

Expiring Data From a Table

Data expiration is a mechanism that allows you to set an expiration interval (in days) for the data in a table. You can use the TD Console to configure table data expiration or Hive and Presto query engines. When an expiration period is set for a specific table, Treasure Data automatically filters out any records with a timestamp older than that of the expiration period and marks the records for purging from internal storage.

Data Expiration in the TD Console

Use Case - Managing the Amount of Storage Consumed by a Table

For example, you can configure a 90-day data expiration policy to limit an events table to hold 90 days of history. Assuming relatively constant event traffic, this policy would in general limit the growth of the table without requiring ongoing management.

Non-Compliance with Data Retention Regulations. Data expiration in Treasure Data is not sufficient to comply with data privacy regulations and other industry-specific data retention regulations. In cases requiring immediate removal of data from TD, contact Treasure Data support for our recommended practices.

Configure Table Data Expiration in the TD Console

1. Navigate to **Workbench > Databases**.
2. Select a database.
3. Select the table.
4. Select ... > **Table settings**.
5. Select **Expire Data?**
6. Type the number of days you would like to keep the data active.
7. Select **Submit**.

Data Expiration in the TD Toolbelt

Use the TD Toolbelt

TD Toolbelt Example

```
$ td table:expire test_db test_table <expiration_days>
```

Where <expiration_days> is an integer number greater than 0.

Disabling Table Data Expiration

To disable the table data expiration, set the expiration days to 0. For example, using TD toolbelt:

```
$ td table:expire test_db test_table 0
```

Data Expiration in Hive and Presto Query Engines

The Hive version that you use might affect data expiration:

In Hive 0.13, data expirations take effect in the partitioning filter stage. However, in Hive 2 and Presto query engines, data expiration takes effect in the record reader. Therefore, Hive 0.13 might include more excluded data. For more information, see [Hive 0.13/Hive 2 Code and Syntax Compatibility](#).

Effects of Data Expiration

Review the following to understand how Treasure Data handles data expiration:

- **Defined expiration period and threshold.** For an expiration period of N days, the expiration threshold is N days prior to the current time when the query starts. Rows with a time column value earlier than the expiration threshold are filtered out. For example, a query that runs at 9:15 AM today referencing a table with an expiration period of 2 days automatically excludes rows from that table with a time earlier than 9:15 AM of two days ago when processing the query.

The time column value may not be related to the actual date when the record was created in Treasure Data.

- **Filtering.** Treasure Data transparently adds time filter clauses to queries as needed, so expired rows are not used in calculating query results. If multiple tables have different thresholds, each table is filtered based on its own expiration policy.
- **Importing records with old timestamps.** When records are ingested or created, if they have a time value older than the expiration threshold, they are filtered out as expired even though they are new.
- **Retaining and purging expired records.** Expired records remain in storage for a period of time before being purged. The rules are:
 - Whenever you change the expiration policy for a table, records are not purged from that table for the following 30 days. This allows time to experiment with different expiration policy values and see their effects and restore data by simply changing the policy again.
 - If the expiration policy has not been changed in the previous 30 days, it is assumed that the policy is working as intended. Purging expired records then happen more frequently. For more details on the policy for purging records, contact Treasure Data support.
 - You can restore access to expired records that have not been purged by setting the expiration policy to zero days. A SELECT statement from the table returns all rows.