

Snowflake

Exam Questions DEA-C01

SnowPro Advanced: Data Engineer Certification Exam



NEW QUESTION 1

Which use case would be BEST suited for the search optimization service?

- A. Analysts who need to perform aggregates over high cardinality columns
- B. Business users who need fast response times using highly selective filters
- C. Data Scientists who seek specific JOIN statements with large volumes of data
- D. Data Engineers who create clustered tables with frequent reads against clustering keys

Answer: B

Explanation:

The use case that would be best suited for the search optimization service is business users who need fast response times using highly selective filters. The search optimization service is a feature that enables faster queries on tables with high cardinality columns by creating inverted indexes on those columns. High cardinality columns are columns that have a large number of distinct values, such as customer IDs, product SKUs, or email addresses. Queries that use highly selective filters on high cardinality columns can benefit from the search optimization service because they can quickly locate the relevant rows without scanning the entire table. The other options are not best suited for the search optimization service. Option A is incorrect because analysts who need to perform aggregates over high cardinality columns will not benefit from the search optimization service, as they will still need to scan all the rows that match the filter criteria. Option C is incorrect because data scientists who seek specific JOIN statements with large volumes of data will not benefit from the search optimization service, as they will still need to perform join operations that may involve shuffling or sorting data across nodes. Option D is incorrect because data engineers who create clustered tables with frequent reads against clustering keys will not benefit from the search optimization service, as they already have an efficient way to organize and access data based on clustering keys.

NEW QUESTION 2

A Data Engineer defines the following masking policy:

```
current_role() IN ('ADMIN') THEN val  
*****!
```

....

must be applied to the full_name column in the customer table:

```
TABLE customer(  
  name VARCHAR,  
  email VARCHAR,  
  full_name VARCHAR AS CONCAT(first_name, ' ', last_name)
```

Which query will apply the masking policy on the full_name column?

- A. ALTER TABLE customer MODIFY COLUMN full_name SET MASKING POLICY name_policy;
- B. ALTER TABLE customer MODIFY COLUMN full_name ADD MASKING POLICY name_policy;
- C. ALTER TABLE customer MODIFY COLUMN first_name SET MASKING POLICY name_policy; last_name SET MASKING POLICY name_policy;
- D. ALTER TABLE customer MODIFY COLUMN first_name ADD MASKING POLICY name_policy;

Answer: A

Explanation:

The query that will apply the masking policy on the full_name column is ALTER TABLE customer MODIFY COLUMN full_name SET MASKING POLICY name_policy;. This query will modify the full_name column and associate it with the name_policy masking policy, which will mask the first and last names of the customers with asterisks. The other options are incorrect because they do not follow the correct syntax for applying a masking policy on a column. Option B is incorrect because it uses ADD instead of SET, which is not a valid keyword for modifying a column. Option C is incorrect because it tries to apply the masking policy on two columns, first_name and last_name, which are not part of the table structure. Option D is incorrect because it uses commas instead of dots to separate the database, schema, and table names.

NEW QUESTION 3

Which methods will trigger an action that will evaluate a DataFrame? (Select TWO)

- A. DataFrame.random_split()
- B. DataFrame.collect()
- C. DataFrame.select()
- D. DataFrame.col()
- E. DataFrame.show()

Answer: BE

Explanation:

The methods that will trigger an action that will evaluate a DataFrame are DataFrame.collect() and DataFrame.show(). These methods will force the execution of any pending transformations on the DataFrame and return or display the results. The other options are not methods that will evaluate a DataFrame. Option A, DataFrame.random_split(), is a method that will split a DataFrame into two or more DataFrames based on random weights. Option C, DataFrame.select(), is a method that will project a set of expressions on a DataFrame and return a new DataFrame. Option D, DataFrame.col(), is a method that will return a Column object based on a column name in a DataFrame.

NEW QUESTION 4

Database XYZ has the data_retention_time_in_days parameter set to 7 days and table xyz.public.ABC has the data_retention_time_in_days set to 10 days.

A Developer accidentally dropped the database containing this single table 8 days ago and just discovered the mistake. How can the table be recovered?

- A. undrop database xyz;
- B. create table abc_restore as select * from xyz.public.abc at (offset => -60*60*24*8);
- C. create table abc_restore clone xyz.public.abc at (offset => -3600*24*3);
- D. Create a Snowflake Support case to restore the database and table from fail-safe

Answer: A

Explanation:

The table can be recovered by using the undrop database xyz; command. This command will restore the database that was dropped within the last 14 days, along with all its schemas and tables, including the customer table. The data_retention_time_in_days parameter does not affect this command, as it only applies to time travel queries that reference historical data versions of tables or databases. The other options are not valid ways to recover the table. Option B is incorrect because creating a table as select * from xyz.public.ABC at (offset => -6060248) will not work, as this query will try to access a historical data version of the ABC table that does not exist anymore after dropping the database. Option C is incorrect because creating a table clone xyz.public.ABC at (offset => -360024*3) will not work, as this query will try to clone a historical data version of the ABC table that does not exist anymore after dropping the database. Option D is incorrect because creating a Snowflake Support case to restore the database and table from fail-safe will not work, as fail-safe is only available for disaster recovery scenarios and cannot be accessed by customers.

NEW QUESTION 5

A Data Engineer has created table t1 with datatype VARIANT: create or replace table t1 (cl variant);
The Engineer has loaded the following JSON data set. which has information about 4 laptop models into the table:

```
{
  "device_model": [
    {
      "manufacturer": "HP",
      "model": "HP 240 G8",
      "model_id": "hp 240 g8",
      "model_name": "240 G8"
    },
    {
      "manufacturer": "HP",
      "model": "HP EliteBook 1030 G1",
      "model_id": "hp elitebook 1030 g1",
      "model_name": "EliteBook 1030 G1"
    },
    {
      "manufacturer": "HP",
      "model": "HP ZBook 15 G2",
      "model_id": "hp zbook 15 g2",
      "model_name": "ZBook 15 G2"
    },
    {
      "manufacturer": "Lenovo",
      "model": "Lenovo B50-70",
      "model_id": "lenovo b50-70",
      "model_name": "B50-70"
    }
  ]
}
```

The Engineer now wants to query that data set so that results are shown as normal structured data. The result should be 4 rows and 4 columns without the double quotes surrounding the data elements in the JSON data.

The result should be similar to the use case where the data was selected from a normal relational table z2 where t2 has string data type columns model id, model, manufacturer, and model_name. and is queried with the SQL clause select * from t2;
Which select command will produce the correct results?

A)

```
select value:model_id::string
, value:model::string
, value:manufacturer::string
, value:model_name::string
from t1
, lateral flatten(input => c1);
```

B)

```
select value:model_id::string
, value:model::string
, value:manufacturer::string
, value:model_name::string
from t1
, lateral flatten(input => c1:device_model);
```

C)

```
select model_id::string
, model::string
, manufacturer::string
, model_name::string
from t1
, lateral flatten(input => c1:device_model);
```

D)

```
select value:model_id
, value:model
, value:manufacturer
, value:model_name
from t1
, lateral flatten(input => c1:device_model);
```

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

Answer: B**NEW QUESTION 6**

A new customer table is created by a data pipeline in a Snowflake schema where MANAGED ACCESS is enabled. Can gran access to the CUSTOMER table? (Select THREE.)

- A. The role that owns the schema
- B. The role that owns the database
- C. The role that owns the customer table
- D. The SYSADMIN role
- E. The SECURITYADMIN role
- F. The USERADMIN role with the manage grants privilege

Answer: ABE**Explanation:**

The roles that can grant access to the CUSTOMER table are the role that owns the schema, the role that owns the database, and the SECURITYADMIN role. These roles have the ownership or the manage grants privilege on the schema or the database level, which allows them to grant access to any object within them. The other options are incorrect because they do not have the necessary privilege to grant access to the CUSTOMER table. Option C is incorrect because the role that owns the customer table cannot grant access to itself or to other roles. Option D is incorrect because the SYSADMIN role does not have the manage grants privilege by default and cannot grant access to objects that it does not own. Option F is incorrect because the USERADMIN role with the manage grants privilege can only grant access to users and roles, not to tables.

NEW QUESTION 7

While running an external function, the following error message is received: Error: function received the wrong number of rows. What is causing this to occur?

- A. External functions do not support multiple rows

- B. Nested arrays are not supported in the JSON response
- C. The JSON returned by the remote service is not constructed correctly
- D. The return message did not produce the same number of rows that it received

Answer: D

Explanation:

The error message “function received the wrong number of rows” is caused by the return message not producing the same number of rows that it received. External functions require that the remote service returns exactly one row for each input row that it receives from Snowflake. If the remote service returns more or fewer rows than expected, Snowflake will raise an error and abort the function execution. The other options are not causes of this error message. Option A is incorrect because external functions do support multiple rows as long as they match the input rows. Option B is incorrect because nested arrays are supported in the JSON response as long as they conform to the return type definition of the external function. Option C is incorrect because the JSON returned by the remote service may be constructed correctly but still produce a different number of rows than expected.

NEW QUESTION 8

A Data Engineer ran a stored procedure containing various transactions During the execution, the session abruptly disconnected preventing one transaction from committing or rolling back. The transaction was left in a detached state and created a lock on resources ...must the Engineer take to immediately run a new transaction?

- A. Call the system function SYSTEM\$ABORT_TRANSACTION.
- B. Call the system function SYSTEM\$CANCEL_TRANSACTION.
- C. Set the LOCK_TIMEOUT to FALSE in the stored procedure
- D. Set the transaction abort on error to true in the stored procedure.

Answer: A

Explanation:

The system function SYSTEM\$ABORT_TRANSACTION can be used to abort a detached transaction that was left in an open state due to a session disconnect or termination. The function takes one argument: the transaction ID of the detached transaction. The function will abort the transaction and release any locks held by it. The other options are incorrect because they do not address the issue of a detached transaction. The system function SYSTEM\$CANCEL_TRANSACTION can be used to cancel a running transaction, but not a detached one. The LOCK_TIMEOUT parameter can be used to set a timeout period for acquiring locks on resources, but it does not affect existing locks. The TRANSACTION_ABORT_ON_ERROR parameter can be used to control whether a transaction should abort or continue when an error occurs, but it does not affect detached transactions.

NEW QUESTION 9

A Data Engineer would like to define a file structure for loading and unloading data Where can the file structure be defined? (Select THREE)

- A. copy command
- B. MERGE command
- C. FILE FORMAT Object
- D. pipe object
- E. stage object
- F. INSERT command

Answer: ACE

Explanation:

The places where the file format can be defined are copy command, file format object, and stage object. These places allow specifying or referencing a file format that defines how data files are parsed and loaded into or unloaded from Snowflake tables. A file format can include various options, such as field delimiter, field enclosure, compression type, date format, etc. The other options are not places where the file format can be defined. Option B is incorrect because MERGE command is a SQL command that can merge data from one table into another based on a join condition, but it does not involve loading or unloading data files. Option D is incorrect because pipe object is a Snowflake object that can load data from an external stage into a Snowflake table using COPY statements, but it does not define or reference a file format. Option F is incorrect because INSERT command is a SQL command that can insert data into a Snowflake table from literal values or subqueries, but it does not involve loading or unloading data files.

NEW QUESTION 10

Which query will show a list of the 20 most recent executions of a specified task ktask, that have been scheduled within the last hour that have ended or are still running's.

A)

```
select * from table(information_schema.task_history(scheduled_time_range_start
=>dateadd('hour',-1,current_timestamp()), result_limit => 20,
task_name=>'MYTASK'))
```

B)

```
select * from table(information_schema.task_history(scheduled_time_range_start
=>dateadd('hour',-1,current_timestamp()), result_limit => 20,
task_name=>'MYTASK')) where query_id IS NOT NULL;
```

C)

```
select * from table(information_schema.task_history(scheduled_time_range_start
=>dateadd('hour',-1,current_timestamp()), result_limit => 20,
task_name=>'MYTASK')) where STATE IN ('EXECUTING', 'SUCCEEDED', 'FAILED')
```

D)

```
select * from table(information_schema.task_history(scheduled_time_range_end
=>dateadd('hour',-1,current_timestamp()), result_limit => 10,
task_name=>'MYTASK')) where STATE IN ('EXECUTING', 'SUCCEEDED')
```

- A. Option A
- B. Option B

- C. Option C
- D. Option D

Answer: B

NEW QUESTION 10

A Data Engineer is working on a continuous data pipeline which receives data from Amazon Kinesis Firehose and loads the data into a staging table which will later be used in the data transformation process. The average file size is 300-500 MB.

The Engineer needs to ensure that Snowpipe is performant while minimizing costs. How can this be achieved?

- A. Increase the size of the virtual warehouse used by Snowpipe.
- B. Split the files before loading them and set the SIZE_LIMIT option to 250 MB.
- C. Change the file compression size and increase the frequency of the Snowpipe loads.
- D. Decrease the buffer size to trigger delivery of files sized between 100 to 250 MB in Kinesis Firehose.

Answer: B

Explanation:

This option is the best way to ensure that Snowpipe is performant while minimizing costs. By splitting the files before loading them, the Data Engineer can reduce the size of each file and increase the parallelism of loading. By setting the SIZE_LIMIT option to 250 MB, the Data Engineer can specify the maximum file size that can be loaded by Snowpipe, which can prevent performance degradation or errors due to large files. The other options are not optimal because:

? Increasing the size of the virtual warehouse used by Snowpipe will increase the performance but also increase the costs, as larger warehouses consume more credits per hour.

? Changing the file compression size and increasing the frequency of the Snowpipe loads will not have much impact on performance or costs, as Snowpipe already supports various compression formats and automatically loads files as soon as they are detected in the stage.

? Decreasing the buffer size to trigger delivery of files sized between 100 to 250 MB in Kinesis Firehose will not affect Snowpipe performance or costs, as Snowpipe does not depend on Kinesis Firehose buffer size but rather on its own SIZE_LIMIT option.

NEW QUESTION 12

Company A and Company B both have Snowflake accounts. Company A's account is hosted on a different cloud provider and region than Company B's account. Companies A and B are not in the same Snowflake organization.

How can Company A share data with Company B? (Select TWO).

- A. Create a share within Company A's account and add Company B's account as a recipient of that share.
- B. Create a share within Company A's account, and create a reader account that is a recipient of the share. Grant Company B access to the reader account.
- C. Use database replication to replicate Company A's data into Company B's account. Create a share within Company B's account and grant users within Company B's account access to the share.
- D. Create a new account within Company A's organization in the same cloud provider and region as Company B's account. Use database replication to replicate Company A's data to the new account. Create a share within the new account and add Company B's account as a recipient of that share.
- E. Create a separate database within Company A's account to contain only those data sets they wish to share with Company B. Create a share within Company A's account and add all the objects within this separate database to the share. Add Company B's account as a recipient of the share.

Answer: AE

Explanation:

The ways that Company A can share data with Company B are:

? Create a share within Company A's account and add Company B's account as a recipient of that share: This is a valid way to share data between different accounts on different cloud platforms and regions. Snowflake supports cross-cloud and cross-region data sharing, which allows users to create shares and grant access to other accounts regardless of their cloud platform or region. However, this option may incur additional costs for network transfer and storage replication.

? Create a separate database within Company A's account to contain only those data sets they wish to share with Company B. Create a share within Company A's account and add all the objects within this separate database to the share. Add Company B's account as a recipient of the share: This is also a valid way to share data between different accounts on different cloud platforms and regions. This option is similar to the previous one, except that it uses a separate database to isolate the data sets that need to be shared. This can improve security and manageability of the shared data. The other options are not valid because:

? Create a share within Company A's account, and create a reader account that is a recipient of the share. Grant Company B access to the reader account: This option is not valid because reader accounts are not supported for cross-cloud or cross-region data sharing. Reader accounts are Snowflake accounts that can only consume data from shares created by their provider account. Reader accounts must be on the same cloud platform and region as their provider account.

? Use database replication to replicate Company A's data into Company B's account. Create a share within Company B's account and grant users within Company B's account access to the share: This option is not valid because database replication cannot be used for cross-cloud or cross-region data sharing. Database replication is a feature in Snowflake that allows users to copy databases across accounts within the same cloud platform and region. Database replication cannot copy databases across different cloud platforms or regions.

? Create a new account within Company A's organization in the same cloud provider and region as Company B's account. Use database replication to replicate Company A's data to the new account. Create a share within the new account and add Company B's account as a recipient of that share: This option is not valid because it involves creating a new account within Company A's organization, which may not be feasible or desirable for Company A. Moreover, this option is unnecessary, as Company A can directly share data with Company B without creating an intermediate account.

NEW QUESTION 14

A Data Engineer wants to centralize grant management to maximize security. A user needs ownership on a table in a new schema. However, this user should not have the ability to make grant decisions.

What is the correct way to do this?

- A. Grant ownership to the user on the table.
- B. Revoke grant decisions from the user on the table.
- C. Revoke grant decisions from the user on the schema.
- D. Add the WITH MANAGED ACCESS parameter on the schema.

Answer: D

Explanation:

The with managed access parameter on the schema enables the schema owner to control the grant and revoke privileges on the objects within the schema. This way, the user who owns the table cannot make grant decisions, but only the schema owner can. This is the best way to centralize grant management and maximize security.

NEW QUESTION 15

A Data Engineer has developed a dashboard that will issue the same SQL select clause to Snowflake every 12 hours.
---will Snowflake use the persisted query results from the result cache provided that the underlying data has not changed^

- A. 12 hours
- B. 24 hours
- C. 14 days
- D. 31 days

Answer: C

Explanation:

Snowflake uses the result cache to store the results of queries that have been executed recently. The result cache is maintained at the account level and is shared across all sessions and users. The result cache is invalidated when any changes are made to the tables or views referenced by the query. Snowflake also has a retention policy for the result cache, which determines how long the results are kept in the cache before they are purged. The default retention period for the result cache is 24 hours, but it can be changed at the account, user, or session level. However, there is a maximum retention period of 14 days for the result cache, which cannot be exceeded. Therefore, if the underlying data has not changed, Snowflake will use the persisted query results from the result cache for up to 14 days.

NEW QUESTION 18

Given the table sales which has a clustering key of column CLOSED_DATE which table function will return the average clustering depth for the SALES_REPRESENTATIVE column for the North American region?

A)

```
select system$clustering_information('Sales', 'sales_representative', 'region = ''North America''');  
  
select system$clustering_depth('Sales', 'sales_representative', 'region = ''North America''');
```

C)

```
select system$clustering_depth('Sales', 'sales_representative') where region = 'North America';
```

D)

```
select system$clustering_information('Sales', 'sales_representative') where region = 'North America';
```

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

Answer: B

Explanation:

The table function SYSTEM\$CLUSTERING_DEPTH returns the average clustering depth for a specified column or set of columns in a table. The function takes two arguments: the table name and the column name(s). In this case, the table name is sales and the column name is SALES_REPRESENTATIVE. The function also supports a WHERE clause to filter the rows for which the clustering depth is calculated. In this case, the WHERE clause is REGION = 'North America'. Therefore, the function call in Option B will return the desired result.

NEW QUESTION 21

What are characteristics of Snowpark Python packages? (Select THREE).

Third-party packages can be registered as a dependency to the Snowpark session using the session.import() method.

- A. Python packages can access any external endpoints
- B. Python packages can only be loaded in a local environment
- C. Third-party supported Python packages are locked down to prevent hitting
- D. The SQL command DESCRIBE FUNCTION will list the imported Python packages of the Python User-Defined Function (UDF).
- E. Querying information schema .packages will provide a list of supported Python packages and versions

Answer: ADE

Explanation:

The characteristics of Snowpark Python packages are:

? Third-party packages can be registered as a dependency to the Snowpark session using the session.import() method.

? The SQL command DESCRIBE FUNCTION will list the imported Python packages of the Python User-Defined Function (UDF).

? Querying information_schema.packages will provide a list of supported Python packages and versions.

These characteristics indicate how Snowpark Python packages can be imported, inspected, and verified in Snowflake. The other options are not characteristics of Snowpark Python packages. Option B is incorrect because Python packages can be loaded in both local and remote environments using Snowpark. Option C is incorrect because third-party supported Python packages are not locked down to prevent hitting external endpoints, but rather restricted by network policies and security settings.

NEW QUESTION 24

A Data Engineer is building a set of reporting tables to analyze consumer requests by region for each of the Data Exchange offerings annually, as well as click-through rates for each listing

Which views are needed MINIMALLY as data sources?

- A. SNOWFLAKE- DATA_SHARING_USAGE - LISTING_EVENTS_BAILY
- B. SNOWFLAKE.DATA_SHARING_USAGE.LISTING_CONSOKE>TION_DAILY
- C. SNOWFLAK
- D. DATA_SHARING_USAG
- E. LISTING_TELEMETRY_DAILY
- F. SNOWFLAKE.ACCOUNT_USAGE.DATA _TRANSFER_HISTORY

Answer: B

Explanation:

The SNOWFLAKE.DATA SHARING _USAGE.LISTING_CONSOKE>TION_DAILY view provides information about consumer requests by region for each of the Data Exchange offerings annually, as well as click- through rates for each listing. This view is the minimal data source needed for building the reporting tables. The other views are not relevant for this use case.

NEW QUESTION 26

A table is loaded using Snowpipe and truncated afterwards Later, a Data Engineer finds that the table needs to be reloaded but the metadata of the pipe will not allow the same files to be loaded again.

How can this issue be solved using the LEAST amount of operational overhead?

- A. Wait until the metadata expires and then reload the file using Snowpipe
- B. Modify the file by adding a blank row to the bottom and re-stage the file
- C. Set the FORCE=TRUE option in the Snowpipe COPY INTO command
- D. Recreate the pipe by using the create or replace pipe command

Answer: C

Explanation:

The FORCE=TRUE option in the Snowpipe COPY INTO command allows Snowpipe to load files that have already been loaded before, regardless of the metadata. This is the easiest way to reload the same files without modifying them or recreating the pipe.

NEW QUESTION 27

A Data Engineer is working on a Snowflake deployment in AWS eu-west-1 (Ireland). The Engineer is planning to load data from staged files into target tables using the copy into command

Which sources are valid? (Select THREE)

- A. Internal stage on GCP us-central1 (Iowa)
- B. Internal stage on AWS eu-central-1 (Frankfurt)
- C. External stage on GCP us-central1 (Iowa)
- D. External stage in an Amazon S3 bucket on AWS eu-west-1 (Ireland)
- E. External stage in an Amazon S3 bucket on AWS eu-central 1 (Frankfurt)
- F. SSO attached to an Amazon EC2 instance on AWS eu-west-1 (Ireland)

Answer: CDE

Explanation:

The valid sources for loading data from staged files into target tables using the copy into command are:

? External stage on GCP us-central1 (Iowa): This is a valid source because Snowflake supports cross-cloud data loading from external stages on different cloud platforms and regions than the Snowflake deployment.

? External stage in an Amazon S3 bucket on AWS eu-west-1 (Ireland): This is a valid source because Snowflake supports data loading from external stages on the same cloud platform and region as the Snowflake deployment.

? External stage in an Amazon S3 bucket on AWS eu-central 1 (Frankfurt): This is a valid source because Snowflake supports cross-region data loading from external stages on different regions than the Snowflake deployment within the same cloud platform. The invalid sources are:

? Internal stage on GCP us-central1 (Iowa): This is an invalid source because internal stages are always located on the same cloud platform and region as the Snowflake deployment. Therefore, an internal stage on GCP us-central1 (Iowa) cannot be used for a Snowflake deployment on AWS eu-west-1 (Ireland).

? Internal stage on AWS eu-central-1 (Frankfurt): This is an invalid source because internal stages are always located on the same region as the Snowflake deployment. Therefore, an internal stage on AWS eu-central-1 (Frankfurt) cannot be used for a Snowflake deployment on AWS eu-west-1 (Ireland).

? SSO attached to an Amazon EC2 instance on AWS eu-west-1 (Ireland): This is an invalid source because SSO stands for Single Sign-On, which is a security integration feature in Snowflake, not a data staging option.

NEW QUESTION 32

A Data Engineer wants to check the status of a pipe named my_pipe. The pipe is inside a database named test and a schema named Extract (case-sensitive). Which query will provide the status of the pipe?

- A. SELECT FROM SYSTEM\$PIPE_STATUS ("test.'extract'.my_pipe"):
- B. SELECT FROM SYSTEM\$PIPE_STATUS (,test,,Extracr,,ny_pipe, i l
- C. SELE2T * FROM SYSTEM\$PIPE_STATUS < ' tes
- D. "Extract", my_pipe');
- E. SELECT * FROM SYSTEM\$PIPE_STATUS ("tes
- F. 'extract' .my_pipe");

Answer: C

Explanation:

The query that will provide the status of the pipe is SELECT * FROM SYSTEM\$PIPE_STATUS('test."Extract".my_pipe');. The SYSTEM\$PIPE_STATUS function returns information about a pipe, such as its name, status, last received message timestamp, etc. The function takes one argument: the pipe name in a qualified form. The pipe name should include the database name, the schema name, and the pipe name, separated by dots. If any of these names are case-sensitive identifiers, they should be enclosed in double quotes. In this case, the schema name Extract is case-sensitive and should be quoted. The other options are incorrect because they do not follow the correct syntax for the pipe name argument. Option A and B use single quotes instead of double quotes for case-sensitive identifiers. Option D uses double quotes instead of single quotes for non-case-sensitive identifiers.

NEW QUESTION 35

A Data Engineer is implementing a near real-time ingestion pipeline to load data into Snowflake using the Snowflake Kafka connector. There will be three Kafka topics created.

.....snowflake objects are created automatically when the Kafka connector starts? (Select THREE)

- A. Tables
- B. Tasks
- C. Pipes
- D. internal stages
- E. External stages
- F. Materialized views

Answer: ACD

Explanation:

The Snowflake objects that are created automatically when the Kafka connector starts are tables, pipes, and internal stages. The Kafka connector will create one table, one pipe, and one internal stage for each Kafka topic that is configured in the connector properties. The table will store the data from the Kafka topic, the pipe will load the data from the stage to the table using COPY statements, and the internal stage will store the files that are produced by the Kafka connector using PUT commands. The other options are not Snowflake objects that are created automatically when the Kafka connector starts. Option B, tasks, are objects that can execute SQL statements on a schedule without requiring a warehouse. Option E, external stages, are objects that can reference locations outside of Snowflake, such as cloud storage services. Option F, materialized views, are objects that can store the precomputed results of a query and refresh them periodically.

NEW QUESTION 37

Which system role is recommended for a custom role hierarchy to be ultimately assigned to?

- A. ACCOUNTADMIN
- B. SECURITYADMIN
- C. SYSTEMADMIN
- D. USERADMIN

Answer: B

Explanation:

The system role that is recommended for a custom role hierarchy to be ultimately assigned to is SECURITYADMIN. This role has the manage grants privilege on all objects in an account, which allows it to grant access privileges to other roles or revoke them as needed. This role can also create or modify custom roles and assign them to users or other roles. By assigning custom roles to SECURITYADMIN, the role hierarchy can be managed centrally and securely. The other options are not recommended system roles for a custom role hierarchy to be ultimately assigned to. Option A is incorrect because ACCOUNTADMIN is the most powerful role in an account, which has full access to all objects and operations. Assigning custom roles to ACCOUNTADMIN can pose a security risk and should be avoided. Option C is incorrect because SYSTEMADMIN is a role that has full access to all objects in the public schema of the account, but not to other schemas or databases. Assigning custom roles to SYSTEMADMIN can limit the scope and flexibility of the role hierarchy. Option D is incorrect because USERADMIN is a role that can manage users and roles in an account, but not grant access privileges to other objects. Assigning custom roles to USERADMIN can prevent the role hierarchy from controlling access to data and resources.

NEW QUESTION 42

Which functions will compute a 'fingerprint' over an entire table, query result, or window to quickly detect changes to table contents or query results? (Select TWO).

- A. HASH (*)
- B. HASH_AGG(*)
- C. HASH_AGG(<expr>, <expr>)
- D. HASH_AGG_COMPARE (*)
- E. HASH_COMPARE(*)

Answer: BC

Explanation:

The functions that will compute a 'fingerprint' over an entire table, query result, or window to quickly detect changes to table contents or query results are:

? HASH_AGG(*): This function computes a hash value over all columns and rows in

a table, query result, or window. The function returns a single value for each group defined by a GROUP BY clause, or a single value for the entire input if no GROUP BY clause is specified.

? HASH_AGG(<expr>, <expr>): This function computes a hash value over two

expressions in a table, query result, or window. The function returns a single value for each group defined by a GROUP BY clause, or a single value for the entire input if no GROUP BY clause is specified. The other functions are not correct because:

? HASH (*): This function computes a hash value over all columns in a single row.

The function returns one value per row, not one value per table, query result, or window.

? HASH_AGG_COMPARE (): This function compares two hash values computed by

HASH_AGG() over two tables or query results and returns true if they are equal or false if they are different. The function does not compute a hash value itself, but rather compares two existing hash values.

? HASH_COMPARE(): This function compares two hash values computed by

HASH() over two rows and returns true if they are equal or false if they are different. The function does not compute a hash value itself, but rather compares two existing hash values.

NEW QUESTION 45

Which Snowflake feature facilitates access to external API services such as geocoders. data transformation, machine Learning models and other custom code?

- A. Security integration
- B. External tables
- C. External functions
- D. Java User-Defined Functions (UDFs)

Answer: C

Explanation:

External functions are Snowflake functions that facilitate access to external API services such as geocoders, data transformation, machine learning models and other custom code. External functions allow users to invoke external services from within SQL queries and pass arguments and receive results as JSON values. External functions require creating an API integration object and an external function object in Snowflake, as well as deploying an external service endpoint that can communicate with Snowflake via HTTPS.

NEW QUESTION 46

A Data Engineer wants to create a new development database (DEV) as a clone of the permanent production database (PROD) There is a requirement to disable Fail-safe for all tables.

Which command will meet these requirements?

- A. CREATE DATABASE DEV CLONE PROD FAIL_SAFE=FALSE;
- B. CREATE DATABASE DEV CLONE PROD;
- C. CREATE TRANSIENT DATABASE DEV CLONE RPOD
- D. CREATE DATABASE DEV CLOSE PRODDATA_RETENTION_TIME_IN_DAYS =0L

Answer: C

Explanation:

This option will meet the requirements of creating a new development database (DEV) as a clone of the permanent production database (PROD) and disabling Fail-safe for all tables. By using the CREATE TRANSIENT DATABASE command, the Data Engineer can create a transient database that does not have Fail-safe enabled by default. Fail-safe is a feature in Snowflake that provides additional protection against data loss by retaining historical data for seven days beyond the time travel retention period. Transient databases do not have Fail-safe enabled, which means that they do not incur additional storage costs for historical data beyond their time travel retention period. By using the CLONE option, the Data Engineer can create an exact copy of the PROD database, including its schemas, tables, views, and other objects.

NEW QUESTION 49

A Data Engineer executes a complex query and wants to make use of Snowflake s query results caching capabilities to reuse the results.

Which conditions must be met? (Select THREE).

- A. The results must be reused within 72 hours.
- B. The query must be executed using the same virtual warehouse.
- C. The USED_CACHED_RESULT parameter must be included in the query.
- D. The table structure contributing to the query result cannot have changed
- E. The new query must have the same syntax as the previously executed query.
- F. The micro-partitions cannot have changed due to changes to other data in the table

Answer: ADE

Explanation:

Snowflake's query results caching capabilities allow users to reuse the results of previously executed queries without re-executing them. For this to happen, the following conditions must be met:

? The results must be reused within 24 hours (not 72 hours), which is the default time-to-live (TTL) for cached results.

? The query must be executed using any virtual warehouse (not necessarily the same one), as long as it is in the same region and account as the original query.

? The USED_CACHED_RESULT parameter does not need to be included in the query, as it is enabled by default at the account level. However, it can be disabled or overridden at the session or statement level.

? The table structure contributing to the query result cannot have changed, such as adding or dropping columns, changing data types, or altering constraints.

? The new query must have the same syntax as the previously executed query, including whitespace and case sensitivity.

? The micro-partitions cannot have changed due to changes to other data in the table, such as inserting, updating, deleting, or merging rows.

NEW QUESTION 52

.....

Thank You for Trying Our Product

We offer two products:

1st - We have Practice Tests Software with Actual Exam Questions

2nd - Questions and Answers in PDF Format

DEA-C01 Practice Exam Features:

- * DEA-C01 Questions and Answers Updated Frequently
- * DEA-C01 Practice Questions Verified by Expert Senior Certified Staff
- * DEA-C01 Most Realistic Questions that Guarantee you a Pass on Your FirstTry
- * DEA-C01 Practice Test Questions in Multiple Choice Formats and Updatesfor 1 Year

100% Actual & Verified — Instant Download, Please Click
[Order The DEA-C01 Practice Test Here](#)