

Snowflake

SnowPro-Advanced-Data-Engineer

SnowPro Advanced Data Engineer QUESTION & ANSWERS

Questions & Answers (Demo Version – Limited Content)

Thank you for Downloading SnowPro-Advanced-Data-Engineer
exam PDF Demo

Download Free Demo:

<https://www.dumpsplanet.com/snowpro-advanced-data-engineer-dumps/>

<https://www.dumpsplanet.com/>

Version: 04

QUESTION 1

Which are the valid options for the validation mode parameter in the COPY COMMAND?

- A. RETURN_ROWS
- B. RETURN_ERROR
- C. RETURN_ERRORS
- D. RETURN_ALL_ERRORS

Correct Answer: A, C, D

Explanation/Reference:

VALIDATION_MODE = RETURN_n_ROWS | RETURN_ERRORS | RETURN_ALL_ERRORS

String (constant) that instructs the COPY command to validate the data files instead of loading them into the specified table; i.e. the COPY command tests the files for errors but does not load them. The command validates the data to be loaded and returns results based on the validation option specified:
Supported Values

Notes:

RETURN_n_ROWS (e.g. RETURN_10_ROWS)

Validates the specified number of rows, if no errors are encountered; otherwise, fails at the first error encountered in the rows.

RETURN_ERRORS

Returns all errors (parsing, conversion, etc.) across all files specified in the COPY statement.

RETURN_ALL_ERRORS

Returns all errors across all files specified in the COPY statement, including files with errors that were partially loaded during an earlier load because the ON_ERROR copy option was set to CONTINUE during the load.

<https://docs.snowflake.com/en/sql-reference/sql/copy-into-table.html#optional-parameters>

QUESTION 2

Which of the below functions are recommended to be used to understand the clustering ratio of a table?

- A. SYSTEM\$CLUSTERING_RATIO
- B. SYSTEM\$CLUSTERING_DEPTH
- C. SYSTEM\$CLUSTERING_INFORMATION

Correct Answer: B,C

Explanation/Reference:

https://docs.snowflake.com/en/sql-reference/functions/system_clustering_ratio.html

QUESTION 3

You have many files which are loaded onto the cloud storage. Most of them are less than 200 MB in size, but a few are 1GB or more. You need to process them using SNOWPIPE. Which of the below options is recommended

- A. Split the 1 GB files into smaller file sizes
- B. Use a dedicated large warehouse for the 1 GB files.
- C. Merger the smaller 200 MB files into a single file.

Correct Answer: A

Explanation/Reference:

Explanation

<https://docs.snowflake.com/en/user-guide/data-load-considerations-prepare.html#general-file-sizing-recommendations>

The number of load operations that run in parallel cannot exceed the number of data files to be loaded. To optimize the number of parallel operations for a load, we recommend aiming to produce data files roughly 100-250 MB (or larger) in size compressed.

Note

Loading very large files (e.g. 100 GB or larger) is not recommended.

If you must load a large file, carefully consider the ON_ERROR copy option value. Aborting or skipping a file due to a small number of errors could result in delays and wasted credits. In addition, if a data loading operation continues beyond the maximum allowed duration of 24 hours, it could be aborted without any portion of the file being committed.

Aggregate smaller files to minimize the processing overhead for each file. Split larger files into a greater number of smaller files to distribute the load among the compute resources in an active warehouse. The number of data files that are processed in parallel is determined by the amount of compute resources in a warehouse. We recommend splitting large files by line to avoid records that span chunks.

If your source database does not allow you to export data files in smaller chunks, you can use a third-party utility to split large CSV files.

QUESTION 4

The employee project details has the project names as array against each employee as shown below

```
66 select * from employee_project_details;
67
```

Results Data Preview ← Open History

✓ Query ID SQL 28ms 1 rows

Filter result... Columns ▾

Row	EMP_ID	EMP_NAME	PROJECT_NAMES
1	1	john	["it", "prod"]

Which of the query below will convert the array into individual rows?

- A.

```
select emp_id,
       emp_name,
       p.value::string as project_names
from   employee_project_details,table(flatten(employee_project_details.project_names)) p
;
```
- B.

```
select emp_id,
       emp_name,
       p.value::string as project_names
from   employee_project_details,lateral(flatten(employee_project_details.project_names)) p
;
```
- C.

```
select emp_id,
       emp_name,
       p.value::string as project_names
from   employee_project_details,lateral flatten(employee_project_details.project_names)) p
```

Correct Answer: A

Explanation/Reference:

Explanation

Try this out in your snowflake instance

Step 1 - Create the table

```
create or replace table employee_project_details(emp_id varchar, emp_name varchar, project_names
array);
```

Step 2 - Insert values

```
insert into employee_project_details
select '1','john',array_cat(to_array('it'),to_array('prod'));
```

Step 3 - Convert to rows

```
select emp_id,
       emp_name,
       p.value::string as project_names
from   employee_project_details,table(flatten(employee_project_details.project_names)) p
```

QUESTION 5

For using Snowflake with Spark, which of the below privileges are required?

- A. USAGE on the schema that contains the table that you will read from or write to
- B. CREATE STAGE on the schema that contains the table that you will read from or write to
- C. Accountadmin
- D. Sysadmin

Correct Answer: A,B

Explanation/Reference:

Explanation

<https://docs.snowflake.com/en/user-guide/spark-connector-install.html#requirements>

Requirements

To install and use Snowflake with Spark, you need the following:

A supported operating system. For a list of supported operating systems, see Operating System Support.

Snowflake Connector for Spark.

Snowflake JDBC Driver (the version compatible with the version of the connector).

Apache Spark environment, either self-hosted or hosted in any of the following:

Qubole Data Service.

Databricks.

Amazon EMR.

In addition, you can use a dedicated Amazon S3 bucket or Azure Blob storage container as a staging zone between the two systems; however, this is not required with version 2.2.0 (and higher) of the connector, which uses a temporary Snowflake internal stage (by default) for all data exchange.

The role used in the connection needs USAGE and CREATE STAGE privileges on the schema that contains the table that you will read from or write to.

QUESTION 6

Which of the below privileges are required to add or remove search optimization?

- A. OWNERSHIP privilege on the table
- B. ADD SEARCH OPTIMIZATION privilege on the schema that contains the table
- C. ACCOUNTADMIN privilege
- D. ALL OF THE ABOVE

Correct Answer: A,B

Explanation/Reference:

Explanation

What Access Control Privileges Are Needed For the Search Optimization Service?

To add or remove search optimization for a table, you must have the following privileges:

You must have OWNERSHIP privilege on the table.

You must have ADD SEARCH OPTIMIZATION privilege on the schema that contains the table.
GRANT ADD SEARCH OPTIMIZATION ON SCHEMA TO ROLE ;
To use the search optimization service for a query, you just need SELECT privileges on the table.
You do not need any additional privileges. Because search optimization is a table property, it is automatically detected and used (if appropriate) when querying a table.

QUESTION 7

While using kafka connector, what charges are applied to your account?

- A. Snowpipe processing time
- B. Data storage
- C. Kafka connector usage

Correct Answer: A,B

Explanation/Reference:

Explanation

Billing Information

There is no direct charge for using the Kafka connector. However, there are indirect costs: Snowpipe is used to load the data that the connector reads from Kafka, and Snowpipe processing time is charged to your account.

Data storage is charged to your account

QUESTION 8

External functions must be scalar functions

- A. TRUE
- B. FALSE

Correct Answer: A

Explanation/Reference:

Explanation

<https://docs.snowflake.com/en/sql-reference/external-functions-introduction.html#execution-time-limitations-and-issues>

Execution-time Limitations and Issues

Because the remote service is opaque to Snowflake, the optimizer might not be able to perform some optimizations that it could perform for equivalent internal functions.

External functions have more overhead than internal functions (both built-in functions and internal UDFs) and usually execute more slowly.

Currently, external functions must be scalar functions. A scalar external function returns a single value for each input row.

Currently, external functions cannot be shared with data consumers via Secure Data Sharing.

The maximum response size per batch is 10MB.

External functions cannot be used in the following situations:

As part of a database object (e.g. table, view, UDF, or masking policy) shared via Secure Data Sharing. For example, you cannot create a shared view that uses an external function. The following is not supported:

```
create view my_shared_view as select my_external_function(x) ...;
```

```
create share things_to_share;
```

```
grant select on view my_shared_view to share things_to_share;
```

A DEFAULT clause of a CREATE TABLE statement. In other words, the default value for a column cannot be an expression that calls an external function. If you try to include an external function in a DEFAULT clause, then the CREATE TABLE statement fails.

A COPY transformation.

External functions can raise additional security issues. For example, if you call a third party's function, that party could keep copies of the data passed to the function.

QUESTION 9

Which option needs to be followed to allow a user to have only OWNERSHIP privilege on the table, but should not be able to manage privilege grants on the object

- A. Create a managed schema
- B. Create a regular schema
- C. By default object owner will not be able to manage privileges on the object

Correct Answer: A

Explanation/Reference:

Explanation

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

WITH MANAGED ACCESS

Specifies a managed schema. Managed access schemas centralize privilege management with the schema owner.

In regular schemas, the owner of an object (i.e. the role that has the OWNERSHIP privilege on the object) can grant further privileges on their objects to other roles. In managed schemas, the schema owner manages all privilege grants, including future grants, on objects in the schema. Object owners retain the OWNERSHIP privileges on the objects; however, only the schema owner can manage privilege grants on the objects.

QUESTION 10

The Snowflake Kafka Connector does not guarantee that rows are inserted in the order that they were originally published.

- A. TRUE
- B. FALSE

Correct Answer: A

Explanation/Reference:

Explanation

There is no guarantee that rows are inserted in the order that they were originally published.
<https://docs.snowflake.com/en/user-guide/kafka-connector-overview.html#fault-tolerance>

QUESTION 11

Which system table will you use to get the total credit consumption over a specific time period?

- A. WAREHOUSE_METERING_HISTORY
- B. WAREHOUSE_CREDIT_USAGE_HISTORY
- C. WAREHOUSE_USAGE_HISTORY

Correct Answer: A

Explanation/Reference:

The WAREHOUSE_METERING_HISTORY table in the ACCOUNT_USAGE Schema can be used to get the desired information. Run the below query to try this out.

```
SELECT WAREHOUSE_NAME,  
       SUM(CREDITS_USED_COMPUTE) AS CREDITS_USED_COMPUTE_SUM  
FROM ACCOUNT_USAGE.WAREHOUSE_METERING_HISTORY  
GROUP BY 1  
ORDER BY 2 DESC;
```

Thank You for trying SnowPro-Advanced-Data-Engineer PDF Demo

To try our SnowPro-Advanced-Data-Engineer practice exam software visit the link below

<https://www.dumpsplanet.com/snowpro-advanced-data-engineer-dumps/>

Start Your SnowPro-Advanced-Data-Engineer Preparation

[Limited Time Offer] Use Coupon “**dp25%**” for extra 25% discount on the purchase of Pdf. Test your SnowPro-Advanced-Data-Engineer preparation with actual exam questions.

<https://www.dumpsplanet.com/>