Snowflake

Exam Questions DEA-C01

SnowPro Advanced: Data Engineer Certification Exam

NEW QUESTION 1
A Data Engineer needs to load JSON output from some software into Snowflake using Snowpipe.
Which recommendations apply to this scenario? (Select THREE)

A. Load large files (1 GB or larger)
B. Ensure that data files are 100-250 MB (or larger) in size compressed
C. Load a single huge array containing multiple records into a single table row
D. Verify each value of each unique element stores a single native data type (string or number)
E. Extract semi-structured data elements containing null values into relational columns before loading
F. Create data files that are less than 100 MB and stage them in cloud storage at a sequence greater than once each minute

Answer: BDF

Explanation:
 The recommendations that apply to this scenario are:
? Ensure that data files are 100-250 MB (or larger) in size compressed: This recommendation will improve Snowpipe performance by reducing the number of files that need to be loaded and increasing the parallelism of loading. Smallerfiles can cause performance degradation or errors due to excessive metadata operations or network latency.
? Verify each value of each unique element stores a single native data type (string or number): This recommendation will improve Snowpipe performance by avoiding data type conversions or errors when loading JSON data into variant columns. Snowflake supports two native data types for JSON elements: string and number. If an element has mixed data types across different files or records, such as string and boolean, Snowflake will either convert them to string or raise an error, depending on the FILE_FORMAT option.
? Create data files that are less than 100 MB and stage them in cloud storage at a sequence greater than once each minute: This recommendation will minimize Snowpipe costs by reducing the number of notifications that need to be sent to Snowpipe for auto-ingestion. Snowpipe charges for notifications based on the number of files per notification and the frequency of notifications. By creating smaller files and staging them at a lower frequency, fewer notifications will be needed.

NEW QUESTION 2
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 3
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 4
What is a characteristic of the use of binding variables in JavaScript stored procedures in Snowflake?

A. All types of JavaScript variables can be bound
B. All Snowflake first-class objects can be bound
C. Only JavaScript variables of type number, string and sf Date can be bound
D. Users are restricted from binding JavaScript variables because they create SQL injection attack vulnerabilities

Answer: C

Explanation:
A characteristic of the use of binding variables in JavaScript stored procedures in Snowflake is that only JavaScript variables of type number, string and sf Date can be bound. Binding variables are a way to pass values from JavaScript variables to SQL statements within a stored procedure. Binding variables can improve the security and performance of the stored procedure by preventing SQL injection attacks and reducing the parsing overhead. However, not all types of JavaScript variables can be bound. Only the primitive types number and string, and the Snowflake-specific type sf Date, can be bound. The other options are incorrect because they do not describe a characteristic of the use of binding variables in JavaScript stored procedures in Snowflake. Option A is incorrect because authenticator is not a type of JavaScript variable, but a parameter of the snowflake.connector.connect function. Option B is incorrect because arrow_number_to_decimal is not a type of JavaScript variable, but a parameter of the snowflake.connector.connect function. Option D is incorrect because users are not restricted from binding JavaScript variables, but encouraged to do so.

NEW QUESTION 5
A Data Engineer defines the following masking policy:

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

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

```
TABLE customer(

ame VARCHAR,
me VARCHAR,
me 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_nam ADD MASKING POLICYname_poiicy;
C. ALTER TABLE customer MODIFY COLUMN first_nane SET MASKING POLICY name_policy; lasT_name SET MASKING POLICY name_pclicy;
D. ALTER TABLE customer MODIFY COLUMN first_name ADD MASKING POLICY name_policy,

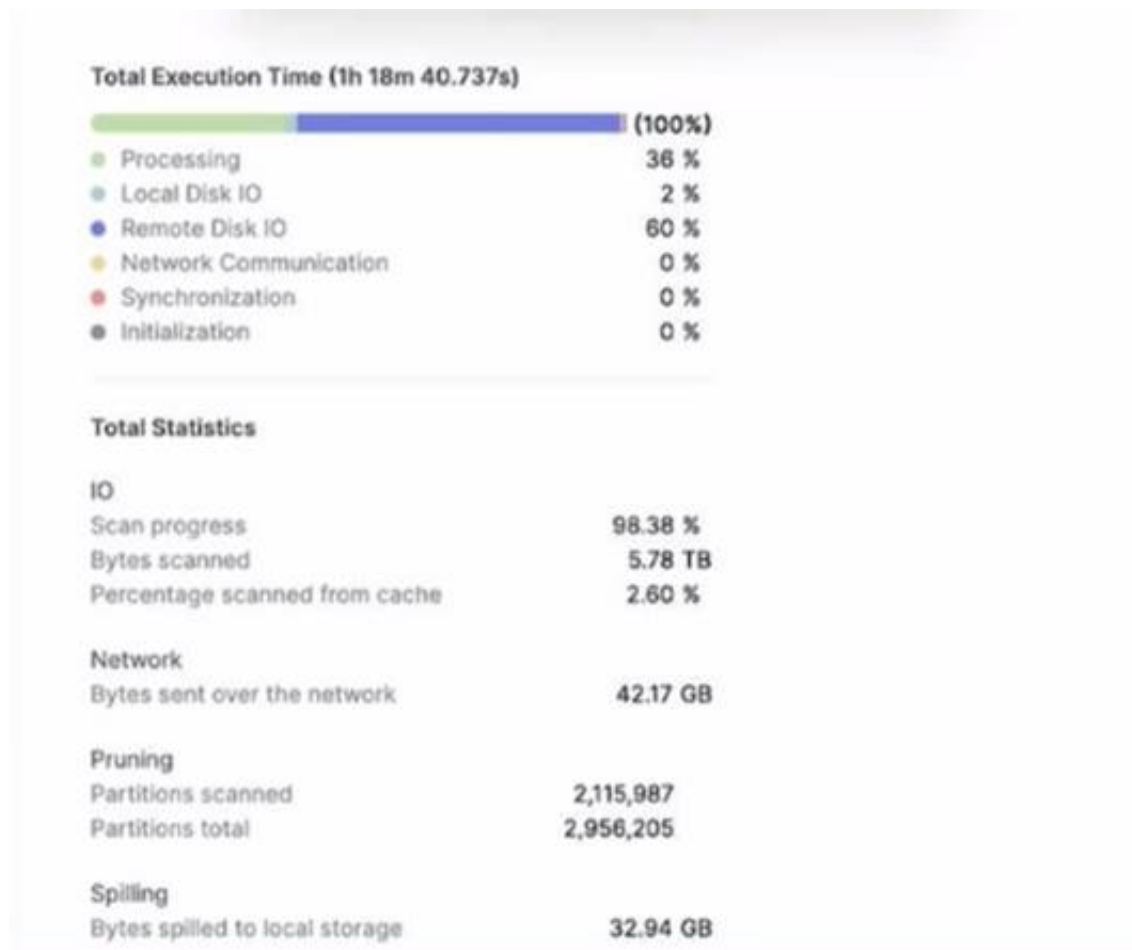**Answer:** A

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 6
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:

Total Execution Time (1h 18m 40.737s)

	(100%)
● Processing	36 %
● Local Disk IO	2 %
● Remote Disk IO	60 %
● Network Communication	0 %
● Synchronization	0 %
● Initialization	0 %

Total Statistics

IO
Scan progress	98.38 %
Bytes scanned	5.78 TB
Percentage scanned from cache	2.60 %

Network
Bytes sent over the network	42.17 GB

Pruning
Partitions scanned	2,115,987
Partitions total	2,956,205

Spilling
Bytes spilled to local storage	32.94 GB

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 7
Which callback function is required within a JavaScript User-Defined Function (UDF) for it to execute successfully?

A. initialize ()
B. processRow ()
C. handler
D. finalize ()

Answer: B

Explanation:
The processRow () callback function is required within a JavaScript UDF for it to execute successfully. This function defines how each row of input data is processed and what output is returned. The other callback functions are optional and can be used for initialization, finalization, or error handling.

NEW QUESTION 8
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 9
The following is returned fromSYSTEMCLUSTERING_INFORMATION () for a tablenamed orders with adate column named O_ORDERDATE:

```
{
  "cluster_by_keys" : "LINEAR(YEAR(O_ORDERDATE))",
  "total_partition_count" : 536,
  "total_constant_partition_count" : 493,
  "average_overlaps" : 0.1716,
  "average_depth" : 1.0914,
  "partition_depth_histogram" : {
    "00000" : 0,
    "00001" : 491,
    "00002" : 41,
    "00003" : 4,
    "00004" : 0,
    "00005" : 0,
    "00006" : 0,
    "00007" : 0,
    "00008" : 0,
    "00009" : 0,
    "00010" : 0,
    "00011" : 0,
    "00012" : 0,
    "00013" : 0,
    "00014" : 0,
    "00015" : 0,
    "00016" : 0
  }
}
```

What does the total_constant_partition_count value indicate about this table?

A. The table is clustered very well on_ORDERDATE, as there are 493 micro-partitions that could not be significantly improved by reclustering
B. The table is not clustered well on O_ORDERDATE, as there are 493 micro-partitions where the range of values in that column overlap with every other micro partition in the table.
C. The data inO_ORDERDATEdoes not change very often as there are 493 micro- partitionscontaining rows where that column has not been modified since the row was created
D. The data inO_ORDERDATEhas a very low cardinality as there are 493 micro-partitions where there is only a single distinct value in that column for all rows in the micro-partition

Answer: B

Explanation:
The total_constant_partition_count value indicates the number of micro- partitions where the clustering key column has a constant value across all rows in the micro-partition. However, this does not necessarily mean that the table is clustered well on that column, as there could be other micro-partitions where the range of values in that column overlap with each other. This is the case for the orders table, as the clustering depth is 1, which means that every micro-partition overlaps with every other micro-partition on O_ORDERDATE. This indicates that the table is not clustered well on O_ORDERDATE and could benefit from reclustering.

NEW QUESTION 10
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
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 andset 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 bySnowpipe, 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 15
A Data Engineer enables a result cache at the session level with the following command: ALTER SESSION SET USE CACHED RESULT = TRUE;
The Engineer then runs the following select query twice without delay:

```
SELECT *
FROM SNOWFLAKE_SAMPLE_DATA.TPCH_SF1.CUSTOMER
SAMPLE(10) SEED (99);
```

The underlying table does not change between executions What are the results of both runs?

A. The first and second run returned the same results because sample is deterministic
B. The first and second run returned the same results, because the specific SEEDvalue was provided.
C. The first and second run returned different results because the query is evaluated each time it is run.
D. The first and second run returned differentresults because the query uses *instead of an explicit column list

Answer: B

Explanation:
 The result cache is enabled at the session level, which means that repeated queries will return cached results if there is no change in the underlying data or session parameters. However, in this case, the result cache is not relevant because the query uses a specific SEED value for sampling, which makes it deterministic. Therefore, both runs will return the same results regardless of caching.

NEW QUESTION 20
A Data Engineer has written a stored procedure that will run with caller's rights. The Engineer has granted ROLEA right to use this stored procedure.
What is a characteristic of the stored procedure being called using ROLEA?

A. The stored procedure must run with caller's rights it cannot be converted later to run with owner's rights
B. If the stored procedure accesses an object that ROLEA does not have access to the stored procedure will fail
C. The stored procedure will run in the context (database and schema) where the owner created the stored procedure
D. ROLEA will not be able to see the source code for the stored procedure even though the role has usage privileges on the stored procedure

Answer: B

Explanation:
 A stored procedure that runs with caller's rights executes with the privileges of the role that calls it. Therefore, if the stored procedure accesses an object that ROLEA does not have access to, such as a table or a view, the stored procedure will fail with an insufficient privileges error. The other options are not correct because:
? A stored procedure can be converted from caller's rights to owner's rights by using the ALTER PROCEDURE command with the EXECUTE AS OWNER option.
? A stored procedure that runs with caller's rights executes in the context (database and schema) of the caller, not the owner.
? ROLEA will be able to see the source code for the stored procedure by using the GET_DDL function or the DESCRIBE command, as long as it has usage privileges on the stored procedure.

NEW QUESTION 22
What is a characteristic of the operations of streams in Snowflake?

A. Whenever a stream is queried, the offset is automatically advanced.
B. When a stream is used to update a target table the offset is advanced to the current time.
C. Querying a stream returns all change records and table rows from the current offset to the current time.
D. Each committed and uncommitted transaction on the source table automatically puts a change record in the stream.

Answer: C

Explanation:
 A stream is a Snowflake object that records the history of changes made to a
table. A stream has an offset, which is a point in time that marks the beginning of the change records to be returned by the stream. Querying a stream returns all change records and table rows from the current offset to the current time. The offset is not automatically advanced by querying the stream, but it can be manually advanced by using the ALTER STREAM command. When a stream is used to update a target table, the offset is advanced to the current time only if the ON UPDATE clause is specified in the stream definition. Each committed transaction on the source table automatically puts a change record in the stream, but uncommitted transactions do not.

NEW QUESTION 24
The following code is executed ina Snowflake environment with the default settings:

```
able customer;

transaction;

table customer
teger,
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 25
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
                            registeredDttm
            t2.value:name::va
            ,t2.value:registe
                            yeColor
            ,t2.value:age::nu
        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 27
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 32
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 33
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 34
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 39
A company is building a dashboard for thousands of Analysts. The dashboard presents the results of a few summary queries on tables that are regularly updated. The query conditions vary by tope according to what data each Analyst needs Responsiveness of the dashboard queries is a top priority, and the data cache should be preserved.

How should the Data Engineer configure the compute resources to support this dashboard?

A. Assign queries to a multi-cluster virtual warehouse with economy auto-scaling Allow the system to automatically start and stop clusters according to demand.
B. Assign all queries to a multi-cluster virtual warehouse set to maximized mode Monitor to determine the smallest suitable number of clusters.
C. Create a virtual warehouse for every 250 Analysts Monitor to determine how many of these virtual warehouses are being utilized at capacity.
D. Create a size XL virtual warehouse to support all the dashboard queries Monitor query runtimes to determine whether the virtual warehouse should be resized.

Answer: B

Explanation:
This option is the best way to configure the compute resources to support this dashboard. By assigning all queries to a multi-cluster virtual warehouse set to maximized mode, the Data Engineer can ensure that there is enough compute capacity to handle thousands of concurrent queries from different analysts. A multi-cluster virtual warehouse can scale up or down by adding or removing clusters based on the load. A maximized scaling policy ensures that there is always at least one cluster running and that new clusters are added as soon as possible whenneeded. By monitoring the utilization and performance of the virtual warehouse, the Data Engineer can determine the smallest suitable number of clusters that can meet the responsiveness requirement and minimize costs.

NEW QUESTION 44
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 45
How can the following relational data be transformed into semi-structured data using the LEAST amount of operational overhead?

```
create table provinces (province varchar, created_date date);
```

Row	PROVINCE	CREATED_DATE
2	Alberta	2020-01-19
1	Manitoba	2020-01-18

A. Use the to_json function
B. Use the PAESE_JSON function to produce a variant value
C. Use the OBJECT_CONSTRUCT function to return a Snowflake object
D. Use the TO_VARIANT function to convert each of the relational columns to VARIANT.

Answer: C

Explanation:
This option is the best way to transform relational data into semi-structured data using the least amount of operational overhead. The OBJECT_CONSTRUCT function takes a variable number of key-value pairs as arguments and returns a Snowflake object, which is a variant type that can store JSON data. The function can be used to convert each row of relational data into a JSON object with the column names as keys and the column values as values.

NEW QUESTION 46
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 47
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 50
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 55
At what isolation level are Snowflake streams?

A. Snapshot
B. Repeatable read
C. Read committed
D. Read uncommitted

Answer: B

Explanation:
The isolation level of Snowflake streams is repeatable read, which means that each transaction sees a consistent snapshot of data that does not change during its execution. Streams use time travel internally to provide this isolation level and ensure that queries on streams return consistent results regardless of concurrent transactions on their source tables.

NEW QUESTION 60
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 65
A Data Engineer is writing a Python script using the Snowflake Connector for Python. The Engineer will use the snowflake. Connector.connect function to connect to Snowflake The requirementsare:
*Raise an exception if the specified database schema or warehouse does not exist
*improve download performance
Whichparameters of the connect function should be used? (Select TWO).

A. authenticator
B. arrow_nunber_to_decimal
C. client_prefetch_threads
D. client_session_keep_alivs
E. validate_default_parameters

Answer: CE

Explanation:

The parameters of the connect function that should be used are client_prefetch_threads and validate_default_parameters. The client_prefetch_threads parameter controls the number of threads used to download query results from Snowflake. Increasing this parameter can improve download performance by parallelizing the download process. The validate_default_parameters parameter controls whether an exception should be raised if the specified database, schema, or warehouse does not exist or is not authorized. Setting this parameter to True can help catch errors early and avoid unexpected results.

NEW QUESTION 69
......

Thank You for Trying Our Product

We offer two products:

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

2nd - Questons 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