SQL TIPS AND TRICKS

PART 38

SLOWLY CHANGING DIMENSIONS

Type 1, 2, 3

MAYURI DANDEKAR

Slowly Changing Dimensions (SCD) are techniques used in data warehousing to manage and track changes in dimension data over time.
Dimensions are structures that categorize facts and measures to enable users to answer business questions.
For example, a customer dimension might include attributes such as customer name, address, and phone number. When these attributes change, different strategies are used to capture and manage these changes.

There are three common types of Slowly Changing Dimensions, each handling changes differently:
Type 1: Overwrite
Type 2: Add a New Row
Type 3: Add a New Column

Type 1: Overwrite

In a Type 1 SCD, when an attribute value changes, the old value is simply overwritten with the new value.

Use Case:
This method is used when tracking historical changes is not necessary. It's straightforward and requires the least amount of storage.

Example:

Original Record:
CustomerID: 1, Name: John Doe, Address: 123 Elm St.

After Update:
CustomerID: 1, Name: John Doe, Address: 456 Oak St.

Type 2: Add a New Row

In a Type 2 SCD, when an attribute value changes, a new record is added to the table with a new surrogate key. The old record is retained to preserve historical data.

Use Case:
This method is used when it is important to track the history of changes. Each change is captured as a new row, allowing analysis of data changes over time.

Example:

Original Record:
CustomerID: 1, Name: John Doe, Address: 123 Elm St., StartDate: 2020-01-01, EndDate: NULL

After Update:
CustomerID: 1, Name: John Doe, Address: 123 Elm St., StartDate: 2020-01-01, EndDate: 2021-05-01
CustomerID: 2, Name: John Doe, Address: 456 Oak St., StartDate: 2021-05-01, EndDate: NULL

Type 3: Add a New Column

In a Type 3 SCD, when an attribute value changes, a new column is added to store the previous value, while the current column is updated with the new value.

Use Case:
This method is used when only the current and previous values of a dimension attribute need to be tracked. It's less complex than Type 2 but still provides some historical context.

Example:

Original Record:
CustomerID: 1, Name: John Doe, CurrentAddress: 123 Elm St., PreviousAddress: NULL

After Update:
CustomerID: 1, Name: John Doe, CurrentAddress: 456 Oak St., PreviousAddress: 123 Elm St.

THANK YOU

MAYURI DANDEKAR