Exam Questions DEA-C01

SnowPro Advanced: Data Engineer Certification Exam

https://www.2passeasy.com/dumps/DEA-C01/

NEW QUESTION 1
Within a Snowflake account permissions have been defined with custom roles and role hierarchies.
To set up column-level masking using a role in the hierarchy of the current user, what command would be used?

A. CORRECT_ROLE
B. IKVOKER_ROLE
C. IS_RCLE_IN_SESSION
D. IS_GRANTED_TO_INVOKER_ROLE

Answer: C

Explanation:
The IS_ROLE_IN_SESSION function is used to set up column-level masking using a role in the hierarchy of the current user. Column-level masking is a feature in Snowflake that allows users to apply dynamic data masking policies to specific columns
based on the roles of the users who access them. The IS_ROLE_IN_SESSION function takes a role name as an argument and returns true if the role is in the current user's session, or false otherwise. The function can be used in a masking policy expression to determine whether to mask or unmask a column value based on the role of the user. For example:
CREATE OR REPLACE MASKING POLICY email_mask AS (val string) RETURNS string -
> CASE WHEN IS_ROLE_IN_SESSION('HR') THEN val ELSE REGEXP_REPLACE(val, '(.).(.@.)', '\1****\2') END;
In this example, the IS_ROLE_IN_SESSION function is used to create a masking policy for an email column. The masking policy returns the original email value if the user has the HR role in their session, or returns a masked email value with asterisks if not.

NEW QUESTION 2
The following chart represents the performance of a virtual warehouse over time:

A DataEngineer notices that the warehouse is queueing queries The warehouse is size X- Smallthe minimum and maximum cluster counts are set to 1 the scaling policy is set to i and auto-suspend is set to 10 minutes.
How canthe performance be improved?

A. Change the cluster settings
B. Increase the size of the warehouse
C. Change the scaling policy to economy
D. Change auto-suspend to a longer time frame

Answer: B

Explanation:
The performance can be improved by increasing the size of the warehouse. The chart shows that the warehouse is queueing queries, which means that there are more queries than the warehouse can handle at its current size. Increasing the size of the warehouse will increase its processing power and concurrency limit, which could reduce the queueing time and improve the performance. The other options are not likely to improve the performance significantly. Option A, changing the cluster settings, will not help unless the minimum and maximum cluster countsare increased to allow for multi-cluster scaling. Option C, changing the scaling policy to economy, will not help because it will reduce the responsiveness of the warehouse to scale up or down based on demand. Option D, changing auto-suspend to a longer time frame, will not help because it will only affect how long the warehouse stays idle before suspending itself.

NEW QUESTION 3
Which methods will trigger an action that will evaluate a DataFrame? (Select TWO)

A. DataFrame.random_split ( )
B. DataFrame.collect ()
C. DateFrame.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
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 =iccisi_r.an=. 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 5
A large table with 200 columns contains two years of historical data. When queried. the table is filtered on a single day Below is the Query Profile:

Using a size 2XL virtual warehouse, this query look over an hour to complete What will improve the query performance the MOST?

A. increase the size of the virtual warehouse.
B. Increase the number of clusters in the virtual warehouse
C. Implement the search optimization service on the table
D. Add a date column as a cluster key on the table

Answer: D

Explanation:
Adding a date column as a cluster key on the table will improve the query performance by reducing the number of micro-partitions that need to be scanned. Since the table is filtered on a single day, clustering by date will make the query more selective and efficient.

NEW QUESTION 6
A Data Engineer ran a stored procedure containing various transactions During the execution, the session abruptly disconnected preventing one transactionfrom committing or rolling hark.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_TIMEOUTto 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 7
A stream called TRANSACTIONS_STM is created on top of a transactions table in a continuous pipeline running in Snowflake. After a couple of months, the TRANSACTIONS table is renamed transactiok3_raw to comply with new naming standards
What will happen to the TRANSACTIONS _STM object?

A. TRANSACTIONS _STMwill keep working as expected
B. TRANSACTIONS _STMwill be stale and will need to be re-created
C. TRANSACTIONS _STMwill be automatically renamedTRANSACTIONS _RAW_STM.
D. Reading from the traksactioks_3T>: stream will succeed for some time after the expected STALE_TIME.

Answer: B

Explanation:
A stream is a Snowflake object that records the history of changes made to a table. A stream is associated with a specific table at the time of creation, and it cannot be altered to point to a different table later. Therefore, if the source table is renamed, the stream will become stale and will need to be re-created with the new table name. The other options are not correct because:
? TRANSACTIONS _STM will not keep working as expected, as it will lose track of the changes made to the renamed table.
? TRANSACTIONS _STM will not be automatically renamed TRANSACTIONS _RAW_STM, as streams do not inherit the name changes of their source tables.
? Reading from the transactions_stm stream will not succeed for some time after the expected STALE_TIME, as streams do not have a STALE_TIME property.

NEW QUESTION 8
What is a characteristic of the use of external tokenization?

A. Secure data sharing can be used with external tokenization
B. External tokenization cannot be used with database replication
C. Pre-loading of unmasked data is supported with external tokenization
D. External tokenization allows (he preservation of analytical values after de-identification

Answer: D

Explanation:
External tokenization is a feature in Snowflake that allows users to replace sensitive data values with tokens that are generated and managed by an external service. External tokenization allows the preservation of analytical values after de-identification, such as preserving the format, length, or range of the original values. This way, users can perform analytics on the tokenized data without compromising the security or privacy of the sensitive data.

NEW QUESTION 9
A Data Engineer wants to centralize grant management to maximize security. A user needs ownership on a table m 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 10
A CSV file around 1 TB in size is generated daily on an on-premise server A corresponding table. Internal stage, and file format have already been created in Snowflake to facilitate the data loading process
How can the process of bringing the CSV file into Snowflake be automated using the LEAST amount of operational overhead?

A. Create a task in Snowflake that executes once a day and runs a copy into statement that references the internal stage The internal stage will read the files directly from the on-premise server and copy the newest file into the table from the on-premise server to the Snowflake table
B. On the on-premise server schedule a SQL file to run using SnowSQL that executes a PUT to push a specific file to the internal stage Create a task that executes once a day m Snowflake and runs a OOPY WTO statement that references the internal stage Schedule the task to start after the file lands in the internal stage
C. On the on-premise server schedule a SQL file to run using SnowSQL that executes a PUT to push a specific file to the internal stag
D. Create a pipe that runs a copy into statement that references the internal stage Snowpipe auto-ingest will automatically load the file from the internal stage when the new file lands in the internal stage.
E. On the on premise server schedule a Python file that uses the Snowpark Python library.The Python script will read the CSV data into a DataFrame and generate an insert into statement that will directly load into the table The script will bypass the need to move a file into an internal stage

Answer: C

Explanation:
This option is the best way to automate the process of bringing the CSV file into Snowflake with the least amount of operational overhead. SnowSQL is a command- line tool that can be used to execute SQL statements and scripts on Snowflake. By scheduling a SQL file that executes a PUT command, the CSV file can be pushed from the on-premise server to the internal stage in Snowflake. Then, by creating a pipe that runs a COPY INTO statement that references the internal stage, Snowpipe can automatically load the file from the internal stage into the table when it detects a new file in the stage. This way, there is no need to manually start or monitor a virtual warehouse or task.

NEW QUESTION 10
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 12
The following code is executed ina Snowflake environment with the default settings:

```
able customer;

transaction;

 table customer
nteger,
 varchar


 into customer values ('1', 'John');

ck;

 $1 from customer;
```

What will be the result of the select statement?

A. SQL compilation error object CUSTOMER' does not exist or is not authorized.
B. John
C. 1
D. 1John

Answer: C

NEW QUESTION 16
A company built a sales reporting system with Python, connecting to Snowflake using the Python Connector. Based on the user's selections, the system generates the SQL queries needed to fetch the data for the report First it gets the customers that meet the given query parameters (on average 1000 customer records for each report run) and then it loops the customer records sequentially Inside that loop it runs the generated SQL clause for the current customer to get the detailed data for that customer number from the sales data table
When the Data Engineer tested the individual SQL clauses they were fast enough (1 second to get the customers 0 5 second to get the sales data for one customer) but the total runtime of the report is too long
How can this situation be improved?

A. Increase the size of the virtual warehouse
B. Increase the number of maximum clusters of the virtual warehouse
C. Define a clustering key for the sales data table
D. Rewrite the report to eliminate the use of the loop construct

Answer: D

Explanation:
This option is the best way to improve the situation, as using a loop construct to run SQL queries for each customer is very inefficient and slow. Instead, the report should be rewritten to use a single SQL query that joins the customer and sales data tables and applies the query parameters as filters. This way, the report can

leverage Snowflake's parallel processing and optimization capabilities and reduce the network overhead and latency.

NEW QUESTION 21
The JSON below is stored in a variant column named v in a table named jCustRaw:

```
id": "6282638561cf48544e2ef7e9",
ompany": "FLYBOYZ",
sActive": true,
ame": "Dean Head",
eamMembers": [
  {
      "age": 29,
      "eyeColor": "green",
      "name": "Dominique Grimes",
      "registered": "2017-02-19T06:12:36 +06:00"
  },
  {
      "age": 39,
      "eyeColor": "green",
      "name": "Pearl Dunlap",
      "registered": "2018-05-12T09:21:42 +05:00"
  },
  {
      "age": 22,
      "eyeColor": "blue",
      "name": "Cardenas Warren",
      "registered": "2019-04-08T01:24:29 +05:00"

      }
  ]
}
```

Which query will return one row per team member (stored in the teamMembers array) along all of the attributes of each team member?
A)

```
select
      t2.name AS memberName
      ,t2.registered AS registeredDttm
      ,t2.age AS age
      ,t2.eyeColor AS eyeColor
from jCustRaw t1
  ateral flatten(v) t2
se
  lect                    Name
     t2.value:name::va     registeredDttm
     ,t2.value:registe
     ,t2.value:age::nu    yeColor
fr   ,t2.value:eyeColo
,l   om jCustRaw t1       ;
  ateral flatten(input
```

C)

```
select
     v:teamMembers.name::varchar AS memberName
     ,v:teamMembers.registered::timestamp AS
     registeredDttm
     ,v:teamMembers.age::number AS age
     ,v:teamMembers.eyeColor::varchar AS eyeColor
from jCustRaw;
```

D)

```
select
     v:teamMembers[0].name::varchar AS memberName
     ,v:teamMembers[0].registered::timestamp AS registeredDttm
     ,v:teamMembers[0].age::number AS age
     ,v:teamMembers[0].eyeColor::varchar AS eyeColor
from jCustRaw;
```

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

Answer: B

NEW QUESTION 23
What is the purpose of the BUILD_FILE_URL function in Snowflake?

A. It generates an encrypted URL foe accessing a file in a stage.
B. It generates a staged URL for accessing a file in a stage.
C. It generates a permanent URL for accessing files in a stage.
D. It generates a temporary URL for accessing a file in a stage.

Answer: B

Explanation:
The BUILD_FILE_URL function in Snowflake generates a temporary URL for accessing a file in a stage. The function takes two arguments: the stage name and the file path. The generated URL is valid for 24 hours and can be used to download or view the file contents. The other options are incorrect because they do not describe the purpose of the BUILD_FILE_URL function.

NEW QUESTION 27
A database contains a table and a stored procedure defined as.

```
CREATE OR REPLACE TABLE log_table(col1 VARCHAR);

CREATE OR REPLACE PROCEDURE insert_log(input VARCHAR)
RETURNS FLOAT
LANGUAGE JAVASCRIPT
RETURNS NULL ON NULL INPUT
AS
'
var rs = snowflake.execute({sqlText:`INSERT INTO log_table(col1) VALUES(:1);`
,binds:[INPUT]});

return 1;
';
```

The `log_table` is initially empty and a Data Engineer issues the following command:

```
CALL insert_log(NULL::VARCHAR);
```

No other operations are affecting the log_table. What will be the outcome of the procedure call?

A. The log_table contains zero records and the stored procedure returned 1 as a return value
B. The log_table contains one record and the stored procedure returned 1 as a return value
C. The log_table contains one record and the stored procedure returned NULL as a return value
D. The log_table contains zero records and the stored procedure returned NULL as a return value

Answer: B

Explanation:
The stored procedure is defined with a FLOAT return type and a JavaScript language. The body of the stored procedure contains a SQL statement that inserts a row into the log_table with a value of '1' for col1. The body also contains a return statement that returns 1 as a float value. When the stored procedure is called with any VARCHAR parameter, it will execute successfully and insert one record into the log_table and return 1 as a return value. The other options are not correct because:
? The log_table will not be empty after the stored procedure call, as it will contain
one record inserted by the SQL statement.
? The stored procedure will not return NULL as a return value, as it has an explicit return statement that returns 1.

NEW QUESTION 29
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.

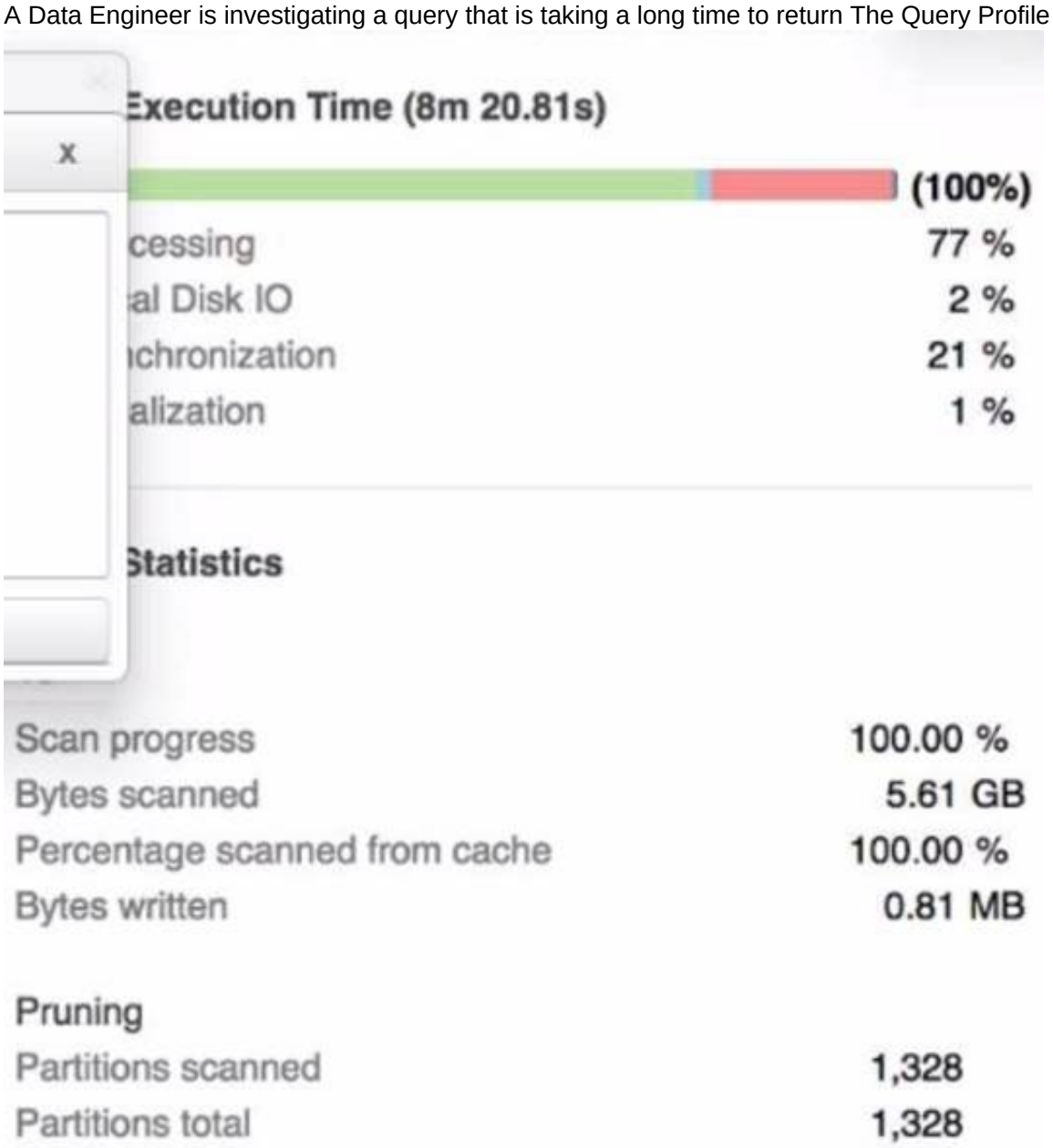
NEW QUESTION 31
Assuming a Data Engineer has all appropriate privileges and context which statements would be used to assess whether the User-Defined Function (UDF), MTBATA3ASZ. SALES .REVENUE_BY_REGION, exists and is secure? (Select TWO)

A. SHOW DS2R FUNCTIONS LIKE 'REVEX'^BYJIESION' IN SCHEMA SALES;
B. SELECT IS_SECURE FROM SNOWFLAK
C. INFCRXATION_SCKZM
D. FUNCTIONS WHERE FUNCTI0N_3CHEMA = 'SALES' AND FUNCTI CN_NAXE =•ftEVEXUE_BY_RKXQH4;
E. SELECT IS_SEC"JRE FROM INFOR>LVTICN_SCHEM
F. FUNCTIONS WHERE FUNCTION_SCHEMA = 'SALES1 AND FUNGTZON_NAME = ' REVENUE_BY_REGION';
G. SHOW EXTERNAL FUNCTIONS LIKE ‘REVENUE_BY_REGION’IB SCHEMA SALES;
H. SHOW SECURE FUNCTIONS LIKE 'REVENUE 3Y REGION' IN SCHEMA SALES;

Answer: AB

Explanation:
The statements that would be used to assess whether the UDF, MTBATA3ASZ. SALES .REVENUE_BY_REGION, exists and is secure are:
? SHOW DS2R FUNCTIONS LIKE ‘REVEX’^BYJIESION’ IN SCHEMA SALES;:
This statement will show information about the UDF, including its name, schema, database, arguments, return type, language, and security option. If the UDF does not exist, the statement will return an empty result set.
? SELECT IS_SECURE FROM SNOWFLAKE. INFCRXATION_SCKZMA.
FUNCTIONS WHERE FUNCTI0N_3CHEMA = ‘SALES’ AND FUNCTI CN_NAXE
= •ftEVEXUE_BY_RKXQH4;: This statement will query the SNOWFLAKE.INFORMATION_SCHEMA.FUNCTIONS view, which contains metadata about the UDFs in the current database. The statement will return the IS_SECURE column, which indicates whether the UDF is secure or not. If the UDF does not exist, the statement will return an empty result set. The other statements are not correct because:
? SELECT IS_SEC"JRE FROM INFOR>LVTICN_SCHEMA. FUNCTIONS WHERE
FUNCTION_SCHEMA = ‘SALES1 AND FUNGTZON_NAME = ’
REVENUE_BY_REGION’;: This statement will query the INFORMATION_SCHEMA.FUNCTIONS view, which contains metadata about the UDFs in the current schema. However, the statement has a typo in the schema name (‘SALES1’ instead of ‘SALES’), which will cause it to fail or return incorrect results.
? SHOW EXTERNAL FUNCTIONS LIKE ‘REVENUE_BY_REGION’ IB SCHEMA
SALES;: This statement will show information about external functions, not UDFs. External functions are Snowflake functions that invoke external services via HTTPS requests and responses. The statement will not return any results for the UDF.
? SHOW SECURE FUNCTIONS LIKE ‘REVENUE 3Y REGION’ IN SCHEMA
SALES;: This statement is invalid because there is no such thing as secure functions in Snowflake. Secure functions are a feature of some other databases, such as PostgreSQL, but not Snowflake. The statement will cause a syntax error.

NEW QUESTION 35
A Data Engineer is investigating a query that is taking a long time to return The Query Profile shows the following:



What step should the Engineer take to increase the query performance?

A. Add additional virtual warehouses.
B. increasethe size of the virtual warehouse.
C. Rewrite the query using Common Table Expressions (CTEs)
D. Change the order of the joins and start with smaller tables first

Answer: B

Explanation:
The step that the Engineer should take to increase the query performance is to increase the size of the virtual warehouse. The Query Profile shows that most of the time was spent on local disk IO, which indicates that the query was reading a lot of data from disk rather than from cache. This could be due to a large amount of data being scanned or a low cache hit ratio. Increasing the size of the virtual warehouse will increase the amount of memory and cache available for the query, which could reduce the disk IO time and improve the query performance. The other options are not likely to increase the query performance significantly. Option A, adding additional virtual warehouses, will not help
unless they are used in a multi-cluster warehouse configuration or for concurrent queries. Option C, rewriting the query using Common Table Expressions (CTEs), will not affect the amount of data scanned or cached by the query. Option D, changing the order of the joins and starting with smaller tables first, will not reduce the disk IO time unless it also reduces the amount of data scanned or cached by the query.

NEW QUESTION 40
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 44
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 querywill provide the status of the pipe?

A. SELECT FROM SYSTEM$PIPE_STATUS ("test.'extract'.my_pipe"i:
B. SELECT FROM SYSTEM$PIPE_STATUS (,test.,,Extracr,,.ny_pipe, i I
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 49
A Data Engineer needs to ingest invoice data in PDF format into Snowflake so that the data can be queried and used in a forecasting solution.
..... recommended way to ingest this data?

A. Use Snowpipe to ingest the files that land in an external stage into a Snowflake table
B. Use a COPY INTO command to ingest the PDF files in an external stage into a Snowflake table with a VARIANT column.
C. Create an external table on the PDF files that are stored in a stage and parse the data nto structured data
D. Create a Java User-Defined Function (UDF) that leverages Java-based PDF parser libraries to parse PDF data into structured data

Answer: D

Explanation:
The recommended way to ingest invoice data in PDF format into Snowflake
is to create a Java User-Defined Function (UDF) that leverages Java-based PDF parser libraries to parse PDF data into structured data. This option allows for more flexibility and control over how the PDF data is extracted and transformed. The other options are not suitable for ingesting PDF data into Snowflake. Option A and B are incorrect because Snowpipe and COPY INTO commands can only ingest files that are in supported file formats, such as CSV, JSON, XML, etc. PDF files are not supported by Snowflake and will cause errors or unexpected results. Option C is incorrect because external tables can only query files that are in supported file formats as well. PDF files cannot be parsed by external tables and will cause errors or unexpected results.

NEW QUESTION 51
Assuming that the session parameter USE_CACHED_RESULT is set to false, what are characteristics of Snowflake virtual warehouses in terms of the use of Snowpark?

A. Creating a DataFrame from a table will start a virtual warehouse
B. Creating a DataFrame from a staged file with the read () method will start a virtual warehouse
C. Transforming a DataFrame with methods like replace () will start a virtual warehouse -
D. Calling a Snowpark stored procedure to query the database with session, call () will start a virtual warehouse

Answer: A

Explanation:
Creating a DataFrame from a table will start a virtual warehouse because it requires reading data from Snowflake. The other options will not start a virtual warehouse because they either operate on local data or use an existing session to query Snowflake.

NEW QUESTION 54
A secure function returns data coming through an inbound share
What will happen if a Data Engineer tries to assign usage privileges on this function to an outbound share?

A. An error will be returned because the Engineer cannot share data that has already been shared
B. An error will be returned because only views and secure stored procedures can be shared
C. An error will be returned because only secure functions can be shared with inboundshares
D. The Engineer will be able to share the secure function with other accounts

Answer: A

Explanation:
An error will be returned because the Engineer cannot share data that has already been shared. A secure function is a Snowflake function that can access data from an inbound share, which is a share that is created by another account and consumed by the current account. A secure function can only be shared with an inbound share, not an outbound share, which is a share that is created by the current account and shared with other accounts. This is to prevent data leakage or unauthorized access to the data from the inbound share.

NEW QUESTION 58
A Data Engineer is implementing a near real-time ingestionpipeline to toad 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 oneinternal 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 59
Which output is provided by both theSYSTEM$CLUSTERING_DEPTHfunction and theSYSTEM$CLUSTERING_INFORMATIONfunction?

A. average_depth
B. notes
C. average_overlaps
D. total_partition_count

Answer: A

Explanation:
The output that is provided by both the SYSTEM$CLUSTERING_DEPTH function and the SYSTEM$CLUSTERING_INFORMATION function is average_depth. This output indicates the average number of micro-partitions that contain data for a given column value or combinationof column values. The other outputs are not common to both functions. The notes output is only provided by the SYSTEM$CLUSTERING_INFORMATION function and it contains additional information or recommendations about the clustering status of the table. The average_overlaps output is only provided by the SYSTEM$CLUSTERING_DEPTH function and it indicates the average number of micro-partitions that overlap with other micro-partitions for a given column value or combination of column values. The total_partition_count output is only provided by the SYSTEM$CLUSTERING_INFORMATION function and it indicates the total number of micro-partitions in the table.

NEW QUESTION 64
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 68
Which methods can be used to create a DataFrame object in Snowpark? (Select THREE)

A. session.jdbc_connection()
B. session.read.json{)
C. session,table()
D. DataFraas.writeO
E. session.builder()
F. session.sql()

Answer: BCF

Explanation:
The methods that can be used to create a DataFrame object in Snowpark are session.read.json(), session.table(), and session.sql(). These methods can create a DataFrame from different sources, such as JSON files, Snowflake tables, or SQL queries.
The other options are not methods that can create a DataFrame object in Snowpark. Option A, session.jdbc_connection(), is a method that can create a JDBC connection object to connect to a database. Option D, DataFrame.write(), is a method that can write a DataFrame to a destination, such as a file or a table. Option E, session.builder(), is a method that can create a SessionBuilder object to configure and build a Snowpark session.

NEW QUESTION 73
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 74
.....

THANKS FOR TRYING THE DEMO OF OUR PRODUCT

Visit Our Site to Purchase the Full Set of Actual DEA-C01 Exam Questions With Answers.

We Also Provide Practice Exam Software That Simulates Real Exam Environment And Has Many Self-Assessment Features. Order the DEA-C01 Product From:

https://www.2passeasy.com/dumps/DEA-C01/

Money Back Guarantee

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