



# SMBCI Data Engineer Technical Test: Slowly Changing Dimension Type 2

---



## Context: Implementing SCD Type 2 for Customer Data

---

You're working with a data warehouse that tracks customer information for a subscription-based service. The business needs to maintain historical records of how customer attributes change over time for accurate reporting and analysis.

You have a staging table `customer_staging` that receives daily updates with the current customer data from the source system:

```
1. customer_staging (  
2.     customer_id INT,  
3.     first_name VARCHAR(50),  
4.     last_name VARCHAR(50),  
5.     email VARCHAR(100),  
6.     phone VARCHAR(20),  
7.     address VARCHAR(200),  
8.     city VARCHAR(50),  
9.     state VARCHAR(2),  
10.    zip VARCHAR(10),  
11.    subscription_tier VARCHAR(20),  
12.    monthly_fee DECIMAL(10,2)  
13.)
```

You need to implement a Slowly Changing Dimension Type 2 approach for the `dim_customer` table to track historical changes to any customer attributes:

```
1. dim_customer (  
2.     customer_sk INT IDENTITY(1,1) PRIMARY KEY,  
3.     customer_id INT,  
4.     first_name VARCHAR(50),  
5.     last_name VARCHAR(50),  
6.     email VARCHAR(100),  
7.     phone VARCHAR(20),  
8.     address VARCHAR(200),  
9.     city VARCHAR(50),  
10.    state VARCHAR(2),
```

```
11.     zip VARCHAR(10),
12.     subscription_tier VARCHAR(20),
13.     monthly_fee DECIMAL(10,2),
14.     effective_date DATE,
15.     expiration_date DATE,
16.     is_current BIT
17. )
```

---

## Deliverables:

1. Write the SQL logic to implement the SCD Type 2 process that will:
    - Insert new customer records
    - Update existing records by expiring the current record and inserting a new one when attributes change
    - Keep track of the validity period for each version of a customer record
    - Maintain a flag for the current version
  2. Explain how your implementation handles the following scenarios:
    - A new customer appears in the staging table
    - An existing customer has changed their address
    - An existing customer has upgraded their subscription tier
    - Multiple attributes change simultaneously for a customer
  3. How would you modify your approach if certain attributes (like name and email) should be tracked historically, but others (like phone number) should just be overwritten (SCD Type 1)?
- 

## Grading Rubric

Criteria	Weight
Correctness of the SCD Type 2 implementation	40%
Clarity of explanations and SQL code	20%
Performance considerations for large datasets	10%
Handling of edge cases	15%
Presentation (clarity)	5%

 **Good luck, have fun!**