# Pass Leader Data-Engineer-Associate Dumps & Instant Data-Engineer-Associate Download


PLPassLeader
Leader of IT Certifications
New VCE and PDF Exam Dumps from PassLeader

- Vendor: Microsoft
- Exam Code: DP-900
- Exam Name: Microsoft Azure Data Fundamentals
- Part of New Questions from PassLeader (Updated in Feb/2021)

Visit PassLeader and Download Full Version DP-900 Exam Dumps

NEW QUESTION 125
You have an application that runs on Windows and requires across to a mapped drive. Which Azure service should you use?

A. Azure Cosmos DB
B. Azure Table storage
C. Azure Files
D. Azure Blob Storage

Answer: A

NEW QUESTION 126
You need to recommend a non-relational data store that is optimized for storing and retrieving files, videos, audio stream, and virtual disk images. The data store must store data, some metadata, and a unique ID for each file. Which type of data store should you recommend?

A. document
B. key/value
C. object
D. columnar

Answer: C
Explanation:
Object storage is optimized for storing and retrieving large binary objects (images, files, video and audio streams, large application data objects and documents, virtual machine disk images). Large data files are also popularly used in this model, for example, delimiter file (CSV), parquet, and ORC. Object stores can manage extremely large amounts of unstructured data.
https://docs.microsoft.com/en-us/azure/architecture/guide/technology-choices/data-store-overview

NEW QUESTION 127
Which statement is an example of Data Manipulation Language (DML)?

A. INSERT
B. DISABLE
C. ALTER
D. DROP

DP-900 Exam Dumps DP-900 Exam Questions DP-900 PDF Dumps DP-900 VCE Dumps
https://www.passleader.com/dp-900.html

BTW, DOWNLOAD part of LatestCram Data-Engineer-Associate dumps from Cloud Storage: https://drive.google.com/open?id=1m2HGW4rimeYrMVMDPFhJa-JLfSVlnU5z

It is universally accepted that in this competitive society in order to get a good job we have no choice but to improve our own capacity and explore our potential constantly, and try our best to get the related Data-Engineer-Associate certification is the best way to show our professional ability, however, the exam is hard nut to crack and there are so many Data-Engineer-Associate Preparation questions related to the exam, it seems impossible for us to systematize all of the key points needed for the exam by ourselves.

It is of no exaggeration to say that sometimes Data-Engineer-Associate certification is exactly a stepping-stone to success, especially when you are hunting for a job. The Data-Engineer-Associate study materials are of great help in this sense. With the Data-Engineer-Associate test training, you can both have the confidence and gumption to ask for better treatment. To earn such a material, you can spend some time to study our Data-Engineer-Associate study torrent. No study can be done successfully without a specific goal and a powerful drive, and here to earn a better living by getting promotion is a good one.

**>> Pass Leader Data-Engineer-Associate Dumps <<**

## Try a Free Demo and Then Buy Amazon Data-Engineer-Associate Exam Dumps

In this Desktop-based Amazon Data-Engineer-Associate practice exam software, you will enjoy the opportunity to self-exam your preparation. The chance to customize the Amazon Data-Engineer-Associate practice exams according to the time and types of

Amazon Data-Engineer-Associate practice test questions will contribute to your ease. This format operates only on Windows-based devices. But what is helpful is that it functions without an active internet connection. It copies the exact pattern and style of the real Amazon Data-Engineer-Associate Exam to make your preparation productive and relevant.

## Amazon AWS Certified Data Engineer - Associate (DEA-C01) Sample Questions (Q83-Q88):

**NEW QUESTION # 83**
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 API calls to access and integrate third-party datasets from AWS Data Exchange.
- B. Use Amazon Kinesis Data Streams to access and integrate third-party datasets from Amazon Elastic Container Registry (Amazon ECR).
- 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

**Answer: A**

Explanation:
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 .
References:
* 1: AWS Data Exchange User Guide
* 2: AWS Data Exchange FAQs
* 3: AWS Glue Developer Guide
* : AWS CodeCommit User Guide
* : Amazon Kinesis Data Streams Developer Guide
* : Amazon Elastic Container Registry User Guide
* : Build a Continuous Delivery Pipeline for Your Container Images with Amazon ECR as Source

**NEW QUESTION # 84**
A company uses an Amazon Redshift cluster as a data warehouse that is shared across two departments. To comply with a security policy, each department must have unique access permissions.
Department A must have access to tables and views for Department A. Department B must have access to tables and views for

Department B.
The company often runs SQL queries that use objects from both departments in one query.
Which solution will meet these requirements with the LEAST operational overhead?

- A. Create an IAM user group for each department. Use identity-based IAM policies to grant table and view permissions based on the IAM user group.
- B. Group tables and views for each department into dedicated databases. Manage permissions at the database level.
- C. Update the names of the tables and views to follow a naming convention that contains the department names. Manage permissions based on the new naming convention.
- D. Group tables and views for each department into dedicated schemas. Manage permissions at the schema level.

**Answer: D**

**NEW QUESTION # 85**
A manufacturing company wants to collect data from sensors. A data engineer needs to implement a solution that ingests sensor data in near real time.
The solution must store the data to a persistent data store. The solution must store the data in nested JSON format. The company must have the ability to query from the data store with a latency of less than 10 milliseconds.
Which solution will meet these requirements with the LEAST operational overhead?

- A. Use Amazon Simple Queue Service (Amazon SQS) to buffer incoming sensor data. Use AWS Glue to store the data in Amazon RDS for querying.
- B. Use AWS Lambda to process the sensor data. Store the data in Amazon S3 for querying.
- C. Use Amazon Kinesis Data Streams to capture the sensor data. Store the data in Amazon DynamoDB for querying.
- D. Use a self-hosted Apache Kafka cluster to capture the sensor data. Store the data in Amazon S3 for querying.

**Answer: C**

Explanation:
Amazon Kinesis Data Streams is a service that enables you to collect, process, and analyze streaming data in real time. You can use Kinesis Data Streams to capture sensor data from various sources, such as IoT devices, web applications, or mobile apps. You can create data streams that can scale up to handle any amount of data from thousands of producers. You can also use the Kinesis Client Library (KCL) or the Kinesis Data Streams API to write applications that process and analyze the data in the streams1.
Amazon DynamoDB is a fully managed NoSQL database service that provides fast and predictable performance with seamless scalability. You can use DynamoDB to store the sensor data in nested JSON format, as DynamoDB supports document data types, such as lists and maps. You can also use DynamoDB to query the data with a latency of less than 10 milliseconds, as DynamoDB offers single-digit millisecond performance for any scale of data. You can use the DynamoDB API or the AWS SDKs to perform queries on the data, such as using key-value lookups, scans, or queries2.
The solution that meets the requirements with the least operational overhead is to use Amazon Kinesis Data Streams to capture the sensor data and store the data in Amazon DynamoDB for querying. This solution has the following advantages:
It does not require you to provision, manage, or scale any servers, clusters, or queues, as Kinesis Data Streams and DynamoDB are fully managed services that handle all the infrastructure for you. This reduces the operational complexity and cost of running your solution.
It allows you to ingest sensor data in near real time, as Kinesis Data Streams can capture data records as they are produced and deliver them to your applications within seconds. You can also use Kinesis Data Firehose to load the data from the streams to DynamoDB automatically and continuously3.
It allows you to store the data in nested JSON format, as DynamoDB supports document data types, such as lists and maps. You can also use DynamoDB Streams to capture changes in the data and trigger actions, such as sending notifications or updating other databases.
It allows you to query the data with a latency of less than 10 milliseconds, as DynamoDB offers single-digit millisecond performance for any scale of data. You can also use DynamoDB Accelerator (DAX) to improve the read performance by caching frequently accessed data.
Option A is incorrect because it suggests using a self-hosted Apache Kafka cluster to capture the sensor data and store the data in Amazon S3 for querying. This solution has the following disadvantages:
It requires you to provision, manage, and scale your own Kafka cluster, either on EC2 instances or on-premises servers. This increases the operational complexity and cost of running your solution.
It does not allow you to query the data with a latency of less than 10 milliseconds, as Amazon S3 is an object storage service that is not optimized for low-latency queries. You need to use another service, such as Amazon Athena or Amazon Redshift Spectrum, to query the data in S3, which may incur additional costs and latency.
Option B is incorrect because it suggests using AWS Lambda to process the sensor data and store the data in Amazon S3 for

querying. This solution has the following disadvantages:
It does not allow you to ingest sensor data in near real time, as Lambda is a serverless compute service that runs code in response to events. You need to use another service, such as API Gateway or Kinesis Data Streams, to trigger Lambda functions with sensor data, which may add extra latency and complexity to your solution.
It does not allow you to query the data with a latency of less than 10 milliseconds, as Amazon S3 is an object storage service that is not optimized for low-latency queries. You need to use another service, such as Amazon Athena or Amazon Redshift Spectrum, to query the data in S3, which may incur additional costs and latency.
Option D is incorrect because it suggests using Amazon Simple Queue Service (Amazon SQS) to buffer incoming sensor data and use AWS Glue to store the data in Amazon RDS for querying. This solution has the following disadvantages:
It does not allow you to ingest sensor data in near real time, as Amazon SQS is a message queue service that delivers messages in a best-effort manner. You need to use another service, such as Lambda or EC2, to poll the messages from the queue and process them, which may add extra latency and complexity to your solution.
It does not allow you to store the data in nested JSON format, as Amazon RDS is a relational database service that supports structured data types, such as tables and columns. You need to use another service, such as AWS Glue, to transform the data from JSON to relational format, which may add extra cost and overhead to your solution.
Reference:
1: Amazon Kinesis Data Streams - Features
2: Amazon DynamoDB - Features
3: Loading Streaming Data into Amazon DynamoDB - Amazon Kinesis Data Firehose
[4]: Capturing Table Activity with DynamoDB Streams - Amazon DynamoDB
[5]: Amazon DynamoDB Accelerator (DAX) - Features
[6]: Amazon S3 - Features
[7]: AWS Lambda - Features
[8]: Amazon Simple Queue Service - Features
[9]: Amazon Relational Database Service - Features
[10]: Working with JSON in Amazon RDS - Amazon Relational Database Service
[11]: AWS Glue - Features

**NEW QUESTION # 86**
A company uses an Amazon Redshift cluster that runs on RA3 nodes. The company wants to scale read and write capacity to meet demand. A data engineer needs to identify a solution that will turn on concurrency scaling.
Which solution will meet this requirement?

- A. Turn on concurrency scaling in workload management (WLM) for Redshift Serverless workgroups.
- B. Turn on concurrency scaling at the workload management (WLM) queue level in the Redshift cluster.
- C. Turn on concurrency scaling for the daily usage quota for the Redshift cluster.
- D. Turn on concurrency scaling in the settings during the creation of and new Redshift cluster.

**Answer: B**

Explanation:
Concurrency scaling is a feature that allows you to support thousands of concurrent users and queries, with consistently fast query performance. When you turn on concurrency scaling, Amazon Redshift automatically adds query processing power in seconds to process queries without any delays. You can manage which queries are sent to the concurrency-scaling cluster by configuring WLM queues. To turn on concurrency scaling for a queue, set the Concurrency Scaling mode value to auto. The other options are either incorrect or irrelevant, as they do not enable concurrency scaling for the existing Redshift cluster on RA3 nodes. Reference:
Working with concurrency scaling - Amazon Redshift
Amazon Redshift Concurrency Scaling - Amazon Web Services
Configuring concurrency scaling queues - Amazon Redshift
AWS Certified Data Engineer - Associate DEA-C01 Complete Study Guide (Chapter 6, page 163)

**NEW QUESTION # 87**
A company is building an analytics solution. The solution uses Amazon S3 for data lake storage and Amazon Redshift for a data warehouse. The company wants to use Amazon Redshift Spectrum to query the data that is in Amazon S3.
Which actions will provide the FASTEST queries? (Choose two.)

- A. Use a columnar storage file format.
- B. Partition the data based on the most common query predicates.

- C. Split the data into files that are less than 10 KB.
- D. Use file formats that are not
- E. Use gzip compression to compress individual files to sizes that are between 1 GB and 5 GB.

**Answer: A,B**

Explanation:
Amazon Redshift Spectrum is a feature that allows you to run SQL queries directly against data in Amazon S3, without loading or transforming the data. Redshift Spectrum can query various data formats, such as CSV, JSON, ORC, Avro, and Parquet. However, not all data formats are equally efficient for querying. Some data formats, such as CSV and JSON, are row-oriented, meaning that they store data as a sequence of records, each with the same fields. Row-oriented formats are suitable for loading and exporting data, but they are not optimal for analytical queries that often access only a subset of columns. Row-oriented formats also do not support compression or encoding techniques that can reduce the data size and improve the query performance.
On the other hand, some data formats, such as ORC and Parquet, are column-oriented, meaning that they store data as a collection of columns, each with a specific data type. Column-oriented formats are ideal for analytical queries that often filter, aggregate, or join data by columns. Column-oriented formats also support compression and encoding techniques that can reduce the data size and improve the query performance. For example, Parquet supports dictionary encoding, which replaces repeated values with numeric codes, and run-length encoding, which replaces consecutive identical values with a single value and a count. Parquet also supports various compression algorithms, such as Snappy, GZIP, and ZSTD, that can further reduce the data size and improve the query performance.
Therefore, using a columnar storage file format, such as Parquet, will provide faster queries, as it allows Redshift Spectrum to scan only the relevant columns and skip the rest, reducing the amount of data read from S3. Additionally, partitioning the data based on the most common query predicates, such as date, time, region, etc., will provide faster queries, as it allows Redshift Spectrum to prune the partitions that do not match the query criteria, reducing the amount of data scanned from S3. Partitioning also improves the performance of joins and aggregations, as it reduces data skew and shuffling.
The other options are not as effective as using a columnar storage file format and partitioning the data. Using gzip compression to compress individual files to sizes that are between 1 GB and 5 GB will reduce the data size, but it will not improve the query performance significantly, as gzip is not a splittable compression algorithm and requires decompression before reading. Splitting the data into files that are less than 10 KB will increase the number of files and the metadata overhead, which will degrade the query performance. Using file formats that are not supported by Redshift Spectrum, such as XML, will not work, as Redshift Spectrum will not be able to read or parse the data. Reference:
Amazon Redshift Spectrum
Choosing the Right Data Format
AWS Certified Data Engineer - Associate DEA-C01 Complete Study Guide, Chapter 4: Data Lakes and Data Warehouses, Section 4.3: Amazon Redshift Spectrum

**NEW QUESTION # 88**
.....

# In the event that you fail the Amazon Data-Engineer-Associate exam, you will receive a refund

The result must go beyond your expectations, Instant Data-Engineer-Associate Download Every buyer can share one year free updates and preparation assist, As we all know, LatestCram's Amazon Data-Engineer-Associate Exam Training materials has very high profile, and it is also well-known in the worldwide.

You can do this easily after passing the AWS Certified Data Engineer - Associate (DEA-C01) Data-Engineer-Associate certification exam.

- New Data-Engineer-Associate Dumps Sheet □ Pdf Data-Engineer-Associate Torrent □ Data-Engineer-Associate Reliable Test Answers □ Easily obtain [ Data-Engineer-Associate ] for free download through 「 www.examcollectionpass.com 」 □Latest Data-Engineer-Associate Dumps Ebook
- Top Amazon Pass Leader Data-Engineer-Associate Dumps - Authoritative Pdfvce - Leading Offer in Qualification Exams □ □ Search for ⇛ Data-Engineer-Associate ⇚ and download it for free immediately on （ www.pdfvce.com ） □Latest Data-Engineer-Associate Dumps Ebook
- Pass Leader Data-Engineer-Associate Dumps Newest Questions Pool Only at www.examcollectionpass.com □ Search for ➡ Data-Engineer-Associate □□□ on " www.examcollectionpass.com " immediately to obtain a free download □Data-Engineer-Associate Key Concepts
- Pdf Data-Engineer-Associate Torrent □ Data-Engineer-Associate Real Exam Questions □ Training Data-Engineer-Associate For Exam ↘ Copy URL （ www.pdfvce.com ） open and search for [ Data-Engineer-Associate ] to download for free □Latest Data-Engineer-Associate Dumps Ebook
- Reliable Data-Engineer-Associate Braindumps Questions □ Latest Data-Engineer-Associate Dumps Ebook □ Data-Engineer-Associate Latest Examprep □ Download 「 Data-Engineer-Associate 」 for free by simply searching on [ www.prepawaypdf.com ] □Exam Data-Engineer-Associate Question
- Free PDF Perfect Amazon - Data-Engineer-Associate - Pass Leader AWS Certified Data Engineer - Associate (DEA-C01) Dumps □ Simply search for ➤ Data-Engineer-Associate □ for free download on ➤ www.pdfvce.com □ □Latest Data-Engineer-Associate Dumps Ebook
- Data-Engineer-Associate Reliable Dumps Questions 〜 Training Data-Engineer-Associate For Exam □ Latest Data-Engineer-Associate Test Question □ Easily obtain □ Data-Engineer-Associate □ for free download through □ www.prepawayexam.com □ □Data-Engineer-Associate Sure Pass
- Data-Engineer-Associate Reliable Dumps Questions □ Latest Data-Engineer-Associate Dumps Ebook □ Exam Data-Engineer-Associate Preparation □ The page for free download of ▸ Data-Engineer-Associate ◂ on （ www.pdfvce.com ） will open immediately □Data-Engineer-Associate Latest Examprep
- Reliable Data-Engineer-Associate Test Book ↖ New Data-Engineer-Associate Dumps Sheet □ Exam Data-Engineer-Associate Question □ The page for free download of 「 Data-Engineer-Associate 」 on ✔ www.vceengine.com □✔□ will open immediately □Latest Data-Engineer-Associate Test Question
- Get Latest Amazon Data-Engineer-Associate Exam Dumps [2026] □ Search for ▹ Data-Engineer-Associate ◃ on " www.pdfvce.com " immediately to obtain a free download □Exam Data-Engineer-Associate Preparation
- Data-Engineer-Associate Reliable Dumps Questions □ Latest Data-Engineer-Associate Test Question □ Exam Data-Engineer-Associate Preparation □ Open website ➥ www.validtorrent.com □ and search for ☀ Data-Engineer-Associate □☀□ for free download □Data-Engineer-Associate Reliable Dumps Questions
- mpgimer.edu.in, pct.edu.pk, 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, www.stes.tyc.edu.tw, shortcourses.russellcollege.edu.au, 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, www.stes.tyc.edu.tw, Disposable vapes

DOWNLOAD the newest LatestCram Data-Engineer-Associate PDF dumps from Cloud Storage for free: https://drive.google.com/open?id=1m2HGW4rimeYrMVMDPFhJa-JLfSVlnU5z