

Amazon Data-Engineer-Associate VCE & Data-Engineer-Associate exam simulator



2026 Latest BraindumpsPass Data-Engineer-Associate PDF Dumps and Data-Engineer-Associate Exam Engine Free Share:
<https://drive.google.com/open?id=1nZ94TGX2Uh9qvT7f0yJalF07mCONsThn>

The clients at home and abroad can both purchase our Data-Engineer-Associate study tool online. Our brand enjoys world-wide fame and influences so many clients at home and abroad choose to buy our Data-Engineer-Associate test guide. Our company provides convenient service to the clients all around the world so that the clients all around the world can use our Data-Engineer-Associate Study Materials efficiently. Our company boosts an entire sale system which provides the links to the clients all around the world so that the clients can receive our Data-Engineer-Associate exam questions timely.

You may be given the Amazon Data-Engineer-Associate practice exam results as soon as they have been saved in the software. BraindumpsPass modified Amazon Data-Engineer-Associate exam dumps allow students to learn effectively about the real Amazon Data-Engineer-Associate Certification Exam. Amazon Data-Engineer-Associate practice exam software allows students to review and refine skills in a preceding test setting.

>> **Reliable Data-Engineer-Associate Exam Blueprint** <<

Data-Engineer-Associate Dumps Cost - Data-Engineer-Associate Pass4sure Pass Guide

Nowadays, flexible study methods become more and more popular with the development of the electronic products. The latest technologies have been applied to our Data-Engineer-Associate actual exam as well since we are at the most leading position in this field. You can get a complete new and pleasant study experience with our Data-Engineer-Associate Study Materials. Besides, you have varied choices for there are three versions of our Data-Engineer-Associate practice materials. At the same time, you are bound to pass the Data-Engineer-Associate exam and get your desired certification for the validity and accuracy of our Data-Engineer-Associate study materials.

Amazon AWS Certified Data Engineer - Associate (DEA-C01) Sample Questions (Q20-Q25):

NEW QUESTION # 20

A data engineer needs to query data from multiple sources to generate an annual report. The analytics team uses Amazon Redshift for analysis. The data engineer needs to integrate Amazon Redshift data with 10 years of historical data from Amazon RDS for PostgreSQL and RDS for MySQL. All the databases are in the same VPC. The data engineer needs a solution that provides seamless data integration with Amazon Redshift.

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

- **A. Use federated queries in Amazon Redshift to fetch data from RDS for PostgreSQL and RDS for MySQL. Apply the necessary transformations within Amazon Redshift.**
- B. Use AWS Database Migration Service (AWS DMS) to ingest data from RDS for PostgreSQL and RDS for MySQL. Implement the necessary transformations within Amazon Redshift.
- C. Use the SELECT INTO OUTFILE S3 statement to export data from Amazon RDS to Amazon S3. Use the COPY command to load the data into Amazon Redshift.
- D. Create a visual extract, transform, and load (ETL) job in AWS Glue to extract the required data and load it to Amazon Redshift.

Answer: A

Explanation:

Option A is the most cost-effective because it enables seamless integration by allowing analysts to access and join external relational data directly from within Amazon Redshift for reporting, without building and operating a separate ingestion pipeline for a one-time (annual) workload. The study material frames Amazon Redshift as the centralized analytics store for complex queries and reporting, making it the natural place to perform the final transformations and analysis.

The alternatives introduce extra operational steps and recurring costs. Option B requires exporting large historical datasets to S3 and then loading them into Redshift, which increases data movement and operational complexity. Option C adds ETL job development, scheduling, retries, and monitoring overhead that is unnecessary if the primary goal is integrated querying for an annual report. Option D (DMS) is best when you need ongoing migration or continuous replication; it is typically more operationally heavy than needed for "query across sources" use cases and still requires additional setup and maintenance.

Because all databases are already in the same VPC and the analytics platform is Redshift, federated querying provides the most direct, lowest-operations path to integrate and analyze the data where it already needs to be consumed.

NEW QUESTION # 21

A transportation company wants to track vehicle movements by capturing geolocation records. The records are 10 bytes in size. The company receives up to 10,000 records every second. Data transmission delays of a few minutes are acceptable because of unreliable network conditions.

The transportation company wants to use Amazon Kinesis Data Streams to ingest the geolocation data. The company needs a reliable mechanism to send data to Kinesis Data Streams. The company needs to maximize the throughput efficiency of the Kinesis shards.

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

- **A. Kinesis Producer Library (KPL)**
- B. Amazon Data Firehose
- C. Kinesis SDK
- D. Kinesis Agent

Answer: A

Explanation:

* Problem Analysis:

* The company ingests geolocation records (10 bytes each) at 10,000 records per second into Kinesis Data Streams.

* Data transmission delays are acceptable, but the solution must maximize throughput efficiency.

* Key Considerations:

* The Kinesis Producer Library (KPL) batches records and uses aggregation to optimize shard throughput.

* Efficiently handles high-throughput scenarios with minimal operational overhead.

* Solution Analysis:

* Option A: Kinesis Agent

* Designed for file-based ingestion; not optimized for geolocation records.

* Option B: KPL

* Aggregates records into larger payloads, significantly improving shard throughput.

* Suitable for applications generating small, high-frequency records.

- * Option C: Kinesis Firehose
- * Firehose is for delivery to destinations like S3 or Redshift and is not optimized for direct ingestion to Kinesis Data Streams.
- * Option D: Kinesis SDK
- * The SDK lacks advanced features like aggregation, resulting in lower throughput efficiency.
- * Final Recommendation:
- * Use Kinesis Producer Library (KPL) for its built-in aggregation and batching capabilities.

:
 Kinesis Producer Library (KPL) Overview
 Best Practices for Amazon Kinesis

NEW QUESTION # 22

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. Create a daily schedule to run the crawler. Specify a database name for the output.**
- 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. Allocate data processing units (DPUs) to run the crawler every day. Configure the output destination to a new path in the existing S3 bucket.
- 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: A

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 # 23

A company is building an inventory management system and an inventory reordering system to automatically reorder products. Both systems use Amazon Kinesis Data Streams. The inventory management system uses the Amazon Kinesis Producer Library (KPL) to publish data to a stream. The inventory reordering system uses the Amazon Kinesis Client Library (KCL) to consume data from the stream. The company configures the stream to scale up and down as needed.

Before the company deploys the systems to production, the company discovers that the inventory reordering system received duplicated data.

Which factors could have caused the reordering system to receive duplicated data? (Select TWO.)

- A. The `max_records` configuration property was set to a number that was too high.
- B. The stream's value for the `IteratorAgeMilliseconds` metric was too high.
- C. The producer experienced network-related timeouts.
- D. There was a change in the number of shards, record processors, or both.
- E. The `AggregationEnabled` configuration property was set to true.

Answer: C,D

Explanation:

Problem Analysis:

The company uses Kinesis Data Streams for both inventory management and reordering.

The Kinesis Producer Library (KPL) publishes data, and the Kinesis Client Library (KCL) consumes data.

Duplicate records were observed in the inventory reordering system.

Key Considerations:

Kinesis streams are designed for durability but may produce duplicates under certain conditions.

Factors such as network timeouts, shard splits, or changes in record processors can cause duplication.

Solution Analysis:

Option A: Network-Related Timeouts

If the producer (KPL) experiences network timeouts, it retries data submission, potentially causing duplicates.

Option B: High `IteratorAgeMilliseconds`

High iterator age suggests delays in processing but does not directly cause duplication.

Option C: Changes in Shards or Processors

Changes in the number of shards or record processors can lead to re-processing of records, causing duplication.

Option D: `AggregationEnabled` Set to True

`AggregationEnabled` controls the aggregation of multiple records into one, but it does not cause duplication.

Option E: High `max_records` Value

A high `max_records` value increases batch size but does not lead to duplication.

Final Recommendation:

Network-related timeouts and changes in shards or processors are the most likely causes of duplicate data in this scenario.

Amazon Kinesis Data Streams Best Practices

Kinesis Producer Library (KPL) Overview

Kinesis Client Library (KCL) Overview

NEW QUESTION # 24

A company needs to use an AWS Glue PySpark job to read specific data from an Amazon DynamoDB table.

The company knows the partition key values for the required records. The existing processing logic of the AWS Glue PySpark job requires the data to be in `DynamicFrame` format. The company needs a solution to ensure that the job reads only the specified data.

Which solution will meet this requirement with the MINIMUM number of read capacity units (RCUs)?

- A. Perform a query on the DynamoDB table in the AWS Glue job by using only the sort key in the key condition expression. Load the data into a `DynamicFrame`.
- B. Use the AWS Glue DynamoDB ETL connector to read the DynamoDB table. Use the filter option to read the required partition key.
- C. Perform a scan on the DynamoDB table in the AWS Glue job. Put the data into a `DynamicFrame`. Filter the `DynamicFrame` on the partition key.
- D. Perform a query on the DynamoDB table in the AWS Glue job. Use the partition key in the key condition expression. Put the data into a `DynamicFrame`.

Answer: D

Explanation:

Comprehensive and Detailed Explanation (150-250 words)

Amazon DynamoDB is optimized for Query operations, which retrieve items efficiently using the partition key and optionally the sort key. Queries consume significantly fewer RCUs than Scan operations because only matching partitions are read.

By performing a Query using the partition key in the key condition expression, the AWS Glue job reads only the required items directly from DynamoDB. This minimizes the number of read capacity units consumed and avoids unnecessary reads.

Using a Scan operation, even when followed by filtering in a DynamicFrame, consumes RCUs for the entire table or index and is highly inefficient. Using only the sort key in a query is invalid without the partition key and would not meet DynamoDB query requirements. The Glue DynamoDB connector filter option still performs a scan under the hood and does not minimize RCUs.

The queried data can be directly loaded into a DynamicFrame, satisfying the existing processing logic requirement without additional transformations.

Therefore, Option D provides the most efficient and cost-effective solution.

NEW QUESTION # 25

.....

The web-based Amazon Data-Engineer-Associate Practice Exam is compatible with all operating systems, including Mac, Linux, iOS, Android, and Windows. It is a browser-based AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) practice exam that works on all major browsers, including Chrome, Firefox, Safari, Internet Explorer, and Opera. This means that you won't have to worry about installing any complicated software or plug-ins.

Data-Engineer-Associate Dumps Cost: <https://www.braindumpsPASS.com/Amazon/Data-Engineer-Associate-practice-exam-dumps.html>

If you prefer to prepare for your exam on paper, then our Data-Engineer-Associate exam materials will be your best choice, Comparing to attending expensive training institution, Data-Engineer-Associate Dumps Cost - AWS Certified Data Engineer - Associate (DEA-C01) updated training questions is more suitable for people who are eager to passing actual test but no time and energy, The first one is a Data-Engineer-Associate Dumps PDF form, and it is printable and portable.

Because each image tells a story, as communicators we need to make Data-Engineer-Associate sure that the story the audience perceives is the story we are telling, See how core components work behind the scenes.

If you prefer to prepare for your exam on paper, then our Data-Engineer-Associate Exam Materials will be your best choice, Comparing to attending expensive training institution, AWS Certified Data Engineer - Associate (DEA-C01) updated training questions Reliable Data-Engineer-Associate Exam Blueprint is more suitable for people who are eager to passing actual test but no time and energy.

Free PDF Quiz Amazon - Data-Engineer-Associate - AWS Certified Data Engineer - Associate (DEA-C01) Fantastic Reliable Exam Blueprint

The first one is a Data-Engineer-Associate Dumps PDF form, and it is printable and portable, The AWS Certified Data Engineer - Associate (DEA-C01) is the first step of your professional IT journey, We will try genuinely and sincerely to meet all the requirements of our customers.

- Data-Engineer-Associate New Test Camp Study Data-Engineer-Associate Reference Free Data-Engineer-Associate Test Questions Simply search for **➡ Data-Engineer-Associate** for free download on **【 www.vceengine.com 】** Data-Engineer-Associate Exams
- Professional Reliable Data-Engineer-Associate Exam Blueprint - Pass Data-Engineer-Associate Exam Search for **▷ Data-Engineer-Associate** on **➡ www.pdfvce.com** immediately to obtain a free download Data-Engineer-Associate Exam Format
- VCE Data-Engineer-Associate Exam Simulator Data-Engineer-Associate New Dumps Free Study Data-Engineer-Associate Reference Search for Data-Engineer-Associate and download it for free immediately on **➤ www.torrentvce.com** Practice Data-Engineer-Associate Exam Fee
- Pass Guaranteed 2026 Data-Engineer-Associate: Pass-Sure Reliable AWS Certified Data Engineer - Associate (DEA-C01) Exam Blueprint Easily obtain free download of **▷ Data-Engineer-Associate** by searching on { **www.pdfvce.com** } Data-Engineer-Associate New Test Camp
- Data-Engineer-Associate Exam Reliable Exam Blueprint - Trustable Data-Engineer-Associate Dumps Cost Pass Success Immediately open **【 www.practicevce.com 】** and search for **✓ Data-Engineer-Associate** to obtain a free download Test Data-Engineer-Associate Online
- Pass Guaranteed Amazon - Data-Engineer-Associate - Professional Reliable Exam Blueprint Search for **➡ Data-Engineer-Associate** and download it for free on **【 www.pdfvce.com 】** website Data-Engineer-Associate Exam

Format

- Pass Guaranteed 2026 Data-Engineer-Associate: Pass-Sure Reliable AWS Certified Data Engineer - Associate (DEA-C01) Exam Blueprint Open [www.testkingpass.com] and search for 「 Data-Engineer-Associate 」 to download exam materials for free Reliable Data-Engineer-Associate Dumps Book
- Valid Test Data-Engineer-Associate Testking Exam Data-Engineer-Associate Training Data-Engineer-Associate New Test Camp Simply search for (Data-Engineer-Associate) for free download on www.pdfvce.com Study Data-Engineer-Associate Reference
- Valid Test Data-Engineer-Associate Testking Data-Engineer-Associate New Dumps Free Study Data-Engineer-Associate Reference Download 《 Data-Engineer-Associate 》 for free by simply entering 【 www.testkingpass.com 】 website Braindump Data-Engineer-Associate Free
- Pass Guaranteed 2026 Data-Engineer-Associate: Pass-Sure Reliable AWS Certified Data Engineer - Associate (DEA-C01) Exam Blueprint Search for ▷ Data-Engineer-Associate ◁ and easily obtain a free download on ➡ www.pdfvce.com Data-Engineer-Associate New Dumps Free
- Latest Data-Engineer-Associate Test Camp Study Data-Engineer-Associate Reference VCE Data-Engineer-Associate Exam Simulator Open www.vceengine.com and search for 「 Data-Engineer-Associate 」 to download exam materials for free Data-Engineer-Associate New Test Camp
- www.stes.tyc.edu.tw, ascntleadershipinstitute.org, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, thotsmithconsulting.com, www.stes.tyc.edu.tw, pixabay.com, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, Disposable vapes

P.S. Free & New Data-Engineer-Associate dumps are available on Google Drive shared by BraindumpsPass:
<https://drive.google.com/open?id=1nZ94TGX2Uh9qvT7f0yJalF07mC0NsThm>