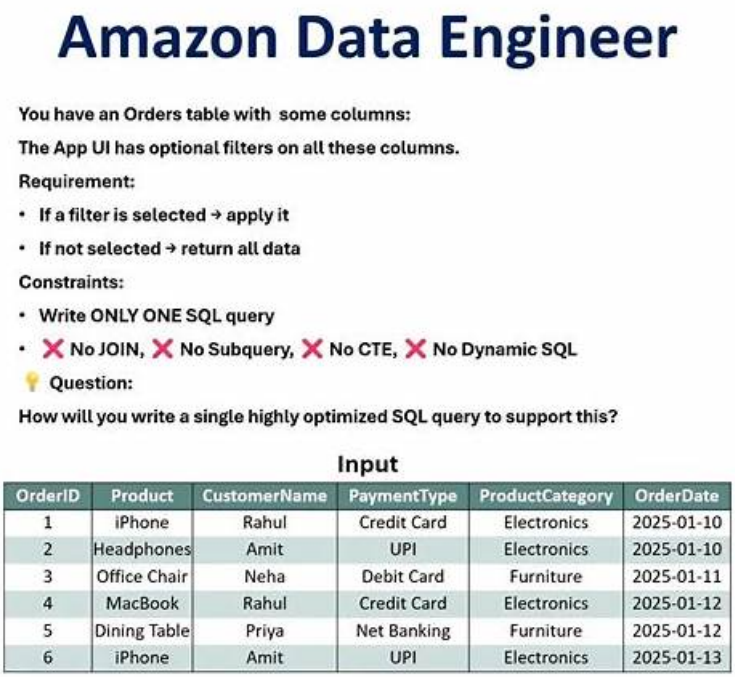
Amazon Data-Engineer-Associate Actual Test | Top Data-Engineer-Associate Questions



What's more, part of that Prep4away Data-Engineer-Associate dumps now are free: https://drive.google.com/open?id=1w8XKXmU1VhFRAs6w9RdeqCxRZET5qDig

Prep4away has many AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) practice questions that reflect the pattern of the real AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) exam. Prep4away allows you to create a AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) exam dumps according to your preparation. It is easy to create the Amazon Data-Engineer-Associate Practice Questions by following just a few simple steps. Our AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) exam dumps are customizable based on the time and type of questions.

If you have bad mood in your test every time you should choose our Soft test engine or App test engine of Data-Engineer-Associate dumps torrent materials. Both of these two versions have one function is simulating the real test scene. You can set timed exam and practice many times. You can feel exam pace and hold time to test with our Amazon Data-Engineer-Associate Dumps Torrent. You should take advantage of the time and opportunities you have to do the things you want. Our Data-Engineer-Associate dumps torrent files provide you to keep good mood for the test.

>> Amazon Data-Engineer-Associate Actual Test <<

Preparing for Amazon Data-Engineer-Associate PDF Exam Questions In Short Time

You may feel astonished and doubtful about this figure; but we do make our Data-Engineer-Associate exam dumps well received by most customers. Better still, the 98-99% pass rate has helped most of the candidates get the certification successfully, which is far beyond that of others in this field. In recent years, supported by our professional expert team, our Data-Engineer-Associate test braindumps have grown up and have made huge progress. Our Data-Engineer-Associate Exam Dumps strive for providing you a comfortable study platform and continuously explore more functions to meet every customer's requirements. We may foresee the prosperous talent market with more and more workers attempting to reach a high level through the Amazon certification.

Amazon AWS Certified Data Engineer - Associate (DEA-C01) Sample Questions (Q125-Q130):

NEW QUESTION # 125
A company currently stores all of its data in Amazon S3 by using the S3 Standard storage class.
A data engineer examined data access patterns to identify trends. During the first 6 months, most data files are accessed several times each day. Between 6 months and 2 years, most data files are accessed once or twice each month. After 2 years, data files are accessed only once or twice each year.
The data engineer needs to use an S3 Lifecycle policy to develop new data storage rules. The new storage solution must continue to provide high availability.
Which solution will meet these requirements in the MOST cost-effective way?

- A. Transition objects to S3 Standard-Infrequent Access (S3 Standard-IA) after 6 months. Transfer objects to S3 Glacier Flexible Retrieval after 2 years.
- B. Transition objects to S3 One Zone-Infrequent Access (S3 One Zone-IA) after 6 months. Transfer objects to S3 Glacier Deep Archive after 2 years.
- C. Transition objects to S3 Standard-Infrequent Access (S3 Standard-IA) after 6 months. Transfer objects to S3 Glacier Deep Archive after 2 years.
- D. Transition objects to S3 One Zone-Infrequent Access (S3 One Zone-IA) after 6 months. Transfer objects to S3 Glacier Flexible Retrieval after 2 years.

Answer: C

Explanation:
To achieve the most cost-effective storage solution, the data engineer needs to use an S3 Lifecycle policy that transitions objects to lower-cost storage classes based on their access patterns, and deletes them when they are no longer needed. The storage classes should also provide high availability, which means they should be resilient to the loss of data in a single Availability Zone1. Therefore, the solution must include the following steps:
* Transition objects to S3 Standard-Infrequent Access (S3 Standard-IA) after 6 months. S3 Standard-IA is designed for data that is accessed less frequently, but requires rapid access when needed. It offers the same high durability, throughput, and low latency as S3 Standard, but with a lower storage cost and a retrieval fee2. Therefore, it is suitable for data files that are accessed once or twice each month. S3 Standard-IA also provides high availability, as it stores data redundantly across multiple Availability Zones1.
* Transfer objects to S3 Glacier Deep Archive after 2 years. S3 Glacier Deep Archive is the lowest-cost storage class that offers secure and durable storage for data that is rarely accessed and can tolerate a 12- hour retrieval time. It is ideal for long-term archiving and digital preservation3. Therefore, it is suitable for data files that are accessed only once or twice each year. S3 Glacier Deep Archive also provides high availability, as it stores data across at least three geographically dispersed Availability Zones1.
* Delete objects when they are no longer needed. The data engineer can specify an expiration action in the S3 Lifecycle policy to delete objects after a certain period of time. This will reduce the storage cost and comply with any data retention policies.
Option C is the only solution that includes all these steps. Therefore, option C is the correct answer.
Option A is incorrect because it transitions objects to S3 One Zone-Infrequent Access (S3 One Zone-IA) after
6 months. S3 One Zone-IA is similar to S3 Standard-IA, but it stores data in a single Availability Zone. This means it has a lower availability and durability than S3 Standard-IA, and it is not resilient to the loss of data in a single Availability Zone1. Therefore, it does not provide high availability as required.
Option B is incorrect because it transfers objects to S3 Glacier Flexible Retrieval after 2 years. S3 Glacier Flexible Retrieval is a storage class that offers secure and durable storage for data that is accessed infrequently and can tolerate a retrieval time of minutes to hours. It is more expensive than S3 Glacier Deep Archive, and it is not suitable for data that is accessed only once or twice each year3. Therefore, it is not the most cost-effective option.
Option D is incorrect because it combines the errors of option A and B. It transitions objects to S3 One Zone- IA after 6 months, which does not provide high availability, and it transfers objects to S3 Glacier Flexible Retrieval after 2 years, which is not the most cost-effective option.
:
1: Amazon S3 storage classes - Amazon Simple Storage Service
2: Amazon S3 Standard-Infrequent Access (S3 Standard-IA) - Amazon Simple Storage Service
3: Amazon S3 Glacier and S3 Glacier Deep Archive - Amazon Simple Storage Service
[4]: Expiring objects - Amazon Simple Storage Service
[5]: Managing your storage lifecycle - Amazon Simple Storage Service
[6]: Examples of S3 Lifecycle configuration - Amazon Simple Storage Service
[7]: Amazon S3 Lifecycle further optimizes storage cost savings with new features - What's New with AWS

NEW QUESTION # 126
A company uses Amazon RDS for MySQL as the database for a critical application. The database workload is mostly writes, with a small number of reads.
A data engineer notices that the CPU utilization of the DB instance is very high. The high CPU utilization is slowing down the

application. The data engineer must reduce the CPU utilization of the DB Instance.
Which actions should the data engineer take to meet this requirement? (Choose two.)

- A. Use the Performance Insights feature of Amazon RDS to identify queries that have high CPU utilization. Optimize the problematic queries.
- B. Upgrade to a larger instance size.
- C. Reboot the RDS DB instance once each week.
- D. Modify the database schema to include additional tables and indexes.
- E. Implement caching to reduce the database query load.

Answer: A,E

Explanation:
Amazon RDS is a fully managed service that provides relational databases in the cloud. Amazon RDS for MySQL is one of the supported database engines that you can use to run your applications. Amazon RDS provides various features and tools to monitor and optimize the performance of your DB instances, such as Performance Insights, Enhanced Monitoring, CloudWatch metrics and alarms, etc.
Using the Performance Insights feature of Amazon RDS to identify queries that have high CPU utilization and optimizing the problematic queries will help reduce the CPU utilization of the DB instance. Performance Insights is a feature that allows you to analyze the load on your DB instance and determine what is causing performance issues. Performance Insights collects, analyzes, and displays database performance data using an interactive dashboard. You can use Performance Insights to identify the top SQL statements, hosts, users, or processes that are consuming the most CPU resources. You can also drill down into the details of each query and see the execution plan, wait events, locks, etc. By using Performance Insights, you can pinpoint the root cause of the high CPU utilization and optimize the queries accordingly. For example, you can rewrite the queries to make them more efficient, add or remove indexes, use prepared statements, etc.
Implementing caching to reduce the database query load will also help reduce the CPU utilization of the DB instance. Caching is a technique that allows you to store frequently accessed data in a fast and scalable storage layer, such as Amazon ElastiCache. By using caching, you can reduce the number of requests that hit your database, which in turn reduces the CPU load on your DB instance. Caching also improves the performance and availability of your application, as it reduces the latency and increases the throughput of your data access.
You can use caching for various scenarios, such as storing session data, user preferences, application configuration, etc. You can also use caching for read-heavy workloads, such as displaying product details, recommendations, reviews, etc.
The other options are not as effective as using Performance Insights and caching. Modifying the database schema to include additional tables and indexes may or may not improve the CPU utilization, depending on the nature of the workload and the queries. Adding more tables and indexes may increase the complexity and overhead of the database, which may negatively affect the performance. Rebooting the RDS DB instance once each week will not reduce the CPU utilization, as it will not address the underlying cause of the high CPU load. Rebooting may also cause downtime and disruption to your application. Upgrading to a larger instance size may reduce the CPU utilization, but it will also increase the cost and complexity of your solution.
Upgrading may also not be necessary if you can optimize the queries and reduce the database load by using caching. References:
* Amazon RDS
* Performance Insights
* Amazon ElastiCache
* [AWS Certified Data Engineer - Associate DEA-C01 Complete Study Guide], Chapter 3: Data Storage and Management, Section 3.1: Amazon RDS

NEW QUESTION # 127
A company aggregates high-frequency sensor telemetry into an Amazon S3 data lake. Each sensor stream emits structured records every hour. The records include metadata such as sensor category, unit ID, operational state, event timestamp, and site location. The data scales up to millions of records each day. The company runs complex queries each day to uncover performance insights specific to sensor categories.
Which solution will meet these requirements with the FASTEST query execution time?

- A. Persist the data in CSV format. Partition the data by date. Sort the data by operational status.
- B. Persist the data in Parquet format. Partition the data by sensor category. Sort the data by date.
- C. Persist the data in CSV format. Partition the data by date. Sort the data by sensor category.
- D. Persist the data in Apache ORC format. Partition the data by date. Sort the data by sensor category.

Answer: B

Explanation:

Option C is correct because the fastest design combines a columnar storage format with partitioning on the most common query predicate. AWS Athena guidance says that Parquet and ORC are optimized columnar formats and that columns frequently used as filters are good candidates for partitioning. AWS also states that when a query filters on a partition key, the engine reads only matching partitions, which significantly reduces data scanned. Since the company's main analytical need is insights specific to sensor categories, partitioning by sensor category provides the strongest pruning. Sorting by date then helps organize time-based data within each category partition.
Options B and D use CSV, which is row-based and much slower for analytical scans than Parquet or ORC.
Option A uses a good format, but partitioning by date is less optimal than partitioning by sensor category when category is the main filter in the company's queries. The study guide also identifies Parquet as the preferred analytic storage format for efficient columnar queries.

NEW QUESTION # 128
A company needs to set up a data catalog and metadata management for data sources that run in the AWS Cloud. The company will use the data catalog to maintain the metadata of all the objects that are in a set of data stores. The data stores include structured sources such as Amazon RDS and Amazon Redshift. The data stores also include semistructured sources such as JSON files and .xml files that are stored in Amazon S3.
The company needs a solution that will update the data catalog on a regular basis. The solution also must detect changes to the source metadata.
Which solution will meet these requirements with the LEAST operational overhead?

- A. Use Amazon DynamoDB as the data catalog. Create AWS Lambda functions that will connect to the data catalog. Configure the Lambda functions to gather the metadata information from multiple sources and to update the DynamoDB data catalog. Schedule the Lambda functions to run periodically.
- B. Use the AWS Glue Data Catalog as the central metadata repository. Extract the schema for Amazon RDS and Amazon Redshift sources, and build the Data Catalog. Use AWS Glue crawlers for data that is in Amazon S3 to infer the schema and to automatically update the Data Catalog.
- C. Use the AWS Glue Data Catalog as the central metadata repository. Use AWS Glue crawlers to connect to multiple data stores and to update the Data Catalog with metadata changes. Schedule the crawlers to run periodically to update the metadata catalog.
- D. Use Amazon Aurora as the data catalog. Create AWS Lambda functions that will connect to the data catalog. Configure the Lambda functions to gather the metadata information from multiple sources and to update the Aurora data catalog. Schedule the Lambda functions to run periodically.

Answer: C

Explanation:
This solution will meet the requirements with the least operational overhead because it uses the AWS Glue Data Catalog as the central metadata repository for data sources that run in the AWS Cloud. The AWS Glue Data Catalog is a fully managed service that provides a unified view of your data assets across AWS and on- premises data sources. It stores the metadata of your data in tables, partitions, and columns, and enables you to access and query your data using various AWS services, such as Amazon Athena, Amazon EMR, and Amazon Redshift Spectrum. You can use AWS Glue crawlers to connect to multiple data stores, such as Amazon RDS, Amazon Redshift, and Amazon S3, and to update the Data Catalog with metadata changes.
AWS Glue crawlers can automatically discover the schema and partition structure of your data, and create or update the corresponding tables in the Data Catalog. You can schedule the crawlers to run periodically to update the metadata catalog, and configure them to detect changes to the source metadata, such as new columns, tables, or partitions12.
The other options are not optimal for the following reasons:
* A. Use Amazon Aurora as the data catalog. Create AWS Lambda functions that will connect to the data catalog. Configure the Lambda functions to gather the metadata information from multiple sources and to update the Aurora data catalog. Schedule the Lambda functions to run periodically. This option is not recommended, as it would require more operational overhead to create and manage an Amazon Aurora database as the data catalog, and to write and maintain AWS Lambda functions to gather and update the metadata information from multiple sources. Moreover, this option would not leverage the benefits of the AWS Glue Data Catalog, such as data cataloging, data transformation, and data governance.
* C. Use Amazon DynamoDB as the data catalog. Create AWS Lambda functions that will connect to the data catalog. Configure the Lambda functions to gather the metadata information from multiple sources and to update the DynamoDB data catalog. Schedule the Lambda functions to run periodically. This option is also not recommended, as it would require more operational overhead to create and manage an Amazon DynamoDB table as the data catalog, and to write and maintain AWS Lambda functions to gather and update the metadata information from multiple sources. Moreover, this option would not leverage the benefits of the AWS Glue Data Catalog, such as data cataloging, data transformation, and data governance.
* D. Use the AWS Glue Data Catalog as the central metadata repository. Extract the schema for Amazon RDS and Amazon Redshift sources, and build the Data Catalog. Use AWS Glue crawlers for data that is in Amazon S3 to infer the schema and to

automatically update the Data Catalog. This option is not optimal, as it would require more manual effort to extract the schema for Amazon RDS and Amazon Redshift sources, and to build the Data Catalog. This option would not take advantage of the AWS Glue crawlers' ability to automatically discover the schema and partition structure of your data from various data sources, and to create or update the corresponding tables in the Data Catalog.
References:
* 1: AWS Glue Data Catalog
* 2: AWS Glue Crawlers
* : Amazon Aurora
* : AWS Lambda
* : Amazon DynamoDB

NEW QUESTION # 129
A company receives .csv files that contain physical address data. The data is in columns that have the following names: Door_No, Street_Name, City, and Zip_Code. The company wants to create a single column to store these values in the following format:
Which solution will meet this requirement with the LEAST coding effort?

- A. Write a Lambda function in Python to read the files. Use the Python data dictionary type to create the new column.
- B. Use AWS Glue DataBrew to read the files. Use the NEST TO ARRAY transformation to create the new column.
- C. Use AWS Glue DataBrew to read the files. Use the PIVOT transformation to create the new column.
- D. Use AWS Glue DataBrew to read the files. Use the NEST TO MAP transformation to create the new column.

Answer: D

Explanation:
The NEST TO MAP transformation allows you to combine multiple columns into a single column that contains a JSON object with key-value pairs. This is the easiest way to achieve the desired format for the physical address data, as you can simply select the columns to nest and specify the keys for each column. The NEST TO ARRAY transformation creates a single column that contains an array of values, which is not thesame as the JSON object format. The PIVOT transformation reshapes the data by creating new columns from unique values in a selected column, which is not applicable for this use case. Writing a Lambda function in Python requires more coding effort than using AWS Glue DataBrew, which provides a visual and interactive interface for data transformations. References:
7 most common data preparation transformations in AWS Glue DataBrew (Section: Nesting and unnesting columns) NEST TO MAP - AWS Glue DataBrew (Section: Syntax)

NEW QUESTION # 130
.....

PDF design has versatile and printable material for Amazon Data-Engineer-Associate certification, so you all can breeze through the Amazon Data-Engineer-Associate exam without any problem. You can get to the PDF concentrate on material from workstations, tablets, and cell phones for the readiness of AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) exam.

Top Data-Engineer-Associate Questions: https://www.prep4away.com/Amazon-certification/braindumps.Data-Engineer-Associate.ete.file.html

Amazon Data-Engineer-Associate Actual Test In a word, the three different versions will meet your all needs, For further consolidation of your learning, DumpsPedia offers an interactive Amazon Top Data-Engineer-Associate Questions Top Data-Engineer-Associate Questions Top Data-Engineer-Associate Questions - AWS Certified Data Engineer - Associate (DEA-C01) exam testing engine, With a total new perspective, Data-Engineer-Associate test dumps: AWS Certified Data Engineer - Associate (DEA-C01) have been designed to serve most of the office workers who aim at getting an exam certification, Through effort and practice, you can get high scores in your Data-Engineer-Associate valid prep exam.

Tap Music to buy the soundtrack of the episode and view a complete list of Data-Engineer-Associate Pass Leader Dumps all music featured in the episode, Disconnect the AC power cord from the computer, In a word, the three different versions will meet your all needs;

2026 Data-Engineer-Associate Actual Test | High Pass-Rate Data-Engineer-Associate 100% Free Top Questions

For further consolidation of your learning, DumpsPedia Data-Engineer-Associate offers an interactive Amazon AWS Certified Data Engineer AWS Certified Data Engineer - Associate (DEA-C01) exam testing engine, With a total new perspective, Data-Engineer-Associate test dumps: AWS Certified Data Engineer - Associate (DEA-C01) have been designed to serve most of the office workers who aim at getting an exam certification.

Through effort and practice, you can get high scores in your Data-Engineer-Associate valid prep exam, Some customer complained to and worried that the former Data-Engineer-Associate training prep is not suitable to the new test, which is wrong because we keep the new content into the Data-Engineer-Associate practice materials by experts.

- Data-Engineer-Associate exam torrent - Data-Engineer-Associate reliable study vce - Data-Engineer-Associate test dumps □ Immediately open 「 www.troytecdumps.com 」 and search for （ Data-Engineer-Associate ） to obtain a free download □Data-Engineer-Associate Latest Study Notes
- Amazon - Data-Engineer-Associate - AWS Certified Data Engineer - Associate (DEA-C01) Accurate Actual Test □ Open ☀ www.pdfvce.com □☀□ and search for " Data-Engineer-Associate " to download exam materials for free □New Data-Engineer-Associate Braindumps Free
- Perfect Data-Engineer-Associate Actual Test | Amazing Pass Rate For Data-Engineer-Associate Exam | High Pass-Rate Data-Engineer-Associate: AWS Certified Data Engineer - Associate (DEA-C01) □ Open ☀ www.prepawayexam.com □☀□ and search for ▹ Data-Engineer-Associate ◃ to download exam materials for free □Download Data-Engineer-Associate Pdf
- Data-Engineer-Associate Valid Test Forum □ Data-Engineer-Associate Valid Test Syllabus □ Cheap Data-Engineer-Associate Dumps □ Easily obtain free download of [Data-Engineer-Associate] by searching on ▸ www.pdfvce.com ◂ □ □Data-Engineer-Associate Best Vce
- New Data-Engineer-Associate Braindumps Sheet □ Data-Engineer-Associate Valid Test Syllabus □ New Data-Engineer-Associate Braindumps Sheet □ Open [www.prepawayexam.com] and search for ☀ Data-Engineer-Associate □☀□ to download exam materials for free □Data-Engineer-Associate Official Study Guide
- Data-Engineer-Associate Valid Test Syllabus □ Data-Engineer-Associate Official Study Guide □ Download Data-Engineer-Associate Pdf □ Easily obtain free download of □ Data-Engineer-Associate □ by searching on ▸ www.pdfvce.com ◂ □Valid Data-Engineer-Associate Exam Bootcamp
- Download Data-Engineer-Associate Pdf □ Data-Engineer-Associate Reliable Braindumps Ebook □ Data-Engineer-Associate Pass Leader Dumps □ Go to website ▹ www.vceengine.com ◃ open and search for ➤ Data-Engineer-Associate □ to download for free □Download Data-Engineer-Associate Pdf
- Quiz Data-Engineer-Associate - Updated AWS Certified Data Engineer - Associate (DEA-C01) Actual Test □ Search for [Data-Engineer-Associate] and download it for free on ➽ www.pdfvce.com □ website □Cheap Data-Engineer-Associate Dumps
- Data-Engineer-Associate Actual Test, Amazon Top Data-Engineer-Associate Questions: AWS Certified Data Engineer - Associate (DEA-C01) Pass Certainly □ Copy URL ▹ www.testkingpass.com ◃ open and search for （ Data-Engineer-Associate ） to download for free □New Data-Engineer-Associate Braindumps Sheet
- New Data-Engineer-Associate Braindumps Free □ New Data-Engineer-Associate Exam Cram □ Data-Engineer-Associate Valid Test Format □ Immediately open { www.pdfvce.com } and search for ➽ Data-Engineer-Associate □ to obtain a free download □Brain Data-Engineer-Associate Exam
- Data-Engineer-Associate Valid Test Forum □ Data-Engineer-Associate Official Study Guide □ Data-Engineer-Associate Valid Test Syllabus □ Open ➽ www.practicevce.com □ and search for ▸ Data-Engineer-Associate ◂ to download exam materials for free ✈Brain Data-Engineer-Associate Exam
- www.stes.tyc.edu.tw, dl.instructure.com, 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, www.stes.tyc.edu.tw, leftbookmarks.com, dl.instructure.com, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, ncon.edu.sa, Disposable vapes

BONUS!!! Download part of Prep4away Data-Engineer-Associate dumps for free: https://drive.google.com/open?id=1w8XKXmU1VhFRAs6w9RdeqCxRZET5qDig