1. What is Zero Copy Cloning?

It makes a copy of a database without duplicating the data it contains. Clone operation takes a snapshot of the source data when the clone is created and makes this data available to the cloned object (Target).

CREATE TABLE <table\_name>

CLONE <source\_table\_name>

1. What is time travel?

It allows to access historical data (updating or deleting the data) at any point in time.

It is used for some tasks like Restoring the data, backing up and duplicating data, Analyzing data manipulation, and consumption.

By default, time travel is enabled with one day retention period. We can configure it also.

1. Transient table vs permanent table vs temporary table vs external table

Transient Table: They are like the permanent table except they do not have fail–safe period and only have a very limited time travel period. Suitable in scenarios where the data is not critical and can be recovered from external means if required.

create transient table employee (id number, name varchar (50));

permanent table: Default table type in snowflake. It comes with time travel and fail-safe.

Temporary table: Exist only within the session in which they were created and available only for the remainder of the session. Will not be visible to other users or sessions. The table is dropped completely and is not recoverable once the session is ended.

External table:

These are not the typical type of tables that can be created directly in Snowflake.

Snowflake External tables allow you to access the data from files stored in an external stage such as Amazon S3, Azure blob storage, or GCP bucket as a regular table. This allows you to query a file as if it is a table in Snowflake.

1. Fail Safe Implementation on all the tables and how it is done.

Fail-safe is not provided as a means for accessing historical data after the Time Travel retention period has ended. It is for use only by Snowflake to recover data that may have been lost or damaged due to extreme operational failures.

Fail-safe cannot be disabled on the tables.

1. What will happen in Zero copy cloning when you change the data on the source and target?

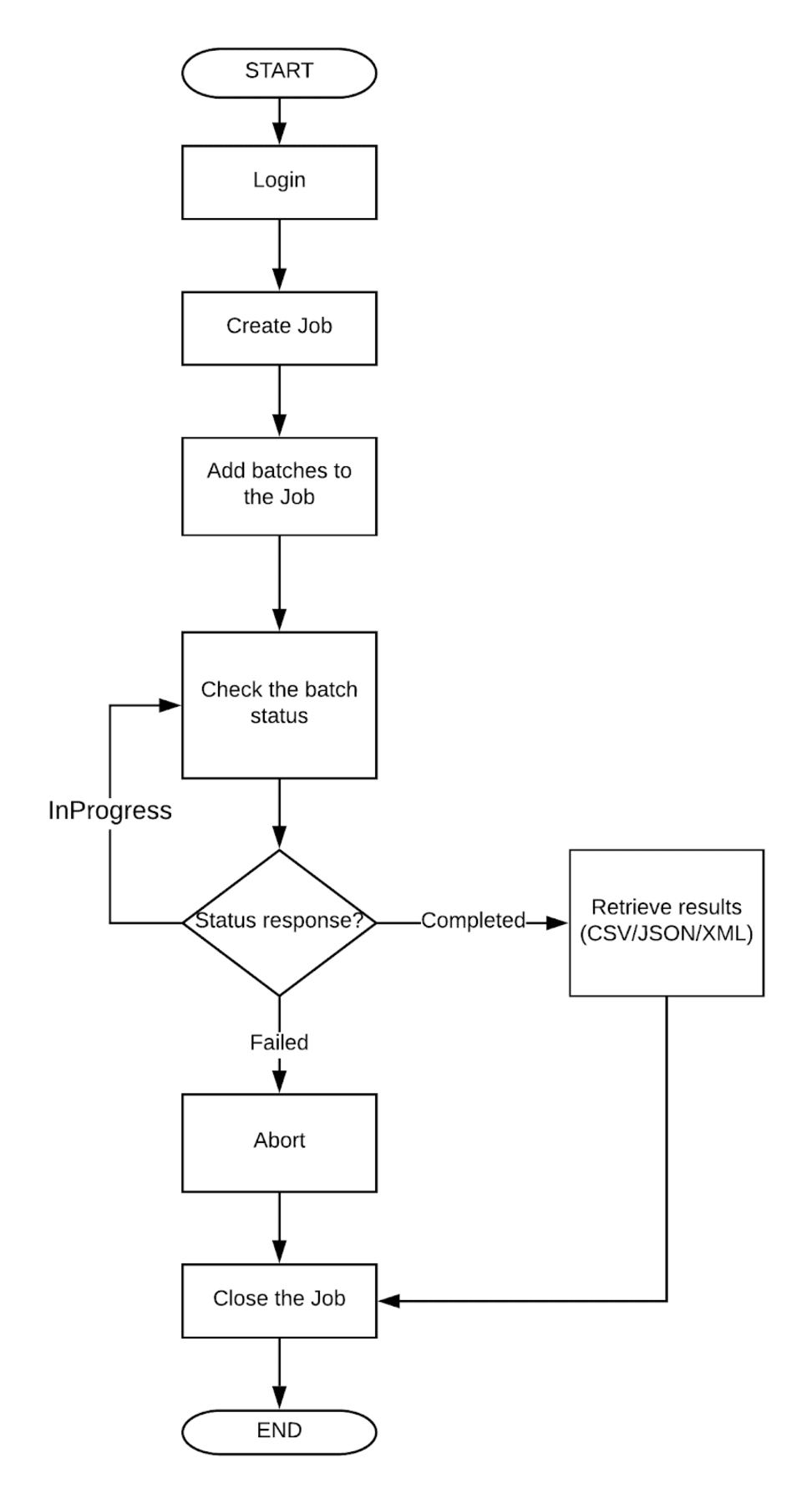
Changes made to the original object, or the copy don't affect the other.

1. Task in snowflake

The task is for executing the SQL Code. CREATE TASK … CLONE

1. How to load 10k records from Salesforce to Snowflake

We can load the data using Salesforce API. Hevo will take care of fetching the data from data sources like Salesforce, etc., and sending it to your destination warehouse for free.



1. How do you perform or increase the performance tuning in or project?

Performance can be improved by partitioning, Bucketing, Cache mechanisms.

Data Partitioning: Ensure that your data is properly partitioned before processing. Spark operates on distributed data partitions, and well-partitioned data can lead to more efficient processing, minimizing data shuffling during operations.

Caching and Persistence: Use caching and data persistence to store intermediate results in memory or disk. This avoids redundant computations and speeds up iterative algorithms.

Broadcasting: For smaller datasets used in join operations, consider broadcasting the data to all nodes instead of shuffling it across the network.

Coalesce and Repartition: Use coalesce() or repartition() to reduce or increase the number of partitions in RDDs/DataFrames, respectively. This can help optimize data distribution and minimize data movement.