Fast2Test
FIRST TEST, FIRST PASS!

Fast2Test

Instant Download

After Payment, our system will send you the products you purchase in mailbox in a minute after payment. If not received within 2 hours, please contact us.

365 Days Free Updates

Free update is available within 365 days after your purchase. After 365 days, you will get 50% discounts for updating.

Money Back Guarantee

Full refund if you fail the corresponding exam in 60 days after purchasing. And Free get any another product.

Security & Privacy

We respect customer privacy. We use McAfee's security service to provide you with utmost security for your personal information & peace of mind.

We're not the only ones **happy** about Fast2test Practice Materials ...

62316+ customers in 100+ countries use Fast2test Self Test Engine. Meet our customers.

https://kr.fast2test.com

유효한 시험자료,최고의 적중율을 자랑하는 시험대비 덤프

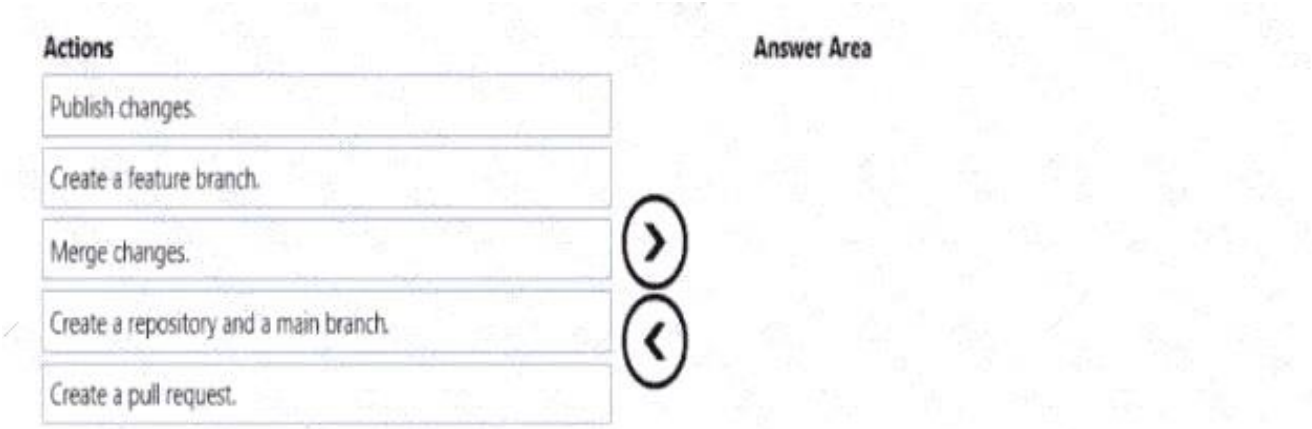
Exam : **DP-203-KR**

Title : Data Engineering on Microsoft Azure (DP-203 Korean Version)

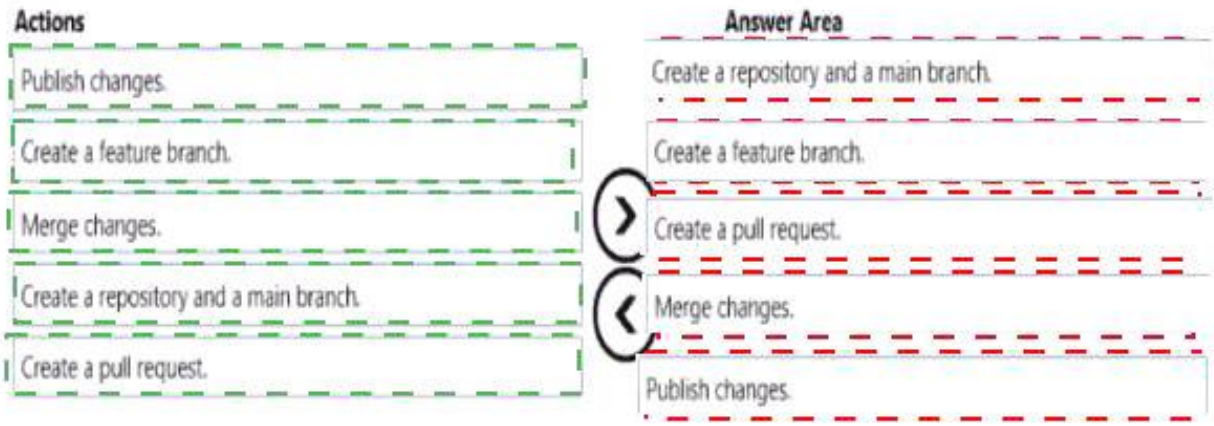
Vendor : Microsoft

Version : DEMO

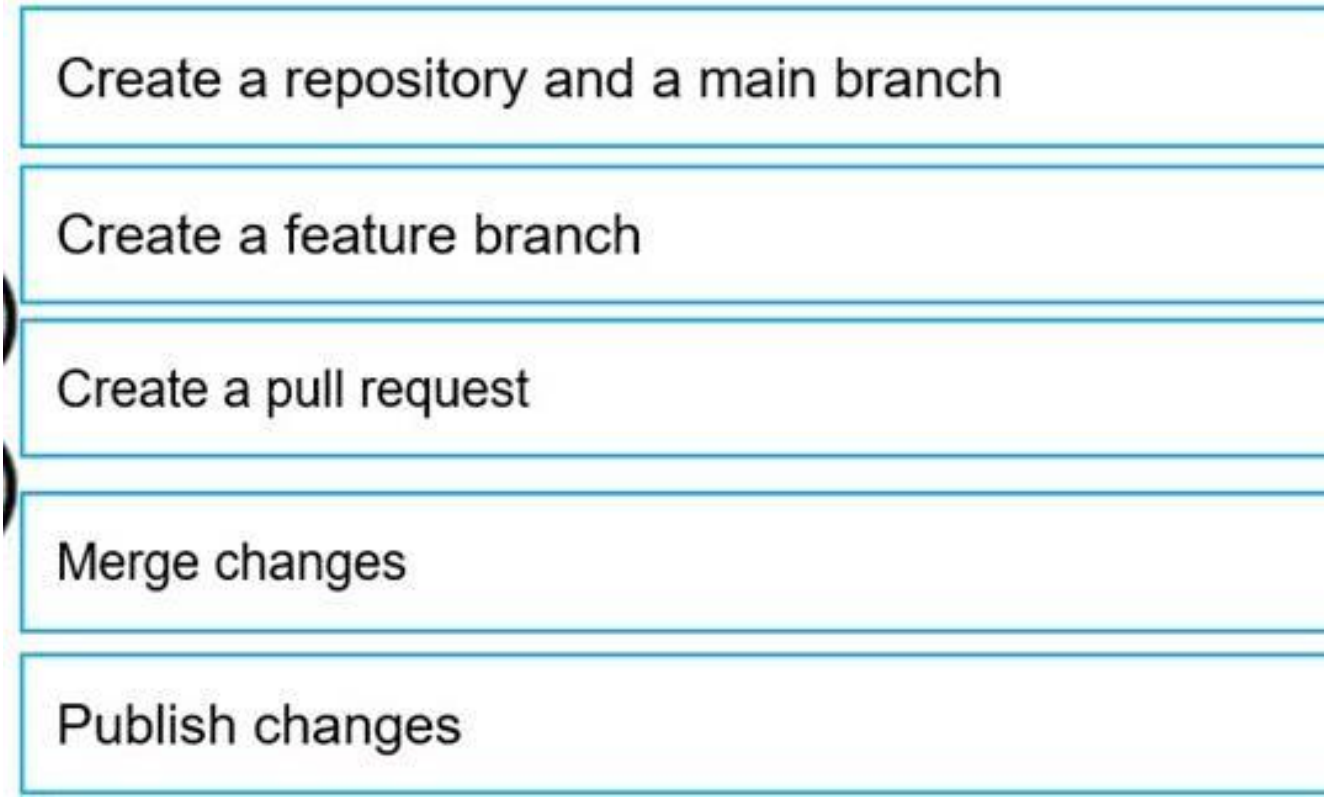
QUESTION NO: 1
통합 파이프라인에 대한 버전 변경 사항을 구현해야 합니다. 솔루션은 데이터 통합 요구 사항을 충족해야 합니다.
어떤 순서로 작업을 수행해야 합니까? 응답하려면 작업 목록에서 모든 작업을 응답 영역으로 이동하고 올바른 순서로 정렬하십시오.



Answer:



Explanation:



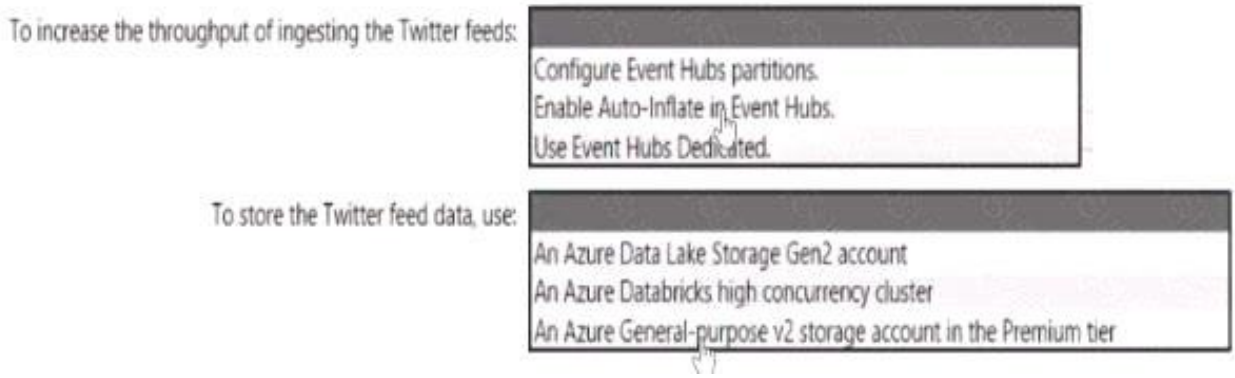
Scenario: Identify a process to ensure that changes to the ingestion and transformation activities can be version-controlled and developed independently by multiple data engineers.
Step 1: Create a repository and a main branch
You need a Git repository in Azure Pipelines, TFS, or GitHub with your app.
Step 2: Create a feature branch
Step 3: Create a pull request
Step 4: Merge changes
Merge feature branches into the main branch using pull requests.
Step 5: Publish changes
Reference:
https://docs.microsoft.com/en-us/azure/devops/pipelines/repos/pipeline-options-for-git
Topic 1, Contoso Case StudyTransactional Date
Contoso has three years of customer, transactional, operation, sourcing, and supplier data comprised of 10 billion records stored across multiple on-premises Microsoft SQL Server servers. The SQL server instances contain data from various operational systems. The data is loaded into the instances by using SQL server integration Services (SSIS) packages.
You estimate that combining all product sales transactions into a company-wide sales transactions dataset will result in a single table that contains 5 billion rows, with one row per transaction.
Most queries targeting the sales transactions data will be used to identify which products were sold in retail stores and which products were sold online during different time period.
Sales transaction data that is older than three years will be removed monthly.
You plan to create a retail store table that will contain the address of each retail store. The table will be approximately 2 MB. Queries for retail store sales will include the retail store addresses.
You plan to create a promotional table that will contain a promotion ID. The promotion ID will

be associated to a specific product. The product will be identified by a product ID. The table will be approximately 5 GB.
Streaming Twitter Data
The ecommerce department at Contoso develops and Azure logic app that captures trending Twitter feeds referencing the company's products and pushes the products to Azure Event Hubs.
Planned Changes
Contoso plans to implement the following changes:
* Load the sales transaction dataset to Azure Synapse Analytics.
* Integrate on-premises data stores with Azure Synapse Analytics by using SSIS packages.
* Use Azure Synapse Analytics to analyze Twitter feeds to assess customer sentiments about products.
Sales Transaction Dataset Requirements
Contoso identifies the following requirements for the sales transaction dataset:
* Partition data that contains sales transaction records. Partitions must be designed to provide efficient loads by month. Boundary values must belong: to the partition on the right.
* Ensure that queries joining and filtering sales transaction records based on product ID complete as quickly as possible.
* Implement a surrogate key to account for changes to the retail store addresses.
* Ensure that data storage costs and performance are predictable.
* Minimize how long it takes to remove old records.
Customer Sentiment Analytics Requirement
Contoso identifies the following requirements for customer sentiment analytics:
* Allow Contoso users to use PolyBase in an A/ure Synapse Analytics dedicated SQL pool to query the content of the data records that host the Twitter feeds. Data must be protected by using row-level security (RLS). The users must be authenticated by using their own A/ureAD credentials.
* Maximize the throughput of ingesting Twitter feeds from Event Hubs to Azure Storage without purchasing additional throughput or capacity units.
* Store Twitter feeds in Azure Storage by using Event Hubs Capture. The feeds will be converted into Parquet files.
* Ensure that the data store supports Azure AD-based access control down to the object level.
* Minimize administrative effort to maintain the Twitter feed data records.
* Purge Twitter feed data records;itftaitJ are older than two years.
Data Integration Requirements
Contoso identifies the following requirements for data integration:
Use an Azure service that leverages the existing SSIS packages to ingest on-premises data into datasets stored in a dedicated SQL pool of Azure Synaps Analytics and transform the data.
Identify a process to ensure that changes to the ingestion and transformation activities can be version controlled and developed independently by multiple data engineers.

QUESTION NO: 2
Twitter 피드에 대한 데이터 수집 및 저장 솔루션을 설계해야 합니다. 솔루션은 고객 감정 분석 요구 사항을 충족해야 합니다.

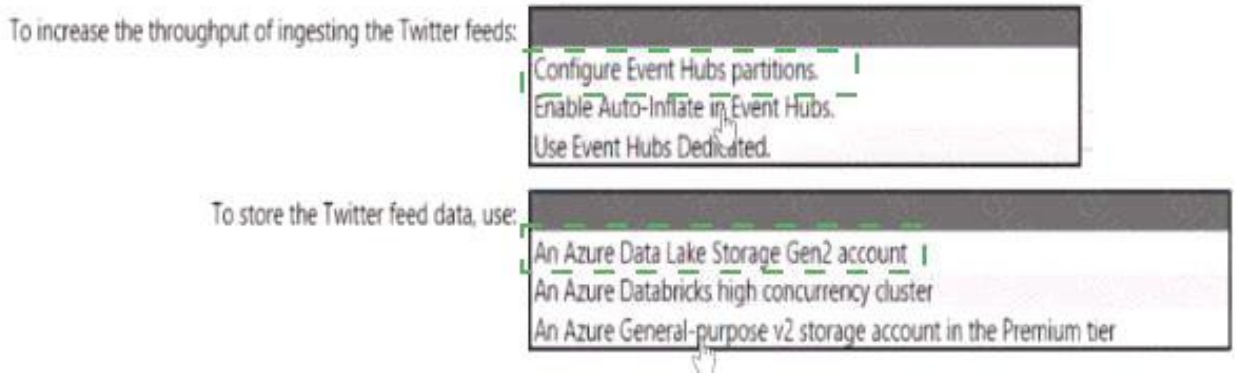
솔루션에 무엇을 포함해야 합니까? 대답하려면 대답 영역에서 적절한 옵션을 선택하십시오.
참고: 각 정답은 1점의 가치가 있습니다.

Answer Area

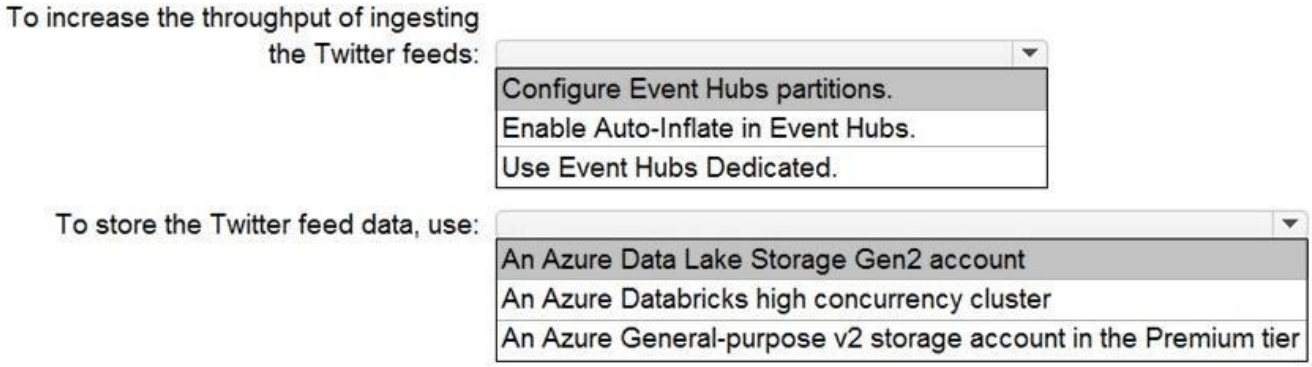


Answer:

Answer Area



Explanation:

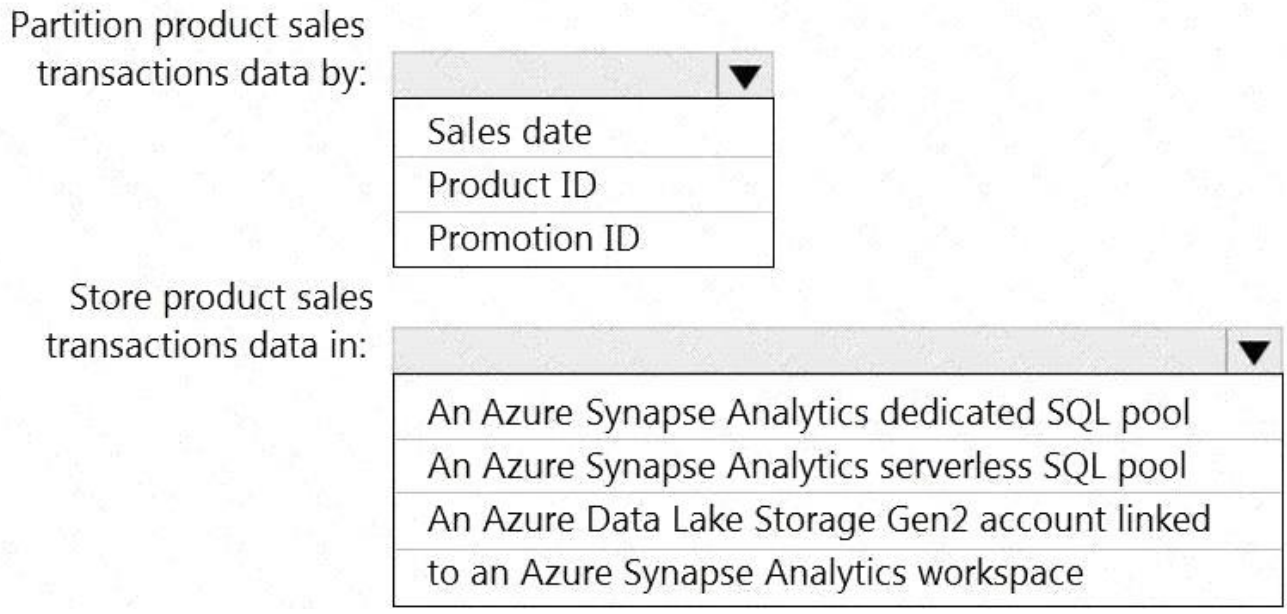


Box 1: Configure Evegent Hubs partitions
Scenario: Maximize the throughput of ingesting Twitter feeds from Event Hubs to Azure Storage without purchasing additional throughput or capacity units.
Event Hubs is designed to help with processing of large volumes of events. Event Hubs throughput is scaled by using partitions and throughput-unit allocations.
Event Hubs traffic is controlled by TUs (standard tier). Auto-inflate enables you to start small with the minimum required TUs you choose. The feature then scales automatically to the maximum limit of TUs you need, depending on the increase in your traffic.
Box 2: An Azure Data Lake Storage Gen2 account
Scenario: Ensure that the data store supports Azure AD-based access control down to the

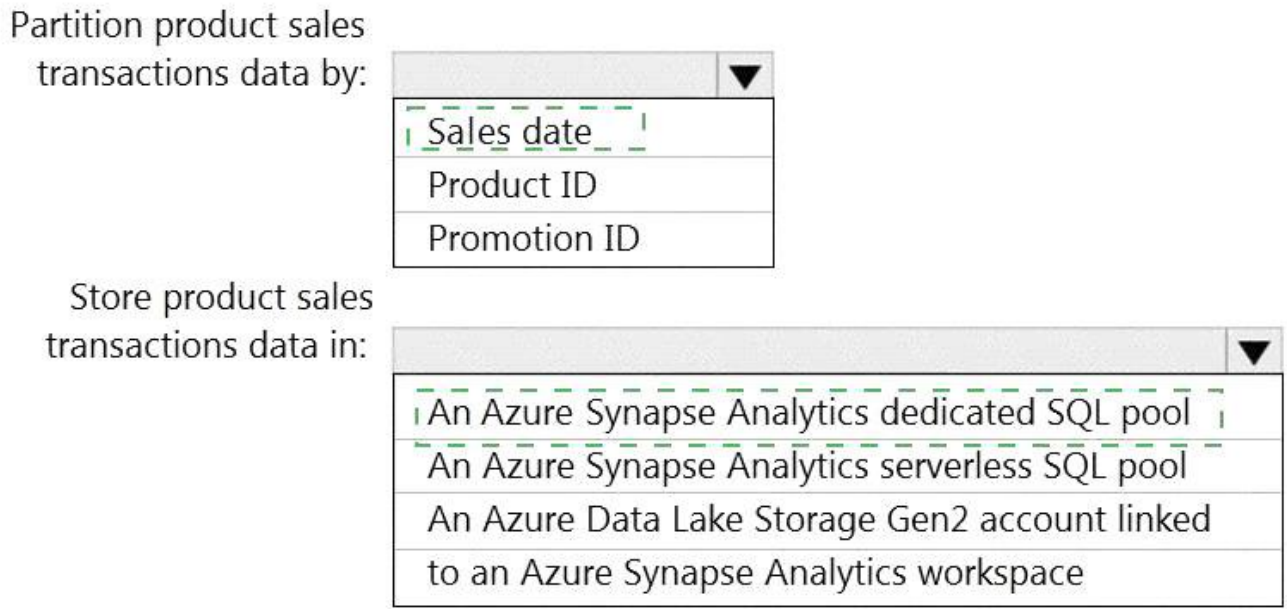
object level.
Azure Data Lake Storage Gen2 implements an access control model that supports both Azure role-based access control (Azure RBAC) and POSIX-like access control lists (ACLs).
Reference:
https://docs.microsoft.com/en-us/azure/event-hubs/event-hubs-features
https://docs.microsoft.com/en-us/azure/storage/blobs/data-lake-storage-access-control

QUESTION NO: 3
온-프레미스 데이터 원본과 Azure Synapse Analytics를 통합해야 합니다. 솔루션은 데이터 통합 요구 사항을 충족해야 합니다.
어떤 유형의 통합 런타임을 사용해야 하나요?
A. Azure-SSIS 통합 런타임
B. 자체 호스팅 통합 런타임
C. Azure 통합 런타임
Answer: C

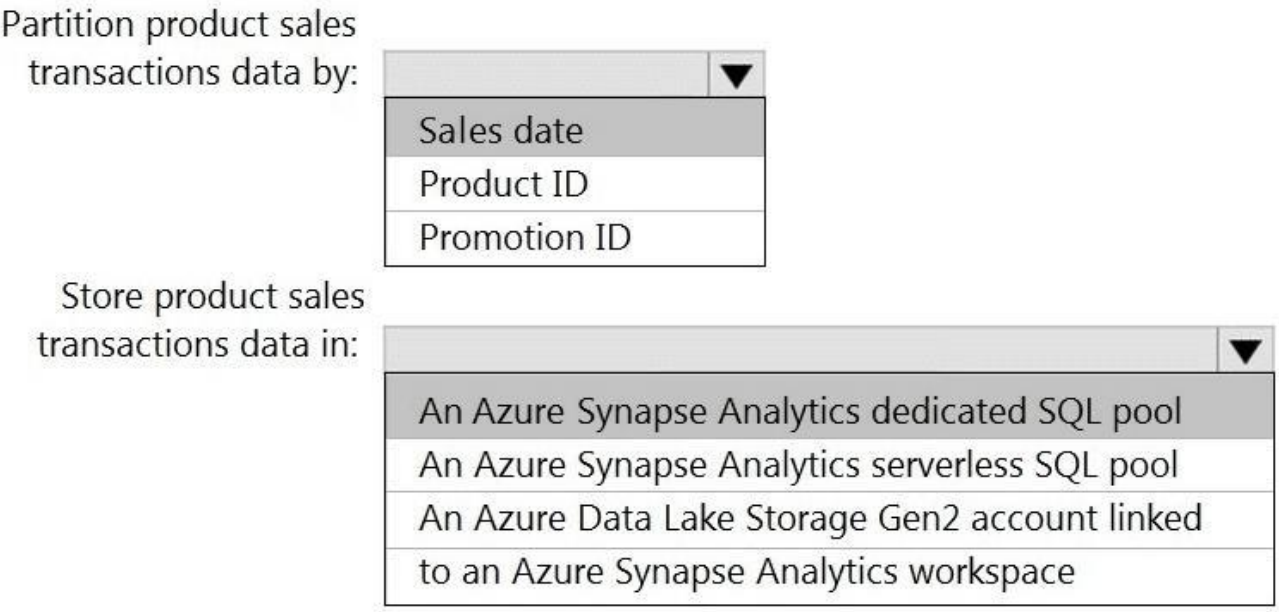
QUESTION NO: 4
제품 판매 거래를 위한 파티션을 설계해야 합니다. 솔루션은 판매 거래 데이터 세트 요구 사항을 충족해야 합니다.
솔루션에 무엇을 포함해야 합니까? 대답하려면 대답 영역에서 적절한 옵션을 선택하십시오.
참고: 각 올바른 선택은 1점의 가치가 있습니다.



Answer:



Explanation:



Box 1: Sales date
Scenario: Contoso requirements for data integration include:
Partition data that contains sales transaction records. Partitions must be designed to provide efficient loads by month. Boundary values must belong to the partition on the right.
Box 2: An Azure Synapse Analytics Dedicated SQL pool
Scenario: Contoso requirements for data integration include:
Ensure that data storage costs and performance are predictable.
The size of a dedicated SQL pool (formerly SQL DW) is determined by Data Warehousing Units (DWU).
Dedicated SQL pool (formerly SQL DW) stores data in relational tables with columnar storage. This format significantly reduces the data storage costs, and improves query performance.
Synapse analytics dedicated sql pool
Reference:

https://docs.microsoft.com/en-us/azure/synapse-analytics/sql-data-warehouse/sql-data-warehouse-overview- what-is

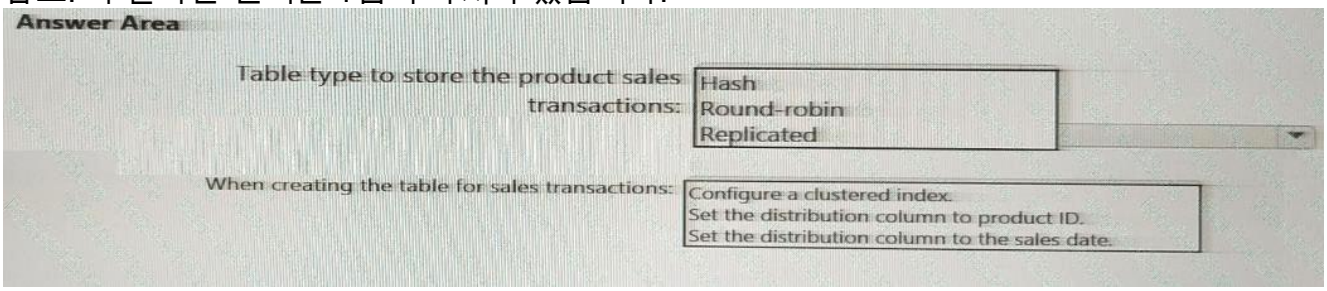
QUESTION NO: 5
Pool1이라는 이름의 Azure Synapse Analytics 전용 SQL 풀이 포함된 Azure 구독이 있습니다.
Azure Monitor를 사용합니다.
Pool1에서 실행된 쿼리의 성능을 모니터링해야 합니다.
어떤 로그를 쿼리해야 하나요?
A. SynapseSqlPoolWaits
B. SynapseSqlPoolSqlRequests
C. SynapseSqlPoolExecRequests
D. SynapseSqlPoolRequestSteps
Answer: C

QUESTION NO: 6
Twitter ted 데이터 레코드에 대한 데이터 보존 솔루션을 설계해야 합니다. 솔루션은 고객 감정 분석 요구 사항을 충족해야 합니다.
솔루션에 어떤 Azure Storage 기능을 포함해야 합니까?
A. 시간 기반 보존
B. 피드 변경
C. 일시 삭제
D. 라이프사이클 관리
Answer: C

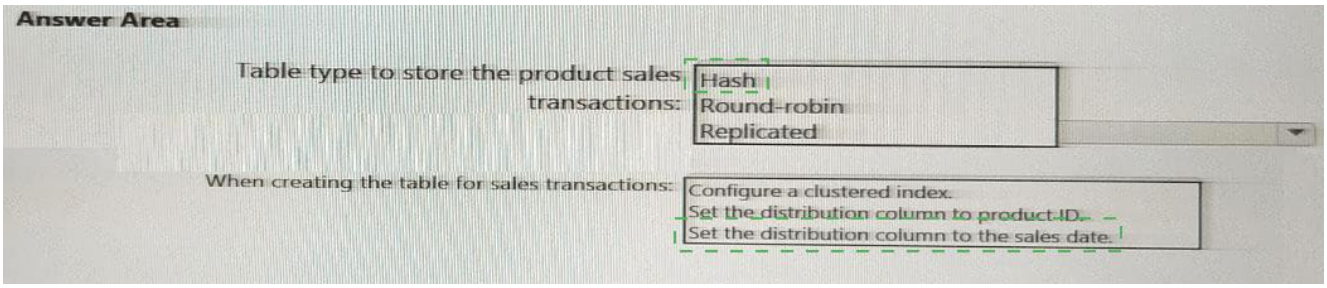
QUESTION NO: 7
소매점 테이블에 대리 키를 구현해야 합니다. 솔루션은 판매 거래 데이터 세트 요구 사항을 충족해야 합니다.
무엇을 만들어야 합니까?
A. IDENTITY 속성이 있는 테이블
B. 시스템 버전 임시 테이블
C. 사용자 정의 SEQUENCE 객체
D. FOREIGN KEY 제약 조건이 있는 테이블
Answer: A
Explanation:
Scenario: Implement a surrogate key to account for changes to the retail store addresses.
A surrogate key on a table is a column with a unique identifier for each row. The key is not generated from the table data. Data modelers like to create surrogate keys on their tables when they design data warehouse models. You can use the IDENTITY property to achieve this goal simply and effectively without affecting load performance.
Reference:
https://docs.microsoft.com/en-us/azure/synapse-analytics/sql-data-warehouse/sql-data-warehouse-tables- identity

QUESTION NO: 8

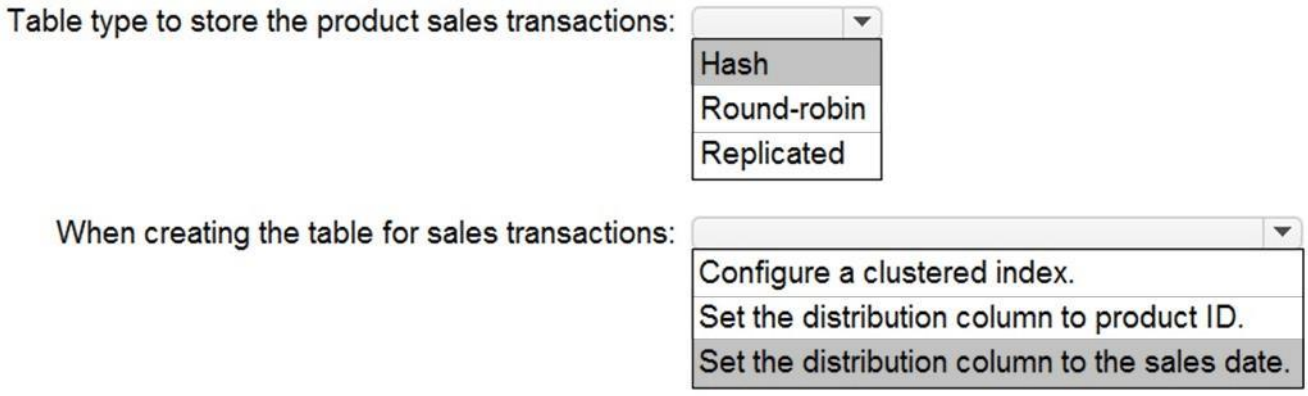
상품 판매 거래를 위한 데이터 저장 구조를 설계해야 합니다. 솔루션은 판매 거래 데이터 세트 요구 사항을 충족해야 합니다.
솔루션에 무엇을 포함해야 합니까? 대답하려면 대답 영역에서 적절한 옵션을 선택하십시오.
참고: 각 올바른 선택은 1점의 가치가 있습니다.



Answer:



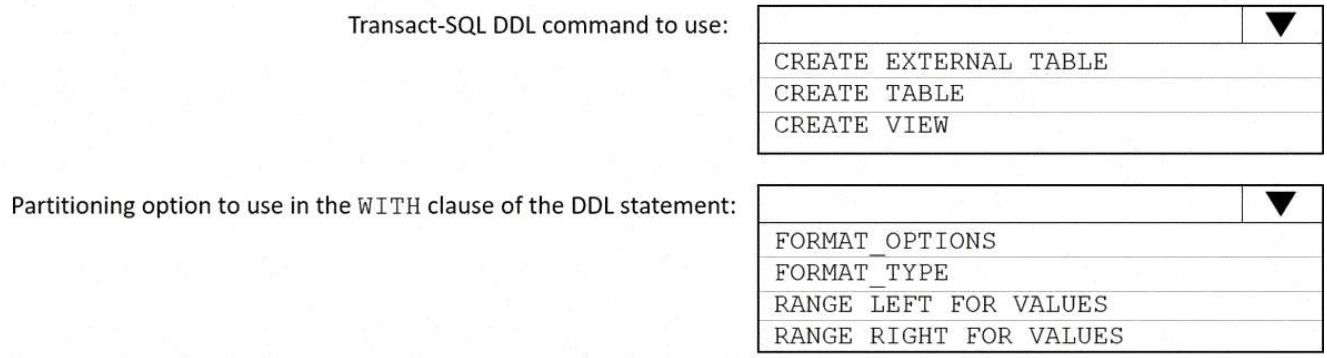
Explanation:



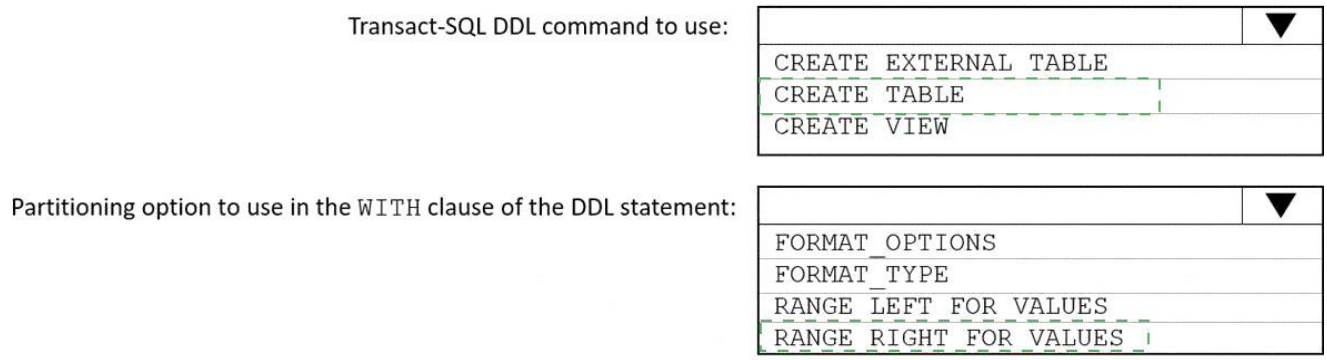
Box 1: Hash
Scenario:
Ensure that queries joining and filtering sales transaction records based on product ID complete as quickly as possible.
A hash distributed table can deliver the highest query performance for joins and aggregations on large tables.
Box 2: Set the distribution column to the sales date.
Scenario: Partition data that contains sales transaction records. Partitions must be designed to provide efficient loads by month. Boundary values must belong to the partition on the right.
Reference:
https://rajanieshkaushikk.com/2020/09/09/how-to-choose-right-data-distribution-strategy-for-azure-synapse/

QUESTION NO: 9
판매 트랜잭션 데이터를 저장하기 위해 Azure Synapse Analytics 데이터베이스 개체를

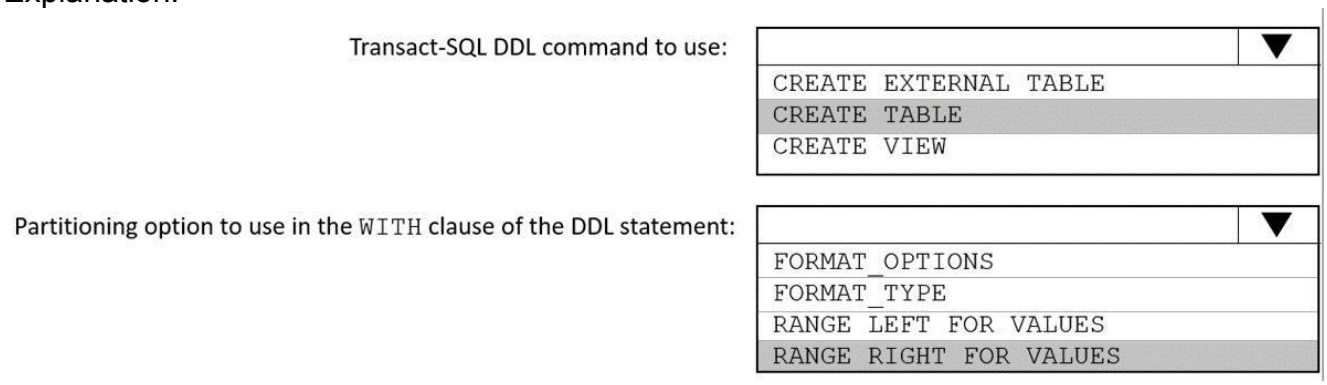
구현해야 합니다. 솔루션은 판매 거래 데이터 세트 요구 사항을 충족해야 합니다.
판매 트랜잭션 데이터 세트 요구 사항을 충족해야 하는 솔루션은 무엇입니까?
어떻게 해야 합니까? 대답하려면 대답 영역에서 적절한 옵션을 선택하십시오.
참고: 각 올바른 선택은 1점의 가치가 있습니다.



Answer:



Explanation:



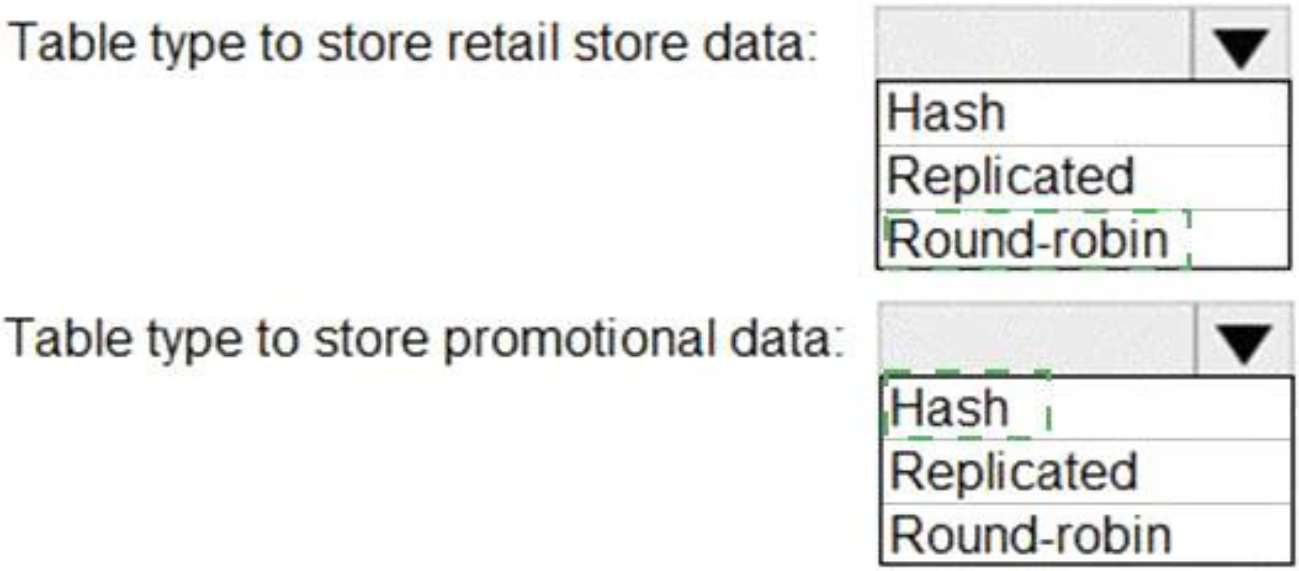
Box 1: Create table
Scenario: Load the sales transaction dataset to Azure Synapse Analytics Box 2: RANGE RIGHT FOR VALUES Scenario: Partition data that contains sales transaction records. Partitions must be designed to provide efficient loads by month. Boundary values must belong to the partition on the right.
RANGE RIGHT: Specifies the boundary value belongs to the partition on the right (higher values).
FOR VALUES (boundary_value [,...n]): Specifies the boundary values for the partition.
Scenario: Load the sales transaction dataset to Azure Synapse Analytics.
Contoso identifies the following requirements for the sales transaction dataset:
Partition data that contains sales transaction records. Partitions must be designed to provide efficient loads by month. Boundary values must belong to the partition on the right.

Ensure that queries joining and filtering sales transaction records based on product ID complete as quickly as possible.
Implement a surrogate key to account for changes to the retail store addresses.
Ensure that data storage costs and performance are predictable.
Minimize how long it takes to remove old records.
Reference:
https://docs.microsoft.com/en-us/sql/t-sql/statements/create-table-azure-sql-data-warehouse

QUESTION NO: 10
Twitter 피드 데이터 레코드에 대한 데이터 보존 솔루션을 설계해야 합니다. 솔루션은 고객 감정 분석 요구 사항을 충족해야 합니다.
솔루션에 어떤 Azure Storage 기능을 포함해야 합니까?
A. 피드 변경
B. 일시 삭제
C. 시간 기반 보존
D. 라이프사이클 관리
Answer: B
Explanation:
Scenario: Purge Twitter feed data records that are older than two years.
Data sets have unique lifecycles. Early in the lifecycle, people access some data often. But the need for access often drops drastically as the data ages. Some data remains idle in the cloud and is rarely accessed once stored. Some data sets expire days or months after creation, while other data sets are actively read and modified throughout their lifetimes. Azure Storage lifecycle management offers a rule-based policy that you can use to transition blob data to the appropriate access tiers or to expire data at the end of the data lifecycle.
Reference:
https://docs.microsoft.com/en-us/azure/storage/blobs/lifecycle-management-overview

QUESTION NO: 11
트랜잭션 데이터 a에 대한 분석 스토리지 솔루션을 설계해야 합니다. 솔루션은 판매 거래 데이터 세트 요구 사항을 충족해야 합니다.
솔루션에 무엇을 포함해야 합니까? 대답하려면 대답 영역에서 적절한 옵션을 선택하십시오.
참고: 각 올바른 선택은 1점의 가치가 있습니다.

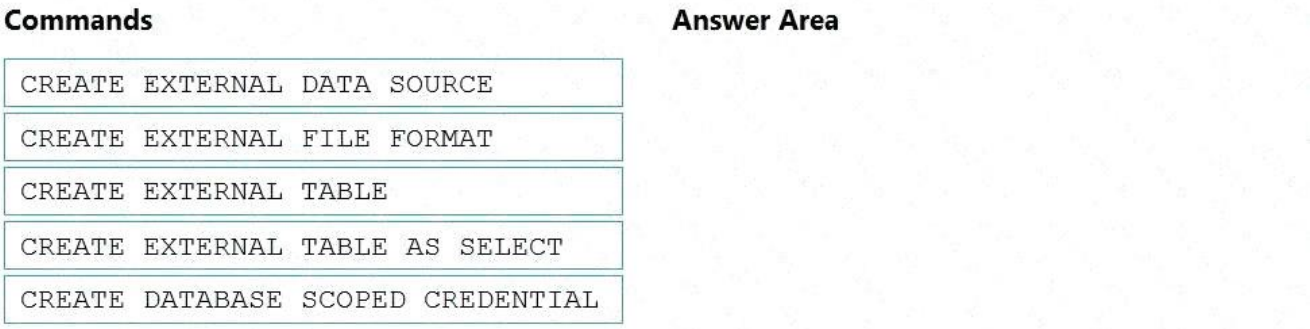
Answer:



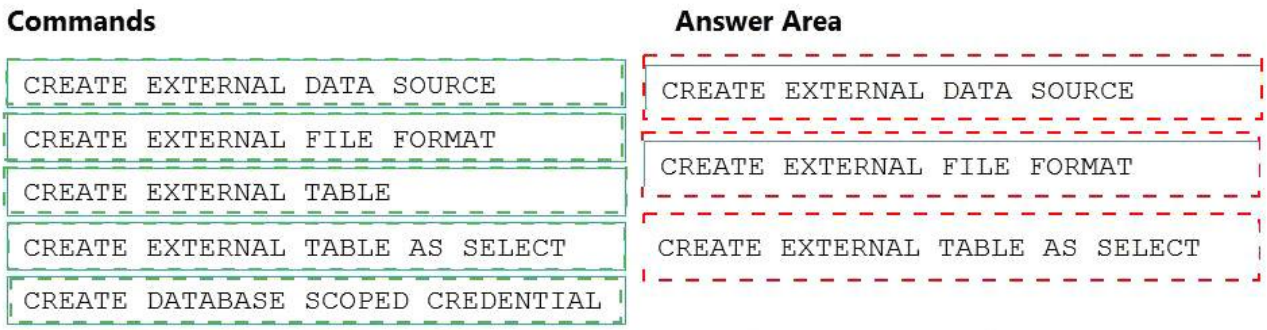
Explanation:

Box 1: Round-robin
Round-robin tables are useful for improving loading speed.
Scenario: Partition data that contains sales transaction records. Partitions must be designed to provide efficient loads by month.
Box 2: Hash
Hash-distributed tables improve query performance on large fact tables.
Reference:
https://docs.microsoft.com/en-us/azure/synapse-analytics/sql-data-warehouse/sql-data-warehouse-tables-distribute

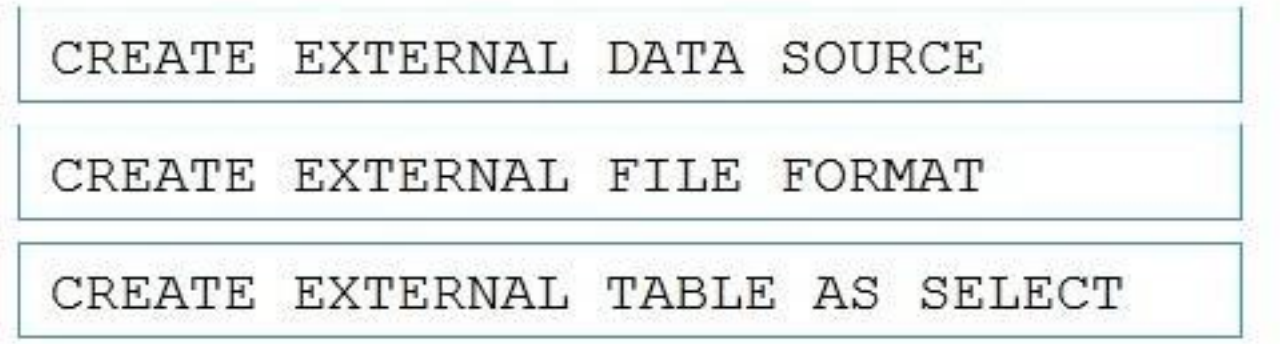
QUESTION NO: 12
전용 SQL 풀에서 Twitter 피드 데이터를 분석할 수 있는지 확인해야 합니다. 솔루션은 고객 감정 분석 요구 사항을 충족해야 합니다.
어떤 세 가지 Transaction-SQL DDL 명령을 순서대로 실행해야 합니까? 응답하려면 명령 목록에서 해당 명령을 응답 영역으로 이동하고 올바른 순서로 정렬하십시오.
참고: 하나 이상의 답변 선택 순서가 정확합니다. 선택한 올바른 주문에 대해 크레딧을 받게 됩니다.



Answer:



Explanation:



Scenario: Allow Contoso users to use PolyBase in an Azure Synapse Analytics dedicated SQL pool to query the content of the data records that host the Twitter feeds. Data must be protected by using row-level security (RLS). The users must be authenticated by using their own Azure AD credentials.
Box 1: CREATE EXTERNAL DATA SOURCE
External data sources are used to connect to storage accounts.
Box 2: CREATE EXTERNAL FILE FORMAT
CREATE EXTERNAL FILE FORMAT creates an external file format object that defines external data stored in Azure Blob Storage or Azure Data Lake Storage. Creating an external file format is a prerequisite for creating an external table.
Box 3: CREATE EXTERNAL TABLE AS SELECT
When used in conjunction with the CREATE TABLE AS SELECT statement, selecting from an external table imports data into a table within the SQL pool. In addition to the COPY statement, external tables are useful for loading data.
Reference:
https://docs.microsoft.com/en-us/azure/synapse-analytics/sql/develop-tables-external-tables

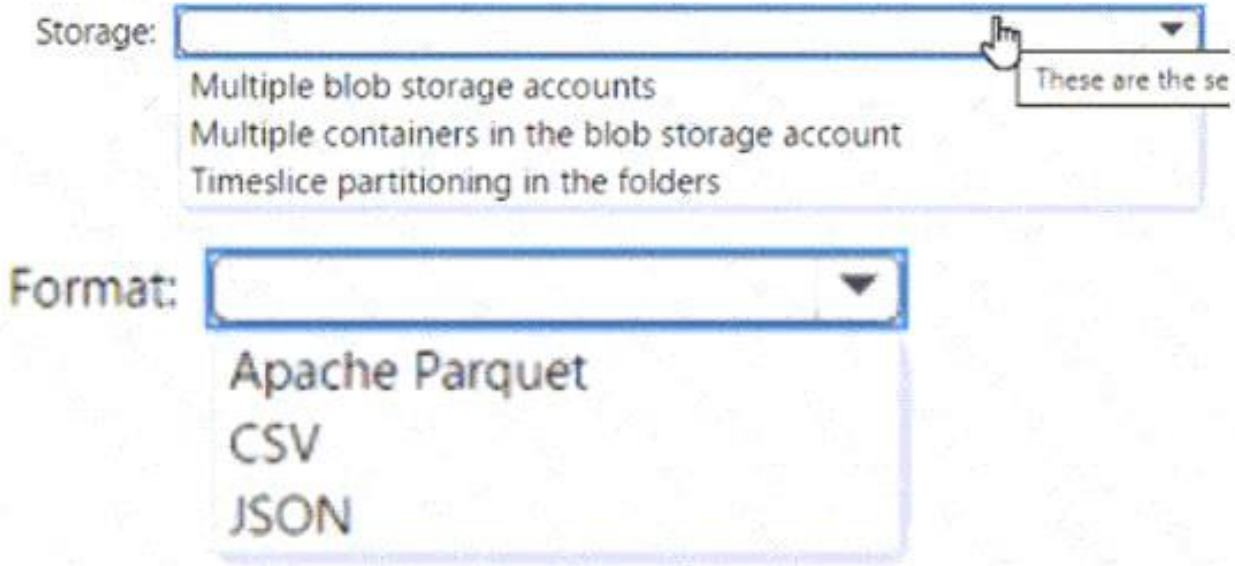
QUESTION NO: 13
Pcol1이라는 Azure Synapse Analytics 전용 SQL 풀이 있습니다. Pool1에는 tablet이라는 테이블이 있습니다. table1에 5TB의 데이터를 로드합니다.
table1에 대해 열 저장소 압축이 최대화되었는지 확인해야 합니다.
어떤 문장을 실행해야 할까요?

A. DBCC IIDEXDEFRAG (pool1, table1)
B. ALTER INDEX ALL on table REORGANIZE
C. DBCC DBREINOEX (table)
D. ALTER INDEX ALL on table REBUILD
Answer: C

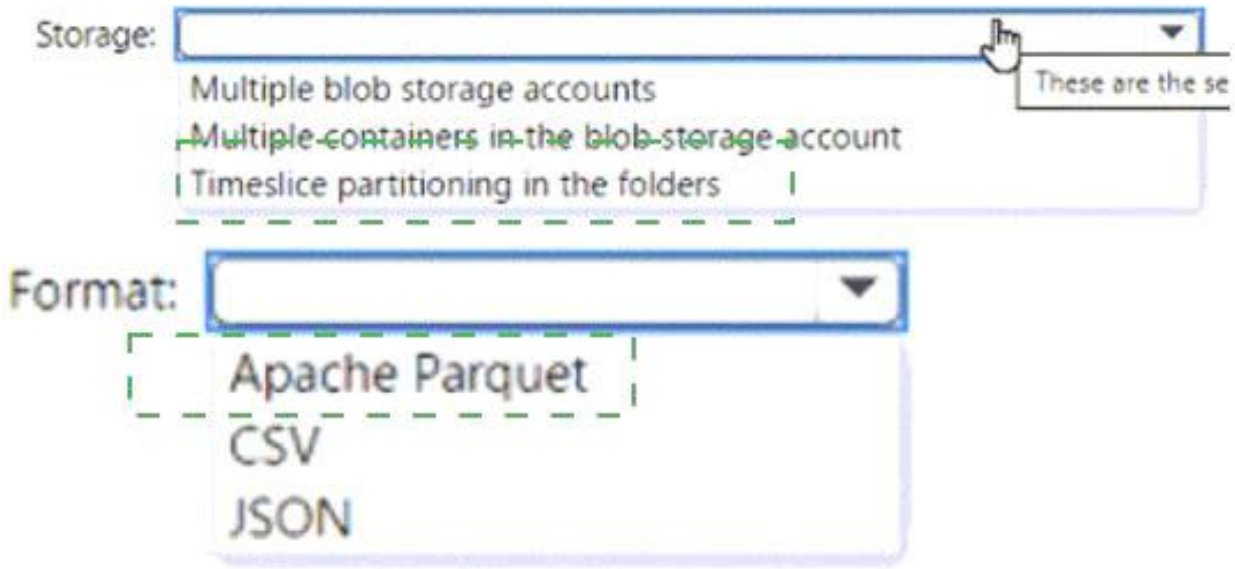
Topic 2, Litware, inc. Case Study
Case study
This is a case study. Case studies are not timed separately. You can use as much exam time as you would like to complete each case. However, there may be additional case studies and sections on this exam. You must manage your time to ensure that you are able to complete all questions included on this exam in the time provided.
To answer the questions included in a case study, you will need to reference information that is provided in the case study. Case studies might contain exhibits and other resources that provide more information about the scenario that is described in the case study. Each question is independent of the other questions in this case study.
At the end of this case study, a review screen will appear. This screen allows you to review your answers and to make changes before you move to the next section of the exam. After you begin a new section, you cannot return to this section.
To start the case study
To display the first question in this case study, click the Next button. Use the buttons in the left pane to explore the content of the case study before you answer the questions. Clicking these buttons displays information such as business requirements, existing environment, and problem statements. If the case study has an All Information tab, note that the information displayed is identical to the information displayed on the subsequent tabs. When you are ready to answer a question, click the Question button to return to the question.
Overview
Litware, Inc. owns and operates 300 convenience stores across the US. The company sells a variety of packaged foods and drinks, as well as a variety of prepared foods, such as sandwiches and pizzas.
Litware has a loyalty club whereby members can get daily discounts on specific items by providing their membership number at checkout.
Litware employs business analysts who prefer to analyze data by using Microsoft Power BI, and data scientists who prefer analyzing data in Azure Databricks notebooks.
Requirements
Business Goals
Litware wants to create a new analytics environment in Azure to meet the following requirements:
See inventory levels across the stores. Data must be updated as close to real time as possible.
Execute ad hoc analytical queries on historical data to identify whether the loyalty club discounts increase sales of the discounted products.
Every four hours, notify store employees about how many prepared food items to produce based on historical demand from the sales data.
Technical Requirements
Litware identifies the following technical requirements:
Minimize the number of different Azure services needed to achieve the business goals.
Use platform as a service (PaaS) offerings whenever possible and avoid having to provision virtual machines that must be managed by Litware.
Ensure that the analytical data store is accessible only to the company's on-premises network and Azure services.

Use Azure Active Directory (Azure AD) authentication whenever possible.
Use the principle of least privilege when designing security.
Stage Inventory data in Azure Data Lake Storage Gen2 before loading the data into the analytical data store.
Litware wants to remove transient data from Data Lake Storage once the data is no longer in use. Files that have a modified date that is older than 14 days must be removed.
Limit the business analysts' access to customer contact information, such as phone numbers, because this type of data is not analytically relevant.
Ensure that you can quickly restore a copy of the analytical data store within one hour in the event of corruption or accidental deletion.
Planned Environment
Litware plans to implement the following environment:
The application development team will create an Azure event hub to receive real-time sales data, including store number, date, time, product ID, customer loyalty number, price, and discount amount, from the point of sale (POS) system and output the data to data storage in Azure.
Customer data, including name, contact information, and loyalty number, comes from Salesforce, a SaaS application, and can be imported into Azure once every eight hours. Row modified dates are not trusted in the source table.
Product data, including product ID, name, and category, comes from Salesforce and can be imported into Azure once every eight hours. Row modified dates are not trusted in the source table.
Daily inventory data comes from a Microsoft SQL server located on a private network.
Litware currently has 5 TB of historical sales data and 100 GB of customer data. The company expects approximately 100 GB of new data per month for the next year.
Litware will build a custom application named FoodPrep to provide store employees with the calculation results of how many prepared food items to produce every four hours.
Litware does not plan to implement Azure ExpressRoute or a VPN between the on-premises network and Azure.

QUESTION NO: 14
폴더가 포함된 Azure Blob 저장소 계정이 있습니다. 폴더에는 120,000개의 파일이 있습니다.
각 파일에는 62개의 열이 있습니다.
매일 1,500개의 새 파일이 폴더에 추가됩니다.
각각의 새 파일에서 Azure Synapse Analytics 작업 영역으로 5개의 데이터 열을 증분식으로 로드할 계획입니다.
증분 로드를 수행하는 데 걸리는 시간을 최소화해야 합니다.
파일과 형식을 저장하기 위해 무엇을 사용해야 합니까?



Answer:



Explanation:
Box 1 = timeslice partitioning in the foldersThis means that you should organize your files into folders based on a time attribute, such as year, month, day, or hour. For example, you can have a folder structure like /yyyy
/mm/dd/file.csv. This way, you can easily identify and load only the new files that are added each day by using a time filter in your Azure Synapse pipeline12. Timeslice partitioning can also improve the performance of data loading and querying by reducing the number of files that need to be scanned Box = 2 Apache Parquet This is because Parquet is a columnar file format that can efficiently store and compress data with many columns. Parquet files can also be partitioned by a time attribute, which can improve the performance of incremental loading

and querying by reducing the number of files that need to be scanned123. Parquet files are supported by both dedicated SQL pool and serverless SQL pool in Azure Synapse Analytics2.

QUESTION NO: 15
Apache Kafka에서 시작되고 Azure Data Lake Storage Gen2로 출력되는 스트리밍 데이터를 집계하는 솔루션을 계획하고 있습니다. 스트림 처리 솔루션을 구현할 개발자는 Java를 사용합니다. 스트리밍 데이터를 처리하려면 어떤 서비스를 사용하는 것이 좋습니까?
A. Azure 데이터 팩터리
B. Azure 스트림 분석
C. Azure Databricks
D. Azure 이벤트 허브
Answer: C
Explanation:
https://docs.microsoft.com/en-us/azure/architecture/data-guide/technology-choices/stream-processing

QUESTION NO: 16
Azure 데이터 팩토리가 포함된 Azure 구독이 있습니다.
Azure Data Factory 활동 JSON을 편집하고 있습니다.
스크립트는 Azure Blob Storage에서 여러 대상으로 파일을 복사해야 합니다. 솔루션은 소스 및 대상 파일에 일관된 폴더 경로가 있는지 확인해야 합니다.
스크립트를 어떻게 완성해야 합니까? 대답하려면 적절한 값을 올바른 타겟으로 드래그합니다. 각 값은 한 번, 여러 번 또는 전혀 사용되지 않을 수 있습니다. 패널 사이의 분할 막대를 드래그하거나 스크롤하여 콘텐츠를 볼 수 있습니다.
참고: 정답은 1점입니다.

Answer:

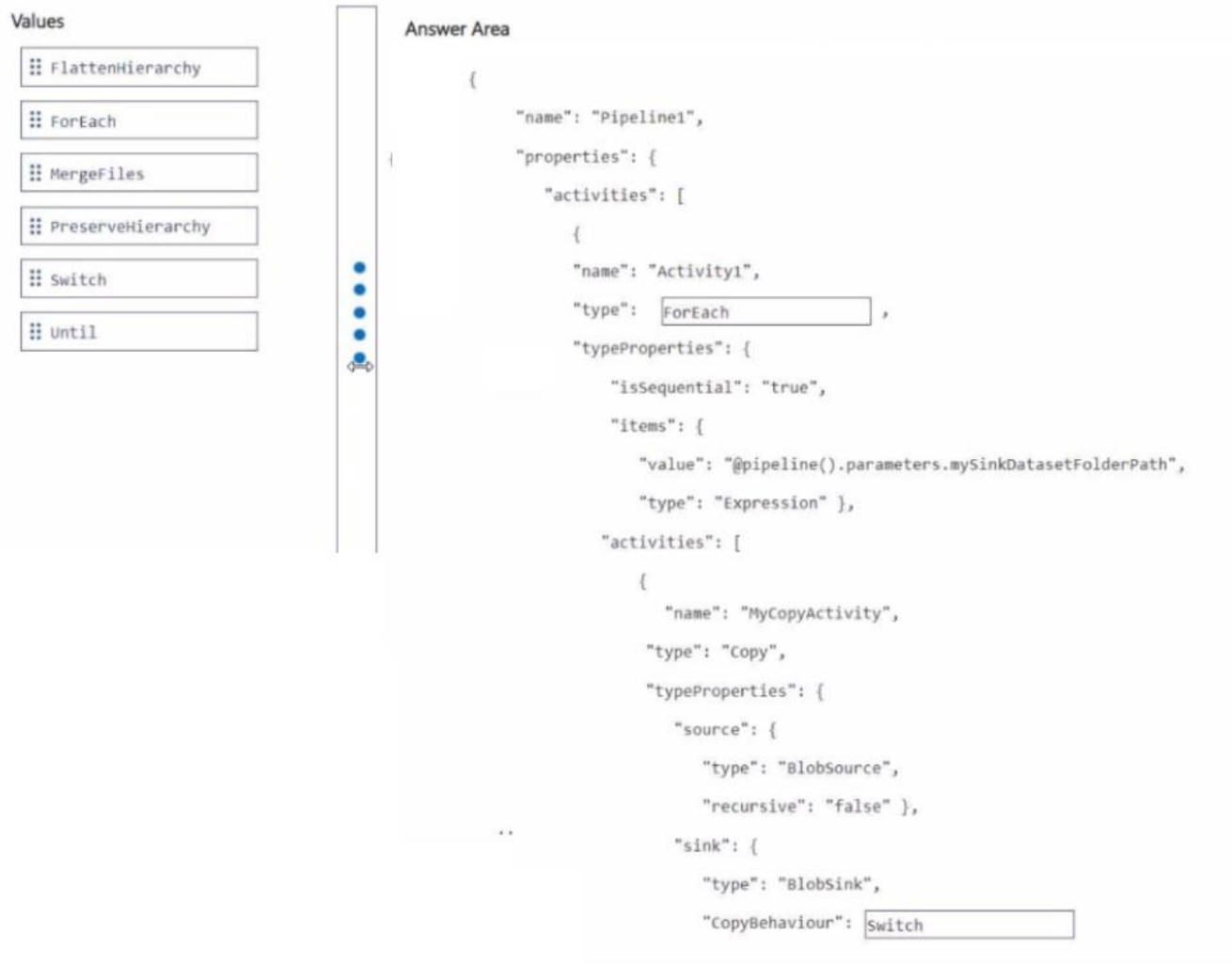
Values

- FlattenHierarchy
- ForEach
- MergeFiles
- PreserveHierarchy
- Switch
- Until

Answer Area

```
{
    "name": "Pipeline1",
    "properties": {
        "activities": [
            {
            "name": "Activity1",
            "type": ForEach,
            "typeProperties": {
                        "isSequential": "true",
                        "items": {
                             "value": "@pipeline
            ().parameters.mySinkDatasetFolderPath",
                             "type": "Expression"},
                        "activities" [
                             {
                                  "name": "MyCopyActivity",
                               "type": "Copy",
                               "typeProperties": {
                                  "source": {
                                       "type": "BlobSource",
                                       "recursive": "false" },
                                  "sink": {
                                       "type": "BlobSink",
                                       "CopyBehavior": Switch
                             …
                             }
```

Explanation:



QUESTION NO: 17
Litware 온프레미스 네트워크 외부의 사용자가 분석 데이터 저장소에 액세스하지 못하도록 하려면 무엇을 권장해야 합니까?
A. 서버 수준 가상 네트워크 규칙
B. 데이터베이스 수준 가상 네트워크 규칙
C. 데이터베이스 수준 방화벽 IP 규칙
D. 서버 수준 방화벽 IP 규칙
Answer: A
Explanation:
Virtual network rules are one firewall security feature that controls whether the database server for your single databases and elastic pool in Azure SQL Database or for your databases in SQL Data Warehouse accepts communications that are sent from particular subnets in virtual networks.
Server-level, not database-level: Each virtual network rule applies to your whole Azure SQL Database server, not just to one particular database on the server. In other words, virtual network rule applies at the serverlevel, not at the database-level.
References:
https://docs.microsoft.com/en-us/azure/sql-database/sql-database-vnet-service-endpoint-rule-overview

QUESTION NO: 18
Azure Synapse Analytics에서 엔터프라이즈 데이터 웨어하우스를 구현합니다.
크기가 10테라바이트(TB)인 대형 팩트 테이블이 있습니다.
들어오는 쿼리는 기본 키 SaleKey 열을 사용하여 다음 표에 표시된 대로 데이터를 검색합니다.

SaleKey	CityKey	CustomerKey	StockItemKey	InvoiceDateKey	Quantity	UnitPrice	TotalExcludingTax
49309	90858	70	69	10/22/13	8	16	128
49313	55710	126	69	10/22/13	2	16	32
49343	44710	234	68	10/22/13	10	16	160
49352	66109	163	70	10/22/13	4	16	64
49488	65312	230	70	10/22/13	8	16	128
49646	85877	271	70	10/24/13	1	16	16
49798	41238	288	69	10/24/13	1	16	16

테이블의 성능을 최적화하려면 대형 팩트 테이블을 여러 노드에 분산해야 합니다.
어떤 기술을 사용해야 합니까?
A. 클러스터형 인덱스가 있는 해시 분산 테이블
B. 클러스터형 Columnstore 인덱스가 있는 해시 분산 테이블
C. 클러스터형 인덱스가 있는 라운드 로빈 분산 테이블
D. 클러스터형 Columnstore 인덱스가 있는 라운드 로빈 분산 테이블
E. 배포 복제가 있는 힙 테이블
Answer: B
Explanation:
Hash-distributed tables improve query performance on large fact tables.
Columnstore indexes can achieve up to 100x better performance on analytics and data warehousing workloads and up to 10x better data compression than traditional rowstore indexes.
Reference:
https://docs.microsoft.com/en-us/azure/sql-data-warehouse/sql-data-warehouse-tables-distribute
https://docs.microsoft.com/en-us/sql/relational-databases/indexes/columnstore-indexes-query-performance

QUESTION NO: 19
Azure Cosmos DB 분석 저장소와 WS 1이라는 Azure Synapse Analytics 작업 영역을 포함하는 Azure 구독이 있습니다. WS1에는 서버리스 SQL 풀 이름 Pool1이 있습니다.
Pool1을 사용하여 다음 쿼리를 실행합니다.

```
WITH IDENTITY = 'SHARED A
SECRET = 'fed43474798724234335636534563456345ddfa==';

SELECT clientID AS ClientID,
       client AS ClientName
FROM OPENROWSET
(
    PROVIDER = 'CosmosDB',
    CONNECTION = 'Account=account1;Database=database1',
                 OBJECT = 'clients',
                 SERVER_CREDENTIAL = 'AccountCred'
)
WITH

(
    clientID int,
    client varchar(50),
    streetAddress varchar(100)
) AS c;
```

다음 각 진술에 대해 해당 진술이 참이면 예를 선택하십시오. 그렇지 않으면 아니요를 선택합니다.
참고: 올바른 선택은 각각 1점의 가치가 있습니다.

Answer Area

Statements	Yes	No
The query returns three columns.	○	○
The container being queried is named `clients`.	○	○
Authentication is performed by using an account key.	○	○

Answer:

Answer Area

Statements	Yes	No
The query returns three columns.	○	○
The container being queried is named `clients`.	○	○
Authentication is performed by using an account key.	○	○

Explanation:

Answer Area

Statements	Yes	No
The query returns three columns.	○	●
The container being queried is named `clients`.	●	○
Authentication is performed by using an account key.	○	●

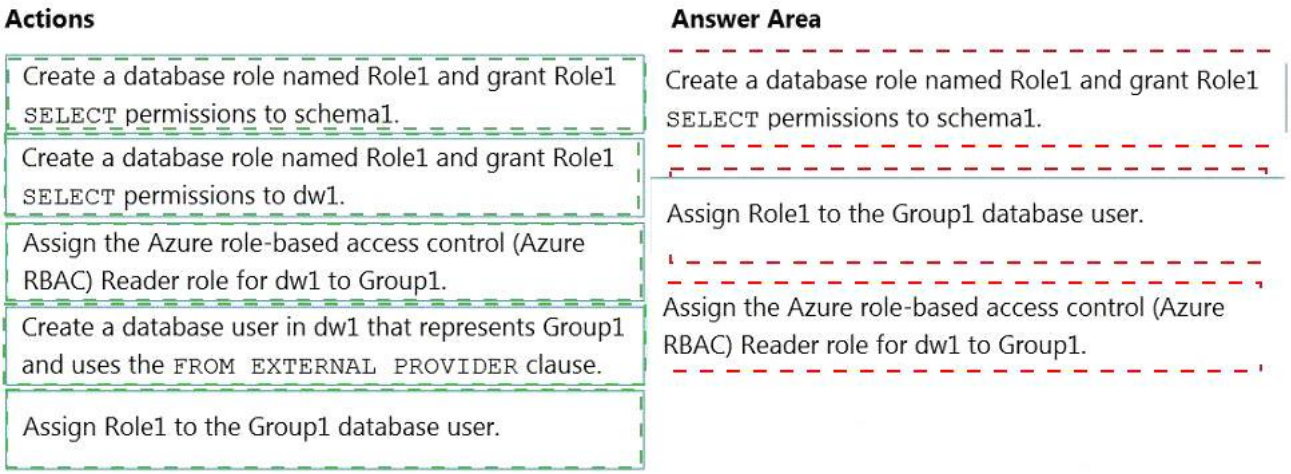
QUESTION NO: 20
Group1이라는 보안 그룹을 포함하는 Azure AD(Azure Active Directory) 테넌트가 있습니다.
schema1이라는 스키마를 포함하는 dw1이라는 Azure Synapse Analytics 전용 SQL 풀이 있습니다.
schema1의 모든 테이블과 보기에 Group1 읽기 전용 권한을 부여해야 합니다. 솔루션은 최소 권한 원칙을 사용해야 합니다.
어떤 세 가지 작업을 순서대로 수행해야 합니까? 응답하려면 작업 목록에서 해당 작업을 응답 영역으로 이동하고 올바른 순서로 정렬하십시오.
참고: 하나 이상의 답변 선택 순서가 정확합니다. 선택한 올바른 주문에 대해 크레딧을 받게 됩니다.

Actions

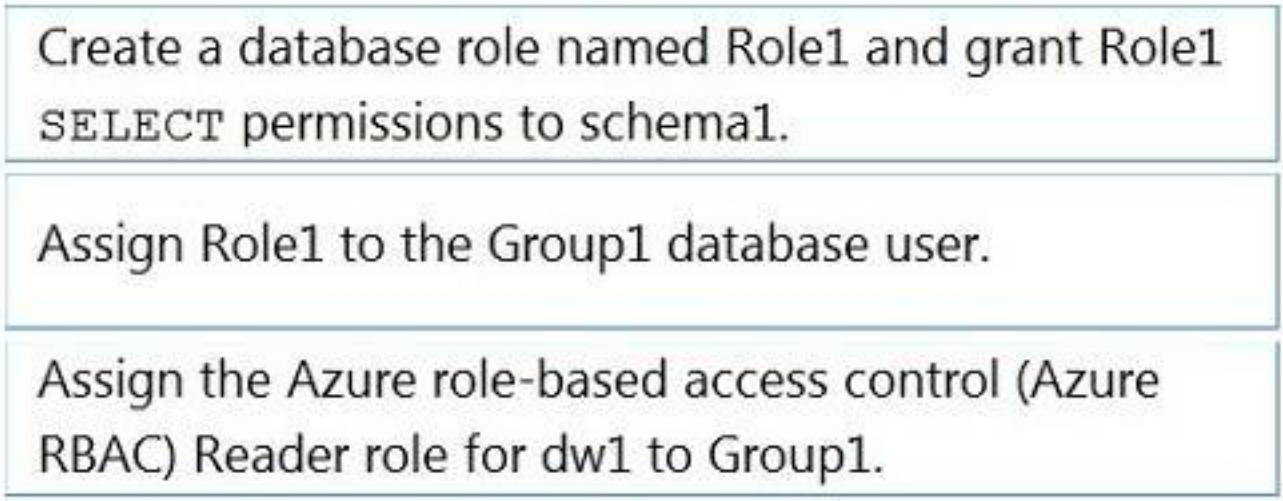
- Create a database role named Role1 and grant Role1 `SELECT` permissions to schema1.
- Create a database role named Role1 and grant Role1 `SELECT` permissions to dw1.
- Assign the Azure role-based access control (Azure RBAC) Reader role for dw1 to Group1.
- Create a database user in dw1 that represents Group1 and uses the `FROM EXTERNAL PROVIDER` clause.
- Assign Role1 to the Group1 database user.

Answer Area

Answer:



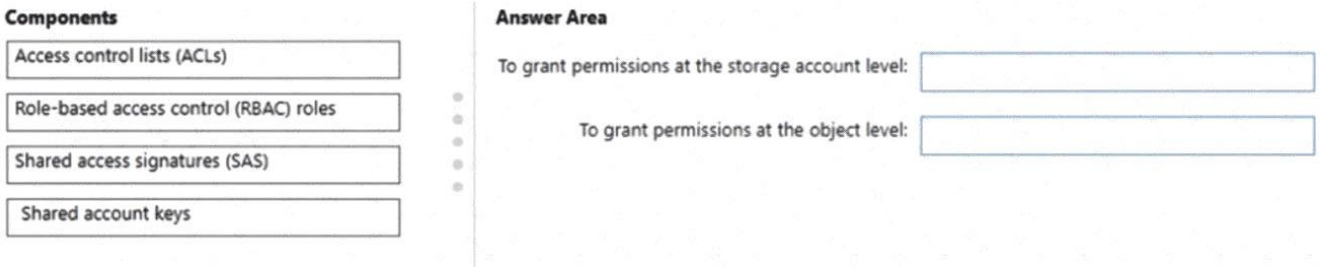
Explanation:



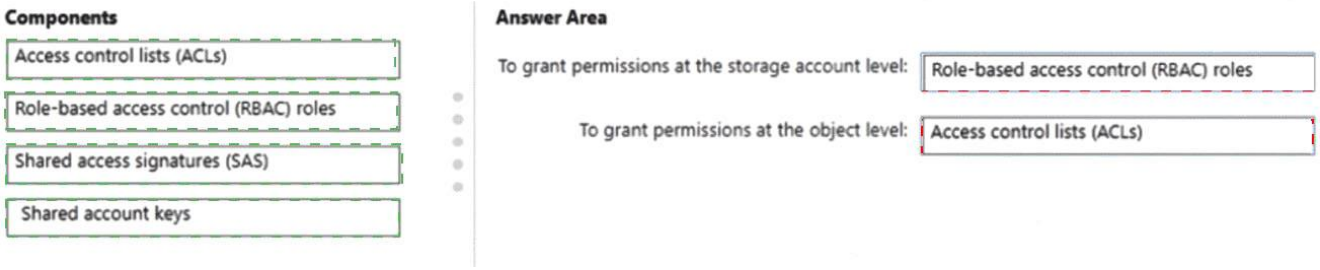
Step 1: Create a database role named Role1 and grant Role1 SELECT permissions to schema You need to grant Group1 read-only permissions to all the tables and views in schema1.
Place one or more database users into a database role and then assign permissions to the database role.
Step 2: Assign Rol1 to the Group database user
Step 3: Assign the Azure role-based access control (Azure RBAC) Reader role for dw1 to Group1 Reference:
https://docs.microsoft.com/en-us/azure/data-share/how-to-share-from-sql

QUESTION NO: 21
storage1이라는 Azure Data Lake Storage Gen 2 계정이 있습니다.
storage1의 콘텐츠에 액세스하기 위한 솔루션을 권장해야 합니다. 솔루션은 다음 요구 사항을 충족해야 합니다.
스토리지 계정 수준에서 나열 및 읽기 권한을 부여해야 합니다.
storage1의 개별 개체에 추가 권한을 적용할 수 있습니다.
Microsoft Entra의 일부인 Microsoft Azure Active Directory(Azure AD)의 보안 주체를 인증에 사용해야 합니다.
무엇을 사용해야 합니까? 대답하려면 적절한 구성 요소를 올바른 요구 사항으로 드래그하십시오. 각 구성 요소는 한 번, 두 번 이상 사용하거나 전혀 사용하지 않을 수

있습니다. 콘텐츠를 보려면 창 사이의 분할 막대를 끌거나 스크롤해야 할 수 있습니다.
참고: 각 올바른 선택은 1점의 가치가 있습니다.



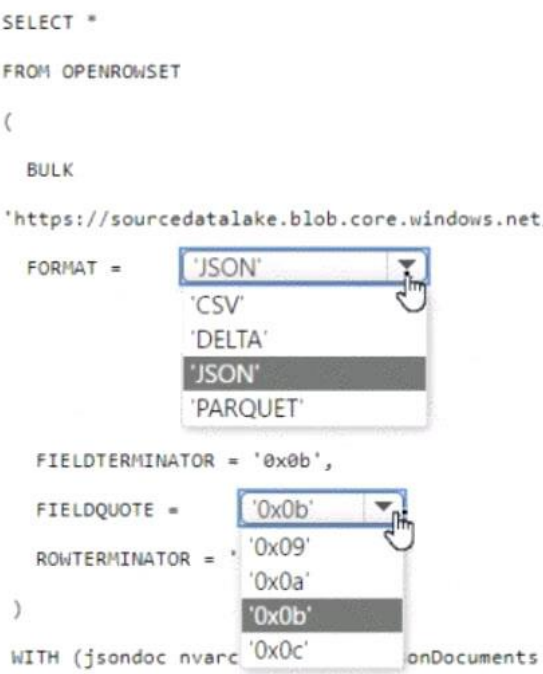
Answer:



Explanation:
Box 1: Role-based access control (RBAC) roles
List and read permissions must be granted at the storage account level.
Security principals from Microsoft Azure Active Directory (Azure AD), part of Microsoft Entra, must be used for authentication.
Role-based access control (Azure RBAC)
Azure RBAC uses role assignments to apply sets of permissions to security principals. A security principal is an object that represents a user, group, service principal, or managed identity that is defined in Azure Active Directory (AD). A permission set can give a security principal a "coarse-grain" level of access such as read or write access to all of the data in a storage account or all of the data in a container.
Box 2: Access control lists (ACLs)
Additional permissions can be applied to individual objects in storage1.
Access control lists (ACLs)
ACLs give you the ability to apply "finer grain" level of access to directories and files. An ACL is a permission construct that contains a series of ACL entries. Each ACL entry associates security principal with an access level.
Reference: https://learn.microsoft.com/en-us/azure/storage/blobs/data-lake-storage-access-control-model

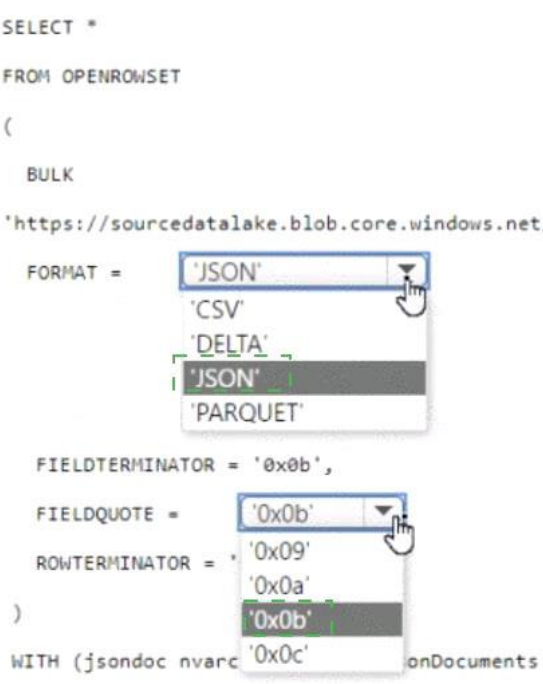
QUESTION NO: 22
Azure Synapse 서버리스 SQL 풀이 있습니다.
OPENROWSET 함수를 사용하여 파일에서 JSON 문서를 읽어야 합니다.
쿼리를 어떻게 완료해야 합니까? 대답하려면 대답 영역에서 적절한 옵션을 선택하십시오.
참고: 각 올바른 선택은 1점의 가치가 있습니다.

Answer Area



Answer:

Answer Area



Explanation:

Answer Area

```
SELECT *
FROM OPENROWSET
(
  BULK
'https://sourcedatalake.blob.core.windows.net/public/docs.json',
  FORMAT =      'JSON'
  FIELDTERMINATOR = '0x0b',
  FIELDQUOTE =      '0x0b'
  ROWTERMINATOR = '0x0b'
)
WITH (jsondoc nvarchar(max) AS JsonDocuments
```

QUESTION NO: 23
storage1이라는 Azure Blob Storage 계정과 Pool1이라는 Azure Synapse Analytics 전용 SQL 풀이 포함된 Azure 구독이 있습니다.
storage1에 데이터를 저장해야 합니다. 데이터는 Pool1에서 읽습니다. 솔루션은 다음 요구 사항을 충족해야 합니다.
쿼리에서 불필요한 열과 행을 건너뛰려면 Pool1을 사용하십시오.
열 통계를 자동으로 생성합니다.
파일 크기를 최소화하십시오.
어떤 유형의 파일을 사용해야 합니까?
A. JSON
B. 마루
C. 아브로
D. CSV
Answer: B
Explanation:
Automatic creation of statistics is turned on for Parquet files. For CSV files, you need to create statistics manually until automatic creation of CSV files statistics is supported.
Reference:
https://docs.microsoft.com/en-us/azure/synapse-analytics/sql/develop-tables-statistics

QUESTION NO: 24
Azure Stream Analytics를 사용하여 스트리밍 소셜 미디어 데이터를 수집할 계획입니다. 데이터는 Azure Data Lake Storage의 파일에 저장되고 Azure Synapse Analytics에서 Azure Dataabricks 및 PolyBase를 사용하여 사용됩니다.
파일에 대한 Databricks 및 PolyBase의 쿼리에서 발생할 수 있는 오류를 최소화하려면 Stream Analytics 데이터 출력 형식을 권장해야 합니다. 솔루션은 타일을 빠르게 쿼리할 수 있고 데이터 유형 정보가 유지되도록 해야 합니다.
무엇을 추천해야 할까요?
A. 쪽모이 세공
B. 아브로

C. CSV
D. JSON
Answer: A
Explanation:
https://docs.microsoft.com/en-us/azure/stream-analytics/stream-analytics-define-outputs

QUESTION NO: 25
Pool1이라는 Azure Synapse Analytics 전용 SQL 풀과 storage1이라는 Azure Data Lake Storage 계정이 포함된 Azure 구독이 있습니다. Storage1에는 안전한 전송이 필요합니다. storage1에서 .orc 파일을 읽는 데 사용할 외부 데이터 원본을 Pool1에 만들어야 합니다. 코드를 어떻게 완성해야 합니까? 대답하려면 대답 영역에서 적절한 옵션을 선택하십시오. 참고: 각 올바른 선택은 1점의 가치가 있습니다.

Answer Area

```
CREATE EXTERNAL DATA SOURCE AzureDataLakeStore

WITH

(  Location1 '  [          ] ://data@newyorktaxidataset.dfs.core.windows.net' ,
                  abfs
                  abfss
                  wasb
                  wasbs

credential = ADLS_credential ,

TYPE -  [          ]
                  BLOB_STORAGE
);                HADOOP
                  RDBMS
                  SHARP MAP MANAGER
```

Answer:

Answer Area

```
CREATE EXTERNAL DATA SOURCE AzureDataLakeStore
WITH
(  Location1 '[abfs|abfss|wasb|wasbs]://data@newyorktaxidataset.dfs.core.windows.net' ,
credential = ADLS_credential ,
TYPE = [BLOB_STORAGE|HADOOP|RDBMS|SHARP MAP MANAGER]
);
```

Explanation:

```
CREATE EXTERNAL DATA SOURCE AzureDataLakeStore
WITH
(  Location1 '[abfs|abfss|wasb|wasbs]://data@newyorktaxidataset.dfs.core.windows.net' ,
credential = ADLS_credential ,
TYPE = [BLOB_STORAGE|HADOOP|RDBMS|SHARP MAP MANAGER]
);
```

Reference:
https://docs.microsoft.com/en-us/sql/t-sql/statements/create-external-data-source-transact-sql?view=azure- sqldw-latest&preserve-view=true&tabs=dedicated

QUESTION NO: 26
Azure Databricks에서 Delta Lake의 테이블을 사용하는 두 가지 솔루션을 디자인하고 있습니다.
다음 작업을 수행하는 데 걸리는 시간을 최소화해야 합니다.
* 분할되지 않은 테이블에 대한 쿼리
* 분할되지 않은 열에 대한 조인
솔루션에 어떤 두 가지 옵션을 포함해야 합니까? 각 정답은 솔루션의 일부를 나타냅니다.
(Microsoft Azure의 데이터 엔지니어링을 기반으로 정답을 선택하고 설명을 제공하고 답변을

뒷받침하는 참고문헌을 제시하세요)
A. Z 순서
B. Apache Spark 캐싱
C. 동적 파일 정리(DFP)
D. 복제 명령
Answer: A C
Explanation:
According to the information I found on the web, two options that you should include in the solution to minimize how long it takes to perform queries and joins on non-partitioned tables are:
Z-Ordering: This is a technique to colocate related information in the same set of files. This co-locality is automatically used by Delta Lake in data-skipping algorithms. This behavior dramatically reduces the amount of data that Delta Lake on Azure Databricks needs to read123.
Apache Spark caching: This is a feature that allows you to cache data in memory or on disk for faster access.
Caching can improve the performance of repeated queries and joins on the same data. You can cache Delta tables using the CACHE TABLE or CACHE LAZY commands.
To minimize the time it takes to perform queries against non-partitioned tables and joins on non-partitioned columns in Delta Lake on Azure Databricks, the following options should be included in the solution:
A:Z-Ordering: Z-Ordering improves query performance by co-locating data that share the same column values in the same physical partitions. This reduces the need for shuffling data across nodes during query execution. By using Z-Ordering, you can avoid full table scans and reduce the amount of data processed.
B:Apache Spark caching: Caching data in memory can improve query performance by reducing the amount of data read from disk. This helps to speed up subsequent queries that need to access the same data. When you cache a table, the data is read from the data source and stored in memory. Subsequent queries can then read the data from memory, which is much faster than reading it from disk.
References:
Delta Lake on Databricks: https://docs.databricks.com/delta/index.html
Best Practices for Delta Lake on Databricks: https://databricks.com/blog/2020/05/14/best-practices-for-delta- lake-on-databricks.html

QUESTION NO: 27
Azure Data Lake Storage Gen2 계정(storage1)과 Azure Synapse Analytics 작업 영역(Workspace1)을 포함하는 Azure 구독이 있습니다. Workspace1에는 서버리스 SQL 풀이 있습니다.
서버리스 SQL 풀을 사용하여 storage1에 있는 파일에서 고객 주문을 쿼리합니다.
다음 쿼리를 실행합니다.
선택하다 *
FROM OPENROWSET(BULK 'https://storage1.blob.core.windows.net/data/orders/year =* /month =* / *.* ', FORMAT = 'parquet') AS customerorders WHERE customerorders.filepath(1) = '2024' AND customerorders.filepath(2) IN ('3','4'); 다음 각 문장에

대해 문장이 참이면 예를 선택합니다. 그렇지 않으면 아니요를 선택합니다.
참고사항: 정답 하나당 1점입니다.

Answer Area

Statements	Yes	No
storage1 provides a hierarchical namespace.	○	○
Files from March 2025 will be included.	○	○
Only files that have a Parquet file extension will be included.	○	○

Answer:

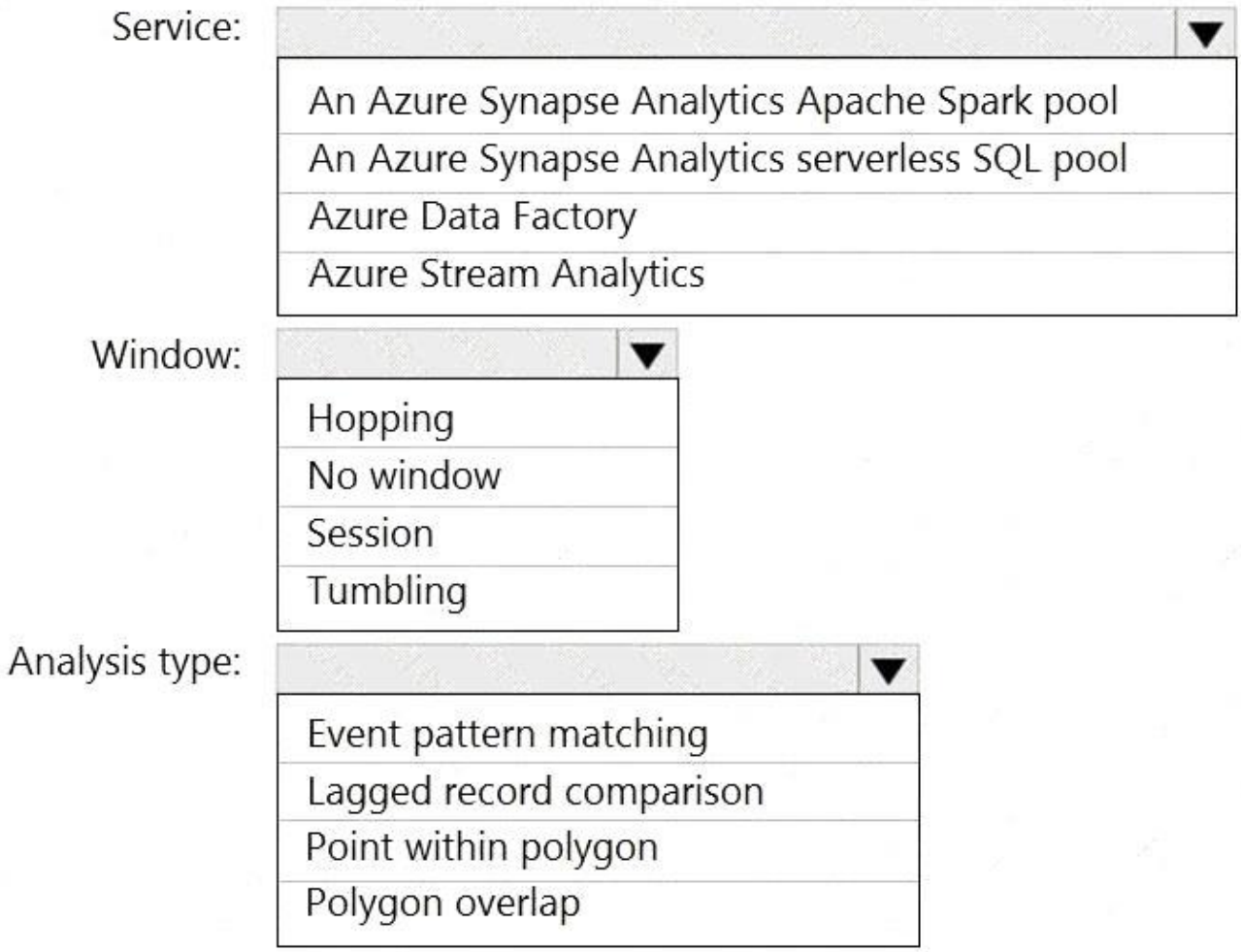
Answer Area

Statements	Yes	No
storage1 provides a hierarchical namespace.	○	○
Files from March 2025 will be included.	○	○
Only files that have a Parquet file extension will be included.	○	○

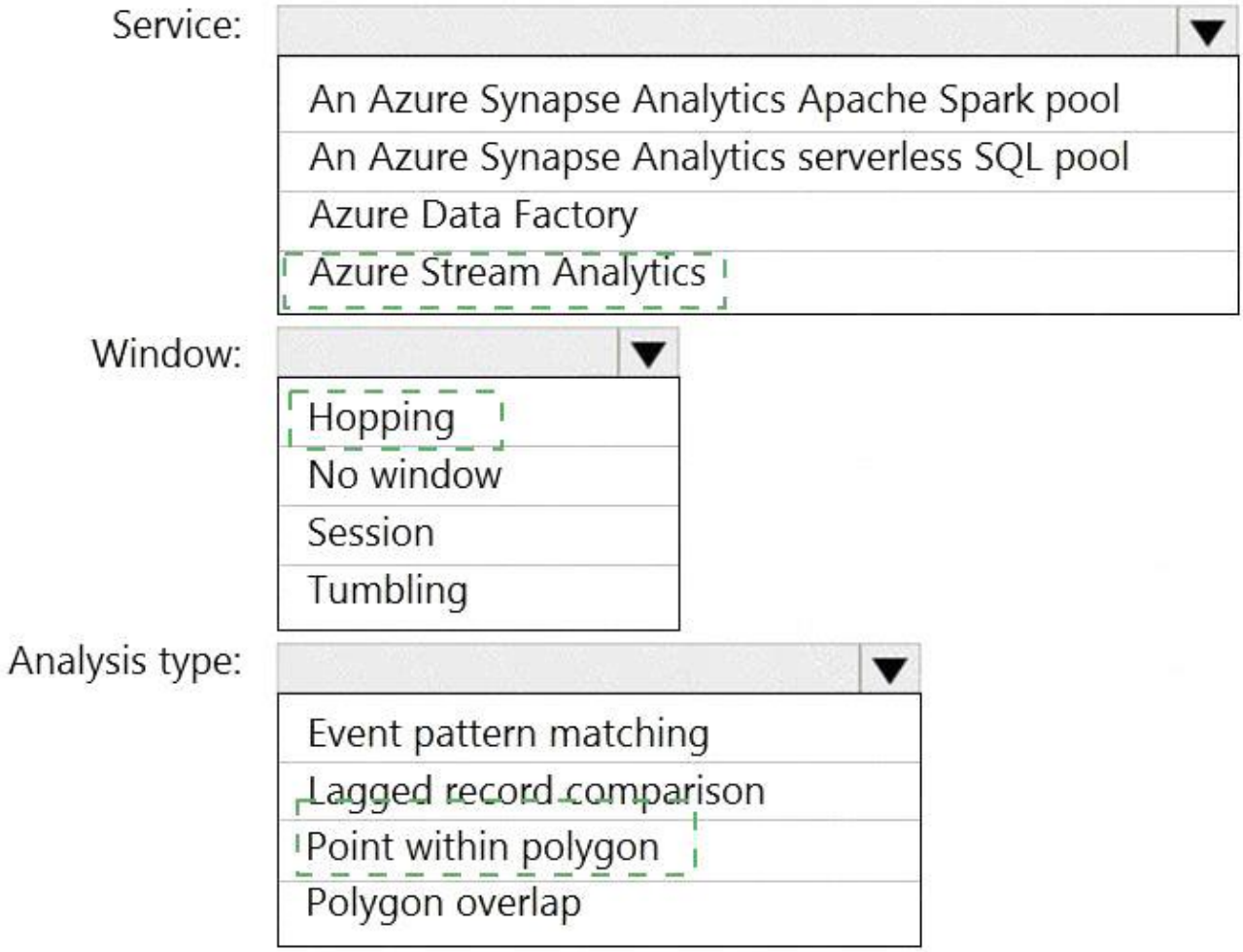
Explanation:
Storage1 provides a hierarchical namespace: Yes
Files from March 2025 will be included: No
Only files that have a Parquet file extension will be included: Yes
Query Breakdown
* Data Source:
* The OPENROWSET function queries data stored in Azure Data Lake Storage Gen2 (storage1) using the serverless SQL pool in Synapse Analytics.
* The data is stored in Parquet files in the folder structure data/orders/year=YYYY/month=MM/.
* Query Filter:
* The filter conditions in the query are:
* customerorders.filepath(1) = '2024': Limits the query to files in the folder year=2024.
* customerorders.filepath(2) IN ('3', '4'): Limits the query to files in the subfolders month=3 or month=4.
* File Format:
* The FORMAT = 'parquet' clause specifies that only Parquet files will be queried.
Statements Analysis
* Storage1 provides a hierarchical namespace.answer: Yes
* Azure Data Lake Storage Gen2 supports a hierarchical namespace, which enables folder-based organization.
* The folder structure (e.g., data/orders/year=2024/month=3/) demonstrates the use of a hierarchical namespace.
* Files from March 2025 will be included.answer: No
* The query explicitly filters for year=2024, so files from 2025 will not be included in the results.
* Only files that have a Parquet file extension will be included.answer: Yes

* The FORMAT = 'parquet' clause in the query ensures that only Parquet files are queried. Files with other extensions (e.g., .csv or .json) will not be included.

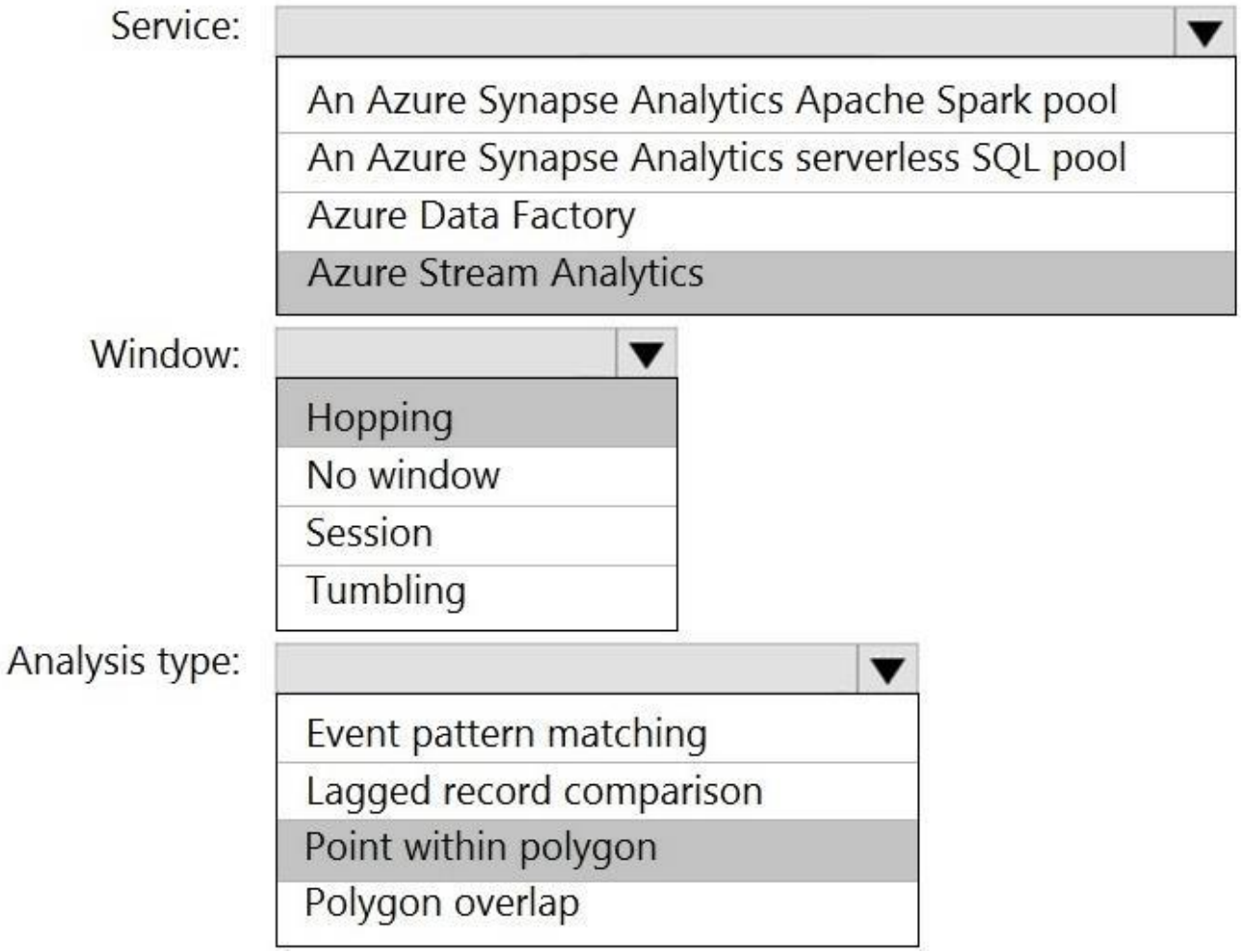
QUESTION NO: 28
귀하는 500대의 차량을 위한 모니터링 솔루션을 설계하고 있습니다. 각 차량에는 분당 한 번씩 Azure 이벤트 허브로 데이터를 전송하는 GPS 추적 장치가 있습니다.
Azure Data Lake Storage Gen2 컨테이너에 CSV 파일이 있습니다. 파일은 각 차량이 있어야 할 예상 지리적 영역을 유지합니다.
GPS 위치가 예상 영역 밖에 있을 때 메시지가 30초 이내에 처리되도록 다른 이벤트 허브에 추가되는지 확인해야 합니다. 솔루션은 비용을 최소화해야 합니다.
솔루션에 무엇을 포함해야 합니까? 대답하려면 대답 영역에서 적절한 옵션을 선택하십시오.
참고: 각 올바른 선택은 1점의 가치가 있습니다.



Answer:



Explanation:



Box 1: Azure Stream Analytics
Box 2: Hopping
Hopping window functions hop forward in time by a fixed period. It may be easy to think of them as Tumbling windows that can overlap and be emitted more often than the window size. Events can belong to more than one Hopping window result set. To make a Hopping window the same as a Tumbling window, specify the hop size to be the same as the window size.
Box 3: Point within polygon
Reference:
https://docs.microsoft.com/en-us/azure/stream-analytics/stream-analytics-window-functions

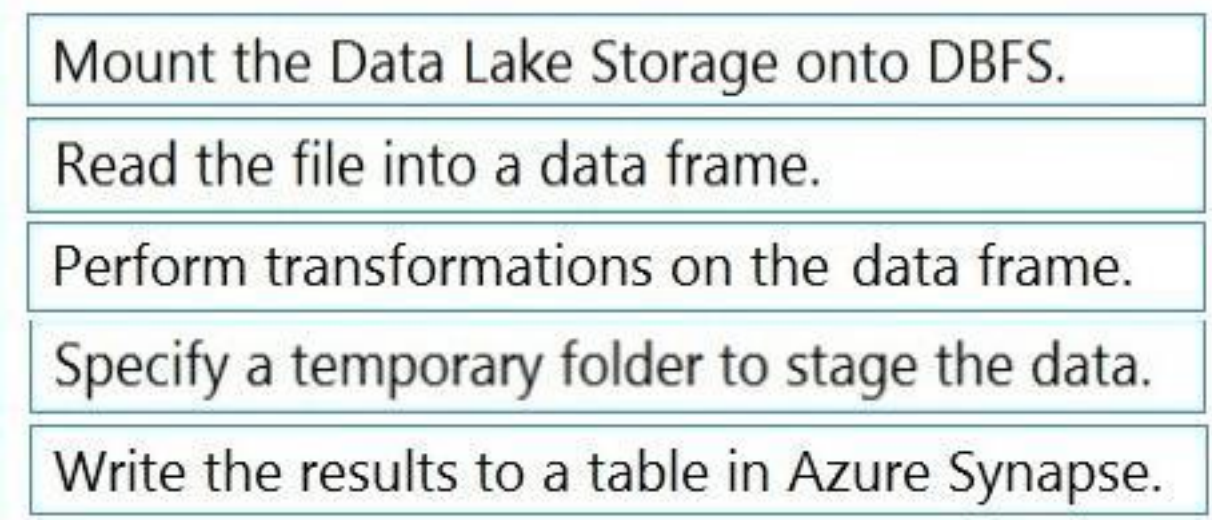
QUESTION NO: 29
고객용 JSON 파일이 포함된 Azure Data Lake Storage Gen2 계정이 있습니다. 파일에는 FirstName 및 LastName이라는 두 가지 특성이 포함되어 있습니다.
Azure Databricks를 사용하여 JSON 파일에서 Azure Synapse Analytics 테이블로 데이터를 복사해야 합니다. FirstName 및 LastName 값을 연결하는 새 열을 만들어야 합니다.
다음 구성 요소를 생성합니다.
Azure Synapse의 대상 테이블
Azure Blob 스토리지 컨테이너
서비스 주체
어떤 순서로 작업을 수행해야 합니까? 응답하려면 작업 목록에서 해당 작업을 응답 영역으로 이동하고 올바른 순서로 정렬하십시오.

Actions	Answer Area
Mount the Data Lake Storage onto DBFS.	
Write the results to a table in Azure Synapse.	
Specify a temporary folder to stage the data.	
Read the file into a data frame.	
Perform transformations on the data frame.	

Answer:

Actions	Answer Area
Mount the Data Lake Storage onto DBFS.	Mount the Data Lake Storage onto DBFS.
Write the results to a table in Azure Synapse.	Read the file into a data frame.
Specify a temporary folder to stage the data.	Perform transformations on the data frame.
Read the file into a data frame.	Specify a temporary folder to stage the data.
Perform transformations on the data frame.	Write the results to a table in Azure Synapse.

Explanation:



Step 1: Mount the Data Lake Storage onto DBFS
Begin with creating a file system in the Azure Data Lake Storage Gen2 account.
Step 2: Read the file into a data frame.
You can load the json files as a data frame in Azure Databricks.
Step 3: Perform transformations on the data frame.
Step 4: Specify a temporary folder to stage the data
Specify a temporary folder to use while moving data between Azure Databricks and Azure Synapse.
Step 5: Write the results to a table in Azure Synapse.
You upload the transformed data frame into Azure Synapse. You use the Azure Synapse connector for Azure Databricks to directly upload a dataframe as a table in a Azure Synapse.
Reference:
https://docs.microsoft.com/en-us/azure/azure-databricks/databricks-extract-load-sql-data-

warehouse

QUESTION NO: 30
Pool1이라는 Azure Synapse Analytics Apache Spark 풀이 있습니다.
Azure Data Lake Storage Gen2 컨테이너에서 Pool1의 테이블로 JSON 파일을 로드할 계획입니다. 구조 및 데이터 유형은 파일에 따라 다릅니다.
파일을 테이블에 로드해야 합니다. 솔루션은 소스 데이터 유형을 유지해야 합니다.
어떻게 해야 합니까?

A. Azure Data Factory에서 메타데이터 가져오기 활동을 사용합니다.

B. Azure Synapse 데이터 흐름에서 조건부 분할 변환을 사용합니다.

C. Azure Synapse Anarytics 서버리스 SQL 풀에서 OPEHROwset Transact-SQL 명령을 사용하여 데이터를 로드합니다.

D. PySpark를 사용하여 데이터를 로드합니다.

Answer: A
Explanation:
Serverless SQL pool can automatically synchronize metadata from Apache Spark. A serverless SQL pool database will be created for each database existing in serverless Apache Spark pools.
Serverless SQL pool enables you to query data in your data lake. It offers a T-SQL query surface area that accommodates semi-structured and unstructured data queries.
To support a smooth experience for in place querying of data that's located in Azure Storage files, serverless SQL pool uses the OPENROWSET function with additional capabilities.
The easiest way to see to the content of your JSON file is to provide the file URL to the OPENROWSET function, specify csv FORMAT.
Reference:
https://docs.microsoft.com/en-us/azure/synapse-analytics/sql/query-json-files
https://docs.microsoft.com/en-us/azure/synapse-analytics/sql/query-data-storage

QUESTION NO: 31
준비 영역이 포함된 Azure Data Lake Storage 계정이 있습니다.
준비 영역에서 증분 데이터를 수집하고 R 스크립트를 실행하여 데이터를 변환한 다음 변환된 데이터를 Azure Synapse Analytics의 데이터 웨어하우스에 삽입하는 일일 프로세스를 설계해야 합니다.
솔루션: Azure Data Factory 일정 트리거를 사용하여 Azure Databricks 노트북을 실행하는 파이프라인을 실행한 다음 데이터를 데이터 웨어하우스에 삽입합니다.
이것이 목표를 달성합니까?

A. 예

B. 아니요

Answer: B
Explanation:
If you need to transform data in a way that is not supported by Data Factory, you can create a custom activity, not an Azure Databricks notebook, with your own data processing logic and use the activity in the pipeline.
You can create a custom activity to run R scripts on your HDInsight cluster with R installed.
Reference:

https://docs.microsoft.com/en-US/azure/data-factory/transform-data

QUESTION NO: 32
분석 워크로드에서 사용할 원시 JSON 파일을 변환하는 Azure Data Lake Storage 솔루션을 설계하고 있습니다.
변환된 파일의 형식을 권장해야 합니다. 솔루션은 다음 요구 사항을 충족해야 합니다.
파일에 있는 각 열의 데이터 유형에 대한 정보를 포함합니다.
파일에서 열의 하위 집합 쿼리를 지원합니다.
읽기가 많은 분석 워크로드를 지원합니다.
파일 크기를 최소화하십시오.
무엇을 추천해야 할까요?
A. JSON
B. CSV
C. 아파치 아브로
D. 아파치 파켓
Answer: D
Explanation:
Parquet, an open-source file format for Hadoop, stores nested data structures in a flat columnar format.
Compared to a traditional approach where data is stored in a row-oriented approach, Parquet file format is more efficient in terms of storage and performance.
It is especially good for queries that read particular columns from a "wide" (with many columns) table since only needed columns are read, and IO is minimized.
Reference: https://www.clairvoyant.ai/blog/big-data-file-formats

QUESTION NO: 33
참고: 이 질문은 동일한 시나리오를 제시하는 일련의 질문 중 일부입니다. 시리즈의 각 질문에는 명시된 목표를 충족할 수 있는 고유한 솔루션이 포함되어 있습니다. 일부 질문 세트에는 하나 이상의 올바른 솔루션이 있을 수 있지만 다른 질문 세트에는 올바른 솔루션이 없을 수 있습니다.
이 시나리오에서 질문에 답한 후에는 해당 질문으로 돌아갈 수 없습니다. 결과적으로 이러한 질문은 검토 화면에 나타나지 않습니다.
100GB의 파일이 포함된 Azure Storage 계정이 있습니다. 파일에는 텍스트와 숫자 값이 포함되어 있습니다. 행의 75%에는 평균 길이가 1.1MB인 설명 데이터가 포함되어 있습니다.
스토리지 계정에서 Azure Synapse Analytics의 엔터프라이즈 데이터 웨어하우스로 데이터를 복사할 계획입니다.
데이터가 빠르게 복사되도록 파일을 준비해야 합니다.
해결 방법: 파일을 구분된 압축된 텍스트 파일로 변환합니다.
이것이 목표를 달성합니까?
A. 예
B. 아니요
Answer: A
Explanation:
All file formats have different performance characteristics. For the fastest load, use compressed delimited text files.

Reference:
https://docs.microsoft.com/en-us/azure/sql-data-warehouse/guidance-for-loading-data

QUESTION NO: 34
온도라는 Apache Spark DataFrame이 있습니다. 데이터 샘플은 다음 표에 나와 있습니다.

Date	Temp
...	...
18-01-2021	3
19-01-2021	4
20-01-2021	2
21-01-2021	2
...	...

Spark SQL 쿼리를 사용하여 다음 테이블을 생성해야 합니다.

Year	JAN	FEB	MAR	APR	MAY
2019	2.3	4.1	5.2	7.6	9.2
2020	2.4	4.2	4.9	7.8	9.1
2021	2.6	5.3	3.4	7.9	9.5

쿼리를 어떻게 완료해야 합니까? 응답하려면 적절한 값을 올바른 대상으로 드래그하십시오. 각 값은 한 번, 두 번 이상 사용되거나 전혀 사용되지 않을 수 있습니다. 콘텐츠를 보려면 창 사이의 분할 막대를 끌거나 스크롤해야 할 수 있습니다.
참고: 각 올바른 선택은 1점의 가치가 있습니다.

Values

CAST
COLLATE
CONVERT
FLATTEN
PIVOT
UNPIVOT

Answer Area

```
SELECT * FROM (
 SELECT YEAR(Date) Year, MONTH(Date) Month, Temp
 FROM temperatures
 WHERE date BETWEEN DATE '2019-01-01' AND DATE '2021-08-31'
)
[          ] (
 AVG ( [          ] (Temp AS DECIMAL(4, 1)))

 FOR Month in (
  1 JAN, 2 FEB, 3 MAR, 4 APR, 5 MAY, 6 JUN,
  7 JUL, 8 AUG, 9 SEP, 10 OCT, 11 NOV, 12 DEC
          )
 )
 ORDER BY Year ASC
```

Answer:

Values

CAST

COLLATE

CONVERT

FLATTEN

PIVOT

UNPIVOT

Answer Area

```
SELECT * FROM (
 SELECT YEAR(Date) Year, MONTH(Date) Month, Temp
 FROM temperatures
 WHERE date BETWEEN DATE '2019-01-01' AND DATE '2021-08-31'
)
PIVOT (
 AVG ( CAST (Temp AS DECIMAL(4, 1)))
 FOR Month in (
  1 JAN, 2 FEB, 3 MAR, 4 APR, 5 MAY, 6 JUN,
  7 JUL, 8 AUG, 9 SEP, 10 OCT, 11 NOV, 12 DEC
            )
 )
 ORDER BY Year ASC
```

Explanation:

```
SELECT * FROM (
 SELECT YEAR(Date) Year, MONTH(Date) Month, Temp
 FROM temperatures
 WHERE date BETWEEN DATE '2019-01-01' AND DATE '2021-08-31'
)
PIVOT (
 AVG ( CAST (Temp AS DECIMAL(4, 1)))
 FOR Month in (
  1 JAN, 2 FEB, 3 MAR, 4 APR, 5 MAY, 6 JUN,
  7 JUL, 8 AUG, 9 SEP, 10 OCT, 11 NOV, 12 DEC
            )
 )
 ORDER BY Year ASC
```

Box 1: PIVOT

PIVOT rotates a table-valued expression by turning the unique values from one column in the expression into multiple columns in the output. And PIVOT runs aggregations where they're required on any remaining column values that are wanted in the final output.

Reference:

https://learnsql.com/cookbook/how-to-convert-an-integer-to-a-decimal-in-sql-server/

https://docs.microsoft.com/en-us/sql/t-sql/queries/from-using-pivot-and-unpivot

QUESTION NO: 35

회사는 Azure Data Lake Storage Gen2 서비스를 사용합니다.

다음 요구 사항을 충족하는 데이터 보관 솔루션을 설계해야 합니다.

5년이 넘은 데이터는 자주 액세스되지 않지만 요청 시 1초 이내에 사용할 수 있어야 합니다.

7년이 넘은 데이터는 액세스되지 않습니다.
필요한 가용성을 유지하면서 비용을 최소화해야 합니다.
데이터를 어떻게 관리해야 할까요? 답변하려면 답변 영역에서 적절한 옵션을 선택하세요.
참고: 올바른 선택은 각각 1점의 가치가 있습니다.

Answer:

Explanation:

Answer Area


Data over five years old:
Delete the blob.
Move to archive storage.
Move to cool storage.
Move to hot storage.
Data over seven years old:
Delete the blob.
Move to archive storage.
Move to cool storage.
Move to hot storage.