

SnowPro Core Certification DEA-C02

SnowPro Advanced Data Engineer

Thank You for Downloading DEA-C02 Updated Exam Questions

<https://certs4sale.com/snowflake/dea-c02-pdf-exam-dumps>

QUESTION: 1

An existing clustering key is copied in which of the below scenarios

Option A : CREATE TABLE...CLONE

Option B : CREATE TABLE...LIKE

Option C : CREATE TABLE...AS SELECT

Correct Answer: A

Explanation/Reference:

Explanation

An existing clustering key is copied when a table is created using CREATE TABLE ... CLONE.

An existing Clustering key is not propagated when a table is created using CREATE TABLE ... LIKE.

An existing clustering key is not supported when a table is created using CREATE TABLE ... AS SELECT; however, you can define a clustering key after the table is created.

QUESTION: 2

Which of the following statements is/are incorrect regarding Fail-safe data recovery?

Option A : Data stored in temporary tables is not recoverable after the table is dropped as they do not have fail-safe.

Option B : Historical data in transient tables can be recovered by Snowflake due to Operation fail-ure after the Time Travel retention period ends using Fail-safe.

Option C : Long-lived tables, such as fact tables, should always be defined as permanent to ensure they are fully protected by Fail-safe.

Option D : Short-lived tables (i.e. <1 day), such as ETL work tables, can be defined as transient to eliminate Fail-safe costs.

Option E : If downtime and the time required to reload lost data are factors, permanent tables, even with their added Fail-safe costs, may offer a better overall solution than transient tables.

Correct Answer: B

QUESTION: 3

Which of the following objects are automatically created by the KAFKA connector?

Option A : One internal stage to temporarily store data files for each topic.
Option B : One pipe to ingest the data files for each topic partition.
Option C : One table for each topic

Option D : One external stage to temporarily store data files for each topic.

Correct Answer: A,B,C

Explanation/Reference:

Explanation

<https://docs.snowflake.com/en/user-guide/kafka-connector-overview.html#workflow-for-the-kafka-connector>

Workflow for the Kafka Connector

The Kafka connector completes the following process to subscribe to Kafka topics and create Snowflake objects:

The Kafka connector subscribes to one or more Kafka topics based on the configuration information provided via the Kafka configuration file or command line (or the Confluent Control Center; Confluent only).

The connector creates the following objects for each topic:

One internal stage to temporarily store data files for each topic.

One pipe to ingest the data files for each topic partition.

One table for each topic. If the table specified for each topic does not exist, the connector creates it; otherwise, the connector creates the RECORD_CONTENT and RECORD_METADATA columns in the existing table and verifies that the other columns are nullable (and produces an error if they are not).

QUESTION: 4

If dataframes are used, the snowflake spark connector does not support the following operations

Option A : SELECT
Option B : SHOW
Option C : DESC
Option D : INSERT

Correct Answer: B,C,D

Explanation/Reference:

Explanation

<https://docs.snowflake.com/en/user-guide/spark-connector-use.html#moving-data-from-snowflake-to-spark>

Moving Data from Snowflake to Spark

Note

When using DataFrames, the Snowflake connector supports SELECT queries only.

To read data from Snowflake into a Spark DataFrame:

Use the read() method of the SqlContext object to construct a DataFrameReader.

Specify SNOWFLAKE_SOURCE_NAME using the format() method. For the definition, see Specifying the Data Source Class Name (in this topic).

Specify the connector options using either the option() or options() method. For more information, see Setting Configuration Options for the Connector (in this topic).

Specify one of the following options for the table data to be read:

dbtable: The name of the table to be read. All columns and records are retrieved (i.e. it is equivalent to SELECT * FROM db_table).

query: The exact query (SELECT statement) to run.

Usage Notes

Currently, the connector does not support other types of queries (e.g. SHOW or DESC, or DML statements) when using DataFrames.

There is an upper limit to the size of an individual row. For more details, see Limits on Query Text Size.

QUESTION: 5

What can you infer from the below query profile?

Profile Overview (Finished)

Total Execution Time (1.299s)



Total Statistics

IO

Scan progress	100.00 %
Bytes scanned	12.08 GB
Percentage scanned from cache	99.47 %

Network

Bytes sent over the network	29.67 MB
-----------------------------	----------

Pruning

Partitions scanned	3,853
Partitions total	3,853

Option A : Less number of bytes of data are read from the cloud storage

Option B : More bytes are read from the cloud storage

Option C : The query did a good partition pruning

Correct Answer: A

Explanation/Reference:

Look at the number against "Percentage read from cache". It is more than 99% which means that most of the data is read from the SSD cache and few bytes were read from the cloud storage

QUESTION: 6

Which of the below statements are true about search optimization?

Option A : The search optimization service does not directly improve the performance of joins.

Option B : It can improve the performance of filtering rows from either table prior to the join, if the table has search optimization enabled and the predicate is selective

Option C : Search optimization supports external tables

Option D : Search optimization can improve range searches

Correct Answer: A,B

Explanation/Reference:

Explanation

<https://docs.snowflake.com/en/user-guide/search-optimization-service.html#how-search-optimization-affects-joins>

How Search Optimization Affects Joins

The search optimization service does not directly improve the performance of joins. However, it can improve the performance of filtering rows from either table prior to the join, if the table has search optimization enabled and the predicate is selective. Both tables do not need to have search optimization enabled. The decision to use search optimization is made for each table independently.

How Search Optimization Affects Views

As is the case with joins, the search optimization service can indirectly improve the performance of views (including secure views). If a base table for a view has search optimization enabled and if the query uses a selective predicate for that table, the search optimization service can improve performance when filtering rows.

All tables in the view do not need to have search optimization enabled. Search optimization is performed on each table independently.

Current Limitations of the Search Optimization Service

The search optimization service does not support the following:

External tables.

Materialized views.

Columns defined with a COLLATE clause.

Column concatenation.

Analytical expressions.

Tables and views protected by row access policies cannot be used with the Search Optimization Service.

Casts on table columns (except for fixed-point numbers cast to strings).

QUESTION: 7

Does sensitive data in Snowflake is modified in an existing table while applying Masking policies?

Option A : YES
Option B : NO

Correct Answer: B

Explanation/Reference:

Snowflake supports masking policies as a schema-level object to protect sensitive data from unauthorized access while allowing authorized users to access sensitive data at query runtime. This means that sensitive data in Snowflake is not modified in an existing table (i.e. no static masking). Rather, when users execute a query in which a masking policy applies, the masking policy conditions determine whether unauthorized users see masked, partially masked, obfuscated, or tokenized data

QUESTION: 8

Which of the below standard API objects are supported by the Snowflake connector for Python

- Option A : Connection
- Option B : Cursor
- Option C : SnowflakeConnection
- Option D : SnowflakeCursor

Correct Answer: A,B

Explanation/Reference:

<https://docs.snowflake.com/en/user-guide/python-connector.html#snowflake-connector-for-python>

The connector supports developing applications using the Python Database API v2 specification (PEP-249), including using the following standard API objects:

Connection objects for connecting to Snowflake.

Cursor objects for executing DDL/DML statements and queries.

SnowSQL, the command line client provided by Snowflake, is an example of an application developed using the connector.

QUESTION: 9

Which of the below statement is true when search access path hasn't been updated yet by search optimization?

- Option A : Queries might run slower but will always return up-to-date results.
Option B : There will be no impact on performance but queries might return stale data
Option C : Queries will fail when the access path is getting updated

Correct Answer: A

Explanation/Reference:

Explanation

How Does the Search Optimization Service Work?

To improve performance for point lookups, the search optimization service relies on a persistent data structure that serves as an optimized search access path.

A maintenance service that runs in the background is responsible for creating and maintaining the search access path:

When you add search optimization to a table, the maintenance service creates and populates the search access path with the data needed to perform the lookups.

The process of populating data can take time, depending on the size of the table. The service does this work in the background and does not block any concurrent operations on the table.

When data in the table is updated (for example, by loading new data sets or through DML operations), the maintenance service automatically updates the search access path to reflect the changes to the data.

If queries are run when the search access path hasn't been updated yet, the queries might run slower but will always return up-to-date results.

QUESTION: 10

Mark the Incorrect Statements with respect to types of streams supported by Snowflake?

- Option A : Standard streams cannot retrieve update data for geospatial data.
Option B : An append-only stream returns the appended rows only and therefore can be much more performant than a standard stream for extract, load, transform (ELT).
Option C : Insert-only Stream supported on external tables only.
Option D : An insert-only stream tracks row inserts & Delete ops only

Correct Answer: D

Explanation/Reference:

Standard Stream:

Supported for streams on tables, directory tables, or views. A standard (i.e. delta) stream tracks all DML changes to the source object, including inserts, updates, and deletes (including table truncates). This stream

type performs a join on inserted and deleted rows in the change set to provide the row level delta. As a net effect, for example, a row that is inserted and then deleted between two transactional points of time in a table is removed in the delta (i.e. is not returned when the stream is queried).

Append-only Stream:

Supported for streams on standard tables, directory tables, or views. An append-only stream tracks row inserts only. Update and delete operations (including table truncates) are not recorded. For example, if 10 rows are inserted into a table and then 5 of those rows are deleted before the offset for an append-only stream is advanced, the stream records 10 rows.

An append-only stream returns the appended rows only and therefore can be much more performant than a standard stream for extract, load, transform (ELT) and similar scenarios that depend exclusively on row inserts. For example, a source table can be truncated immediately after the rows in an append-only stream are consumed, and the record deletions do not contribute to the overhead the next time the stream is queried or consumed.

Insert-only Stream:

Supported for streams on external tables only. An insert-only stream tracks row inserts only; they do not record delete operations that remove rows from an inserted set (i.e. no-ops). For example, in-between any two offsets, if File1 is removed from the cloud storage location referenced by the external table, and File2 is added, the stream returns records for the rows in File2 only. Unlike when tracking CDC data for standard tables, Snowflake cannot access the historical records for files in cloud storage.

QUESTION: 11

Snowflake does not provide which of following set of SQL functions to support retrieving information about tasks?

- Option A : SYSTEM\$CURRENT_USER_TASK_NAME
- Option B : TASK_HISTORY
- Option C : TASK_DEPENDENTS
- Option D : TASK_QUERY_HISTORY
- Option E : SYSTEM\$TASK_DEPENDENTS_ENABLE

Correct Answer: C

Explanation/Reference:

Explanation

SYSTEM\$CURRENT_USER_TASK_NAME

Returns the name of the task currently executing when invoked from the statement or stored procedure defined by the task.

SYSTEM\$TASK_DEPENDENTS_ENABLE

Recursively resumes all dependent tasks tied to a specified root task.

TASK_DEPENDENTS

This table function returns the list of child tasks for a given root task in a DAG of tasks.

TASK_HISTORY

This table function can be used to query the history of task usage within a specified date range.

QUESTION: 12

Let us say you have List of 50 Source files, which needs to be loaded into Snowflake internal stage. All these Source system files are already Brotli-compressed files. Which statement is correct with respect to Compression of Staged Files?

Option A : Even though Source files are already compressed, Snowflake do apply default gzip2 Compression to optimize the storage cost.

Option B : Snowflake automatically detect Brotli Compression, will skip further compression of all 50 files.

Option C : Auto-detection is not yet supported for Brotli-compressed files; when staging or loading Brotli-compressed files, you must explicitly specify the compression method that was used

Option D : When staging 50 compressed files in a Snowflake stage, the files are automatically compressed using gzip.

Correct Answer: C

Explanation/Reference:

Auto-detection is not yet supported for Brotli-compressed files; when staging or loading Brotli-compressed files, you must explicitly specify the compression method that was used. To Know more about Compression of Staged Files, please refer the link: <https://docs.snowflake.com/en/user-guide/intro-summary-loading.html#compression-of-staged-files>

QUESTION: 13

If an external function is secure, the URL, the HTTP headers, and the context headers are hidden from all users who are not owners of the function.

Option A : TRUE

Option B : FALSE

Correct Answer: A

Explanation/Reference:

Explanation

<https://docs.snowflake.com/en/sql-reference/sql/create-external-function.html#optional-parameters>

Optional Parameters

SECURE

Specifies that the function is secure. If a function is secure, the URL, the HTTP headers, and the context headers are hidden from all users who are not owners of the function.

QUESTION: 14

In order to improve query performance, you can bypass data conversions from the Snowflake internal data type to the native Python data type. Which class in the `snowflake.connector.converter_null` module do you use for this feature?

Option A : `SnowflakeNoConverterToPython`

Option B : `PythonNoConverterToSnowflake`

Option C : `ByPassDataConversion`

Correct Answer: A

Explanation/Reference:

<https://docs.snowflake.com/en/user-guide/python-connector-example.html#improving-query-performance-by-bypassing-data-conversion>

To improve query performance, use the `SnowflakeNoConverterToPython` class in the `snowflake.connector.converter_null` module to bypass data conversions from the Snowflake internal data type to the native Python data type, e.g.:

```
from snowflake.connector.converter_null import SnowflakeNoConverterToPython
```

```
con = snowflake.connector.connect( ...
```

```
    converter_class=SnowflakeNoConverterToPython
```

```
for rec in con.cursor().execute("SELECT * FROM large_table"):
```

```
    # rec includes raw Snowflake data
```

As a result, all data is represented in string form such that the application is responsible for converting it to the native Python data types. For example, `TIMESTAMP_NTZ` and `TIMESTAMP_LTZ` data are the epoch time represented in string form, and `TIMESTAMP_TZ` data is the epoch time followed by a space followed by the offset to UTC in minutes represented in string form. No impact is made to binding data; Python native data can still be bound for updates.

QUESTION: 15

Each micro partition contains between 50 mb and 500 MB of uncompressed data

Option A : TRUE

Option B : FALSE

Correct Answer: B

Explanation/Reference:

What are Micro-partitions?

All data in Snowflake tables is automatically divided into micro-partitions, which are contiguous units of storage. Each micro-partition contains between 50 MB and 500 MB of uncompressed data (note that the actual size in Snowflake is smaller because data is always stored compressed). Groups of rows in tables are mapped into individual micro-partitions, organized in a columnar fashion. This size and structure allows for extremely granular pruning of very large tables, which can be comprised of millions, or even hundreds of millions, of micro-partitions.

Snowflake stores metadata about all rows stored in a micro-partition, including:

1. The range of values for each of the columns in the micro-partition.
2. The number of distinct values.
3. Additional properties used for both optimization and efficient query processing.

QUESTION: 16

Changing the retention period for your account or individual objects changes the value for all lower-level objects that do not have a retention period explicitly set?

Option A : TRUE

Option B : FALSE

Correct Answer: A

QUESTION: 17

Mark the Correct Statements: Statement 1. Enable failover for a primary database to one or more accounts in your organization using an ALTER DATABASE ... ENABLE FAILOVER TO ACCOUNTS statement. Statement 2. Enabling failover for a primary database can be done by Data Engineer either before or after a replica of the primary database has been created in a specified account.

Option A : Statement 1

Option B : Statement 2
Option C : Both are False.
Option D : Both are Correct.

Correct Answer: D

QUESTION: 18

Search optimization is a table-level property and applies to all columns with supported data types

Option A : TRUE
Option B : FALSE

Correct Answer: A

Explanation/Reference:

Explanation

Understanding the Search Optimization Service

The search optimization service aims to significantly improve the performance of selective point lookup queries on tables. A point lookup query returns only one or a small number of distinct rows. Use case examples include:

Business users who need fast response times for critical dashboards with highly selective filters.

Data scientists who are exploring large data volumes and looking for specific subsets of data.

A user can register one or more tables to the search optimization service. Search optimization is a table-level property and applies to all columns with supported data types

QUESTION: 19

You can execute zero, one, or more transactions inside a stored procedure?

Option A : TRUE
Option B : FALSE

Correct Answer: A

QUESTION: 20

What will happen if you try to create a materialized view as below
create or replace materialized view employee_view as
select id, role, name
from employee where role = current_role();

Option A : Only users with the desired role will be able to retrieve results from the view

Option B : Only users with the correct role assigned and SYSADMIN will be able to retrieve the results from the view

Option C : The operation will error out since the function used is a non-deterministic one.

Correct Answer: C

Explanation/Reference:

It will error out with the below message

SQL compilation error: error line 3 at position 33 Invalid materialized view definition. non deterministic function 'CURRENT_ROLE' not allowed in view definition.

<https://docs.snowflake.com/en/user-guide/views-materialized.html#limitations-on-creating-materialized-views>

Functions used in a materialized view must be deterministic. For example, using CURRENT_TIME or CURRENT_TIMESTAMP is not permitted.

QUESTION: 21

If using a JavaScript UDF in a masking policy, Data Engineer needs to ensure the data type of the column, UDF, and masking policy match irrespective of case-sensitivity?

Option A : TRUE

Option B : FALSE

Correct Answer: B

Explanation/Reference:

Please note JavaScript is case sensitive but if we are using a JavaScript UDF in a masking policy, ensure the data type of the column, UDF, and masking policy match.

QUESTION: 22

Stuart, a Lead Data Engineer in MACRO Data Company created streams on set of External tables. He has been asked to extend the data retention period of the stream for 90 days, which parameter he can utilize to enable this extension?

Option A : MAX_DATA_EXTENSION_TIME_IN_DAYS

Option B : DATA_RETENTION_TIME_IN_DAYS

Option C : DATA_EXTENSION_TIME_IN_DAYS
Option D : None of the above

Correct Answer: D

Explanation/Reference:

Explanation

External tables do not have data retention period applicable.

Good to Understand other Options available.

DATA_RETENTION_TIME_IN_DAYS

Type: Object (for databases, schemas, and tables) — Can be set for Account » Database » Schema » Table

Description: Number of days for which Snowflake retains historical data for performing Time Travel actions (SELECT, CLONE, UNDROP) on the object. A value of 0 effectively disables Time Travel for the specified database, schema, or table.

Values:

0 or 1 (for Standard Edition)

0 to 90 (for Enterprise Edition or higher)

Default:

1

MAX_DATA_EXTENSION_TIME_IN_DAYS

Type: Object (for databases, schemas, and tables) — Can be set for Account » Database » Schema » Table

Description: Maximum number of days for which Snowflake can extend the data retention period for tables to prevent streams on the tables from becoming stale. By default, if the

DATA_RETENTION_TIME_IN_DAYS setting for a source table is less than 14 days, and a stream has not been consumed, Snowflake temporarily extends this period to the stream's offset, up to a maximum of 14 days, regardless of the Snowflake Edition for your account. The

MAX_DATA_EXTENSION_TIME_IN_DAYS parameter enables you to limit this automatic extension period to control storage costs for data retention or for compliance reasons.

This parameter can be set at the account, database, schema, and table levels. Note that setting the parameter at the account or schema level only affects tables for which the parameter has not already been explicitly set at a lower level (e.g. at the table level by the table owner). A value of 0 effectively disables the automatic extension for the specified database, schema, or table.

Values:

0 to 90 (i.e. 90 days) — a value of 0 disables the automatic extension of the data retention period. To increase the maximum value for tables in your account, Client needs to contact Snowflake Support.

Default:14

THANK YOU FOR DOWNLOADING DEA-C02 UPDATED EXAM QUESTIONS

Note: Thanks For Trying The Demo Of Our DEA-C02 Exam Product

Visit Our Site to Purchase the Full Set of Actual DEA-C02 Exam Questions With Answers.

Money Back Guarantee



Click The Link Below

<https://certs4sale.com/snowflake/dea-c02-pdf-exam-dumps>