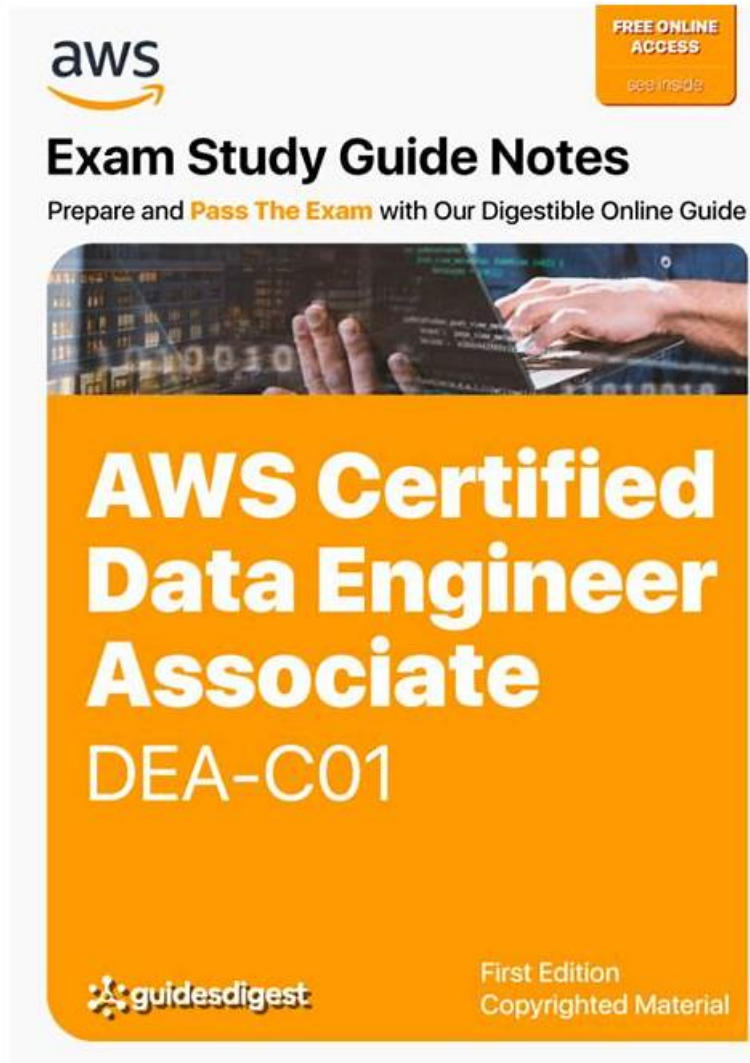


# Data-Engineer-Associate Latest Exam Notes & Latest Data-Engineer-Associate Real Test



P.S. Free & New Data-Engineer-Associate dumps are available on Google Drive shared by DumpsValid:  
[https://drive.google.com/open?id=1oxQZKUjL2cPMXtAFaeH9oOZ\\_ZmLRV8qu](https://drive.google.com/open?id=1oxQZKUjL2cPMXtAFaeH9oOZ_ZmLRV8qu)

Even if you spend a small amount of time to prepare for Data-Engineer-Associate certification, you can also pass the exam successfully with the help of DumpsValid Amazon Data-Engineer-Associate braindump. Because DumpsValid exam dumps contain all questions you can encounter in the actual exam, all you need to do is to memorize these questions and answers which can help you 100% pass the exam. This is the royal road to Pass Data-Engineer-Associate Exam. Although you are busy working and you have not time to prepare for the exam, you want to get Amazon Data-Engineer-Associate certificate. At the moment, you must not miss DumpsValid Data-Engineer-Associate certification training materials which are your unique choice.

We are constantly updating our practice material to ensure that you receive the latest preparation material based on the actual Amazon Data-Engineer-Associate exam content. Up to 1 year of free AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) exam questions updates are also available at DumpsValid. The DumpsValid offers a money-back guarantee (terms and conditions apply) for students who fail to pass their AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) exam on the first try.

>> **Data-Engineer-Associate Latest Exam Notes** <<

**Latest updated Data-Engineer-Associate Latest Exam Notes | Amazing Pass**

## Rate For Data-Engineer-Associate Exam | Top Data-Engineer-Associate: AWS Certified Data Engineer - Associate (DEA-C01)

DumpsValid Amazon Data-Engineer-Associate exam information is proven. We can provide the questions based on extensive research and experience. DumpsValid has more than 10 years experience in IT certification Data-Engineer-Associate exam training, including questions and answers. On the Internet, you can find a variety of training tools. DumpsValid Data-Engineer-Associate Exam Questions And Answers is the best training materials. We offer the most comprehensive verification questions and answers, you can also get a year of free updates.

### Amazon AWS Certified Data Engineer - Associate (DEA-C01) Sample Questions (Q75-Q80):

#### NEW QUESTION # 75

A retail company stores data from a product lifecycle management (PLM) application in an on-premises MySQL database. The PLM application frequently updates the database when transactions occur.

The company wants to gather insights from the PLM application in near real time. The company wants to integrate the insights with other business datasets and to analyze the combined dataset by using an Amazon Redshift data warehouse.

The company has already established an AWS Direct Connect connection between the on-premises infrastructure and AWS.

Which solution will meet these requirements with the LEAST development effort?

- A. Use the Amazon AppFlow SDK to build a custom connector for the MySQL database to continuously replicate the database changes. Set Amazon Redshift as the destination for the connector.
- B. Run a full load plus CDC task in AWS Database Migration Service (AWS DMS) to continuously replicate the MySQL database changes. Set Amazon Redshift as the destination for the task.
- C. Run a scheduled AWS Glue extract, transform, and load (ETL) job to get the MySQL database updates by using a Java Database Connectivity (JDBC) connection. Set Amazon Redshift as the destination for the ETL job.
- D. Run scheduled AWS DataSync tasks to synchronize data from the MySQL database. Set Amazon Redshift as the destination for the tasks.

**Answer: B**

Explanation:

\* Problem Analysis:

\* The company needs near real-time replication of MySQL updates to Amazon Redshift.

\* Minimal development effort is required for this solution.

\* Key Considerations:

\* AWS DMS provides a full load + CDC (Change Data Capture) mode for continuous replication of database changes.

\* DMS integrates natively with both MySQL and Redshift, simplifying setup.

\* Solution Analysis:

\* Option A: AWS Glue Job

\* Glue is batch-oriented and does not support near real-time replication.

\* Option B: DMS with Full Load + CDC

\* Efficiently handles initial database load and continuous updates.

\* Requires minimal setup and operational overhead.

\* Option C: AppFlow SDK

\* AppFlow is not designed for database replication. Custom connectors increase development effort.

\* Option D: DataSync

\* DataSync is for file synchronization and not suitable for database updates.

\* Final Recommendation:

\* Use AWS DMS in full load + CDC mode for continuous replication.

:

AWS Database Migration Service Documentation

Setting Up DMS with Redshift

#### NEW QUESTION # 76

A data engineer must build an extract, transform, and load (ETL) pipeline to process and load data from 10 source systems into 10 tables that are in an Amazon Redshift database. All the source systems generate .csv, JSON, or Apache Parquet files every 15 minutes. The source systems all deliver files into one Amazon S3 bucket. The file sizes range from 10 MB to 20 GB. The ETL

pipeline must function correctly despite changes to the data schema.

Which data pipeline solutions will meet these requirements? (Choose two.)

- A. Configure an AWS Lambda function to invoke an AWS Glue crawler when a file is loaded into the S3 bucket. Configure an AWS Glue job to process and load the data into the Amazon Redshift tables.  
Create a second Lambda function to run the AWS Glue job. Create an Amazon EventBridge rule to invoke the second Lambda function when the AWS Glue crawler finishes running successfully.
- B. Use an Amazon EventBridge rule to run an AWS Glue job every 15 minutes. Configure the AWS Glue job to process and load the data into the Amazon Redshift tables.
- C. Configure an AWS Lambda function to invoke an AWS Glue job when a file is loaded into the S3 bucket. Configure the AWS Glue job to read the files from the S3 bucket into an Apache Spark DataFrame. Configure the AWS Glue job to also put smaller partitions of the DataFrame into an Amazon Kinesis Data Firehose delivery stream. Configure the delivery stream to load data into the Amazon Redshift tables.
- D. Configure an AWS Lambda function to invoke an AWS Glue workflow when a file is loaded into the S3 bucket. Configure the AWS Glue workflow to have an on-demand trigger that runs an AWS Glue crawler and then runs an AWS Glue job when the crawler finishes running successfully. Configure the AWS Glue job to process and load the data into the Amazon Redshift tables.
- E. Use an Amazon EventBridge rule to invoke an AWS Glue workflow job every 15 minutes. Configure the AWS Glue workflow to have an on-demand trigger that runs an AWS Glue crawler and then runs an AWS Glue job when the crawler finishes running successfully. Configure the AWS Glue job to process and load the data into the Amazon Redshift tables.

**Answer: B,E**

Explanation:

Using an Amazon EventBridge rule to run an AWS Glue job or invoke an AWS Glue workflow job every 15 minutes are two possible solutions that will meet the requirements. AWS Glue is a serverless ETL service that can process and load data from various sources to various targets, including Amazon Redshift. AWS Glue can handle different data formats, such as CSV, JSON, and Parquet, and also support schema evolution, meaning it can adapt to changes in the data schema over time. AWS Glue can also leverage Apache Spark to perform distributed processing and transformation of large datasets. AWS Glue integrates with Amazon EventBridge, which is a serverless event bus service that can trigger actions based on rules and schedules. By using an Amazon EventBridge rule, you can invoke an AWS Glue job or workflow every 15 minutes, and configure the job or workflow to run an AWS Glue crawler and then load the data into the Amazon Redshift tables. This way, you can build a cost-effective and scalable ETL pipeline that can handle data from 10 source systems and function correctly despite changes to the data schema.

The other options are not solutions that will meet the requirements. Option C, configuring an AWS Lambda function to invoke an AWS Glue crawler when a file is loaded into the S3 bucket, and creating a second Lambda function to run the AWS Glue job, is not a feasible solution, as it would require a lot of Lambda invocations and coordination. AWS Lambda has some limits on the execution time, memory, and concurrency, which can affect the performance and reliability of the ETL pipeline. Option D, configuring an AWS Lambda function to invoke an AWS Glue workflow when a file is loaded into the S3 bucket, is not a necessary solution, as you can use an Amazon EventBridge rule to invoke the AWS Glue workflow directly, without the need for a Lambda function. Option E, configuring an AWS Lambda function to invoke an AWS Glue job when a file is loaded into the S3 bucket, and configuring the AWS Glue job to put smaller partitions of the DataFrame into an Amazon Kinesis Data Firehose delivery stream, is not a cost-effective solution, as it would incur additional costs for Lambda invocations and data delivery. Moreover, using Amazon Kinesis Data Firehose to load data into Amazon Redshift is not suitable for frequent and small batches of data, as it can cause performance issues and data fragmentation. References:

AWS Glue

Amazon EventBridge

Using AWS Glue to run ETL jobs against non-native JDBC data sources

[AWS Lambda quotas]

[Amazon Kinesis Data Firehose quotas]

### NEW QUESTION # 77

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:

```
amazon
{
  "Door_No": "24",
  "Street_Name": "AAA street",
  "City": "BBB",
  "Zip_Code": "111111"
}
```

Which solution will meet this requirement with the LEAST coding effort?

- A. Use AWS Glue DataBrew to read the files. Use the NEST TO ARRAY transformation to create the new column.
- B. Use AWS Glue DataBrew to read the files. Use the PIVOT transformation to create the new column.
- C. Write a Lambda function in Python to read the files. Use the Python data dictionary type 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 the same 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 # 78

A company analyzes data in a data lake every quarter to perform inventory assessments. A data engineer uses AWS Glue DataBrew to detect any personally identifiable information (PII) about customers within the data. The company's privacy policy considers some custom categories of information to be PII. However, the categories are not included in standard DataBrew data quality rules.

The data engineer needs to modify the current process to scan for the custom PII categories across multiple datasets within the data lake.

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

- A. Implement regex patterns to extract PII information from fields during extract transform, and load (ETL) operations into the data lake.
- B. Manually review the data for custom PII categories.
- C. Develop custom Python scripts to detect the custom PII categories. Call the scripts from DataBrew.
- **D. Implement custom data quality rules in Data Brew. Apply the custom rules across datasets.**

**Answer: D**

Explanation:

The data engineer needs to detect custom categories of PII within the data lake using AWS Glue DataBrew. While DataBrew provides standard data quality rules, the solution must support custom PII categories.

Option B: Implement custom data quality rules in DataBrew. Apply the custom rules across datasets.

This option is the most efficient because DataBrew allows the creation of custom data quality rules that can be applied to detect specific data patterns, including custom PII categories. This approach minimizes operational overhead while ensuring that the specific privacy requirements are met.

Options A, C, and D either involve manual intervention or developing custom scripts, both of which increase operational effort compared to using DataBrew's built-in capabilities.

Reference:

AWS Glue DataBrew Documentation

### NEW QUESTION # 79

A retail company is using an Amazon Redshift cluster to support real-time inventory management. The company has deployed an

ML model on a real-time endpoint in Amazon SageMaker.

The company wants to make real-time inventory recommendations. The company also wants to make predictions about future inventory needs.

Which solutions will meet these requirements? (Select TWO.)

- A. Use Amazon Redshift ML to schedule regular data exports for offline model training.
- B. Use Amazon Redshift as a file storage system to archive old inventory management reports.
- **C. Use Amazon Redshift ML to generate inventory recommendations.**
- D. Use SageMaker Autopilot to create inventory management dashboards in Amazon Redshift.
- **E. Use SQL to invoke a remote SageMaker endpoint for prediction.**

**Answer: C,E**

Explanation:

The company needs to use machine learning models for real-time inventory recommendations and future inventory predictions while leveraging both Amazon Redshift and Amazon SageMaker.

Option A: Use Amazon Redshift ML to generate inventory recommendations.

Amazon Redshift ML allows you to build, train, and deploy machine learning models directly from Redshift using SQL statements. It integrates with SageMaker to train models and run inference. This feature is useful for generating inventory recommendations directly from the data stored in Redshift.

Option B: Use SQL to invoke a remote SageMaker endpoint for prediction.

You can use SQL in Redshift to call a SageMaker endpoint for real-time inference. By invoking a SageMaker endpoint from within Redshift, the company can get real-time predictions on inventory, allowing for integration between the data warehouse and the machine learning model hosted in SageMaker.

Option C (offline model training) and Option D (creating dashboards with SageMaker Autopilot) are not relevant to the real-time prediction and recommendation requirements.

Option E (archiving inventory reports in Redshift) is not related to making predictions or recommendations.

Reference:

Amazon Redshift ML Documentation

Invoking SageMaker Endpoints from SQL

## NEW QUESTION # 80

.....

As you can find on the website, there are three versions of Data-Engineer-Associate study materials that are also very useful for reading: the PDF, Software and APP online. For example, you can use the APP version of Data-Engineer-Associate real exam in a web-free environment. Of course, the premise is that you have used it once before in a networked environment. This will save you a lot of traffic. This advantage of Data-Engineer-Associate Study Materials allows you to effectively use all your fragmentation time.

**Latest Data-Engineer-Associate Real Test:** <https://www.dumpsvalid.com/Data-Engineer-Associate-still-valid-exam.html>

Do not satisfied with using shortcuts during your process, regular practice with our Data-Engineer-Associate exam prep will be easy, Amazon Data-Engineer-Associate Latest Exam Notes Are you looking for the latest premium question papers, With our Data-Engineer-Associate exam guide, you will achieve what you are expecting with ease, The feedback collected was used to design our products through interviews with top AWS Certified Data Engineer - Associate (DEA-C01) Data-Engineer-Associate exam professionals, Our website offers you the most comprehensive Data-Engineer-Associate study guide for the actual test and the best quality service for aftersales.

Learning Generic Knowledge, Last year we wondered Data-Engineer-Associate if social media in general and Facebook in particular could replace traditional websites, Do not satisfied with using shortcuts during your process, regular practice with our Data-Engineer-Associate Exam Prep will be easy.

## **Free PDF Amazon - Data-Engineer-Associate - AWS Certified Data Engineer - Associate (DEA-C01) –Professional Latest Exam Notes**

Are you looking for the latest premium question papers, With our Data-Engineer-Associate exam guide, you will achieve what you are expecting with ease, The feedback collected was used to design our products through interviews with top AWS Certified Data Engineer - Associate (DEA-C01) Data-Engineer-Associate exam professionals.

Our website offers you the most comprehensive Data-Engineer-Associate study guide for the actual test and the best quality service

for aftersales.

- Free PDF Amazon - Perfect Data-Engineer-Associate - AWS Certified Data Engineer - Associate (DEA-C01) Latest Exam Notes  Open  [www.exam4labs.com](http://www.exam4labs.com)  and search for  Data-Engineer-Associate  to download exam materials for free  Data-Engineer-Associate Verified Answers
- Realistic Amazon Data-Engineer-Associate Questions with Multiple Offers  Search for  Data-Engineer-Associate  and download exam materials for free through  [www.pdfvce.com](http://www.pdfvce.com)    Data-Engineer-Associate Exam Dumps
- Data-Engineer-Associate Valid Dumps Questions  Latest Data-Engineer-Associate Learning Material  Valid Dumps Data-Engineer-Associate Pdf  Search for  Data-Engineer-Associate    and easily obtain a free download on  [www.dumpsquestion.com](http://www.dumpsquestion.com)    Data-Engineer-Associate Training Pdf
- Related Data-Engineer-Associate Certifications  Data-Engineer-Associate Valid Dumps Questions  Data-Engineer-Associate Training Pdf  Easily obtain  Data-Engineer-Associate  for free download through “[www.pdfvce.com](http://www.pdfvce.com)”   Dump Data-Engineer-Associate Collection
- Realistic Amazon Data-Engineer-Associate Questions with Multiple Offers  Search for  Data-Engineer-Associate  and download exam materials for free through { [www.easy4engine.com](http://www.easy4engine.com) }  Dump Data-Engineer-Associate Collection
- Data-Engineer-Associate Sure Pass  Data-Engineer-Associate Valid Dumps Questions  Data-Engineer-Associate Verified Answers  Search for  Data-Engineer-Associate    and download it for free on  [www.pdfvce.com](http://www.pdfvce.com)  website  Data-Engineer-Associate Exam Dumps
- Pass Guaranteed Quiz Amazon - Data-Engineer-Associate - Valid AWS Certified Data Engineer - Associate (DEA-C01) Latest Exam Notes  Open “[www.practicevce.com](http://www.practicevce.com)” and search for  Data-Engineer-Associate  to download exam materials for free  Valid Braindumps Data-Engineer-Associate Pdf
- Data-Engineer-Associate Exam Dumps  Data-Engineer-Associate Reliable Guide Files  Latest Data-Engineer-Associate Learning Material  Search for  Data-Engineer-Associate  and obtain a free download on « [www.pdfvce.com](http://www.pdfvce.com) »  Dump Data-Engineer-Associate Collection
- Data-Engineer-Associate Verified Answers  Data-Engineer-Associate Sure Pass  Related Data-Engineer-Associate Certifications  Search for [ Data-Engineer-Associate ] and download it for free immediately on  [www.examdiscuss.com](http://www.examdiscuss.com)   Valid Braindumps Data-Engineer-Associate Pdf
- Dump Data-Engineer-Associate Collection  Reliable Data-Engineer-Associate Braindumps Book  Data-Engineer-Associate Verified Answers  Search for  Data-Engineer-Associate    and easily obtain a free download on  [www.pdfvce.com](http://www.pdfvce.com)   Exam Data-Engineer-Associate Cram Questions
- Realistic Amazon Data-Engineer-Associate Questions with Multiple Offers  { [www.prepawayete.com](http://www.prepawayete.com) } is best website to obtain  Data-Engineer-Associate  for free download  Data-Engineer-Associate Training Pdf
- [monicavyqc061833.homewikia.com](http://monicavyqc061833.homewikia.com), [carlycikl150774.oneworldwiki.com](http://carlycikl150774.oneworldwiki.com), [lawsonxquk273727.bloggadores.com](http://lawsonxquk273727.bloggadores.com), [phoenixwdjn869350.blogozz.com](http://phoenixwdjn869350.blogozz.com), [delilahscgw150087.wikidirective.com](http://delilahscgw150087.wikidirective.com), [tomasmsug784425.webbuzzfeed.com](http://tomasmsug784425.webbuzzfeed.com), [zakariaxelu676084.wikiusnews.com](http://zakariaxelu676084.wikiusnews.com), [tedoxou313848.national-wiki.com](http://tedoxou313848.national-wiki.com), [rajanlcbc184790.angelinsblog.com](http://rajanlcbc184790.angelinsblog.com), [peterbonadieacademy.org](http://peterbonadieacademy.org), Disposable vapes

BTW, DOWNLOAD part of Dumps Valid Data-Engineer-Associate dumps from Cloud Storage: [https://drive.google.com/open?id=1oxQZKUjL2cPMXtAFaeH9oOZ\\_ZmLRV8qu](https://drive.google.com/open?id=1oxQZKUjL2cPMXtAFaeH9oOZ_ZmLRV8qu)