

Databricks-Certified-Professional-Data-Engineer 100% 시험패스덤프 & Databricks-Certified-Professional-Data-Engineer시험정보



그 외, KoreaDumps Databricks-Certified-Professional-Data-Engineer 시험 문제집 일부가 지금은 무료입니다:
https://drive.google.com/open?id=14x7-mBwr4PNTog1W6x_PfCmNZyaHNI-n

KoreaDumps에서 Databricks Certified-Professional-Data-Engineer 덤프를 구매하시면 일년무료 업데이트서비스를 받을수 있습니다. 일년무료 업데이트서비스란 구매일로부터 1년동안 구매한 덤프가 업데이트될때마다 구매시 사용한 메일주소로 가장 최신버전을 보내드리는것을 의미합니다. Databricks Certified-Professional-Data-Engineer 덤프에는 가장 최신시험문제의 기출문제가 포함되어있어 높은 적주율을 자랑하고 있습니다.

Databricks Certified Professional Data Engineer 시험은 데이터 엔지니어링에 대한 광범위한 지식과 경험이 필요한 엄격한 인증 시험입니다. 응시자는 데이터 모델링, 데이터웨어 하우징, ETL, 데이터 거버넌스 및 데이터 보안과 같은 데이터 엔지니어링 개념을 깊이 이해해야 합니다. 또한 Apache Spark, Delta Lake 및 MLFlow와 같은 Databricks 도구 및 기술을 사용한 경험이 있어야 합니다. 이 시험을 통과하면 응시자는 데이터 사업 플랫폼에서 데이터 파이프 라인을 구축하고 최적화하는 데 필요한 기술과 지식이 있음을 보여줍니다.

Databricks 인증 전문 데이터 엔지니어가 되기 위해서는, 후보자들은 데이터 엔지니어링과 관련된 다양한 주제를 다루는 엄격한 시험을 통과해야 합니다. 이 시험은 Databricks 도구와 기술을 사용하여 데이터 파이프라인을 디자인, 구축, 유지 관리하는 능력을 시험하는 것뿐만 아니라, 데이터 모델링, 데이터 웨어하우징 및 데이터 통합에 대한 이해를 시험합니다.

>> Databricks-Certified-Professional-Data-Engineer 100% 시험패스 덤프 <<

시험패스에 유효한 Databricks-Certified-Professional-Data-Engineer 100% 시험패스 덤프 덤프데모

KoreaDumps의 Databricks 인증 Databricks-Certified-Professional-Data-Engineer 덤프를 구매하여 공부한지 일주일만에 바로 시험을 보았는데 고득점으로 시험을 패스했습니다. 이는 KoreaDumps의 Databricks 인증 Databricks-Certified-Professional-Data-Engineer 덤프를 구매한 분이 전해온 회소식입니다. 다른 자료 필요없이 단지 저희 Databricks 인증 Databricks-Certified-Professional-Data-Engineer 덤프로 이렇게 어려운 시험을 일주일만에 패스하고 자격증을 취득할 수 있습니다. 덤프 가격도 다른 사이트보다 만만하여 부담없이 덤프 마련이 가능합니다. 구매전 무료샘플을 다운받아 보시면 믿음을 느낄 것입니다.

최신 Databricks Certification Databricks-Certified-Professional-Data-Engineer 무료샘플문제 (Q22-Q27):

질문 # 22

Which of the statements are incorrect when choosing between lakehouse and Datawarehouse?

- A. Lakehouse cannot serve low query latency with high reliability for BI workloads, only suitable for batch workloads.

- B. Lakehouse uses standard data formats like Parquet.
- C. Lakehouse can be accessed through various API's including but not limited to Py-thon/R/SQL
- D. Lakehouse can have special indexes and caching which are optimized for Machine learning
- E. Traditional Data warehouses have storage and compute are coupled.

정답: A

설명:

Explanation

The answer is Lakehouse cannot serve low query latency with high reliability for BI workloads, only suitable for batch workloads. Lakehouse can replace traditional warehouses by leveraging storage and compute optimizations like caching to serve them with low query latency with high reliability.

Focus on comparisons between Spark Cache vs Delta Cache.

<https://docs.databricks.com/delta/optimizations/delta-cache.html>

What Is a Lakehouse? - The Databricks Blog

Graphical user interface, text, application Description automatically generated

A lakehouse has the following key features:

- **Transaction support:** In an enterprise lakehouse many data pipelines will often be reading and writing data concurrently. Support for ACID transactions ensures consistency as multiple parties concurrently read or write data, typically using SQL.
- **Schema enforcement and governance:** The Lakehouse should have a way to support schema enforcement and evolution, supporting DW schema architectures such as star/snowflake-schemas. The system should be able to **reason about data integrity**, and it should have robust governance and auditing mechanisms.
- **BI support:** Lakehouses enable using BI tools directly on the source data. This reduces staleness and improves recency, reduces latency, and lowers the cost of having to operationalize two copies of the data in both a data lake and a warehouse.
- **Storage is decoupled from compute:** In practice this means storage and compute use separate clusters, thus these systems are able to scale to many more concurrent users and larger data sizes. Some modern data warehouses also have this property.
- **Openness:** The storage formats they use are open and standardized, such as Parquet, and they provide an API so a variety of tools and engines, including machine learning and Python/R libraries, can efficiently access the data **directly**.
- **Support for diverse data types ranging from unstructured to structured data:** The lakehouse can be used to store, refine, analyze, and access data types needed for many new data applications, including images, video, audio, semi-structured data, and text.
- **Support for diverse workloads:** including data science, machine learning, and SQL and analytics. Multiple tools might be needed to support all these workloads but they all rely on the same data repository.
- **End-to-end streaming:** Real-time reports are the norm in many enterprises. Support for streaming eliminates the need for separate systems dedicated to serving real-time data applications.

Bottom of Form

Top of Form

질문 # 23

Which statement characterizes the general programming model used by Spark Structured Streaming?

- A. Structured Streaming leverages the parallel processing of GPUs to achieve highly parallel data throughput.
- B. Structured Streaming is implemented as a messaging bus and is derived from Apache Kafka.

- C. Structured Streaming models new data arriving in a data stream as new rows appended to an unbounded table.
- D. Structured Streaming relies on a distributed network of nodes that hold incremental state values for cached stages.
- E. Structured Streaming uses specialized hardware and I/O streams to achieve sub-second latency for data transfer.

정답: B

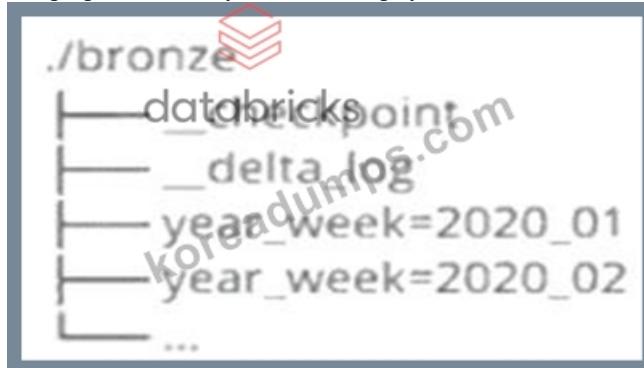
설명:

This is the correct answer because it characterizes the general programming model used by Spark Structured Streaming, which is to treat a live data stream as a table that is being continuously appended. This leads to a new stream processing model that is very similar to a batch processing model, where users can express their streaming computation using the same Dataset/DataFrame API as they would use for static data. The Spark SQL engine will take care of running the streaming query incrementally and continuously and updating the final result as streaming data continues to arrive. Verified References: [Databricks Certified Data Engineer Professional], under "Structured Streaming" section; Databricks Documentation, under "Overview" section.

질문 # 24

A data architect has designed a system in which two Structured Streaming jobs will concurrently write to a single bronze Delta table. Each job is subscribing to a different topic from an Apache Kafka source, but they will write data with the same schema. To keep the directory structure simple, a data engineer has decided to nest a checkpoint directory to be shared by both streams.

The proposed directory structure is displayed below:



Which statement describes whether this checkpoint directory structure is valid for the given scenario and why?

- A. No; each of the streams needs to have its own checkpoint directory.
- B. No; Delta Lake manages streaming checkpoints in the transaction log.
- C. No; only one stream can write to a Delta Lake table.
- D. Yes; both of the streams can share a single checkpoint directory.
- E. Yes; Delta Lake supports infinite concurrent writers.

정답: A

설명:

This is the correct answer because checkpointing is a critical feature of Structured Streaming that provides fault tolerance and recovery in case of failures. Checkpointing stores the current state and progress of a streaming query in a reliable storage system, such as DBFS or S3. Each streaming query must have its own checkpoint directory that is unique and exclusive to that query. If two streaming queries share the same checkpoint directory, they will interfere with each other and cause unexpected errors or data loss. Verified References: [Databricks Certified Data Engineer Professional], under "Structured Streaming" section; Databricks Documentation, under "Checkpointing" section.

질문 # 25

The data engineer team has been tasked with configured connections to an external database that does not have a supported native connector with Databricks. The external database already has data security configured by group membership. These groups map directly to user group already created in Databricks that represent various teams within the company.

A new login credential has been created for each group in the external database. The Databricks Utilities Secrets module will be used to make these credentials available to Databricks users.

Assuming that all the credentials are configured correctly on the external database and group membership is properly configured on Databricks, which statement describes how teams can be granted the minimum necessary access to using these credentials?

- A. "Manage" permission should be set on a secret scope containing only those credentials that will be used by a given team.

- B. "Read" permissions should be set on a secret scope containing only those credentials that will be used by a given team
- C. No additional configuration is necessary as long as all users are configured as administrators in the workspace where secrets have been added.
- D. "Read" permissions should be set on a secret key mapped to those credentials that will be used by a given team.

정답: B

설명:

In Databricks, using the Secrets module allows for secure management of sensitive information such as database credentials. Granting 'Read' permissions on a secret key that maps to database credentials for a specific team ensures that only members of that team can access these credentials. This approach aligns with the principle of least privilege, granting users the minimum level of access required to perform their jobs, thus enhancing security.

Reference:

Databricks Documentation on Secret Management: Secrets

질문 # 26

The DevOps team has configured a production workload as a collection of notebooks scheduled to run daily using the Jobs UI. A new data engineering hire is onboarding to the team and has requested access to one of these notebooks to review the production logic.

What are the maximum notebook permissions that can be granted to the user without allowing accidental changes to production code or data?

- A. Can Run
- B. Can Edit
- C. Can Manage
- D. No permissions
- E. Can Read

정답: D

설명:

This is the correct answer because it is the maximum notebook permissions that can be granted to the user without allowing accidental changes to production code or data. Notebook permissions are used to control access to notebooks in Databricks workspaces. There are four types of notebook permissions: Can Manage, Can Edit, Can Run, and Can Read. Can Manage allows full control over the notebook, including editing, running, deleting, exporting, and changing permissions. Can Edit allows modifying and running the notebook, but not changing permissions or deleting it. Can Run allows executing commands in an existing cluster attached to the notebook, but not modifying or exporting it. Can Read allows viewing the notebook content, but not running or modifying it. In this case, granting Can Read permission to the user will allow them to review the production logic in the notebook without allowing them to make any changes to it or run any commands that may affect production data. Verified References: [Databricks Certified Data Engineer Professional], under "Databricks Workspace" section; Databricks Documentation, under "Notebook permissions" section.

질문 # 27

.....

Databricks 인증 Databricks-Certified-Professional-Data-Engineer 시험에 도전해보려고 결정하셨다면 KoreaDumps 덤프 공부 가이드를 추천해드립니다. KoreaDumps 덤프는 고객님께서 필요한 것이 무엇인지 너무나도 잘 알고 있답니다. KoreaDumps의 Databricks 인증 Databricks-Certified-Professional-Data-Engineer 덤프는 Databricks 인증 Databricks-Certified-Professional-Data-Engineer 시험을 쉽게 만들니다.

Databricks-Certified-Professional-Data-Engineer 시험 정보: https://www.koreadumps.com/Databricks-Certified-Professional-Data-Engineer_exam-braindumps.html

- 최신버전 Databricks-Certified-Professional-Data-Engineer 100% 시험패스 덤프 덤프는 Databricks Certified Professional Data Engineer Exam 시험의 높은 적중율을 자랑 ☐ 무료로 쉽게 다운로드 하려면 “www.pass4test.net”에서 「 Databricks-Certified-Professional-Data-Engineer 」를 검색하세요 Databricks-Certified-Professional-Data-Engineer 최신버전 시험대비 공부자료
- 높은 통과율 Databricks-Certified-Professional-Data-Engineer 100% 시험패스 덤프 시험공부자료 ☐ 무료로 다운로드 하려면 ➡ www.itdumpskr.com ☐로 이동하여 ☀ Databricks-Certified-Professional-Data-Engineer ☐☀☐를 검색

하십시오 Databricks-Certified-Professional-Data-Engineer 최신버전 시험덤프문제

BONUS!!! KoreaDumps Databricks-Certified-Professional-Data-Engineer 시험 문제집 전체 버전을 무료로 다운로드하세요 : https://drive.google.com/open?id=14x7-mBwr4PNT0g1W6x_PfCmNZyaHNI-n