Snowflake

Exam Questions ARA-C01

SnowPro Advanced: Architect Certification Exam

About Exambible

Your Partner of IT Exam

Found in 1998

Exambible is a company specialized on providing high quality IT exam practice study materials, especially Cisco CCNA, CCDA, CCNP, CCIE, Checkpoint CCSE, CompTIA A+, Network+ certification practice exams and so on. We guarantee that the candidates will not only pass any IT exam at the first attempt but also get profound understanding about the certificates they have got. There are so many alike companies in this industry, however, Exambible has its unique advantages that other companies could not achieve.

Our Advances

* 99.9% Uptime

All examinations will be up to date.

* 24/7 Quality Support

We will provide service round the clock.

* 100% Pass Rate

Our guarantee that you will pass the exam.

* Unique Gurantee

If you do not pass the exam at the first time, we will not only arrange FULL REFUND for you, but also provide you another exam of your claim, ABSOLUTELY FREE!

NEW QUESTION 1
A media company needs a data pipeline that will ingest customer review data into a Snowflake table, and apply some transformations. The company also needs to use Amazon Comprehend to do sentiment analysis and make the de-identified final data set
available publicly for advertising companies who use different cloud providers in different regions.
The data pipeline needs to run continuously and efficiently as new records arrive in the object storage leveraging event notifications. Also, the operational complexity, maintenance of the infrastructure, including platform upgrades and security, and the development effort should be minimal.
Which design will meet these requirements?

A. Ingest the data using copy into and use streams and tasks to orchestrate transformation
B. Export the data into Amazon S3 to do model inference with Amazon Comprehend and ingest the data back into a Snowflake tabl
C. Then create a listing in the Snowflake Marketplace to make the data available to other companies.
D. Ingest the data using Snowpipe and use streams and tasks to orchestrate transformation
E. Create an external function to do model inference with Amazon Comprehend and write the final records to a Snowflake tabl
F. Then create a listing in the Snowflake Marketplace to make the data available to other companies.
G. Ingest the data into Snowflake using Amazon EMR and PySpark using the Snowflake Spark connecto
H. Apply transformations using another Spark jo
I. Develop a python program to do model inference by leveraging the Amazon Comprehend text analysis AP
J. Then write the results to a Snowflake table and create a listing in the Snowflake Marketplace to make the data available to other companies.
K. Ingest the data using Snowpipe and use streams and tasks to orchestrate transformation
L. Export the data into Amazon S3 to do model inference with Amazon Comprehend and ingest the data back into a Snowflake tabl
M. Then create a listing in the Snowflake Marketplace to make the data available to other companies.

Answer: B

Explanation:
Option B is the best design to meet the requirements because it uses Snowpipe to ingest the data continuously and efficiently as new records arrive in the object storage, leveraging event notifications. Snowpipe is a service that automates the loading of data from external sources into Snowflake tables1. It also uses streams and tasks to orchestrate transformations on the ingested data. Streams are objects that store the change history of a table, and tasks are objects that execute SQL statements on a
schedule or when triggered by another task2. Option B also uses an external function to do model inference with Amazon Comprehend and write the final records to a Snowflake table. An external function is a user-defined function that calls an external API, such as Amazon Comprehend, to perform computations that are not natively supported by Snowflake3. Finally, option B uses the Snowflake Marketplace to make the de-identified final data set available publicly for advertising companies who use different cloud providers
in different regions. The Snowflake Marketplace is a platform that enables data providers to list and share their data sets with data consumers, regardless of the cloud platform or region they use4.
Option A is not the best design because it uses copy into to ingest the data, which is not as efficient and continuous as Snowpipe. Copy into is a SQL command that loads data from files into a table in a single transaction. It also exports the data into Amazon S3 to do model inference with Amazon Comprehend, which adds an extra step and increases the operational complexity and maintenance of the infrastructure.
Option C is not the best design because it uses Amazon EMR and PySpark to ingest and transform the data, which also increases the operational complexity and maintenance of the infrastructure. Amazon EMR is a cloud service that provides a managed Hadoop framework to process and analyze large-scale data sets. PySpark is a Python API for Spark, a distributed computing framework that can run on Hadoop. Option C also develops a python program to do model inference by leveraging the Amazon Comprehend text analysis API, which increases the development effort.
Option D is not the best design because it is identical to option A, except for the ingestion method. It still exports the data into Amazon S3 to do model inference with Amazon Comprehend, which adds an extra step and increases the operational complexity and maintenance of the infrastructure.
References: 1: Snowpipe Overview 2: Using Streams and Tasks to Automate Data Pipelines 3: External Functions Overview 4: Snowflake Data Marketplace Overview : [Loading Data Using COPY INTO] : [What is Amazon EMR?] : [PySpark Overview]

NEW QUESTION 2
An Architect runs the following SQL query:

```
SELECT
    METADATA$FILENAME,
    METADATA$FILE_ROW_NUMBER
FROM @FILEROWS/Food_Reviews.csv
    (file_format=CSV_N)
```

How can this query be interpreted?

A. FILEROWS is a stag
B. FILE_ROW_NUMBER is line number in file.
C. FILEROWS is the tabl
D. FILE_ROW_NUMBER is the line number in the table.
E. FILEROWS is a fil
F. FILE_ROW_NUMBER is the file format location.
G. FILERONS is the file format locatio
H. FILE_ROW_NUMBER is a stage.

Answer: A

Explanation:
? A stage is a named location in Snowflake that can store files for data loading and unloading. A stage can be internal or external, depending on where the files are stored.
? The query in the question uses the LIST function to list the files in a stage named FILEROWS. The function returns a table with various columns, including FILE_ROW_NUMBER, which is the line number of the file in the stage.
? Therefore, the query can be interpreted as listing the files in a stage named FILEROWS and showing the line number of each file in the stage.
References:

? : Stages
? : LIST Function

NEW QUESTION 3
Is it possible for a data provider account with a Snowflake Business Critical edition to share data with an Enterprise edition data consumer account?

A. A Business Critical account cannot be a data sharing provider to an Enterprise consume
B. Any consumer accounts must also be Business Critical.
C. If a user in the provider account with role authority to create or alter share adds an Enterprise account as a consumer, it can import the share.
D. If a user in the provider account with a share owning role sets share_restrictions to False when adding an Enterprise consumer account, it can import the share.
E. If a user in the provider account with a share owning role which also has override share restrictions privilege share_restrictions set to False when adding an Enterprise consumer account, it can import the share.

Answer: B

Explanation:
In Snowflake, data sharing capabilities allow a Business Critical edition account to share data with an Enterprise edition consumer account. The ability to share data is contingent upon the role permissions within the provider account. If a user has the necessary role authority (like ACCOUNTADMIN or a role with similar privileges to create or manage shares), they can add an Enterprise edition account as a consumer. This feature enables flexibility in data sharing across different Snowflake account editions, facilitating broader data collaboration and accessibility.References: Snowflake's data sharing documentation and the specifics of edition-based capabilities discussed in SnowPro Advanced: Architect certification materials.

NEW QUESTION 4
An Architect needs to grant a group of ORDER_ADMIN users the ability to clean old data in an ORDERS table (deleting all records older than 5 years), without granting any privileges on the table. The group??s manager (ORDER_MANAGER) has full DELETE privileges on the table.
How can the ORDER_ADMIN role be enabled to perform this data cleanup, without needing the DELETE privilege held by the ORDER_MANAGER role?

A. Create a stored procedure that runs with caller??s rights, including the appropriate "> 5 years" business logic, and grant USAGE on this procedure to ORDER_ADMI
B. The ORDER_MANAGER role owns the procedure.
C. Create a stored procedure that can be run using both caller??s and owner??s rights (allowing the user to specify which rights are used during execution), and grant USAGE on this procedure to ORDER_ADMI
D. The ORDER_MANAGER role owns the procedure.
E. Create a stored procedure that runs with owner??s rights, including the appropriate "> 5 years" business logic, and grant USAGE on this procedure to ORDER_ADMI
F. The ORDER_MANAGER role owns the procedure.
G. This scenario would actually not be possible in Snowflake – any user performing a DELETE on a table requires the DELETE privilege to be granted to the role they are using.

Answer: C

Explanation:
This is the correct answer because it allows the ORDER_ADMIN role to perform the data cleanup without needing the DELETE privilege on the ORDERS table. A stored procedure is a feature that allows scheduling and executing SQL statements or
stored procedures in Snowflake. A stored procedure can run with either the caller??s rights or the owner??s rights. A caller??s rights stored procedure runs with the privileges of the role that called the stored procedure, while an owner??s rights stored procedure runs with the privileges of the role that created the stored procedure. By creating a stored procedure that runs with owner??s rights, the ORDER_MANAGER role can delegate the specific task of deleting old data to the ORDER_ADMIN role, without granting the ORDER_ADMIN role more general privileges on the ORDERS table. The stored procedure must include the appropriate business logic to delete only the records older than 5 years, and the ORDER_MANAGER role must grant the USAGE privilege on the stored procedure to the ORDER_ADMIN role. The ORDER_ADMIN role can then execute the stored procedure to perform the data cleanup12.
References:
? Snowflake Documentation: Stored Procedures
? Snowflake Documentation: Understanding Caller??s Rights and Owner??s Rights Stored Procedures

NEW QUESTION 5
A Snowflake Architect is designing a multi-tenant application strategy for an organization in the Snowflake Data Cloud and is considering using an Account Per Tenant strategy.
Which requirements will be addressed with this approach? (Choose two.)

A. There needs to be fewer objects per tenant.
B. Security and Role-Based Access Control (RBAC) policies must be simple to configure.
C. Compute costs must be optimized.
D. Tenant data shape may be unique per tenant.
E. Storage costs must be optimized.

Answer: BD

Explanation:
The Account Per Tenant strategy involves creating separate Snowflake accounts for each tenant within the multi-tenant application. This approach offers a number of advantages.
Option B: With separate accounts, each tenant's environment is isolated, making security and RBAC policies simpler to configure and maintain. This is because each account can have its own set of roles and privileges without the risk of cross-tenant access or the complexity of maintaining a highly granular permission model within a shared environment. Option D: This approach also allows for each tenant to have a unique data shape, meaning that the database schema can be tailored to the specific needs of each tenant without affecting others. This can be essential when tenants have different data models,
usage patterns, or application customizations.

NEW QUESTION 6
An Architect is designing a file ingestion recovery solution. The project will use an internal named stage for file storage. Currently, in the case of an ingestion failure, the Operations team must manually download the failed file and check for errors.

Which downloading method should the Architect recommend that requires the LEAST amount of operational overhead?

A. Use the Snowflake Connector for Python, connect to remote storage and download the file.
B. Use the get command in SnowSQL to retrieve the file.
C. Use the get command in Snowsight to retrieve the file.
D. Use the Snowflake API endpoint and download the file.

Answer: B

Explanation:
The get command in SnowSQL is a convenient way to download files from an internal stage to a local directory. The get command can be used in interactive mode or in a script, and it supports wildcards and parallel downloads. The get command also allows specifying the overwrite option, which determines how to handle existing files with the same name2
The Snowflake Connector for Python, the Snowflake API endpoint, and the get command in Snowsight are not recommended methods for downloading files from an internal stage, because they require more operational overhead than the get command in SnowSQL. The Snowflake Connector for Python and the Snowflake API endpoint require writing and maintaining code to handle the connection, authentication, and file transfer. The get command in Snowsight requires using the web interface and manually selecting the files to download34 References:
? 1: SnowPro Advanced: Architect | Study Guide
? 2: Snowflake Documentation | Using the GET Command
? 3: Snowflake Documentation | Using the Snowflake Connector for Python
? 4: Snowflake Documentation | Using the Snowflake API
? : Snowflake Documentation | Using the GET Command in Snowsight
? : SnowPro Advanced: Architect | Study Guide
? : Using the GET Command
? : Using the Snowflake Connector for Python
? : Using the Snowflake API
? : [Using the GET Command in Snowsight]

NEW QUESTION 7
A company wants to deploy its Snowflake accounts inside its corporate network with no visibility on the internet. The company is using a VPN infrastructure and Virtual Desktop Infrastructure (VDI) for its Snowflake users. The company also wants to re-use the login credentials set up for the VDI to eliminate redundancy when managing logins.
What Snowflake functionality should be used to meet these requirements? (Choose two.)

A. Set up replication to allow users to connect from outside the company VPN.
B. Provision a unique company Tri-Secret Secure key.
C. Use private connectivity from a cloud provider.
D. Set up SSO for federated authentication.
E. Use a proxy Snowflake account outside the VPN, enabling client redirect for user logins.

Answer: CD

Explanation:
According to the SnowPro Advanced: Architect documents and learning resources, the Snowflake functionality that should be used to meet these requirements are:
? Use private connectivity from a cloud provider. This feature allows customers to connect to Snowflake from their own private network without exposing their data to the public Internet. Snowflake integrates with AWS PrivateLink, Azure Private Link, and Google Cloud Private Service Connect to offer private connectivity from customers?? VPCs or VNets to Snowflake endpoints. Customers can control how traffic reaches the Snowflake endpoint and avoid the need for proxies or public IP addresses123.
? Set up SSO for federated authentication. This feature allows customers to use their existing identity provider (IdP) to authenticate users for SSO access to Snowflake. Snowflake supports most SAML 2.0-compliant vendors as an IdP, including Okta, Microsoft AD FS, Google G Suite, Microsoft Azure Active Directory, OneLogin, Ping Identity, and PingOne. By setting up SSO for federated authentication, customers can leverage their existing user credentials and profile information, and provide stronger security than username/password authentication4.
The other options are incorrect because they do not meet the requirements or are not feasible. Option A is incorrect because setting up replication does not allow users to connect from outside the company VPN. Replication is a feature of Snowflake that enables copying databases across accounts in different regions and cloud platforms. Replication does not affect the connectivity or visibility of the accounts5. Option B is incorrect because provisioning a unique company Tri-Secret Secure key does not affect the network or authentication requirements. Tri-Secret Secure is a feature of Snowflake that allows customers to manage their own encryption keys for data at rest in Snowflake, using a combination of three secrets: a master key, a service key, and a security password. Tri- Secret Secure provides an additional layer of security and control over the data encryption and decryption process, but it does not enable private connectivity or SSO6. Option E is incorrect because using a proxy Snowflake account outside the VPN, enabling client redirect for user logins, is not a supported or recommended way of meeting the requirements. Client redirect is a feature of Snowflake that allows customers to connect to a different Snowflake account than the one specified in the connection string. This feature is useful for scenarios such as cross-region failover, data sharing, and account migration, but it does not provide private connectivity or SSO7. References: AWS PrivateLink & Snowflake | Snowflake Documentation, Azure Private Link & Snowflake | Snowflake Documentation, Google Cloud Private Service Connect & Snowflake | Snowflake Documentation, Overview of Federated Authentication and SSO | Snowflake Documentation, Replicating Databases Across Multiple Accounts | Snowflake Documentation, Tri-Secret Secure | Snowflake Documentation, Redirecting Client Connections | Snowflake Documentation

NEW QUESTION 8
An Architect has designed a data pipeline that Is receiving small CSV files from multiple sources. All of the files are landing in one location. Specific files are filtered for loading into Snowflake tables using the copy command. The loading performance is poor.
What changes can be made to Improve the data loading performance?

A. Increase the size of the virtual warehouse.
B. Create a multi-cluster warehouse and merge smaller files to create bigger files.
C. Create a specific storage landing bucket to avoid file scanning.
D. Change the file format from CSV to JSON.

Answer: B

Explanation:
According to the Snowflake documentation, the data loading performance can be improved by following some best practices and guidelines for preparing and staging the data files. One of the recommendations is to aim for data files that are roughly 100-250 MB (or larger) in size compressed, as this will optimize the number of parallel operations for a load. Smaller files should be aggregated and larger files should be split to achieve this size range. Another recommendation is to use a multi-cluster warehouse for loading, as this will allow for scaling up or out the compute resources depending on the load demand.
A single-cluster warehouse may not be able to handle the load concurrency and throughput efficiently. Therefore, by creating a multi-cluster warehouse and merging smaller files to create bigger files, the data loading performance can be improved. References:
? Data Loading Considerations
? Preparing Your Data Files
? Planning a Data Load

NEW QUESTION 9
An Architect needs to meet a company requirement to ingest files from the company's AWS storage accounts into the company's Snowflake Google Cloud Platform (GCP) account. How can the ingestion of these files into the company's Snowflake account be initiated? (Select TWO).

A. Configure the client application to call the Snowpipe REST endpoint when new files have arrived in Amazon S3 storage.
B. Configure the client application to call the Snowpipe REST endpoint when new files have arrived in Amazon S3 Glacier storage.
C. Create an AWS Lambda function to call the Snowpipe REST endpoint when new files have arrived in Amazon S3 storage.
D. Configure AWS Simple Notification Service (SNS) to notify Snowpipe when new files have arrived in Amazon S3 storage.
E. Configure the client application to issue a COPY INTO <TABLE> command to Snowflake when new files have arrived in Amazon S3 Glacier storage.

Answer: AC

Explanation:
Snowpipe is a feature that enables continuous, near-real-time data ingestion from external sources into Snowflake tables. Snowpipe can ingest files from Amazon S3, Google Cloud Storage, or Azure Blob Storage into Snowflake tables on any cloud platform. Snowpipe can be triggered in two ways: by using the Snowpipe REST API or by using cloud notifications2
To ingest files from the company??s AWS storage accounts into the company??s Snowflake GCP account, the Architect can use either of these methods:
? Configure the client application to call the Snowpipe REST endpoint when new
files have arrived in Amazon S3 storage. This method requires the client application to monitor the S3 buckets for new files and send a request to the Snowpipe REST API with the list of files to ingest. The client application must also handle authentication, error handling, and retry logic3
? Create an AWS Lambda function to call the Snowpipe REST endpoint when new
files have arrived in Amazon S3 storage. This method leverages the AWS Lambda service to execute a function that calls the Snowpipe REST API whenever an S3 event notification is received. The AWS Lambda function must be configured with the appropriate permissions, triggers, and code to invoke the Snowpipe REST API4
The other options are not valid methods for triggering Snowpipe:
? Configure the client application to call the Snowpipe REST endpoint when new files have arrived in Amazon S3 Glacier storage. This option is not feasible because Snowpipe does not support ingesting files from Amazon S3 Glacier storage, which is a long-term archival storage service. Snowpipe only supports ingesting files from Amazon S3 standard storage classes5
? Configure AWS Simple Notification Service (SNS) to notify Snowpipe when new files have arrived in Amazon S3 storage. This option is not applicable because Snowpipe does not support cloud notifications from AWS SNS. Snowpipe only supports cloud notifications from AWS SQS, Google Cloud Pub/Sub, or Azure Event Grid6
? Configure the client application to issue a COPY INTO <TABLE> command to Snowflake when new files have arrived in Amazon S3 Glacier storage. This option is not relevant because it does not use Snowpipe, but rather the standard COPY command, which is a batch loading method. Moreover, the COPY command also does not support ingesting files from Amazon S3 Glacier storage7 References:
? 1: SnowPro Advanced: Architect | Study Guide 8
? 2: Snowflake Documentation | Snowpipe Overview 9
? 3: Snowflake Documentation | Using the Snowpipe REST API 10
? 4: Snowflake Documentation | Loading Data Using Snowpipe and AWS Lambda 11
? 5: Snowflake Documentation | Supported File Formats and Compression for Staged Data Files 12
? 6: Snowflake Documentation | Using Cloud Notifications to Trigger Snowpipe 13
? 7: Snowflake Documentation | Loading Data Using COPY into a Table
? : SnowPro Advanced: Architect | Study Guide
? : Snowpipe Overview
? : Using the Snowpipe REST API
? : Loading Data Using Snowpipe and AWS Lambda
? : Supported File Formats and Compression for Staged Data Files
? : Using Cloud Notifications to Trigger Snowpipe
? : Loading Data Using COPY into a Table

NEW QUESTION 10
A retail company has 2000+ stores spread across the country. Store Managers report that they are having trouble running key reports related to inventory management, sales targets, payroll, and staffing during business hours. The Managers report that performance is poor and time-outs occur frequently. Currently all reports share the same Snowflake virtual warehouse. How should this situation be addressed? (Select TWO).

A. Use a Business Intelligence tool for in-memory computation to improve performance.
B. Configure a dedicated virtual warehouse for the Store Manager team.
C. Configure the virtual warehouse to be multi-clustered.
D. Configure the virtual warehouse to size 4-XL
E. Advise the Store Manager team to defer report execution to off-business hours.

Answer: BC

Explanation:
The best way to address the performance issues and time-outs faced by the Store Manager team is to configure a dedicated virtual warehouse for them and make it multi-clustered. This will allow them to run their reports independently from other workloads and scale up or down the compute resources as needed. A dedicated virtual warehouse will also enable them to apply specific security and access policies for their data. A multi- clustered virtual warehouse will provide high availability and concurrency for their queries and avoid queuing or throttling.
Using a Business Intelligence tool for in-memory computation may improve performance, but it will not solve the underlying issue of insufficient compute resources in the shared virtual warehouse. It will also introduce additional costs and complexity for the data architecture.
Configuring the virtual warehouse to size 4-XL may increase the performance, but it will also increase the cost and may not be optimal for the workload. It will also

not address the concurrency and availability issues that may arise from sharing the virtual warehouse with other workloads.
Advising the Store Manager team to defer report execution to off-business hours may reduce the load on the shared virtual warehouse, but it will also reduce the timeliness and usefulness of the reports for the business. It will also not guarantee that the performance issues and time-outs will not occur at other times.
References:
? Snowflake Architect Training
? Snowflake SnowPro Advanced Architect Certification - Preparation Guide
? SnowPro Advanced: Architect Exam Study Guide

NEW QUESTION 10
An Architect is designing a data lake with Snowflake. The company has structured, semi- structured, and unstructured data. The company wants to save the data inside the data lake within the Snowflake
system. The company is planning on sharing data among its corporate branches using Snowflake data sharing.
What should be considered when sharing the unstructured data within Snowflake?

A. A pre-signed URL should be used to save the unstructured data into Snowflake in order to share data over secure views, with no time limit for the URL.
B. A scoped URL should be used to save the unstructured data into Snowflake in order to share data over secure views, with a 24-hour time limit for the URL.
C. A file URL should be used to save the unstructured data into Snowflake in order to share data over secure views, with a 7-day time limit for the URL.
D. A file URL should be used to save the unstructured data into Snowflake in order to share data over secure views, with the "expiration_time" argument defined for the URL time limit.

Answer: B

Explanation:
When sharing unstructured data within Snowflake, using a scoped URL is recommended. Scoped URLs provide temporary access to staged files without granting privileges to the stage itself, enhancing security. The URL expires when the persisted query result period ends, which is currently set to 24 hours. This approach is suitable for sharing unstructured data over secure views within Snowflake??s data sharing framework. References: The answer is based on Snowflake??s official documentation regarding the sharing of unstructured data and the use of scoped URLs1.

NEW QUESTION 13
Which technique will efficiently ingest and consume semi-structured data for Snowflake data lake workloads?

A. IDEF1X
B. Schema-on-write
C. Schema-on-read
D. Information schema

Answer: C

Explanation:
Option C is the correct answer because schema-on-read is a technique that allows Snowflake to ingest and consume semi-structured data without requiring a predefined schema. Snowflake supports various semi-structured data formats such as JSON, Avro, ORC, Parquet, and XML, and provides native data types (ARRAY, OBJECT, and VARIANT) for storing them. Snowflake also provides native support for querying semi- structured data using SQL and dot notation. Schema-on-read enables Snowflake to query semi-structured data at the same speed as performing relational queries while preserving the flexibility of schema-on-read. Snowflake??s near-instant elasticity rightsizes compute resources, and consumption-based pricing ensures you only pay for what you use.
Option A is incorrect because IDEF1X is a data modeling technique that defines the structure and constraints of relational data using diagrams and notations. IDEF1X is not suitable for ingesting and consuming semi-structured data, which does not have a fixed schema or structure.
Option B is incorrect because schema-on-write is a technique that requires defining a schema before loading and processing data. Schema-on-write is not efficient for ingesting and consuming semi-structured data, which may have varying or complex structures that are difficult to fit into a predefined schema. Schema-on-write also introduces additional overhead and complexity for data transformation and validation.
Option D is incorrect because information schema is a set of metadata views that provide information about the objects and privileges in a Snowflake database. Information schema is not a technique for ingesting and consuming semi-structured data, but rather a way of accessing metadata about the data.
References:
? Semi-structured Data
? Snowflake for Data Lake

NEW QUESTION 17
What are some of the characteristics of result set caches? (Choose three.)

A. Time Travel queries can be executed against the result set cache.
B. Snowflake persists the data results for 24 hours.
C. Each time persisted results for a query are used, a 24-hour retention period is reset.
D. The data stored in the result cache will contribute to storage costs.
E. The retention period can be reset for a maximum of 31 days.
F. The result set cache is not shared between warehouses.

Answer: BCF

Explanation:
In Snowflake, the characteristics of result set caches include persistence of data results for 24 hours (B), each use of persisted results resets the 24-hour retention period (C), and result set caches are not shared between different warehouses (F). The result set cache is specifically designed to avoid repeated execution of the same query within this timeframe, reducing computational overhead and speeding up query responses. These caches do not contribute to storage costs, and their retention period cannot be extended beyond the default duration nor up to 31 days, as might be misconstrued.References: Snowflake Documentation on Result Set Caching.

NEW QUESTION 22
What considerations need to be taken when using database cloning as a tool for data lifecycle management in a development environment? (Select TWO).

A. Any pipes in the source are not cloned.
B. Any pipes in the source referring to internal stages are not cloned.

C. Any pipes in the source referring to external stages are not cloned.
D. The clone inherits all granted privileges of all child objects in the source object, including the database.
E. The clone inherits all granted privileges of all child objects in the source object, excluding the database.

Answer: AC

NEW QUESTION 24
How is the change of local time due to daylight savings time handled in Snowflake tasks? (Choose two.)

A. A task scheduled in a UTC-based schedule will have no issues with the time changes.
B. Task schedules can be designed to follow specified or local time zones to accommodate the time changes.
C. A task will move to a suspended state during the daylight savings time change.
D. A frequent task execution schedule like minutes may not cause a problem, but will affect the task history.
E. A task schedule will follow only the specified time and will fail to handle lost or duplicated hours.

Answer: AB

Explanation:
According to the Snowflake documentation1 and the web search results2, these two statements are true about how the change of local time due to daylight savings time is handled in Snowflake tasks. A task is a feature that allows scheduling and executing
SQL statements or stored procedures in Snowflake. A task can be scheduled using a cron expression that specifies the frequency and time zone of the task execution.
? A task scheduled in a UTC-based schedule will have no issues with the time
changes. UTC is a universal time standard that does not observe daylight savings time. Therefore, a task that uses UTC as the time zone will run at the same time throughout the year, regardless of the local time changes1.
? Task schedules can be designed to follow specified or local time zones to
accommodate the time changes. Snowflake supports using any valid IANA time zone identifier in the cron expression for a task. This allows the task to run according to the local time of the specified time zone, which may include daylight savings time adjustments. For example, a task that uses Europe/London as the time zone will run one hour earlier or later when the local time switches between GMT and BST12.
References:
? Snowflake Documentation: Scheduling Tasks
? Snowflake Community: Do the timezones used in scheduling tasks in Snowflake adhere to daylight savings?

NEW QUESTION 25
What is a valid object hierarchy when building a Snowflake environment?

A. Account --> Database --> Schema --> Warehouse
B. Organization --> Account --> Database --> Schema --> Stage
C. Account --> Schema > Table --> Stage
D. Organization --> Account --> Stage --> Table --> View

Answer: B

Explanation:
This is the valid object hierarchy when building a Snowflake environment, according to the Snowflake documentation and the web search results. Snowflake is a cloud data platform that supports various types of objects, such as databases, schemas, tables, views, stages, warehouses, and more. These objects are organized in a hierarchical structure, as follows:
? Organization: An organization is the top-level entity that represents a group of
Snowflake accounts that are related by business needs or ownership. An organization can have one or more accounts, and can enable features such as cross-account data sharing, billing and usage reporting, and single sign-on across accounts12.
? Account: An account is the primary entity that represents a Snowflake customer.
An account can have one or more databases, schemas, stages, warehouses, and other objects. An account can also have one or more users, roles, and security integrations. An account is associated with a specific cloud platform, region, and Snowflake edition34.
? Database: A database is a logical grouping of schemas. A database can have one or more schemas, and can store structured, semi-structured, or unstructured data. A database can also have properties such as retention time, encryption, and ownership56.
? Schema: A schema is a logical grouping of tables, views, stages, and other
objects. A schema can have one or more objects, and can define the namespace and access control for the objects. A schema can also have properties such as ownership and default warehouse .
? Stage: A stage is a named location that references the files in external or internal
storage. A stage can be used to load data into Snowflake tables using the COPY INTO command, or to unload data from Snowflake tables using the COPY INTO LOCATION command. A stage can be created at the account, database, or schema level, and can have properties such as file format, encryption, and credentials .
The other options listed are not valid object hierarchies, because they either omit or misplace some objects in the structure. For example, option A omits the organization level and places the warehouse under the schema level, which is incorrect. Option C omits the organization, account, and stage levels, and places the table under the schema level, which is incorrect. Option D omits the database level and places the stage and table under the account level, which is incorrect.
References:
? Snowflake Documentation: Organizations
? Snowflake Blog: Introducing Organizations in Snowflake
? Snowflake Documentation: Accounts
? Snowflake Blog: Understanding Snowflake Account Structures
? Snowflake Documentation: Databases
? Snowflake Blog: How to Create a Database in Snowflake
? [Snowflake Documentation: Schemas]
? [Snowflake Blog: How to Create a Schema in Snowflake]
? [Snowflake Documentation: Stages]
? [Snowflake Blog: How to Use Stages in Snowflake]

NEW QUESTION 29
The following DDL command was used to create a task based on a stream:

```
CREATE TASK ts_insert_new_customers
  WAREHOUSE = MY_WH
  Schedule = '5 minute'
WHEN
  System$STREAM_HAS_DATA('MYSTREAM')
AS
  INSERT INTO new customers(id, name) SELECT id, name
FROM mystream WHERE METADATA$ACTION = 'INSERT';
```

Assuming MY_WH is set to auto_suspend – 60 and used exclusively for this task, which statement is true?

A. The warehouse MY_WH will be made active every five minutes to check the stream.
B. The warehouse MY_WH will only be active when there are results in the stream.
C. The warehouse MY_WH will never suspend.
D. The warehouse MY_WH will automatically resize to accommodate the size of the stream.

Answer: B

Explanation:
The warehouse MY_WH will only be active when there are results in the stream. This is because the task is created based on a stream, which means that the task will only be executed when there are new data in the stream. Additionally, the warehouse is set to auto_suspend - 60, which means that the warehouse will automatically suspend after 60 seconds of inactivity. Therefore, the warehouse will only be active when there are results in the stream. References:
? [CREATE TASK | Snowflake Documentation]
? [Using Streams and Tasks | Snowflake Documentation]
? [CREATE WAREHOUSE | Snowflake Documentation]

NEW QUESTION 31
A table, EMP_ TBL has three records as shown:

```
create or replace TABLE EMP_TBL (
ID NUMBER(38,0),
NAME VARCHAR(16777216)
);
```

ID	NAME
1	Name1
2	Name2
3	Name3

The following variables are set for the session:

```
set tbl_ref = 'EMP_TBL';
set col_ref = 'NAME';
set (var1, var2, var3) = (select 'Name1', 'Name2', 'Name3');
```

Which SELECT statements will retrieve all three records? (Select TWO).

A. Select * FROM Stbl_ref WHERE Scol_ref IN ('Name1','Nam2','Name3');
B. SELECT * FROM EMP_TBL WHERE identifier(Scol_ref) IN ('Namel','Name2', 'Name3');
C. SELECT * FROM identifier<Stbl_ref> WHERE NAME IN ($var1, $var2, $var3);
D. SELECT * FROM identifier($tbl_ref) WHERE ID IN Cvarl','var2','var3');

E. SELECT * FROM $tb1_ref WHERE $col_ref IN ($var1, Svar2, Svar3);

Answer: BE

Explanation:
? The correct answer is B and E because they use the correct syntax and values for the identifier function and the session variables.
? The identifier function allows you to use a variable or expression as an identifier (such as a table name or column name) in a SQL statement. It takes a single argument and returns it as an identifier. For example, identifier($tbl_ref) returns
EMP_TBL as an identifier.
? The session variables are set using the SET command and can be referenced using the $ sign. For example, $var1 returns Name1 as a value.
? Option A is incorrect because it uses Stbl_ref and Scol_ref, which are not valid session variables or identifiers. They should be $tbl_ref and $col_ref instead.
? Option C is incorrect because it uses identifier<Stbl_ref>, which is not a valid syntax for the identifier function. It should be identifier($tbl_ref) instead.
? Option D is incorrect because it uses Cvarl, var2, and var3, which are not valid session variables or values. They should be $var1, $var2, and $var3
instead. References:
? Snowflake Documentation: Identifier Function
? Snowflake Documentation: Session Variables
? Snowflake Learning: SnowPro Advanced: Architect Exam Study Guide

NEW QUESTION 32
The Data Engineering team at a large manufacturing company needs to engineer data coming from many sources to support a wide variety of use cases and data consumer requirements which include:
1) Finance and Vendor Management team members who require reporting and visualization
2) Data Science team members who require access to raw data for ML model development
3) Sales team members who require engineered and protected data for data monetization What Snowflake data modeling approaches will meet these requirements? (Choose two.)

A. Consolidate data in the company??s data lake and use EXTERNAL TABLES.
B. Create a raw database for landing and persisting raw data entering the data pipelines.
C. Create a set of profile-specific databases that aligns data with usage patterns.
D. Create a single star schema in a single database to support all consumers?? requirements.
E. Create a Data Vault as the sole data pipeline endpoint and have all consumers directly access the Vault.

Answer: BC

Explanation:
To accommodate the diverse needs of different teams and use cases within a company, a flexible and multi-faceted approach to data modeling is required.
Option B: By creating a raw database for landing and persisting raw data, you ensure that the Data Science team has access to unprocessed data for machine learning model development. This aligns with the best practices of having a staging area or raw data zone
in a modern data architecture where raw data is ingested before being transformed or processed for different use cases.
Option C: Having profile-specific databases means creating targeted databases that are designed to meet the specific requirements of each user profile or team within the company. For the Finance and Vendor Management teams, the data can be structured and optimized for reporting and visualization. For the Sales team, the database can include engineered and protected data that is suitable for data monetization efforts. This strategy not only aligns data with usage patterns but also helps in managing data access and security policies effectively.

NEW QUESTION 33
Which command will create a schema without Fail-safe and will restrict object owners from passing on access to other users?

A. create schema EDW.ACCOUNTING WITH MANAGED ACCESS;
B. create schema EDW.ACCOUNTING WITH MANAGED ACCESS DATA_RETENTION_TIME_IN_DAYS - 7;
C. create TRANSIENT schema EDW.ACCOUNTING WITH MANAGED ACCESS DATA_RETENTION_TIME_IN_DAYS = 1;
D. create TRANSIENT schema EDW.ACCOUNTING WITH MANAGED ACCESS DATA_RETENTION_TIME_IN_DAYS = 7;

Answer: D

Explanation:
A transient schema in Snowflake is designed without a Fail-safe period, meaning it does not incur additional storage costs once it leaves Time Travel, and it is not protected by Fail-safe in the event of a data loss. The WITH MANAGED ACCESS option ensures that all privilege grants, including future grants on objects within the schema, are managed by the schema owner, thus restricting object owners from passing on access to other users1.
References =
•Snowflake Documentation on creating schemas1
•Snowflake Documentation on configuring access control2
•Snowflake Documentation on understanding and viewing Fail-safe3

NEW QUESTION 34
A company has a source system that provides JSON records for various loT operations. The JSON Is loading directly into a persistent table with a variant field. The data Is quickly growing to 100s of millions of records and performance to becoming an issue. There is a generic access pattern that Is used to filter on the create_date key within the variant field.
What can be done to improve performance?

A. Alter the target table to Include additional fields pulled from the JSON record
B. This would Include a create_date field with a datatype of time stam
C. When this field Is used in the filter, partition pruning will occur.
D. Alter the target table to include additional fields pulled from the JSON record
E. This would include a create_date field with a datatype of varcha
F. When this field is used in the filter, partition pruning will occur.
G. Validate the size of the warehouse being use
H. If the record count is approaching 100s of millions, size XL will be the minimum size required to process this amount of data.
I. Incorporate the use of multiple tables partitioned by date range
J. When a user or process needs to query a particular date range, ensure the appropriate base table Is used.

Answer: A

Explanation:
? The correct answer is A because it improves the performance of queries by reducing the amount of data scanned and processed. By adding a create_date field with a timestamp data type, Snowflake can automatically cluster the table based on this field and prune the micro-partitions that do not match the filter condition. This avoids the need to parse the JSON data and access the variant field for every record.
? Option B is incorrect because it does not improve the performance of queries. By adding a create_date field with a varchar data type, Snowflake cannot automatically cluster the table based on this field and prune the micro-partitions that do not match the filter condition. This still requires parsing the JSON data and accessing the variant field for every record.
? Option C is incorrect because it does not address the root cause of the performance issue. By validating the size of the warehouse being used, Snowflake can adjust the compute resources to match the data volume and parallelize the query execution. However, this does not reduce the amount of data scanned and processed, which is the main bottleneck for queries on JSON data.
? Option D is incorrect because it adds unnecessary complexity and overhead to the data loading and querying process. By incorporating the use of multiple tables partitioned by date ranges, Snowflake can reduce the amount of data scanned and processed for queries that specify a date range. However, this requires creating and maintaining multiple tables, loading data into the appropriate table based on
the date, and joining the tables for queries that span multiple date ranges. References:
? Snowflake Documentation: Loading Data Using Snowpipe: This document
explains how to use Snowpipe to continuously load data from external sources into Snowflake tables. It also describes the syntax and usage of the COPY INTO command, which supports various options and parameters to control the loading behavior, such as ON_ERROR, PURGE, and SKIP_FILE.
? Snowflake Documentation: Date and Time Data Types and Functions: This
document explains the different data types and functions for working with date and time values in Snowflake. It also describes how to set and change the session timezone and the system timezone.
? Snowflake Documentation: Querying Metadata: This document explains how to
query the metadata of the objects and operations in Snowflake using various functions, views, and tables. It also describes how to access the copy history information using the COPY_HISTORY function or the COPY_HISTORY view.
? Snowflake Documentation: Loading JSON Data: This document explains how to
load JSON data into Snowflake tables using various methods, such as the COPY INTO command, the INSERT command, or the PUT command. It also describes how to access and query JSON data using the dot notation, the FLATTEN function, or the LATERAL join.
? Snowflake Documentation: Optimizing Storage for Performance: This document
explains how to optimize the storage of data in Snowflake tables to improve the performance of queries. It also describes the concepts and benefits of automatic clustering, search optimization service, and materialized views.

NEW QUESTION 37
What are characteristics of Dynamic Data Masking? (Select TWO).

A. A masking policy that Is currently set on a table can be dropped.
B. A single masking policy can be applied to columns in different tables.
C. A masking policy can be applied to the value column of an external table.
D. The role that creates the masking policy will always see unmasked data In query results
E. A masking policy can be applied to a column with the GEOGRAPHY data type.

Answer: AB

Explanation:
Dynamic Data Masking is a feature that allows masking sensitive data in query results based on the role of the user who executes the query. A masking policy is a user-defined function that specifies the masking logic and can be applied to one or more columns in one or more tables. A masking policy that is currently set on a table can be dropped using the ALTER TABLE command. A single masking policy can be applied to columns in different tables using the ALTER TABLE command with the SET MASKING POLICY clause. The other options are either incorrect or not supported by Snowflake. A masking policy cannot be applied to the value column of an external table, as external tables do not support column-level security. The role that creates the masking policy will not always see unmasked data in query results, as the masking policy can be applied to the owner role as well. A masking policy cannot be applied to a column with the GEOGRAPHY data type, as Snowflake only supports masking policies for scalar data types. References: Snowflake Documentation: Dynamic Data Masking, Snowflake Documentation: ALTER TABLE

NEW QUESTION 40
What step will im the performance of queries executed against an external table?

A. Partition the external table.
B. Shorten the names of the source files.
C. Convert the source files' character encoding to UTF-8.
D. Use an internal stage instead of an external stage to store the source files.

Answer: A

Explanation:
Partitioning an external table is a technique that improves the performance of queries executed against the table by reducing the amount of data scanned. Partitioning an external table involves creating one or more partition columns that define how the table is logically divided into subsets of data based on the values in those columns. The partition columns can be derived from the file metadata (such as file name, path, size, or modification time) or from the file content (such as a column value or a JSON
attribute). Partitioning an external table allows the query optimizer to prune the files that do not match the query predicates, thus avoiding unnecessary data scanning and processing2 The other options are not effective steps for improving the performance of queries
executed against an external table:
? Shorten the names of the source files. This option does not have any impact on the query performance, as the file names are not used for query processing. The file names are only used for creating the external table and displaying the query results3
? Convert the source files?? character encoding to UTF-8. This option does not affect the query performance, as Snowflake supports various character encodings for external table files, such as UTF-8, UTF-16, UTF-32, ISO-8859-1, and Windows- 1252. Snowflake automatically detects the character encoding of the files and converts them to UTF-8 internally for query processing4
? Use an internal stage instead of an external stage to store the source files. This
option is not applicable, as external tables can only reference files stored in external stages, such as Amazon S3, Google Cloud Storage, or Azure Blob Storage. Internal stages are used for loading data into internal tables, not external tables5 References:
? 1: SnowPro Advanced: Architect | Study Guide
? 2: Snowflake Documentation | Partitioning External Tables

? 3: Snowflake Documentation | Creating External Tables
? 4: Snowflake Documentation | Supported File Formats and Compression for Staged Data Files
? 5: Snowflake Documentation | Overview of Stages
? : SnowPro Advanced: Architect | Study Guide
? : Partitioning External Tables
? : Creating External Tables
? : Supported File Formats and Compression for Staged Data Files
? : Overview of Stages

NEW QUESTION 42
Which feature provides the capability to define an alternate cluster key for a table with an existing cluster key?

A. External table
B. Materialized view
C. Search optimization
D. Result cache

Answer: B

Explanation:
A materialized view is a feature that provides the capability to define an alternate cluster key for a table with an existing cluster key. A materialized view is a pre-computed result set that is stored in Snowflake and can be queried like a regular table. A materialized view can have a different cluster key than the base table, which can improve the performance and efficiency of queries on the materialized view. A materialized view can also support aggregations, joins, and filters on the base table data. A materialized view is automatically refreshed when the underlying data in the base table changes, as long as the AUTO_REFRESH parameter is set to true1.
References:
? Materialized Views | Snowflake Documentation

NEW QUESTION 44
A company is following the Data Mesh principles, including domain separation, and chose one Snowflake account for its data platform.
An Architect created two data domains to produce two data products. The Architect needs a third data domain that will use both of the data products to create an aggregate data product. The read access to the data products will be granted through a separate role.
Based on the Data Mesh principles, how should the third domain be configured to create the aggregate product if it has been granted the two read roles?

A. Use secondary roles for all users.
B. Create a hierarchy between the two read roles.
C. Request a technical ETL user with the sysadmin role.
D. Request that the two data domains share data using the Data Exchange.

Answer: D

Explanation:
In the scenario described, where a third data domain needs access to two existing data products in a Snowflake account structured according to Data Mesh principles, the best approach is to utilize Snowflake??s Data Exchange functionality. Option D is correct as it facilitates the sharing and governance of data across different domains efficiently and securely. Data Exchange allows domains to publish and subscribe to live data products, enabling real-time data collaboration and access management in a governed manner. This approach is in line with Data Mesh principles, which advocate for decentralized data ownership and architecture, enhancing agility and scalability across the organization.References:
? Snowflake Documentation on Data Exchange
? Articles on Data Mesh Principles in Data Management

NEW QUESTION 45
Which system functions does Snowflake provide to monitor clustering information within a table (Choose two.)

A. SYSTEM$CLUSTERING_INFORMATION
B. SYSTEM$CLUSTERING_USAGE
C. SYSTEM$CLUSTERING_DEPTH
D. SYSTEM$CLUSTERING_KEYS
E. SYSTEM$CLUSTERING_PERCENT

Answer: AC

Explanation:
According to the Snowflake documentation, these two system functions are provided by Snowflake to monitor clustering information within a table. A system function is a type of function that allows executing actions or returning information about the system. A clustering key is a feature that allows organizing data across micro-partitions based on one or more columns in the table. Clustering can improve query performance by reducing the number of files to scan.
? SYSTEM$CLUSTERING_INFORMATION is a system function that returns
clustering information, including average clustering depth, for a table based on one or more columns in the table. The function takes a table name and an optional column name or expression as arguments, and returns a JSON string with the clustering information. The clustering information includes the cluster by keys, the total partition count, the total constant partition count, the average overlaps, and the average depth1.
? SYSTEM$CLUSTERING_DEPTH is a system function that returns the clustering
depth for a table based on one or more columns in the table. The function takes a table name and an optional column name or expression as arguments, and returns an integer value with the clustering depth. The clustering depth is the maximum number of overlapping micro-partitions for any micro-partition in the table. A lower clustering depth indicates a better clustering2.
References:
? SYSTEM$CLUSTERING_INFORMATION | Snowflake Documentation
? SYSTEM$CLUSTERING_DEPTH | Snowflake Documentation

NEW QUESTION 47
A company is designing high availability and disaster recovery plans and needs to maximize redundancy and minimize recovery time objectives for their critical

application processes. Cost is not a concern as long as the solution is the best available. The plan so far consists of the following steps:
* 1. Deployment of Snowflake accounts on two different cloud providers.
* 2. Selection of cloud provider regions that are geographically far apart.
* 3. The Snowflake deployment will replicate the databases and account data between both cloud provider accounts.
* 4. Implementation of Snowflake client redirect.
What is the MOST cost-effective way to provide the HIGHEST uptime and LEAST application disruption if there is a service event?

A. Connect the applications using the <organization_name>-<connection_name> UR
B. Use the Business Critical Snowflake edition.
C. Connect the applications using the <organization_name>-<connection_name> UR
D. Use the Virtual Private Snowflake (VPS) edition.
E. Connect the applications using the <organization_name>-<accountLocator> UR
F. Use the Enterprise Snowflake edition.
G. Connect the applications using the <organization_name>-<accountLocator> UR
H. Use the Business Critical Snowflake edition.

Answer: D

Explanation:
To provide the highest uptime and least application disruption in case of a service event, the best option is to use the Business Critical Snowflake edition and connect the applications using the <organization_name>-<accountLocator> URL. The Business Critical Snowflake edition offers the highest level of security, performance, and availability for Snowflake accounts. It includes features such as customer-managed encryption keys, HIPAA compliance, and 4-hour RPO and RTO SLAs. It also supports account replication and failover across regions and cloud platforms, which enables business continuity and disaster recovery. By using the <organization_name>-<accountLocator> URL, the applications can leverage the Snowflake Client Redirect feature, which automatically redirects the client connections to the secondary account in case of a failover. This way, the applications can seamlessly switch to the backup account without any manual intervention or configuration changes. The other options are less cost-effective or less reliable because they either use a lower edition of Snowflake, which does not support account replication and failover, or they use the <organization_name>-<connection_name> URL, which does not support client redirect and requires manual updates to the connection string in case of a failover. References:
? [Snowflake Editions] 1
? [Replication and Failover/Failback] 2
? [Client Redirect] 3
? [Snowflake Account Identifiers] 4

NEW QUESTION 48
A company has a table with that has corrupted data, named Data. The company wants to recover the data as it was 5 minutes ago using cloning and Time Travel. What command will accomplish this?

A. CREATE CLONE TABLE Recover_Data FROM Data AT(OFFSET => -60*5);
B. CREATE CLONE Recover_Data FROM Data AT(OFFSET => -60*5);
C. CREATE TABLE Recover_Data CLONE Data AT(OFFSET => -60*5);
D. CREATE TABLE Recover Data CLONE Data AT(TIME => -60*5);

Answer: C

Explanation:
This is the correct command to create a clone of the table Data as it was 5 minutes ago using cloning and Time Travel. Cloning is a feature that allows creating a copy of a database, schema, table, or view without duplicating the data or metadata. Time Travel is a feature that enables accessing historical data (i.e. data that has been changed or deleted) at any point within a defined period. To create a clone of a table at a point in time in the past, the syntax is:
CREATE TABLE <clone_name> CLONE <source_table> AT (OFFSET =>
<offset_in_seconds>);
The OFFSET parameter specifies the time difference in seconds from the present time. A negative value indicates a point in the past. For example, -60*5 means 5 minutes ago. Alternatively, the TIMESTAMP parameter can be used to specify an exact timestamp in the past. The clone will contain the data as it existed in the source table at the specified point in time12.
References:
? Snowflake Documentation: Cloning Objects
? Snowflake Documentation: Cloning Objects at a Point in Time in the Past

NEW QUESTION 50
A company??s daily Snowflake workload consists of a huge number of concurrent queries triggered between 9pm and 11pm. At the individual level, these queries are smaller
statements that get completed within a short time period.
What configuration can the company??s Architect implement to enhance the performance of this workload? (Choose two.)

A. Enable a multi-clustered virtual warehouse in maximized mode during the workload duration.
B. Set the MAX_CONCURRENCY_LEVEL to a higher value than its default value of 8 at the virtual warehouse level.
C. Increase the size of the virtual warehouse to size X-Large.
D. Reduce the amount of data that is being processed through this workload.
E. Set the connection timeout to a higher value than its default.

Answer: AB

Explanation:
These two configuration options can enhance the performance of the workload that consists of a huge number of concurrent queries that are smaller and faster.
? Enabling a multi-clustered virtual warehouse in maximized mode allows the
warehouse to scale out automatically by adding more clusters as soon as the current cluster is fully loaded, regardless of the number of queries in the queue. This can improve the concurrency and throughput of the workload by minimizing or preventing queuing. The maximized mode is suitable for workloads that require high performance and low latency, and are less sensitive to credit consumption1.
? Setting the MAX_CONCURRENCY_LEVEL to a higher value than its default value
of 8 at the virtual warehouse level allows the warehouse to run more queries concurrently on each cluster. This can improve the utilization and efficiency of the warehouse resources, especially for smaller and faster queries that do not require a lot of processing power. The MAX_CONCURRENCY_LEVEL parameter can be set when creating or modifying a warehouse, and it can be changed at any time2.

References:
? Snowflake Documentation: Scaling Policy for Multi-cluster Warehouses
? Snowflake Documentation: MAX_CONCURRENCY_LEVEL

NEW QUESTION 55
A Developer is having a performance issue with a Snowflake query. The query receives up to 10 different values for one parameter and then performs an aggregation over the majority of a fact table. It then
joins against a smaller dimension table. This parameter value is selected by the different query users when they execute it during business hours. Both the fact and dimension tables are loaded with new data in an overnight import process.
On a Small or Medium-sized virtual warehouse, the query performs slowly. Performance is acceptable on a size Large or bigger warehouse. However, there is no budget to increase costs. The Developer
needs a recommendation that does not increase compute costs to run this query. What should the Architect recommend?

A. Create a task that will run the 10 different variations of the query corresponding to the 10 different parameters before the users come in to wor
B. The query results will then be cached and ready to respond quickly when the users re-issue the query.
C. Create a task that will run the 10 different variations of the query corresponding to the 10 different parameters before the users come in to wor
D. The task will be scheduled to align with the users' working hours in order to allow the warehouse cache to be used.
E. Enable the search optimization service on the tabl
F. When the users execute the query, the search optimization service will automatically adjust the query execution plan based on the frequently-used parameters.
G. Create a dedicated size Large warehouse for this particular set of querie
H. Create a new role that has USAGE permission on this warehouse and has the appropriate read permissions over the fact and dimension table
I. Have users switch to this role and use this warehouse when they want to access this data.

Answer: C

Explanation:
Enabling the search optimization service on the table can improve the performance of queries that have selective filtering criteria, which seems to be the case here. This service optimizes the execution of queries by creating a persistent data structure called a search access path, which allows some micro-partitions to be skipped during the scanning process. This can significantly speed up query performance without increasing compute costs1.
References
•Snowflake Documentation on Search Optimization Service1.

NEW QUESTION 56
An Architect needs to design a data unloading strategy for Snowflake, that will be used with the COPY INTO <location> command.
Which configuration is valid?

A. Location of files: Snowflake internal locatio
B. File formats: CSV, XM
C. File encoding: UTF-8. Encryption: 128-bit
D. Location of files: Amazon S3. File formats: CSV, JSO
E. File encoding: Latin-1 (ISO-8859). Encryption: 128-bit
F. Location of files: Google Cloud Storag
G. File formats: Parque
H. File encoding: UTF-8· Compression: gzip
I. Location of files: Azure ADL
J. File formats: JSON, XML, Avro, Parquet, OR
K. Compression: bzip2. Encryption: User-supplied key

Answer: C

Explanation:
For the configuration of data unloading in Snowflake, the valid option among the provided choices is "C." This is because Snowflake supports unloading data into Google Cloud Storage using the COPY INTO <location> command with specific configurations. The configurations listed in option C, such as Parquet file format with UTF-8 encoding and gzip compression, are all supported by Snowflake. Notably, Parquet is a columnar storage file format, which is optimal for high-performance data processing tasks in Snowflake. The UTF-8 file encoding and gzip compression are both standard and widely used settings that are compatible with Snowflake??s capabilities for data unloading to cloud storage platforms.References:
? Snowflake Documentation on COPY INTO command
? Snowflake Documentation on Supported File Formats
? Snowflake Documentation on Compression and Encoding Options

NEW QUESTION 57
Which columns can be included in an external table schema? (Select THREE).

A. VALUE
B. METADATASROW_ID
C. METADATASISUPDATE
D. METADAT A$ FILENAME
E. METADATAS FILE_ROW_NUMBER
F. METADATASEXTERNAL TABLE PARTITION

Answer: ADE

Explanation:
An external table schema defines the columns and data types of the data stored in an external stage. All external tables include the following columns by default:
? VALUE: A VARIANT type column that represents a single row in the external file.
? METADATA$FILENAME: A pseudocolumn that identifies the name of each staged data file included in the external table, including its path in the stage.
? METADATA$FILE_ROW_NUMBER: A pseudocolumn that shows the row number for each record in a staged data file.
You can also create additional virtual columns as expressions using the VALUE column and/or the pseudocolumns. However, the following columns are not valid for external tables and cannot be included in the schema:
? METADATASROW_ID: This column is only available for internal tables and shows

the unique identifier for each row in the table.
? METADATASISUPDATE: This column is only available for internal tables and shows whether the row was inserted or updated by a merge operation.
? METADATASEXTERNAL TABLE PARTITION: This column is not a valid column name and does not exist in Snowflake.
References: Introduction to External Tables, CREATE EXTERNAL TABLE

NEW QUESTION 60
What is a characteristic of Role-Based Access Control (RBAC) as used in Snowflake?

A. Privileges can be granted at the database level and can be inherited by all underlying objects.
B. A user can use a "super-user" access along with securityadmin to bypass authorization checks and access all databases, schemas, and underlying objects.
C. A user can create managed access schemas to support future grants and ensure only schema owners can grant privileges to other roles.
D. A user can create managed access schemas to support current and future grants and ensure only object owners can grant privileges to other roles.

Answer: C

Explanation:
Role-Based Access Control (RBAC) is the Snowflake Access Control Framework that allows privileges to be granted by object owners to roles, and roles, in turn, can be assigned to users to restrict or allow actions to be performed on objects. A characteristic of RBAC as used in Snowflake is:
? Privileges can be granted at the database level and can be inherited by all
underlying objects. This means that a role that has a certain privilege on a database, such as CREATE SCHEMA or USAGE, can also perform the same action on any schema, table, view, or other object within that database, unless explicitly revoked. This simplifies the access control management and reduces the number of grants required.
? A user can create managed access schemas to support future grants and ensure
only schema owners can grant privileges to other roles. This means that a user can create a schema with the MANAGED ACCESS option, which changes the default behavior of object ownership and privilege granting within the schema. In a managed access schema, object owners lose the ability to grant privileges on their objects to other roles, and only the schema owner or a role with the MANAGE GRANTS privilege can do so. This enhances the security and governance of the schema and its objects.
The other options are not characteristics of RBAC as used in Snowflake:
? A user can use a ??super-user?? access along with securityadmin to bypass authorization checks and access all databases, schemas, and underlying objects. This is not true, as there is no such thing as a ??super-user?? access in Snowflake. The securityadmin role is a predefined role that can manage users and roles, but it does not have any privileges on any database objects by default. To access any object, the securityadmin role must be explicitly granted the appropriate privilege by the object owner or another role with the grant option.
? A user can create managed access schemas to support current and future grants and ensure only object owners can grant privileges to other roles. This is not true, as this contradicts the definition of a managed access schema. In a managed access schema, object owners cannot grant privileges on their objects to other roles, and only the schema owner or a role with the MANAGE GRANTS privilege can do so.
References:
? Overview of Access Control
? A Functional Approach For Snowflake??s Role-Based Access Controls
? Snowflake Role-Based Access Control simplified
? Snowflake RBAC security prefers role inheritance to role composition
? Overview of Snowflake Role Based Access Control

NEW QUESTION 64
Company A has recently acquired company B. The Snowflake deployment for company B is located in the Azure West Europe region.
As part of the integration process, an Architect has been asked to consolidate company B's sales data into company A's Snowflake account which is located in the AWS us-east-1 region.
How can this requirement be met?

A. Replicate the sales data from company B's Snowflake account into company A's Snowflake account using cross-region data replication within Snowflak
B. Configure a direct share from company B's account to company A's account.
C. Export the sales data from company B's Snowflake account as CSV files, and transfer the files to company A's Snowflake accoun
D. Import the data using Snowflake's data loading capabilities.
E. Migrate company B's Snowflake deployment to the same region as company A's Snowflake deployment, ensuring data localit
F. Then perform a direct database-to-database merge of the sales data.
G. Build a custom data pipeline using Azure Data Factory or a similar tool to extract the sales data from company B's Snowflake accoun
H. Transform the data, then load it into company A's Snowflake account.

Answer: A

Explanation:
The best way to meet the requirement of consolidating company B??s sales data into company A??s Snowflake account is to use cross-region data replication within
Snowflake. This feature allows data providers to securely share data with data consumers across different regions and cloud platforms. By replicating the sales data from company B??s account in Azure West Europe region to company A??s account in AWS us-east-1 region, the data will be synchronized and available for consumption. To enable data replication, the accounts must be linked and replication must be enabled by a user with the ORGADMIN role. Then, a replication group must be created and the sales database must be added to the group. Finally, a direct share must be configured from company B??s account to company A??s account to grant access to the replicated data. This option is more efficient and secure than exporting and importing data using CSV files or migrating the entire Snowflake deployment to another region or cloud platform. It also does not require building a custom data pipeline using external tools.
References:
? Sharing data securely across regions and cloud platforms
? Introduction to replication and failover
? Replication considerations
? Replicating account objects

NEW QUESTION 65
Why might a Snowflake Architect use a star schema model rather than a 3NF model when designing a data architecture to run in Snowflake? (Select TWO).

A. Snowflake cannot handle the joins implied in a 3NF data model.
B. The Architect wants to remove data duplication from the data stored in Snowflake.
C. The Architect is designing a landing zone to receive raw data into Snowflake.

D. The BI tool needs a data model that allows users to summarize facts across different dimensions, or to drill down from the summaries.
E. The Architect wants to present a simple flattened single view of the data to a particular group of end users.

Answer: DE

Explanation:
A star schema model is a type of dimensional data model that consists of a single fact table and multiple dimension tables. A 3NF model is a type of relational data model that follows the third normal form, which eliminates data redundancy and ensures referential integrity. A Snowflake Architect might use a star schema model rather than a 3NF model when designing a data architecture to run in Snowflake for the following reasons:
? A star schema model is more suitable for analytical queries that require
aggregating and slicing data across different dimensions, such as those performed by a BI tool. A 3NF model is more suitable for transactional queries that require inserting, updating, and deleting individual records.
? A star schema model is simpler and faster to query than a 3NF model, as it involves fewer joins and less complex SQL statements. A 3NF model is more complex and slower to query, as it involves more joins and more complex SQL statements.
? A star schema model can provide a simple flattened single view of the data to a
particular group of end users, such as business analysts or data scientists, who need to explore and visualize the data. A 3NF model can provide a more detailed and normalized view of the data to a different group of end users, such as application developers or data engineers, who need to maintain and update the data.
The other options are not valid reasons for choosing a star schema model over a 3NF model in Snowflake:
? Snowflake can handle the joins implied in a 3NF data model, as it supports ANSI
SQL and has a powerful query engine that can optimize and execute complex queries efficiently.
? The Architect can use both star schema and 3NF models to remove data
duplication from the data stored in Snowflake, as both models can enforce data integrity and avoid data anomalies. However, the trade-off is that a star schema model may have more data redundancy than a 3NF model, as it denormalizes the data for faster query performance, while a 3NF model may have less data redundancy than a star schema model, as it normalizes the data for easier data maintenance.
? The Architect can use both star schema and 3NF models to design a landing zone
to receive raw data into Snowflake, as both models can accommodate different types of data sources and formats. However, the choice of the model may depend on the purpose and scope of the landing zone, such as whether it is a temporary or permanent storage, whether it is a staging area or a data lake, and whether it is a single source or a multi-source integration.
References:
? Snowflake Architect Training
? Data Modeling: Understanding the Star and Snowflake Schemas
? Data Vault vs Star Schema vs Third Normal Form: Which Data Model to Use?
? Star Schema vs Snowflake Schema: 5 Key Differences
? Dimensional Data Modeling - Snowflake schema
? Star schema vs Snowflake Schema

NEW QUESTION 66
A user has the appropriate privilege to see unmasked data in a column.
If the user loads this column data into another column that does not have a masking policy, what will occur?

A. Unmasked data will be loaded in the new column.
B. Masked data will be loaded into the new column.
C. Unmasked data will be loaded into the new column but only users with the appropriate privileges will be able to see the unmasked data.
D. Unmasked data will be loaded into the new column and no users will be able to see the unmasked data.

Answer: A

Explanation:
According to the SnowPro Advanced: Architect documents and learning resources, column masking policies are applied at query time based on the privileges of the user who runs the query. Therefore, if a user has the privilege to see unmasked data in a column, they will see the original data when they query that column. If they load this
column data into another column that does not have a masking policy, the unmasked data will be loaded in the new column, and any user who can query the new column will see the unmasked data as well. The masking policy does not affect the underlying data in the column, only the query results.
References:
? Snowflake Documentation: Column Masking
? Snowflake Learning: Column Masking

NEW QUESTION 71
A new table and streams are created with the following commands: CREATE OR REPLACE TABLE LETTERS (ID INT, LETTER STRING) ;
CREATE OR REPLACE STREAM STREAM_1 ON TABLE LETTERS;
CREATE OR REPLACE STREAM STREAM_2 ON TABLE LETTERS APPEND_ONLY = TRUE;
The following operations are processed on the newly created table: INSERT INTO LETTERS VALUES (1, 'A');
INSERT INTO LETTERS VALUES (2, 'B'); INSERT INTO LETTERS VALUES (3, 'C');
TRUNCATE TABLE LETTERS;
INSERT INTO LETTERS VALUES (4, 'D'); INSERT INTO LETTERS VALUES (5, 'E'); INSERT INTO LETTERS VALUES (6, 'F'); DELETE FROM LETTERS WHERE ID = 6;
What would be the output of the following SQL commands, in order? SELECT COUNT (*) FROM STREAM_1;
SELECT COUNT (*) FROM STREAM_2;

A. 2 & 6
B. 2 & 3
C. 4 & 3
D. 4 & 6

Answer: C

Explanation:
In Snowflake, a stream records data manipulation language (DML) changes to its base table since the stream was created or last consumed. STREAM_1 will show all changes including the TRUNCATE operation, while STREAM_2, being APPEND_ONLY, will not show deletions like TRUNCATE. Therefore, STREAM_1 will count the three inserts, the TRUNCATE (counted as a single operation), and the subsequent two inserts before the delete, totaling 4. STREAM_2 will only count the three initial inserts and the two after the TRUNCATE, totaling 3, as it does not count the TRUNCATE or the delete operation. References: The explanation is

based on the Snowflake documentation on streams, which details how streams track changes and the difference between standard and APPEND_ONLY streams12.

NEW QUESTION 73
Which SQL alter command will MAXIMIZE memory and compute resources for a Snowpark stored procedure when executed on the snowpark_opt_wh warehouse?

A)
```
alter warehouse snowpark_opt_wh set max_concurrency_level = 1;
```

B)
```
alter warehouse snowpark_opt_wh set max_concurrency_level = 2;
```

C)
```
alter warehouse snowpark_opt_wh set max_concurrency_level = 8;
```

D)
```
alter warehouse snowpark_opt_wh set max_concurrency_level = 16;
```

A. Option A
B. Option B
C. Option C
D. Option D

Answer: A

Explanation:
To maximize memory and compute resources for a Snowpark stored procedure, you need to set the MAX_CONCURRENCY_LEVEL parameter for the warehouse that executes the stored procedure. This parameter determines the maximum number of concurrent queries that can run on a single warehouse. By setting it to 16, you ensure that the warehouse can use all the available CPU cores and memory on a single node, which is the optimal configuration for Snowpark-optimized warehouses. This will improve the performance and efficiency of the stored procedure, as it will not have to share resources with other queries or nodes. The other options are incorrect because they either do not change the MAX_CONCURRENCY_LEVEL parameter, or they set it to a lower value than 16, which will reduce the memory and compute resources for the stored procedure. References:
? [Snowpark-optimized Warehouses] 1
? [Training Machine Learning Models with Snowpark Python] 2
? [Snowflake Shorts: Snowpark Optimized Warehouses] 3

NEW QUESTION 76
An Architect is troubleshooting a query with poor performance using the QUERY_HIST0RY function. The Architect observes that the COMPILATIONJHME is greater than the EXECUTIONJTIME.
What is the reason for this?

A. The query is processing a very large dataset.
B. The query has overly complex logic.
C. The query is queued for execution.
D. The query is reading from remote storage.

Answer: B

Explanation:
Compilation time is the time it takes for the optimizer to create an optimal query plan for the efficient execution of the query. It also involves some pruning of partition files, making the query execution efficient2
If the compilation time is greater than the execution time, it means that the optimizer spent more time analyzing the query than actually running it. This could indicate that the query has overly complex logic, such as multiple joins, subqueries, aggregations, or
expressions. The complexity of the query could also affect the size and quality of the query plan, which could impact the performance of the query3
To reduce the compilation time, the Architect can try to simplify the query logic, use views or common table expressions (CTEs) to break down the query into smaller parts, or use
hints to guide the optimizer. The Architect can also use the EXPLAIN command to examine the query plan and identify potential bottlenecks or inefficiencies4
References:
? 1: SnowPro Advanced: Architect | Study Guide 5
? 2: Snowflake Documentation | Query Profile Overview 6
? 3: Understanding Why Compilation Time in Snowflake Can Be Higher than Execution Time 7
? 4: Snowflake Documentation | Optimizing Query Performance 8
? : SnowPro Advanced: Architect | Study Guide
? : Query Profile Overview
? : Understanding Why Compilation Time in Snowflake Can Be Higher than Execution Time
? : Optimizing Query Performance

NEW QUESTION 77
An Architect has a design where files arrive every 10 minutes and are loaded into a primary database table using Snowpipe. A secondary database is refreshed every hour with the latest data from the primary database.
Based on this scenario, what Time Travel query options are available on the secondary database?

A. A query using Time Travel in the secondary database is available for every hourly table version within the retention window.
B. A query using Time Travel in the secondary database is available for every hourly table version within and outside the retention window.
C. Using Time Travel, secondary database users can query every iterative version within each hour (the individual Snowpipe loads) in the retention window.

D. Using Time Travel, secondary database users can query every iterative version within each hour (the individual Snowpipe loads) and outside the retention window.

Answer: A

Explanation:
Snowflake??s Time Travel feature allows users to query historical data within a defined retention period. In the given scenario, since the secondary database is refreshed every hour, Time Travel can be used to query each hourly version of the table as long as it falls within the retention window. This does not include individual Snowpipe loads within each hour unless they coincide with the hourly refresh.
References: The answer is verified using Snowflake??s official documentation, which provides detailed information on Time Travel and its usage within the retention period123.

NEW QUESTION 82
Two queries are run on the customer_address table:
create or replace TABLE CUSTOMER_ADDRESS (CA_ADDRESS_SK NUMBER(38,0), CA_ADDRESS_ID VARCHAR(16), CA_STREET_NUMBER VARCHAR(IO) CA_STREET_NAME VARCHAR(60), CA_STREET_TYPE VARCHAR(15), CA_SUITE_NUMBER VARCHAR(10), CA_CITY VARCHAR(60), CA_COUNTY
VARCHAR(30), CA_STATE VARCHAR(2), CA_ZIP VARCHAR(10), CA_COUNTRY VARCHAR(20), CA_GMT_OFFSET NUMBER(5,2), CA_LOCATION_TYPE VARCHAR(20));
ALTER TABLE DEMO_DB.DEMO_SCH.CUSTOMER_ADDRESS ADD SEARCH OPTIMIZATION ON SUBSTRING(CA_ADDRESS_ID);
Which queries will benefit from the use of the search optimization service? (Select TWO).

A. select * from DEMO_DB.DEMO_SCH.CUSTOMER_ADDRESS Where substring(CA_ADDRESS_ID,1,8)= substring('AAAAAAAAPHPPLBAAASKDJHASLKDJHASKJD',1,8);
B. select * from DEMO_DB.DEMO_SCH.CUSTOMER_ADDRESS Where CA_ADDRESS_ID= substring('AAAAAAAAPHPPLBAAASKDJHASLKDJHASKJD',1,16);
C. select*fromDEMO_DB.DEMO_SCH.CUSTOMER_ADDRESSWhereCA_ADDRESS_IDLIKE ??%BAAASKD%';
D. select*fromDEMO_DB.DEMO_SCH.CUSTOMER_ADDRESSWhereCA_ADDRESS_IDLIK E '%PHPP%';
E. select*fromDEMO_DB.DEMO_SCH.CUSTOMER_ADDRESSWhereCA_ADDRESS_IDNO T LIKE '%AAAAAAAAPHPPL%';

Answer: AB

Explanation:
The use of the search optimization service in Snowflake is particularly effective when queries involve operations that match exact substrings or start from the beginning of a string. The ALTER TABLE command adding search optimization specifically for substrings on the CA_ADDRESS_ID field allows the service to create an optimized search path for queries using substring matches.
? Option A benefits because it directly matches a substring from the start of the
CA_ADDRESS_ID, aligning with the optimization's capability to quickly locate records based on the beginning segments of strings.
? Option B also benefits, despite performing a full equality check, because it
essentially compares the full length of CA_ADDRESS_ID to a substring, which can leverage the substring index for efficient retrieval.Options C, D, and E involve patterns that do not start from the beginning of the string or use negations, which are not optimized by the search optimization service configured for starting substring matches.References: Snowflake's documentation on the use of search optimization for substring matching in SQL queries.

NEW QUESTION 83
When loading data into a table that captures the load time in a column with a default value of either CURRENT_TIME () or CURRENT_TIMESTAMP () what will occur?

A. All rows loaded using a specific COPY statement will have varying timestamps based on when the rows were inserted.
B. Any rows loaded using a specific COPY statement will have varying timestamps based on when the rows were read from the source.
C. Any rows loaded using a specific COPY statement will have varying timestamps based on when the rows were created in the source.
D. All rows loaded using a specific COPY statement will have the same timestamp value.

Answer: D

Explanation:
When using the COPY command to load data into Snowflake, if a column has a default value set to CURRENT_TIME() or CURRENT_TIMESTAMP(), all rows loaded by that specific COPY command will have the same timestamp. This is because the
default value for the timestamp is evaluated at the start of the COPY operation, and that same value is applied to all rows loaded by that operation.
References: This behavior is consistent with Snowflake??s documentation on the CURRENT_TIMESTAMP function, which specifies that the timestamp is captured at the time the statement is executed1.

NEW QUESTION 85
A healthcare company is deploying a Snowflake account that may include Personal Health Information (PHI). The company must ensure compliance with all relevant privacy standards.
Which best practice recommendations will meet data protection and compliance requirements? (Choose three.)

A. Use, at minimum, the Business Critical edition of Snowflake.
B. Create Dynamic Data Masking policies and apply them to columns that contain PHI.
C. Use the Internal Tokenization feature to obfuscate sensitive data.
D. Use the External Tokenization feature to obfuscate sensitive data.
E. Rewrite SQL queries to eliminate projections of PHI data based on current_role().
F. Avoid sharing data with partner organizations.

Answer: ABD

Explanation:
? A healthcare company that handles PHI data must ensure compliance with relevant privacy standards, such as HIPAA, HITRUST, and GDPR. Snowflake provides several features and best practices to help customers meet their data protection and compliance requirements1.
? One best practice recommendation is to use, at minimum, the Business Critical edition of Snowflake. This edition provides the highest level of data protection and security, including end-to-end encryption with customer-managed keys, enhanced object-level security, and HIPAA and HITRUST compliance2. Therefore, option A is correct.

? Another best practice recommendation is to create Dynamic Data Masking policies and apply them to columns that contain PHI. Dynamic Data Masking is a feature that allows masking or redacting sensitive data based on the current user??s
role. This way, only authorized users can view the unmasked data, while others will see masked values, such as NULL, asterisks, or random characters3. Therefore, option B is correct.
? A third best practice recommendation is to use the External Tokenization feature to
obfuscate sensitive data. External Tokenization is a feature that allows replacing
sensitive data with tokens that are generated and stored by an external service, such as Protegrity. This way, the original data is never stored or processed by Snowflake, and only authorized users can access the tokenized data through the external service4. Therefore, option D is correct.
? Option C is incorrect, because the Internal Tokenization feature is not available in
Snowflake. Snowflake does not provide any native tokenization functionality, but only supports integration with external tokenization services4.
? Option E is incorrect, because rewriting SQL queries to eliminate projections of
PHI data based on current_role() is not a best practice. This approach is error- prone, inefficient, and hard to maintain. A better alternative is to use Dynamic Data Masking policies, which can automatically mask data based on the user??s role without modifying the queries3.
? Option F is incorrect, because avoiding sharing data with partner organizations is
not a best practice. Snowflake enables secure and governed data sharing with internal and external consumers, such as business units, customers, or partners. Data sharing does not involve copying or moving data, but only granting access privileges to the shared objects. Data sharing can also leverage Dynamic Data Masking and External Tokenization features to protect sensitive data5.
References: : Snowflake??s Security & Compliance Reports : Snowflake Editions : Dynamic Data Masking : External Tokenization : Secure Data Sharing

NEW QUESTION 88
A company has a Snowflake environment running in AWS us-west-2 (Oregon). The company needs to share data privately with a customer who is running their Snowflake environment in Azure East US 2 (Virginia).
What is the recommended sequence of operations that must be followed to meet this requirement?

A. * 1. Create a share and add the database privileges to the share* 2. Create a new listing on the Snowflake Marketplace* 3. Alter the listing and add the share* 4. Instruct the customer to subscribe to the listing on the Snowflake Marketplace
B. * 1. Ask the customer to create a new Snowflake account in Azure EAST US 2 (Virginia)* 2. Create a share and add the database privileges to the share* 3. Alter the share and add the customer's Snowflake account to the share
C. * 1. Create a new Snowflake account in Azure East US 2 (Virginia)* 2. Set up replication between AWS us-west-2 (Oregon) and Azure East US 2 (Virginia) for the database objects to be shared* 3. Create a share and add the database privileges to the share* 4. Alter the share and add the customer's Snowflake account to the share
D. * 1. Create a reader account in Azure East US 2 (Virginia)* 2. Create a share and add the database privileges to the share* 3. Add the reader account to the share* 4. Share the reader account's URL and credentials with the customer

Answer: C

Explanation:
Option C is the correct answer because it allows the company to share data privately with the customer across different cloud platforms and regions. The company can create a new Snowflake account in Azure East US 2 (Virginia) and set up replication between AWS us-west-2 (Oregon) and Azure East US 2 (Virginia) for the database objects to be shared. This way, the company can ensure that the data is always up to date and consistent in both accounts. The company can then create a share and add the database privileges to the share, and alter the share and add the customer??s Snowflake account to the share. The customer can then access the shared data from their own Snowflake account in Azure East US 2 (Virginia).
Option A is incorrect because the Snowflake Marketplace is not a private way of sharing data. The Snowflake Marketplace is a public data exchange platform that allows anyone to browse and subscribe to data sets from various providers. The company would not be able to control who can access their data if they use the Snowflake Marketplace.
Option B is incorrect because it requires the customer to create a new Snowflake account in Azure East US 2 (Virginia), which may not be feasible or desirable for the customer. The customer may already have an existing Snowflake account in a different cloud platform or region, and may not want to incur additional costs or complexity by creating a new account. Option D is incorrect because it involves creating a reader account in Azure East US 2 (Virginia), which is a limited and temporary way of sharing data. A reader account is a special type of Snowflake account that can only access data from a single share, and has a fixed duration of 30 days. The company would have to manage the reader account??s URL and credentials, and renew the account every 30 days. The customer would not be able to use their own Snowflake account to access the shared data, and would have to rely on the company??s reader account.
References:
? Snowflake Replication
? Secure Data Sharing Overview
? Snowflake Marketplace Overview
? Reader Account Overview

NEW QUESTION 90
Which of the following are characteristics of how row access policies can be applied to external tables? (Choose three.)

A. An external table can be created with a row access policy, and the policy can be applied to the VALUE column.
B. A row access policy can be applied to the VALUE column of an existing external table.
C. A row access policy cannot be directly added to a virtual column of an external table.
D. External tables are supported as mapping tables in a row access policy.
E. While cloning a database, both the row access policy and the external table will be cloned.
F. A row access policy cannot be applied to a view created on top of an external table.

Answer: ABC

Explanation:
These three statements are true according to the Snowflake documentation and the web search results. A row access policy is a feature that allows filtering rows based on user-defined conditions. A row access policy can be applied to an external table, which is a table that reads data from external files in a stage. However, there are some limitations and considerations for using row access policies with external tables.
? An external table can be created with a row access policy by using the WITH ROW
ACCESS POLICY clause in the CREATE EXTERNAL TABLE statement. The policy can be applied to the VALUE column, which is the column that contains the raw data from the external files in a VARIANT data type1.
? A row access policy can also be applied to the VALUE column of an existing
external table by using the ALTER TABLE statement with the SET ROW ACCESS POLICY clause2.
? A row access policy cannot be directly added to a virtual column of an external
table. A virtual column is a column that is derived from the VALUE column using
an expression. To apply a row access policy to a virtual column, the policy must be applied to the VALUE column and the expression must be repeated in the

policy definition3.
? External tables are not supported as mapping tables in a row access policy. A
mapping table is a table that is used to determine the access rights of users or roles based on some criteria. Snowflake does not support using an external table as a mapping table because it may cause performance issues or errors4.
? While cloning a database, Snowflake clones the row access policy, but not the
external table. Therefore, the policy in the cloned database refers to a table that is not present in the cloned database. To avoid this issue, the external table must be manually cloned or recreated in the cloned database4.
? A row access policy can be applied to a view created on top of an external table.
The policy can be applied to the view itself or to the underlying external
table. However, if the policy is applied to the view, the view must be a secure view, which is a view that hides the underlying data and the view definition from unauthorized users5.
References:
? CREATE EXTERNAL TABLE | Snowflake Documentation
? ALTER EXTERNAL TABLE | Snowflake Documentation
? Understanding Row Access Policies | Snowflake Documentation
? Snowflake Data Governance: Row Access Policy Overview
? Secure Views | Snowflake Documentation

NEW QUESTION 91
Which of the following ingestion methods can be used to load near real-time data by using the messaging services provided by a cloud provider?

A. Snowflake Connector for Kafka
B. Snowflake streams
C. Snowpipe
D. Spark

Answer: A

Explanation:
Snowflake Connector for Kafka and Snowpipe are two ingestion methods that can be used to load near real-time data by using the messaging services provided by a cloud provider. Snowflake Connector for Kafka enables you to stream structured and semi- structured data from Apache Kafka topics into Snowflake tables. Snowpipe enables you to load data from files that are continuously added to a cloud storage location, such as Amazon S3 or Azure Blob Storage. Both methods leverage Snowflake??s micro-partitioning and columnar storage to optimize data ingestion and query performance. Snowflake streams and Spark are not ingestion methods, but rather components of the Snowflake architecture. Snowflake streams provide change data capture (CDC) functionality by tracking data changes in a table. Spark is a distributed computing framework that can be used to process large-scale data and write it to Snowflake using the Snowflake Spark Connector.
References:
? Snowflake Connector for Kafka
? Snowpipe
? Snowflake Streams
? Snowflake Spark Connector

NEW QUESTION 96
What Snowflake system functions are used to view and or monitor the clustering metadata for a table? (Select TWO).

A. SYSTEMSCLUSTERING
B. SYSTEMSTABLE_CLUSTERING
C. SYSTEMSCLUSTERING_DEPTH
D. SYSTEMSCLUSTERING_RATIO
E. SYSTEMSCLUSTERING_INFORMATION

Answer: CE

Explanation:
? The SYSTEM$CLUSTERING_INFORMATION function in Snowflake returns a variety of clustering information for a specified table. This information includes the average clustering depth, total number of micro-partitions, total constant partition count, average overlaps, average depth, and a partition depth histogram. This function allows you to specify either one or multiple columns for which the clustering information is returned, and it returns this data in JSON format.
? The SYSTEM$CLUSTERING_DEPTH function computes the average depth of a table based on specified columns or the clustering key defined for the table. A lower average depth indicates that the table is better clustered with respect to the specified columns. This function also allows specifying columns to calculate the depth, and the values need to be enclosed in single quotes.
References:
? SYSTEM$CLUSTERING_INFORMATION: Snowflake Documentation
? SYSTEM$CLUSTERING_DEPTH: Snowflake Documentation

NEW QUESTION 97
A company needs to share its product catalog data with one of its partners. The product catalog data is stored in two database tables: product_category, and product_details. Both tables can be joined by the product_id column. Data access should be governed, and only the partner should have access to the records. The partner is not a Snowflake customer. The partner uses Amazon S3 for cloud storage. Which design will be the MOST cost-effective and secure, while using the required
Snowflake features?

A. Use Secure Data Sharing with an S3 bucket as a destination.
B. Publish product_category and product_details data sets on the Snowflake Marketplace.
C. Create a database user for the partner and give them access to the required data sets.
D. Create a reader account for the partner and share the data sets as secure views.

Answer: D

Explanation:
A reader account is a type of Snowflake account that allows external users to access data shared by a provider account without being a Snowflake customer. A reader account can be created and managed by the provider account, and can use the Snowflake web interface or JDBC/ODBC drivers to query the shared data.

A reader account is billed to the provider account based on the credits consumed by the queries1. A secure view is a type of view that applies row-level security filters to the underlying tables, and masks the data that is not accessible to the user. A secure view can be shared with a reader account to provide granular and governed access to the data2. In this scenario, creating a reader account for the partner and sharing the data sets as secure views would be the most cost-effective and secure design, while using the required Snowflake features, because:
? It would avoid the data transfer and storage costs of using an S3 bucket as a destination, and the potential security risks of exposing the data to unauthorized access or modification.
? It would avoid the complexity and overhead of publishing the data sets on the Snowflake Marketplace, and the potential loss of control over the data ownership and pricing.
? It would avoid the need to create a database user for the partner and grant them access to the required data sets, which would require the partner to have a Snowflake account and consume the provider??s resources.
References:
? Reader Accounts
? Secure Views

NEW QUESTION 102
How does a standard virtual warehouse policy work in Snowflake?

A. It conserves credits by keeping running clusters fully loaded rather than starting additional clusters.
B. It starts only if the system estimates that there is a query load that will keep the cluster busy for at least 6 minutes.
C. It starts only f the system estimates that there is a query load that will keep the cluster busy for at least 2 minutes.
D. It prevents or minimizes queuing by starting additional clusters instead of conserving credits.

Answer: D

Explanation:
A standard virtual warehouse policy is one of the two scaling policies available for multi-cluster warehouses in Snowflake. The other policy is economic. A standard policy aims to prevent or minimize queuing by starting additional clusters as soon as the current cluster is fully loaded, regardless of the number of queries in the queue. This policy can improve query performance and concurrency, but it may also consume more credits than an economic policy, which tries to conserve credits by keeping the running clusters fully loaded before starting additional clusters. The scaling policy can be set when creating or modifying a warehouse, and it can be changed at any time.
References:
? Snowflake Documentation: Multi-cluster Warehouses
? Snowflake Documentation: Scaling Policy for Multi-cluster Warehouses

NEW QUESTION 103
Which security, governance, and data protection features require, at a MINIMUM, the Business Critical edition of Snowflake? (Choose two.)

A. Extended Time Travel (up to 90 days)
B. Customer-managed encryption keys through Tri-Secret Secure
C. Periodic rekeying of encrypted data
D. AWS, Azure, or Google Cloud private connectivity to Snowflake
E. Federated authentication and SSO

Answer: BD

Explanation:
According to the SnowPro Advanced: Architect documents and learning resources, the security, governance, and data protection features that require, at a minimum, the Business Critical edition of Snowflake are:
? Customer-managed encryption keys through Tri-Secret Secure. This feature allows customers to manage their own encryption keys for data at rest in Snowflake, using a combination of three secrets: a master key, a service key, and a security password. This provides an additional layer of security and control over the data encryption and decryption process1.
? Periodic rekeying of encrypted data. This feature allows customers to periodically rotate the encryption keys for data at rest in Snowflake, using either Snowflake-managed keys or customer-managed keys. This enhances the security and protection of the data by reducing the risk of key compromise or exposure2.
The other options are incorrect because they do not require the Business Critical edition of Snowflake. Option A is incorrect because extended Time Travel (up to 90 days) is available with the Enterprise edition of Snowflake3. Option D is incorrect because AWS, Azure, or Google Cloud private connectivity to Snowflake is available with the Standard edition of Snowflake4. Option E is incorrect because federated authentication and SSO are available with the Standard edition of Snowflake5. References: Tri-Secret Secure | Snowflake Documentation, Periodic Rekeying of Encrypted Data | Snowflake Documentation, Snowflake Editions | Snowflake Documentation, Snowflake Network Policies | Snowflake Documentation, Configuring Federated Authentication and SSO | Snowflake Documentation

NEW QUESTION 104
An Architect needs to design a Snowflake account and database strategy to store and analyze large amounts of structured and semi-structured data. There are many business units and departments within the company. The requirements are scalability, security, and cost efficiency.
What design should be used?

A. Create a single Snowflake account and database for all data storage and analysis needs, regardless of data volume or complexity.
B. Set up separate Snowflake accounts and databases for each department or business unit, to ensure data isolation and security.
C. Use Snowflake's data lake functionality to store and analyze all data in a central location, without the need for structured schemas or indexes
D. Use a centralized Snowflake database for core business data, and use separate databases for departmental or project-specific data.

Answer: D

Explanation:
The best design to store and analyze large amounts of structured and semi- structured data for different business units and departments is to use a centralized Snowflake database for core business data, and use separate databases for departmental or project-specific data. This design allows for scalability, security, and cost efficiency by leveraging Snowflake??s features such as:
? Database cloning: Cloning a database creates a zero-copy clone that shares the same data files as the original database, but can be modified independently. This reduces storage costs and enables fast and consistent data replication for different purposes.
? Database sharing: Sharing a database allows granting secure and governed access to a subset of data in a database to other Snowflake accounts or consumers. This enables data collaboration and monetization across different business units or external partners.

? Warehouse scaling: Scaling a warehouse allows adjusting the size and concurrency of a warehouse to match the performance and cost requirements of different workloads. This enables optimal resource utilization and flexibility for different data analysis needs. References: Snowflake Documentation: Database Cloning, Snowflake Documentation: Database Sharing, [Snowflake Documentation: Warehouse Scaling]

NEW QUESTION 109
When loading data into a table that captures the load time in a column with a default value of either CURRENT_TIME () or CURRENT_TIMESTAMP() what will occur?

A. All rows loaded using a specific COPY statement will have varying timestamps based on when the rows were inserted.
B. Any rows loaded using a specific COPY statement will have varying timestamps based on when the rows were read from the source.
C. Any rows loaded using a specific COPY statement will have varying timestamps based on when the rows were created in the source.
D. All rows loaded using a specific COPY statement will have the same timestamp value.

Answer: D

Explanation:
According to the Snowflake documentation, when loading data into a table that captures the load time in a column with a default value of either CURRENT_TIME () or CURRENT_TIMESTAMP(), the default value is evaluated once per COPY statement, not once per row. Therefore, all rows loaded using a specific COPY statement will have the same timestamp value. This behavior ensures that the timestamp value reflects the time when the data was loaded into the table, not when the data was read from the source or created in the source. References:
? Snowflake Documentation: Loading Data into Tables with Default Values
? Snowflake Documentation: COPY INTO table

NEW QUESTION 111
Which organization-related tasks can be performed by the ORGADMIN role? (Choose three.)

A. Changing the name of the organization
B. Creating an account
C. Viewing a list of organization accounts
D. Changing the name of an account
E. Deleting an account
F. Enabling the replication of a database

Answer: BCF

Explanation:
According to the SnowPro Advanced: Architect documents and learning resources, the organization-related tasks that can be performed by the ORGADMIN role are:
? Creating an account in the organization. A user with the ORGADMIN role can use the CREATE ACCOUNT command to create a new account that belongs to the same organization as the current account1.
? Viewing a list of organization accounts. A user with the ORGADMIN role can use the SHOW ORGANIZATION ACCOUNTS command to view the names and properties of all accounts in the organization2. Alternatively, the user can use the Admin » Accounts page in the web interface to view the organization name and account names3.
? Enabling the replication of a database. A user with the ORGADMIN role can use the SYSTEM$GLOBAL_ACCOUNT_SET_PARAMETER function to enable database replication for an account in the organization. This allows the user to replicate databases across accounts in different regions and cloud platforms for data availability and durability4.
The other options are incorrect because they are not organization-related tasks that can be performed by the ORGADMIN role. Option A is incorrect because changing the name of the organization is not a task that can be performed by the ORGADMIN role. To change the name of an organization, the user must contact Snowflake Support3. Option D is incorrect because changing the name of an account is not a task that can be performed by the ORGADMIN role. To change the name of an account, the user must contact Snowflake Support5. Option E is incorrect because deleting an account is not a task that can be performed by the ORGADMIN role. To delete an account, the user must contact Snowflake Support. References: CREATE ACCOUNT | Snowflake Documentation, SHOW ORGANIZATION ACCOUNTS | Snowflake Documentation, Getting Started with Organizations | Snowflake Documentation, SYSTEM$GLOBAL_ACCOUNT_SET_PARAMETER | Snowflake Documentation, ALTER ACCOUNT | Snowflake Documentation, [DROP ACCOUNT | Snowflake Documentation]

NEW QUESTION 112
A company is using a Snowflake account in Azure. The account has SAML SSO set up using ADFS as a SCIM identity provider. To validate Private Link connectivity, an Architect performed the following steps:
* Confirmed Private Link URLs are working by logging in with a username/password account
* Verified DNS resolution by running nslookups against Private Link URLs
* Validated connectivity using SnowCD
* Disabled public access using a network policy set to use the company??s IP address range However, the following error message is received when using SSO to log into the company
account:
IP XX.XXX.XX.XX is not allowed to access snowflake. Contact your local security administrator.
What steps should the Architect take to resolve this error and ensure that the account is accessed using only Private Link? (Choose two.)

A. Alter the Azure security integration to use the Private Link URLs.
B. Add the IP address in the error message to the allowed list in the network policy.
C. Generate a new SCIM access token using system$generate_scim_access_token and save it to Azure AD.
D. Update the configuration of the Azure AD SSO to use the Private Link URLs.
E. Open a case with Snowflake Support to authorize the Private Link URLs?? access to the account.

Answer: BD

Explanation:
The error message indicates that the IP address in the error message is not allowed to access Snowflake because it is not in the allowed list of the network policy. The network policy is a feature that allows restricting access to Snowflake based on IP addresses or ranges. To resolve this error, the Architect should take the following steps:
? Add the IP address in the error message to the allowed list in the network policy.

This will allow the IP address to access Snowflake using the Private Link URLs. Alternatively, the Architect can disable the network policy if it is not required for security reasons.
? Update the configuration of the Azure AD SSO to use the Private Link URLs. This
will ensure that the SSO authentication process uses the Private Link URLs instead of the public URLs. The configuration can be updated by following the steps in the Azure documentation1.
These two steps should resolve the error and ensure that the account is accessed using only Private Link. The other options are not necessary or relevant for this scenario. Altering the Azure security integration to use the Private Link URLs is not required because the security integration is used for SCIM provisioning, not for SSO authentication. Generating a new SCIM access token using system$generate_scim_access_token and saving it to
Azure AD is not required because the SCIM access token is used for SCIM provisioning, not for SSO authentication. Opening a case with Snowflake Support to authorize the Private Link URLs?? access to the account is not required because the authorization can be done by the account administrator using the SYSTEM$AUTHORIZE_PRIVATELINK function2.

NEW QUESTION 115
A Snowflake Architect Is working with Data Modelers and Table Designers to draft an ELT framework specifically for data loading using Snowpipe. The Table Designers will add a timestamp column that Inserts the current tlmestamp as the default value as records are loaded into a table. The Intent is to capture the time when each record gets loaded into the table; however, when tested the timestamps are earlier than the loae_take column values returned by the copy_history function or the Copy_HISTORY view (Account Usage).
Why Is this occurring?

A. The timestamps are different because there are parameter setup mismatche
B. The parameters need to be realigned
C. The Snowflake timezone parameter Is different from the cloud provider's parameters causing the mismatch.
D. The Table Designer team has not used the localtimestamp or systimestamp functions in the Snowflake copy statement.
E. The CURRENT_TIMEis evaluated when the load operation is compiled in cloud services rather than when the record is inserted into the table.

Answer: D

Explanation:
? The correct answer is D because the CURRENT_TIME function returns the current timestamp at the start of the statement execution, not at the time of the record insertion. Therefore, if the load operation takes some time to complete, the CURRENT_TIME value may be earlier than the actual load time.
? Option A is incorrect because the parameter setup mismatches do not affect the timestamp values. The parameters are used to control the behavior and performance of the load operation, such as the file format, the error handling, the purge option, etc.
? Option B is incorrect because the Snowflake timezone parameter and the cloud provider??s parameters are independent of each other. The Snowflake timezone parameter determines the session timezone for displaying and converting timestamp values, while the cloud provider??s parameters determine the physical location and configuration of the storage and compute resources.
? Option C is incorrect because the localtimestamp and systimestamp functions are not relevant for the Snowpipe load operation. The localtimestamp function returns the current timestamp in the session timezone, while the systimestamp function
returns the current timestamp in the system timezone. Neither of them reflect the actual load time of the records. References:
? Snowflake Documentation: Loading Data Using Snowpipe: This document
explains how to use Snowpipe to continuously load data from external sources into Snowflake tables. It also describes the syntax and usage of the COPY INTO command, which supports various options and parameters to control the loading behavior.
? Snowflake Documentation: Date and Time Data Types and Functions: This
document explains the different data types and functions for working with date and time values in Snowflake. It also describes how to set and change the session timezone and the system timezone.
? Snowflake Documentation: Querying Metadata: This document explains how to
query the metadata of the objects and operations in Snowflake using various functions, views, and tables. It also describes how to access the copy history information using the COPY_HISTORY function or the COPY_HISTORY view.

NEW QUESTION 118
An Architect is using SnowCD to investigate a connectivity issue.
Which system function will provide a list of endpoints that the network must be able to access to use a specific Snowflake account, leveraging private connectivity?

A. SYSTEMSALLOWLIST ()
B. SYSTEMSGET_PRIVATELINK
C. SYSTEMSAUTHORIZE_PRIVATELINK
D. SYSTEMSALLOWLIST_PRIVATELINK ()

Answer: B

Explanation:
The SYSTEM$GET_PRIVATELINK function is used to retrieve the list of Snowflake service endpoints that need to be accessible when configuring private connectivity (such as AWS PrivateLink or Azure Private Link) for a Snowflake account. The function returns information necessary for setting up the networking infrastructure that allows secure and private access to Snowflake without using the public internet. SnowCD can then be used to verify connectivity to these endpoints.

NEW QUESTION 120
What are characteristics of the use of transactions in Snowflake? (Select TWO).

A. Explicit transactions can contain DDL, DML, and query statements.
B. The autocommit setting can be changed inside a stored procedure.
C. A transaction can be started explicitly by executing a begin work statement and end explicitly by executing a commit work statement.
D. A transaction can be started explicitly by executing a begin transaction statement and end explicitly by executing an end transaction statement.
E. Explicit transactions should contain only DML statements and query statement
F. All DDL statements implicitly commit active transactions.

Answer: AC

Explanation:
A. Snowflake's transactions can indeed include DDL (Data Definition Language), DML (Data Manipulation Language), and query statements. When executed within a transaction block, they all contribute to the atomicity of the transaction—either all of them commit together or none at all.C. Snowflake supports explicit

transaction control through the use of the BEGIN TRANSACTION (or simply BEGIN) and COMMIT statements. Alternatively, the BEGIN WORK and COMMIT WORK syntax is also supported, which is a standard SQL syntax for initiating and ending transactions, respectively.Note: The END TRANSACTION statement is not used in Snowflake to end a transaction; the correct statement is COMMIT or COMMIT WORK.

NEW QUESTION 122
A table contains five columns and it has millions of records. The cardinality distribution of the columns is shown below:

Column	Number of Distinct Values
C1	10,790
C2	108
C3	302,605
C4	1.117,736
C5	2.205,400

Column C4 and C5 are mostly used by SELECT queries in the GROUP BY and ORDER BY clauses. Whereas columns C1, C2 and C3 are heavily used in filter and join conditions of SELECT queries.
The Architect must design a clustering key for this table to improve the query performance. Based on Snowflake recommendations, how should the clustering key columns be ordered
while defining the multi-column clustering key?

A. C5, C4, C2
B. C3, C4, C5
C. C1, C3, C2
D. C2, C1, C3

Answer: D

Explanation:
According to the Snowflake documentation, the following are some considerations for choosing clustering for a table1:
? Clustering is optimal when either:
? Clustering is most effective when the clustering key is used in the following types of query predicates:
? Clustering is less effective when the clustering key is not used in any of the above
query predicates, or when the clustering key is used in a predicate that requires a function or expression to be applied to the key (e.g. DATE_TRUNC, TO_CHAR, etc.).
? For most tables, Snowflake recommends a maximum of 3 or 4 columns (or
expressions) per key. Adding more than 3-4 columns tends to increase costs more than benefits.
Based on these considerations, the best option for the clustering key columns is C. C1, C3, C2, because:
? These columns are heavily used in filter and join conditions of SELECT queries,
which are the most effective types of predicates for clustering.
? These columns have high cardinality, which means they have many distinct values and can help reduce the clustering skew and improve the compression ratio.
? These columns are likely to be correlated with each other, which means they can help co-locate similar rows in the same micro-partitions and improve the scan efficiency.
? These columns do not require any functions or expressions to be applied to them, which means they can be directly used in the predicates without affecting the clustering.
References: 1: Considerations for Choosing Clustering for a Table | Snowflake Documentation

NEW QUESTION 126
.....

Relate Links

100% Pass Your ARA-C01 Exam with Exambible Prep Materials

https://www.exambible.com/ARA-C01-exam/

Contact us

We are proud of our high-quality customer service, which serves you around the clock 24/7.

Viste - https://www.exambible.com/