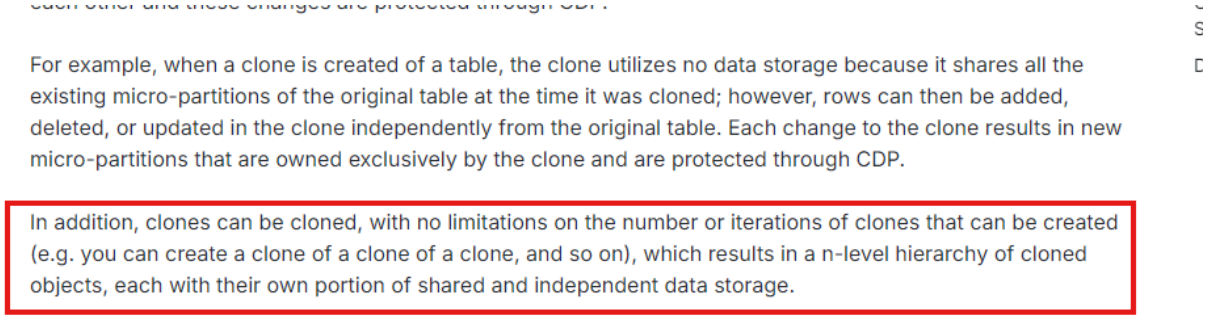
Certification Questions Practice Set (Storage & Data Protection)

1. Clones can be cloned, with no limitations on the number or iterations of clones that can be created (e.g., you can create a clone of a clone of a clone, and so on), which results in an n-level hierarchy of cloned objects, each with their own portion of shared and independent data storage?

A. TRUE
B. FALSE

Answer : TRUE
Explanation : https://docs.snowflake.com/en/user-guide/tables-storage-considerations



For example, when a clone is created of a table, the clone utilizes no data storage because it shares all the existing micro-partitions of the original table at the time it was cloned; however, rows can then be added, deleted, or updated in the clone independently from the original table. Each change to the clone results in new micro-partitions that are owned exclusively by the clone and are protected through CDP.

In addition, clones can be cloned, with no limitations on the number or iterations of clones that can be created (e.g. you can create a clone of a clone of a clone, and so on), which results in a n-level hierarchy of cloned objects, each with their own portion of shared and independent data storage.

---

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

A. Data stored in temporary tables is not recoverable after the table is dropped as they do not have fail-safe.
B. Historical data in transient tables can be recovered by Snowflake due to Operation failure after the Time Travel retention period ends using Fail-safe.
C. Long-lived tables, such as fact tables, should always be defined as permanent to ensure they are fully protected by Fail-safe.
D. Short-lived tables (i.e., <1 day), such as ETL work tables, can be defined as transient to eliminate Fail-safe costs.

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.

Answer : B

Explanation : Temporary and Transient tables do not have any fail safe period. The time travel max retention period is 1 day that's it.

---

3. Lisa is working on a project where she needs to maintain historical versions of data for auditing purposes. Which Snowflake feature should she use to achieve this?

A. Data Cloning

B. Time Travel

C. Secure Data Sharing

D. Materialized Views

Answer : B

---

4. David, a Data Engineer, has been asked to create a backup strategy for critical tables in Snowflake. Which of the following actions should he include in his strategy?

A. Use the COPY INTO command to export data to external storage regularly.

B. Implement fail-safe mechanisms for all critical tables.

C. Schedule regular snapshots of the tables using the CLONE command.

D. Enable multi-cluster warehouses for high availability.

Answer : A

Explanation - B is out because fail safe is available by default and no setting change is required to enable this.

C is out as CLONE just creates a virtual copy and not physical copies

D is out because multi-cluster warehouses are there for performance improvement and not back up strategy

---

5. 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 PROD;

D. CREATE DATABASE DEV -

CLONE PROD -

DATA_RETENTION_TIME_IN DAYS = 0

Answer : C

Explanation : Option C will work because the requirement is to disable fail safe. Transient tables have fail safe disabled by default

---

6. A Data Engineer creates a stream object on a transaction source table to capture new records in the source table and insert the records into a dimension table. The source table gets updated daily. An `INSERT` statement is configured to run a task on a schedule once a week, on Sunday at 8:05 PM Eastern Standard Time, using CRON syntax: 5 20 * * 0

   After 10 days of creating the stream and task objects, the Engineer realized that the task never ran because the task was not resumed after creation.

   If the data retention period for the source table is set to 1 day, how many days of records since the creation of the source table will be read from the stream object once the task is resumed? (NOTE: `MAX_DATA_EXTENSION_TIME_IN_DAYS` parameter is set to 14 days.)

1. The last 1 day of records
2. The last 7 days of records
3. The last 10 days of records
4. The last 14 days of records

   Answer : 4

   Explanation :

A fact table has 700 GB of data, but the overall storage for the table is larger than 100 TB. Time Travel is set to 30 days. The table has the column `BUSINESS_DATE` as a cluster key.

What are the reasons for this excess storage? (Select TWO).

1. The table has multiple streams created against it and each stream is accumulating a large amount of storage.
2. The table is subjected to frequent large updates and has many micro-partition versions.

3. The table is loaded daily from files placed on an external stage. Each file contains data from a single business date.
4. Multiple materialized views have been built on top of this table and each view requires storage.
5. The table is subjected to daily `CREATE OR REPLACE` table functions, which means Time Travel is maintained for every table version.
