

Is Your Data Telling the Whole Story?



What is SCD?

- Dimensions hold descriptive data like customer or product info
- This data is not static and can change over time
- A customer might move to a new address
- A product might be reclassified
- Slowly Changing Dimensions track these historical changes
- Without SCD, you lose the context of past events



Type 1: Overwrite

- The simplest method to handle data changes
- The old value is simply overwritten with the new one
- No history is preserved whatsoever
- Useful for correcting simple errors or typos
- Not suitable for any analysis requiring historical context
- Your report will only ever show the current state



Type 2: Add New Row

- The most common method for tracking history
- Creates a new row for the dimension record when a change occurs
- The old row remains, marked as the former active state
- Each row has a unique surrogate key and effective dates
- A current flag identifies the active record for filtering
- This allows accurate historical reporting at any point in time



Why Type 2 Matters

- Enables accurate time-based analysis
- See what a customer's status was when an order was placed
- Track sales performance by a product's old classification
- Analyze trends based on historical regional territories
- Reports reflect the truth at the time of the transaction
- Essential for compliance and auditable historical records



Implementing in Power BI

- Use a unique key for each row version (surrogate key)
- Add columns for start and end dates for each record
- Include a boolean column like IsCurrent for easy filtering
- Load all historical versions into your dimension table
- Use Power Query to help manage and transform the data
- Model relationships on the surrogate key, not the natural key



The Type 2 Challenge

- Can lead to very large dimension tables over time
- Requires careful management of effective dates
- DAX calculations must account for the changing attributes
- Filtering for the "current" view requires a simple filter
- Filtering for a historical view requires a date-based lookup
- Star schema design becomes even more critical



Other SCD Types

- **Type 3: Add New Column**
- **Adds a new column to store a previous value**
- **Only tracks a limited number of changes (usually the last one)**
- **A compromise between Type 1 and Type 2**
- **Less common due to its limited historical tracking**
- **Type 0: Attributes never change**



Choose Wisely

- **Select the SCD type based on your business needs**
- **Use Type 1 for unimportant attributes or corrections**
- **Type 2 is the gold standard for true historical analysis**
- **Consider the trade-off between complexity and insight**
- **Properly implemented SCDs make your data a true asset**
- **They ensure your reports accurately reflect the past**



Follow for More



Like this post? Repost!

Share your thoughts in the comments!

I share little hacks and tips I actually use in
Data & AI — follow if you wanna see more!



Shoaib Akthar
@theshoaibakthar

Follow for more Data Analytics, Data Science and AI insights