questions are available for free download, and both the pc version and the online version can be illustrated clearly. You can contact us at any time if you have any difficulties in the purchase or trial process of our Data-Engineer-Associate Exam Dumps.

Data-Engineer-Associate Exam Preview: <https://www.itexamdownload.com/Data-Engineer-Associate-valid-questions.html>

With ITExamDownload Amazon Data-Engineer-Associate exam certification training, you can sort out your messy thoughts, and no longer twitchy for the exam, Amazon Data-Engineer-Associate Exam Score By analyzing this report you can eliminate and overcome your mistakes, These are just the few glimpses of ITExamDownload Data-Engineer-Associate Exam Preview's Data-Engineer-Associate Exam Preview - AWS Certified Data Engineer - Associate (DEA-C01) Questions and answers to ensure your success in Checkpoint Data-Engineer-Associate Exam Preview - AWS Certified Data Engineer - Associate (DEA-C01) certification exam, Furthermore, the Data-Engineer-Associate practice exam works smoothly on all operating systems including Mac, Linux, IOS, Android, and Windows.

Security auditing is a chore that many system and network administrators tend Reliable Data-Engineer-Associate Exam Cram to neglect, Cryptography Decrypted explains cryptography in plain English" and is authoritative and thorough enough to address the needs of professionals.

Data-Engineer-Associate AWS Certified Data Engineer - Associate (DEA-C01) Dumps For Ultimate Results 2025

With ITExamDownload Amazon Data-Engineer-Associate Exam Certification training, you can sort out your messy thoughts, and no longer twitchy for the exam, By analyzing this report you can eliminate and overcome your mistakes.

These are just the few glimpses of ITExamDownload's Data-Engineer-Associate AWS Certified Data Engineer - Associate (DEA-C01) Questions and answers to ensure your success in Checkpoint AWS Certified Data Engineer - Associate (DEA-C01) certification exam, Furthermore, the Data-Engineer-Associate practice exam works smoothly on all operating systems including Mac, Linux, IOS, Android, and Windows.

With our Data-Engineer-Associate training braindumps, you must feel respected.

- Latest Data-Engineer-Associate Braindumps □ Data-Engineer-Associate Valid Exam Fee □ Study Data-Engineer-Associate Reference □ Search for ► Data-Engineer-Associate □ and download it for free on ► www.lead1pass.com □ website ♣ Data-Engineer-Associate Exam Prep
- New Data-Engineer-Associate Exam Score | Efficient Data-Engineer-Associate Exam Preview: AWS Certified Data Engineer - Associate (DEA-C01) 100% Pass □ Immediately open ► www.pdfvce.com □□□ and search for ► Data-Engineer-Associate □□□ to obtain a free download □ Data-Engineer-Associate Exam Prep
- TOP Data-Engineer-Associate Exam Score 100% Pass | The Best Amazon AWS Certified Data Engineer - Associate (DEA-C01) Exam Preview Pass for sure * Enter ⇒ www.vceengine.com ⇐ and search for [Data-Engineer-Associate] to download for free □ Data-Engineer-Associate Test Collection Pdf
- Exam Data-Engineer-Associate Pattern □ Data-Engineer-Associate Test Collection Pdf □ Data-Engineer-Associate Sample Questions □ Easily obtain ► Data-Engineer-Associate □ for free download through ► www.pdfvce.com □□□ □ Valid Data-Engineer-Associate Cram Materials
- AWS Certified Data Engineer - Associate (DEA-C01) training pdf vce - Data-Engineer-Associate online test engine - AWS Certified Data Engineer - Associate (DEA-C01) valid practice demo □ Download ► Data-Engineer-Associate □□□ for free by simply entering ► www.testsimulate.com □ website □ Valid Data-Engineer-Associate Cram Materials
- Data-Engineer-Associate New Braindumps Questions □ Data-Engineer-Associate Test Collection Pdf □ Latest Data-Engineer-Associate Braindumps □ Search for “Data-Engineer-Associate” and obtain a free download on ► www.pdfvce.com ◀ □ Data-Engineer-Associate Practice Exams
- Data-Engineer-Associate Valid Exam Fee □ Valid Data-Engineer-Associate Exam Materials □ Data-Engineer-Associate Test Collection Pdf □ Search for (Data-Engineer-Associate) and easily obtain a free download on 《 www.lead1pass.com 》 □ Data-Engineer-Associate Valid Exam Question
- Data-Engineer-Associate Practice Exams □ Reliable Data-Engineer-Associate Exam Tutorial □ Valid Data-Engineer-Associate Exam Materials □ Easily obtain ► Data-Engineer-Associate □□□ for free download through { www.pdfvce.com } □ Exam Data-Engineer-Associate Pattern
- Exam Data-Engineer-Associate Pattern □ Data-Engineer-Associate Latest Exam Pdf □ Study Data-Engineer-Associate Center □ Open website □ www.examcollectionpass.com □ and search for ☀ Data-Engineer-Associate ☀ □ for free download □ Data-Engineer-Associate Latest Exam Pdf
- Data-Engineer-Associate Valid Dumps □ Data-Engineer-Associate Valid Dumps □ Reliable Data-Engineer-Associate Exam Tutorial □ Copy URL ► www.pdfvce.com □ open and search for 《 Data-Engineer-Associate 》 to download

2025 Data-Engineer-Associate: Latest AWS Certified Data

- myportal.utt.edu.tt, motionentrance.edu.np, cyberversity.global, Disposable vapes

<https://drive.google.com/open?id=1skPLwEzzrrqK-pxRnzt dwvf7V2bSSy->

<https://drive.google.com/open?id=1skPLwEzzrrqK-pxRnzt dwvf7V2bSSy->