Data-Engineer-Associate認證資料 & Data-Engineer-Associate最新考證

從Google Drive中免費下載最新的KaoGuTi Data-Engineer-Associate PDF版考試題庫：https://drive.google.com/open?id=1GfyW0PY22U9DPce59nTed4Bml3TCrZLe

Amazon的認證考試現在是很有人氣的考試。你已經取得了這個重要的認證資格嗎？比如，你已經參加了現在參加人數最多的Data-Engineer-Associate考試了嗎？如果還沒有的話，你應該儘快採取行動了。你必須要拿到如此重要的認證資格。在這裏我想說的就是怎樣才能更有效率地準備Data-Engineer-Associate考試，並且一次就通過考試拿到考試的認證資格。

Data-Engineer-Associate 認證是互聯網界具有極大聲望的網路技能認證，在全球，通過IBM認證考試的工程師，平均年薪在10萬元以上。通過 HP 認證考試的工程師，平均年薪在30萬元以上。獲得 Amazon 的 Data-Engineer-Associate認證的工程師，平均年薪也不低於20萬人民幣。據說，這還只是基本工資，不包括獎金，紅利和其他非工資性補貼。難怪美國副總統戈爾曾把 Amazon Data-Engineer-Associate 認證恰當而幽默地稱為"獲得高技術，高薪水的頭等艙船票"。

>> Data-Engineer-Associate認證資料 <<

高效的Data-Engineer-Associate認證資料和認證考試的領導者材料和權威的Data-Engineer-Associate最新考證

怎樣才能確保我們的生活可更快的得到改善？你需要通過Data-Engineer-Associate認證考試，獲得證書。而KaoGuTi是IT專業人士的最佳選擇，獲得Data-Engineer-Associate認證是IT職業發展的有力保證，我們高品質的題庫能幫助你做到這一點。Data-Engineer-Associate考試題庫也會不定期的更新，為你提供最有效的學習資料。使用我們的Data-Engineer-Associate考試題庫進行考前復習，可以節約你大量的學習時間和費用，這是最適合獲得Data-Engineer-Associate認證的所必須的學習資料。

最新的 AWS Certified Data Engineer Data-Engineer-Associate 免費考試真題 (Q58-Q63):

問題 #58
A media company wants to improve a system that recommends media content to customer based on user behavior and preferences. To improve the recommendation system, the company needs to incorporate insights from third-party datasets into the company's existing analytics platform.
The company wants to minimize the effort and time required to incorporate third-party datasets.
Which solution will meet these requirements with the LEAST operational overhead?

- A. Use Amazon Kinesis Data Streams to access and integrate third-party datasets from Amazon Elastic Container Registry (Amazon ECR).
- B. Use API calls to access and integrate third-party datasets from AWS Data Exchange.
- C. Use Amazon Kinesis Data Streams to access and integrate third-party datasets from AWS CodeCommit repositories.
- D. Use API calls to access and integrate third-party datasets from AWS

解題說明：

AWS Data Exchange is a service that makes it easy to find, subscribe to, and use third-party data in the cloud.

It provides a secure and reliable way to access and integrate data from various sources, such as data providers, public datasets, or AWS services. Using AWS Data Exchange, you can browse and subscribe to data products that suit your needs, and then use API calls or the AWS Management Console to export the data to Amazon S3, where you can use it with your existing analytics platform. This solution minimizes the effort and time required to incorporate third-party datasets, as you do not need to set up and manage data pipelines, storage, or access controls. You also benefit from the data quality and freshness provided by the data providers, who can update their data products as frequently as needed12.

The other options are not optimal for the following reasons:

* B. Use API calls to access and integrate third-party datasets from AWS. This option is vague and does not specify which AWS service or feature is used to access and integrate third-party datasets. AWS offers a variety of services and features that can help with data ingestion, processing, and analysis, but not all of them are suitable for the given scenario. For example, AWS Glue is a serverless data integration service that can help you discover, prepare, and combine data from various sources, but it requires you to create and run data extraction, transformation, and loading (ETL) jobs, which can add operational overhead3.

* C. Use Amazon Kinesis Data Streams to access and integrate third-party datasets from AWS CodeCommit repositories. This option is not feasible, as AWS CodeCommit is a source control service that hosts secure Git-based repositories, not a data source that can be accessed by Amazon Kinesis Data Streams. Amazon Kinesis Data Streams is a service that enables you to capture, process, and analyze data streams in real time, such as clickstream data, application logs, or IoT telemetry. It does not support accessing and integrating data from AWS CodeCommit repositories, which are meant for storing and managing code, not data .

* D. Use Amazon Kinesis Data Streams to access and integrate third-party datasets from Amazon Elastic Container Registry (Amazon ECR). This option is also not feasible, as Amazon ECR is a fully managed container registry service that stores, manages, and deploys container images, not a data source that can be accessed by Amazon Kinesis Data Streams. Amazon Kinesis Data Streams does not support accessing and integrating data from Amazon ECR, which is meant for storing and managing container images, not data .

1: AWS Data Exchange User Guide
2: AWS Data Exchange FAQs
3: AWS Glue Developer Guide
4: AWS CodeCommit User Guide
5: Amazon Kinesis Data Streams Developer Guide
6: Amazon Elastic Container Registry User Guide
7: Build a Continuous Delivery Pipeline for Your Container Images with Amazon ECR as Source

問題 #59

A data engineer is configuring Amazon SageMaker Studio to use AWS Glue interactive sessions to prepare data for machine learning (ML) models.

The data engineer receives an access denied error when the data engineer tries to prepare the data by using SageMaker Studio. Which change should the engineer make to gain access to SageMaker Studio?

- A. Add a policy to the data engineer's IAM user that allows the sts:AddAssociation action for the AWS Glue and SageMaker service principals in the trust policy.
- B. Add the AWSGlueServiceRole managed policy to the data engineer's IAM user.
- C. Add the AmazonSageMakerFullAccess managed policy to the data engineer's IAM user.
- D. Add a policy to the data engineer's IAM user that includes the sts:AssumeRole action for the AWS Glue and SageMaker service principals in the trust policy.

解題說明：

This solution meets the requirement of gaining access to SageMaker Studio to use AWS Glue interactive sessions. AWS Glue interactive sessions are a way to use AWS Glue DataBrew and AWS Glue Data Catalog from within SageMaker Studio. To use AWS Glue interactive sessions, the data engineer's IAM user needs to have permissions to assume the AWS Glue service role and the SageMaker execution role. By adding a policy to the data engineer's IAM user that includes the sts:AssumeRole action for the AWS Glue and SageMaker service principals in the trust policy, the data engineer can grant these permissions and avoid the access denied error. The other options are not sufficient or necessary to resolve the error. Reference:

Get started with data integration from Amazon S3 to Amazon Redshift using AWS Glue interactive sessions Troubleshoot Errors - Amazon SageMaker AccessDeniedException on sagemaker:CreateDomain in AWS SageMaker Studio, despite having SageMakerFullAccess

問題 #60

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 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.
- B. 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.
- 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. 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.
- E. 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.

答案：B,D

解題說明：

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. Reference:

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]

問題 #61

A data engineer needs to build an extract, transform, and load (ETL) job. The ETL job will process daily incoming .csv files that users upload to an Amazon S3 bucket. The size of each S3 object is less than 100 MB.

Which solution will meet these requirements MOST cost-effectively?

- A. Write an AWS Glue Python shell job. Use pandas to transform the data.
- B. Write an AWS Glue PySpark job. Use Apache Spark to transform the data.
- C. Write a PySpark ETL script. Host the script on an Amazon EMR cluster.
- D. Write a custom Python application. Host the application on an Amazon Elastic Kubernetes Service (Amazon EKS) cluster.

答案：A

解題說明：
AWS Glue is a fully managed serverless ETL service that can handle various data sources and formats, including .csv files in Amazon S3. AWS Glue provides two types of jobs: PySpark and Python shell. PySpark jobs use Apache Spark to process large-scale data in parallel, while Python shell jobs use Python scripts to process small-scale data in a single execution environment. For this requirement, a Python shell job is more suitable and cost-effective, as the size of each S3 object is less than 100 MB, which does not require distributed processing. A Python shell job can use pandas, a popular Python library for data analysis, to transform the .csv data as needed. The other solutions are not optimal or relevant for this requirement. Writing a custom Python application and hosting it on an Amazon EKS cluster would require more effort and resources to set up and manage the Kubernetes environment, as well as to handle the data ingestion and transformation logic. Writing a PySpark ETL script and hosting it on an Amazon EMR cluster would also incur more costs and complexity to provision and configure the EMR cluster, as well as to use Apache Spark for processing small data files. Writing an AWS Glue PySpark job would also be less efficient and economical than a Python shell job, as it would involve unnecessary overhead and charges for using Apache Spark for small data files. Reference: AWS Glue
Working with Python Shell Jobs
pandas
[AWS Certified Data Engineer - Associate DEA-C01 Complete Study Guide]

問題 #62
A company wants to migrate data from an Amazon RDS for PostgreSQL DB instance in the eu-east-1 Region of an AWS account named Account_A. The company will migrate the data to an Amazon Redshift cluster in the eu-west-1 Region of an AWS account named Account_B.
Which solution will give AWS Database Migration Service (AWS DMS) the ability to replicate data between two data stores?

- A. Set up an AWS DMS replication instance in a new AWS account in eu-west-1
- B. Set up an AWS DMS replication instance in Account_B in eu-west-1.
- C. Set up an AWS DMS replication instance in Account_A in eu-east-1.
- D. Set up an AWS DMS replication instance in Account_B in eu-east-1.

答案：B

解題說明：
To migrate data from an Amazon RDS for PostgreSQL DB instance in the eu-east-1 Region (Account_A) to an Amazon Redshift cluster in the eu-west-1 Region (Account_B), AWS DMS needs a replication instance located in the target region (in this case, eu-west-1) to facilitate the data transfer between regions.
* Option A: Set up an AWS DMS replication instance in Account_B in eu-west-1.Placing the DMS replication instance in the target account and region (Account_B in eu-west-1) is the most efficient solution. The replication instance can connect to the source RDS PostgreSQL in eu-east-1 and migrate the data to the Redshift cluster in eu-west-1. This setup ensures data is replicated across AWS accounts and regions.
Options B, C, and D place the replication instance in either the wrong account or region, which increases complexity without adding any benefit.
References:
* AWS Database Migration Service (DMS) Documentation
* Cross-Region and Cross-Account Replication

問題 #63
......

Amazon的Data-Engineer-Associate考試的考生都知道，Amazon的Data-Engineer-Associate考試是比較不容易通過的，但是它又是通往成功的必經之路，所以不得不選擇，為了提過高你的職業價值，你有權通過測試認證，我們KaoGuTi設計的考試試題及答案包含不同的針對性，覆蓋面廣，沒有任何其他書籍或者別的資料方式可以超越它，KaoGuTi絕對是幫助你通過測試的王牌考試試題及答案。經過眾人多人的使用結果證明，KaoGuTi通過率高達

100%，KaoGuTi是唯一適合你通過考試的方式，選擇了它，等於創建將了一個美好的未來。

Data-Engineer-Associate最新考證：https://www.kaoguti.com/Data-Engineer-Associate_exam-pdf.html

KaoGuTi Data-Engineer-Associate最新考證從使用過考古題的人們那裏得到了很多的好評，隨著21世紀資訊時代的洪流到來，人們不斷提高自己的知識來適應這個時代，但遠遠不夠，就IT行業來說，Amazon的Data-Engineer-Associate考試認證是IT行業必不可少的認證，想要通過這項考試培訓是必須的，因為這項考試是有所困難的，通過了它，就可以受到國際的認可及接受，你將有一個美好的前程及拿著受人矚目的高薪，KaoGuT網站有全世界最可靠的IT認證培訓資料，有了它你就可以實現你美好的計畫，我們保證你100%通過認證，參加Amazon的Data-Engineer-Associate考試認證的考生們，你們還在猶豫什麼呢，趕緊行動吧，該課程讓考生掌握Amazon AWS Certified Data Engineer各種工具功能,Data-Engineer-Associate是Amazon證照中的AWS Certified Data Engineer Exam考試科目。

倒不是他不想反抗，完全就是心有余而力不足，張如苟陰沈著臉道，KaoGuT從使用過考古題的人們那裏得到了很多的好評，隨著21世紀資訊時代的洪流到來，人們不斷提高自己的知識來適應這個時代，但遠遠不夠，就IT行業來說，Amazon的Data-Engineer-Associate考試認證是IT行業必不可少的認證，想要通過這項考試培訓是必須的，因為這項考試是有所困難的，通過了它，就可以受到國際的認可及接受，你將有一個美好的前程及拿著受人矚目的高薪，KaoGuT網站有全世界最可靠的IT認證培訓資料，有了它你就可以實現你美好的計畫，我們保證你100%通過認證，參加Amazon的Data-Engineer-Associate考試認證的考生們，你們還在猶豫什麼呢，趕緊行動吧！

熱門的Data-Engineer-Associate認證資料，由Amazon權威專家撰寫

該課程讓考生掌握Amazon AWS Certified Data Engineer各種工具功能,Data-Engineer-Associate是Amazon證照中的AWS Certified Data Engineer Exam考試科目，這樣在考試時你就可以輕鬆自如地应对了，在這個人才濟濟的社會裏，你不覺得壓力很大嗎，不管你的學歷有多高，它永遠不代表實力。

- Data-Engineer-Associate試題 🔗 Data-Engineer-Associate題庫更新 🔗 Data-Engineer-Associate考題資訊 🔗 🔗 www.vcesoft.com 🔗網站搜索▷ Data-Engineer-Associate ◁並免費下載Data-Engineer-Associate熱門證照
- 新版Data-Engineer-Associate考古題 🔗 Data-Engineer-Associate信息資訊 🔗 Data-Engineer-Associate考題寶典 🔗 🔗 在▷ www.newdumpspdf.com ◁網站上查找🔗 Data-Engineer-Associate 🔗的最新題庫Data-Engineer-Associate考題套裝
- 完成Data-Engineer-Associate認證資料 |第一次嘗試輕鬆學習並通過考試 - 最近更正的Data-Engineer-Associate：AWS Certified Data Engineer - Associate (DEA-C01) 🔗 ✔ www.testpdf.net 🔗✔🔗上搜索【 Data-Engineer-Associate 】輕鬆獲取免費下載Data-Engineer-Associate認證題庫
- 一流的Data-Engineer-Associate認證資料和資格考試的領導者和實用的Data-Engineer-Associate：AWS Certified Data Engineer - Associate (DEA-C01) 🔗 在（ www.newdumpspdf.com ）上搜索☀ Data-Engineer-Associate 🔗☀🔗並獲取免費下載Data-Engineer-Associate考題套裝
- Data-Engineer-Associate熱門證照 🔗 Data-Engineer-Associate考題寶典 🔗 Data-Engineer-Associate考題資訊 🖥 立即到➡ www.kaoguti.com 🔗🔗🔗上搜索" Data-Engineer-Associate "以獲取免費下載新版Data-Engineer-Associate考古題
- 最好的Data-Engineer-Associate認證資料擁有模擬真實考試環境與場境的軟件VCE版本＆精準的Data-Engineer-Associate：AWS Certified Data Engineer - Associate (DEA-C01) 🔗 免費下載✔ Data-Engineer-Associate 🔗✔🔗只需在《 www.newdumpspdf.com 》上搜索Data-Engineer-Associate認證資料
- 熱門的Data-Engineer-Associate認證資料，覆蓋全真 {examanme} Data-Engineer-Associate考試考題 🔗 透過🔗 www.newdumpspdf.com 🔗搜索⇒ Data-Engineer-Associate ⇐免費下載考試資料Data-Engineer-Associate在線考題
- 準確的Data-Engineer-Associate認證資料 |高通過率的考試材料|免費下載Data-Engineer-Associate：AWS Certified Data Engineer - Associate (DEA-C01) 🔗 立即在{ www.newdumpspdf.com }上搜尋▶ Data-Engineer-Associate ◀並免費下載Data-Engineer-Associate通過考試
- Data-Engineer-Associate通過考試 🔗 Data-Engineer-Associate在線考題 🔗 新版Data-Engineer-Associate考古題 🔗 🔗 進入《 www.newdumpspdf.com 》搜尋「 Data-Engineer-Associate 」免費下載Data-Engineer-Associate信息資訊
- 最新Data-Engineer-Associate考古題 🔗 Data-Engineer-Associate考題套裝 🔗 最新Data-Engineer-Associate題庫資訊 🔗 ☀ www.newdumpspdf.com 🔗☀🔗最新➦ Data-Engineer-Associate 🔗問題集合Data-Engineer-Associate熱門證照
- 新版Data-Engineer-Associate考古題 🔗 Data-Engineer-Associate考題套裝 🔗 Data-Engineer-Associate在線題庫 🔗 🔗⇒ www.pdfexamdumps.com ⇐提供免費⇒ Data-Engineer-Associate ⇐問題收集Data-Engineer-Associate考題套裝
- webanalyticsbd.com, 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, myportal.utt.edu.tt, www.stes.tyc.edu.tw, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, www.stes.tyc.edu.tw, study.stcs.edu.np, 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, 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,

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, myportal.utt.edu.tt, Disposable vapes

順便提一下，可以從雲存儲中下載KaoGuTi Data-Engineer-Associate考試題庫的完整版：https://drive.google.com/open?id=1GfyW0PY22U9DPce59nTed4Bml3TCrZLe