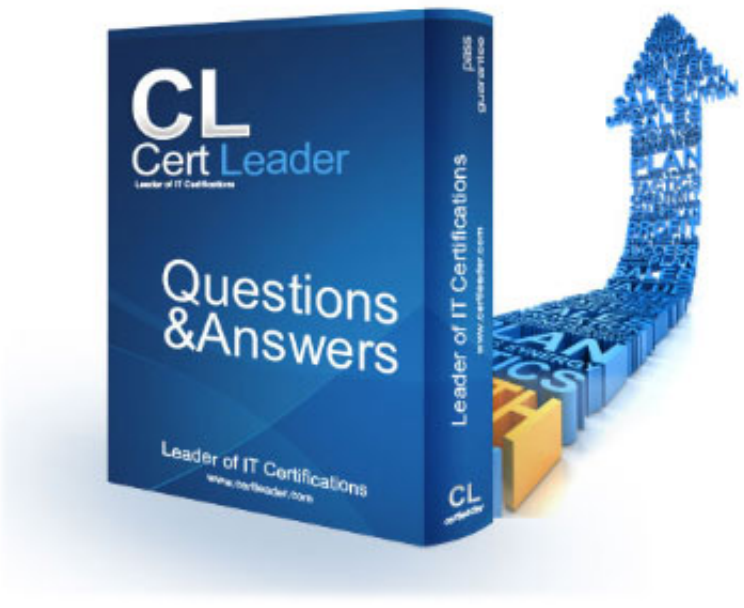
DEA-C01 Dumps

SnowPro Advanced: Data Engineer Certification Exam

https://www.certleader.com/DEA-C01-dumps.html



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
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 3
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_daysset 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 -able 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 => -3€0G*24*3);
D. Create a Snowflake Support case lo restore the database and tab e from "a i-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 4
A new customer table is created by a data pipeline in a Snowflake schema where MANAGED ACCESSenabled.
…. 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 5
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 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 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 8
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 9
What kind of Snowflake integration is required when defining an external function in Snowflake?

A. API integration
B. HTTP integration
C. Notification integration
D. Security integration

Answer: A

Explanation:
An API integration is required when defining an external function in Snowflake. An API integration is a Snowflake object that defines how Snowflake communicates with an externalservice via HTTPS requests and responses. An API integration specifies parameters such as URL, authentication method, encryption settings, request headers, and timeout values. An API integration is used to create an external function object that invokes the external service from within SQL queries.

NEW QUESTION 10
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 10
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 CompanyA'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 setsthey 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'saccount 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 12
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 14
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 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 17
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 21
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 22
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 23
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
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 33
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 36
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 40
When would a Data engineer use table with the flatten function instead of the lateral flatten combination?

A. When TABLE with FLATTENrequires another source in the from clause to refer to
B. WhenTABLE with FLATTENrequires no additional source m the from clause to refer to
C. Whenthe LATERALFLATTENcombination requires no other source m the from clause to refer to
D. When table withFLATTENis acting like a sub-query executed for each returned row

Answer: A

Explanation:
The TABLE function with the FLATTEN function is used to flatten semi- structured data, such as JSON or XML, into a relational format. The TABLE function returns a table expression that can be used in the FROM clause of a query. The TABLE function with the FLATTEN function requires another source in the FROM clause to refer to, such as a table, view, or subquery that contains the semi-structured data. For example: SELECT t.value:city::string AS city, f.value AS population FROM cities t, TABLE(FLATTEN(input => t.value:population)) f;
In this example, the TABLE function with the FLATTEN function refers to the cities table in the FROM clause, which contains JSON data in a variant column named value. The FLATTEN function flattens the population array within each JSON object and returns a table expression with two columns: key and value. The query then selects the city and population values from the table expression.

NEW QUESTION 42
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 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 50
A company has an extensive script in Scala that transforms data by leveraging DataFrames. A Data engineer needs to move these transformations to Snowpark. …characteristics of data transformations in Snowpark should be considered to meet this requirement? (Select TWO)

A. It is possible to join multiple tables using DataFrames.
B. Snowpark operations are executed lazily on the server.
C. User-Defined Functions (UDFs) are not pushed down to Snowflake
D. Snowpark requires a separate cluster outside of Snowflake for computations
E. Columns in different DataFrames with the same name should be referred to with squared brackets

Answer: AB

Explanation:
The characteristics of data transformations in Snowpark that should be considered to meet this requirement are:
? It is possible to join multiple tables using DataFrames.
? Snowpark operations are executed lazily on the server.
These characteristics indicate how Snowpark can perform data transformations using DataFrames, which are similar to the ones used in Scala. DataFrames are distributed collections of rows that can be manipulated using various operations, such as joins, filters, aggregations, etc. DataFrames can be created from different sources, such as tables, files, or SQL queries. Snowpark operations are executed lazily on the server, which means that they are not performed until an action is triggered, such as a write or a collect operation. This allows Snowpark to optimize the execution plan and reduce the amount of data transferred between the client and the server.
The other options are not characteristics of data transformations in Snowpark that should be considered to meet this requirement. Option C is incorrect because User-Defined Functions (UDFs) are pushed down to Snowflake and executed on the server. Option D is incorrect because Snowpark does not require a separate cluster outside of Snowflake for computations, but rather uses virtual warehouses within Snowflake. Option E is incorrect because columns in different DataFrames with the same name should be referred to with dot notation, not squared brackets.

NEW QUESTION 55
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 60
Which Snowflake objects does the Snowflake Kafka connector use? (Select THREE).

A. Pipe
B. Serverless task
C. Internal user stage
D. Internal table stage
E. Internal named stage
F. Storage integration

Answer: ADE

Explanation:
The Snowflake Kafka connector uses three Snowflake objects: pipe, internal table stage, and internal named stage. The pipe object is used to load data from an external stage into a Snowflake table using COPY statements. The internal table stage is used to store files that are loaded from Kafka topics into Snowflake using PUT commands. The internal named stage is used to store files that are rejected by the COPY statements due to errors or invalid data. The other options are not objects that are used by the Snowflake Kafka connector. Option B, serverless task, is an object that can execute SQL statements on a schedule without requiring a warehouse. Option C, internal user stage, is an object that can store files for a specific user in Snowflake using PUT commands. Option F, storage integration, is an object that can enable secure access to external cloud storage services without exposing credentials.

NEW QUESTION 61
.....

Thank You for Trying Our Product

* 100% Pass or Money Back

All our products come with a 90-day Money Back Guarantee.

* One year free update

You can enjoy free update one year. 24x7 online support.

* Trusted by Millions

We currently serve more than 30,000,000 customers.

* Shop Securely

All transactions are protected by VeriSign!

100% Pass Your DEA-C01 Exam with Our Prep Materials Via below:

https://www.certleader.com/DEA-C01-dumps.html