



Kampus
Merdeka
INDONESIA JAYA

BUSINESS INTELLIGENCE

SLOWLY CHANGING DIMENSION (SCD)

FAST: FOCUS, ADAPTIVE, SINCERE, TRUSTWORTHY

JURUSAN TEKNOLOGI INFORMASI
POLITEKNIK NEGERI MALANG



OUTLINE

1. Keys
2. Slowly Changing Dimension



Bagian 1

KEYS



Natural Key

- Unique Key yang berasal dari dunia nyata — artinya, nilainya sudah ada secara alami dalam proses bisnis
- Contoh: NIM, No KTP, Email, Kode SKU (Stock Keeping Unit)

Surrogate Key

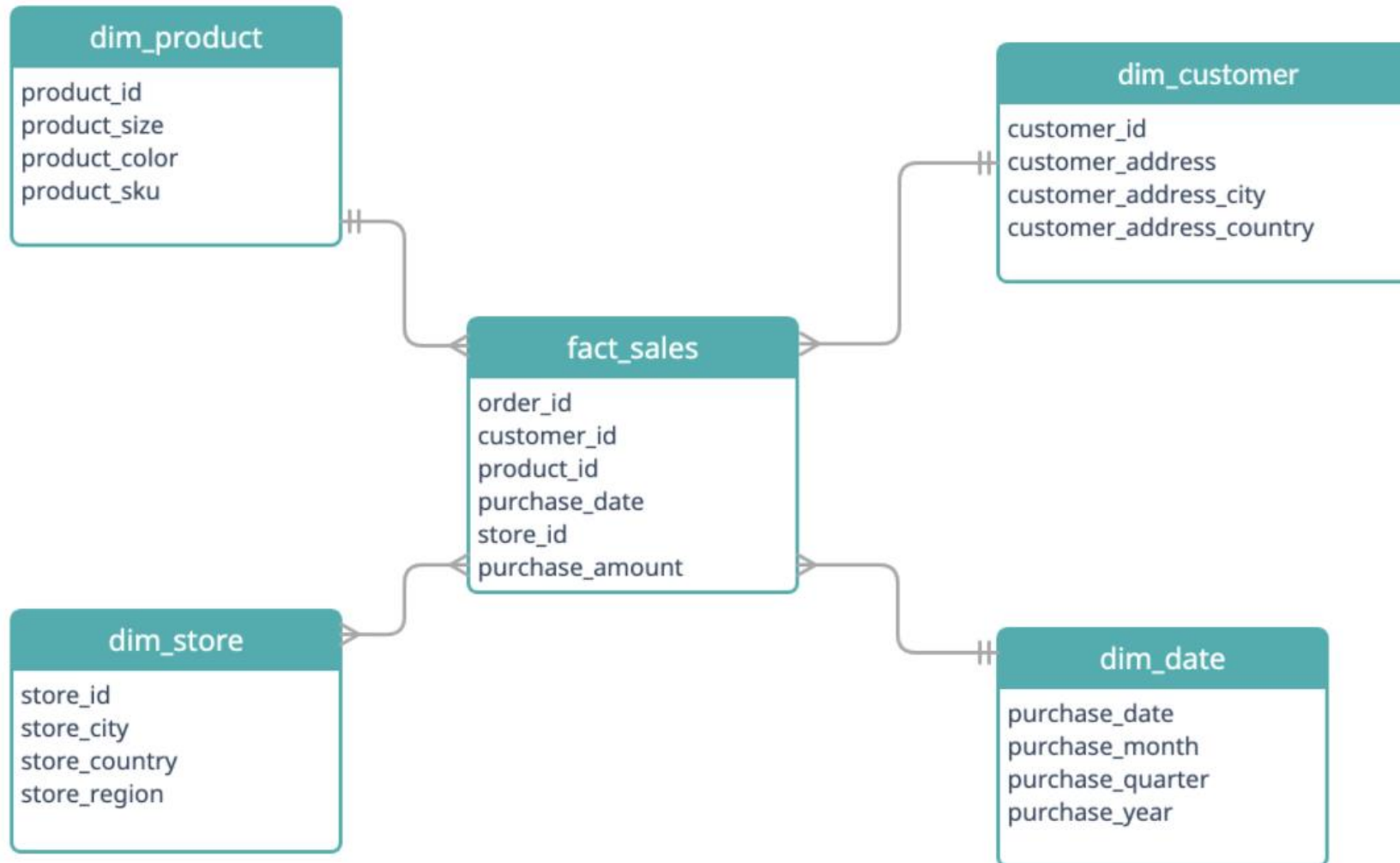
- Sering disebut dengan meaningless keys, integer keys, non-natural keys, artificial keys, atau synthetic keys
- Tidak ada di dalam proses bisnis, tidak memiliki makna di dunia nyata
- Key yang dibuat oleh sistem, biasanya dengan auto-increment
- Contoh: ProductID, CustomerID



Bagian 1

SLOWLY CHANGING DIMENSION

Star Schema



Slowly Changing Dimension (SCD)

- Sebelumnya, diasumsikan bahwa dimension table independen terhadap waktu
- Walaupun relatif statis, perubahan atribut/kolom pada dimension table bisa saja terjadi
- Slowly Changing Dimension → dimensi yang berubah secara perlahan dari waktu ke waktu

SCD Type

- SCD Type → mendefinisikan strategi yang dilakukan untuk meng-handle perubahan nilai atribut pada dimension table
 - ✓ Type 0 : Retain original
 - ✓ Type 1 : Overwrite
 - ✓ Type 2 : Add new row
 - ✓ Type 3 : Add new column
 - ✓ Type 4 : Add mini dimension
 - ✓ Type 5: Combine Type 1 + 4
 - ✓ Type 6: Combine Type 1 + 2 + 3

SCD Type 0 – Retain Original

- **Do nothing:** tidak meng-handle perubahan data pada dimension table.
- Tidak perlu ada yang dilakukan jika terjadi perubahan pada suatu dimensi. Nilainya tetap sama seperti pertama kali diinputkan.
- Contoh:
 - ✓ DateDim
 - ✓ Natural key seperti Email
 - ✓ Field apapun yang ingin dilaporkan berdasarkan versi awalnya

SCD Type 1 – Overwrite

- **Overwrite** → update nilai yang lama dengan nilai yang baru
- Tidak ada riwayat perubahan data
- Mudah di-maintain
- Umumnya digunakan untuk perubahan data yang merupakan proses koreksi
- Contoh: Nama, Tempat Lahir, Tanggal Lahir, Gender

SCD Type 1 – Overwrite

Original row in Product dimension:

| Product Key | SKU (NK) | Product Description | Department Name |
|-------------|----------|---------------------|-----------------|
| 12345 | ABC922-Z | IntelliKidz | Education |

Updated row in Product dimension:

| Product Key | SKU (NK) | Product Description | Department Name |
|-------------|----------|---------------------|-----------------|
| 12345 | ABC922-Z | IntelliKidz | Strategy |

SCD Type 2 – Add New Row

- Paling umum digunakan
- Perubahan data ditambahkan sebagai **baris baru** pada dimension table
- **Tambahkan primary key baru** (surrogate) yang berbeda dengan primary key pada OLTP database (baik OLTP-nya menggunakan surrogate atau natural key)
- Sebaiknya tambahkan atribut penanda, *misalnya*:
 - ✓ Effective Date - Expiration Date - Current Row Indicator
 - ✓ Start Date – Current Date – IsCurrent

SCD Type 2 – Add New Row

Original row in Product dimension:

| Product Key | SKU (NK) | Product Description | Department Name | ... | Row Effective Date | Row Expiration Date | Current Row Indicator |
|-------------|----------|---------------------|-----------------|-----|--------------------|---------------------|-----------------------|
| 12345 | ABC922-Z | IntelliKidz | Education | ... | 2012-01-01 | 9999-12-31 | Current |

Rows in Product dimension following department reassignment:

| Product Key | SKU (NK) | Product Description | Department Name | ... | Row Effective Date | Row Expiration Date | Current Row Indicator |
|-------------|----------|---------------------|-----------------|-----|--------------------|---------------------|-----------------------|
| 12345 | ABC922-Z | IntelliKidz | Education | ... | 2012-01-01 | 2013-01-31 | Expired |
| 25984 | ABC922-Z | IntelliKidz | Strategy | ... | 2013-02-01 | 9999-12-31 | Current |



PK DimProduct

PK Table Product di OLTP

SCD Type 3 – Add New Column

- Jarang digunakan
- Perubahan data disimpan dalam **kolom yang berbeda**,
- Metode ini ideal digunakan jika jumlah perubahan yang ingin disimpan sudah diketahui secara pasti.
- Umumnya digunakan jika hanya perlu menyimpan perubahan terakhir, untuk membandingkan nilai saat ini dan nilai sebelumnya. Gunakan *Current* dan/atau *Prior/Previous*

SCD Type 3 – Add New Column

Original row in Product dimension:

| Product Key | SKU (NK) | Product Description | Department Name |
|-------------|----------|---------------------|-----------------|
| 12345 | ABC922-Z | IntelliKidz | Education |

Updated row in Product dimension:

| Product Key | SKU (NK) | Product Description | Department Name | Prior Department Name |
|-------------|----------|---------------------|-----------------|-----------------------|
| 12345 | ABC922-Z | IntelliKidz | Strategy | Education |

SCD Type 3 – Add New Column (2)

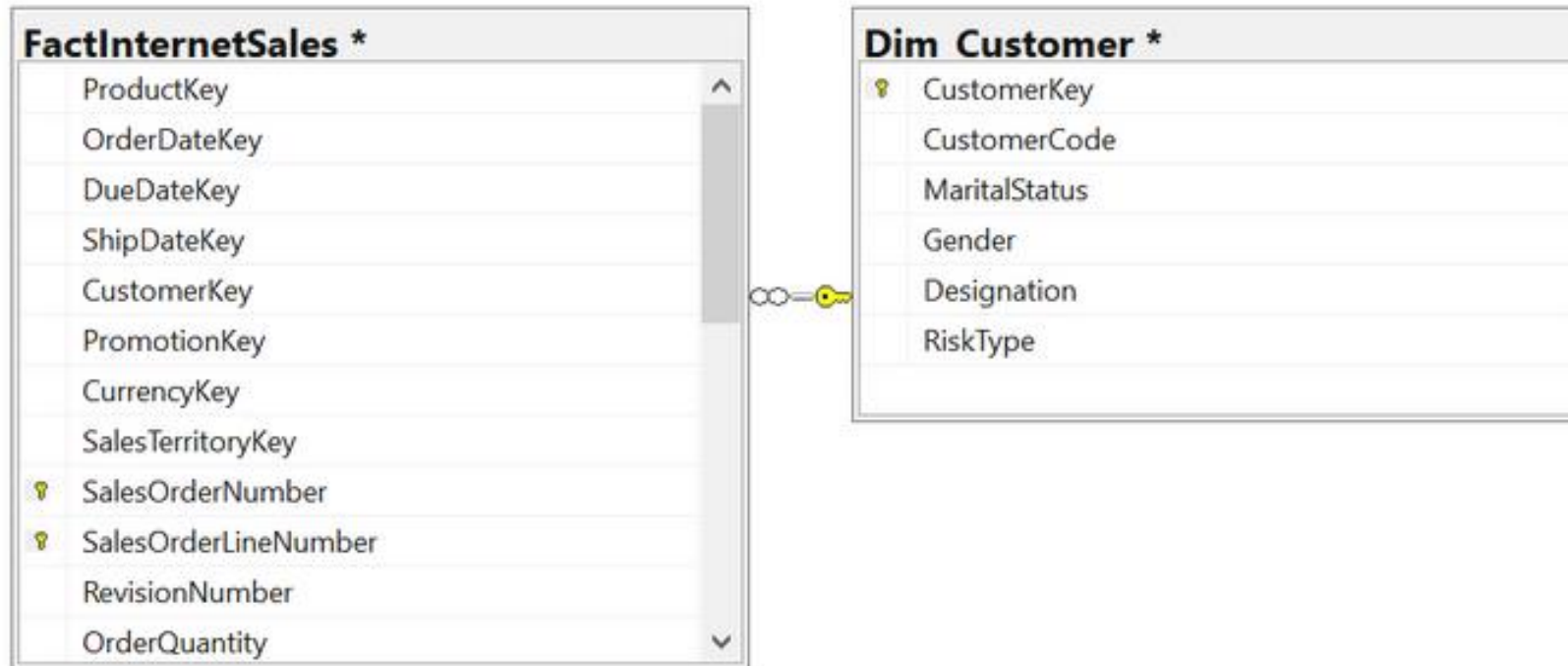
Updated row in Product dimension:

| Product Key | SKU (NK) | Product Description | Current Department Name | 2012 Department Name | 2011 Department Name |
|-------------|----------|---------------------|-------------------------|----------------------|----------------------|
| 12345 | ABC922-Z | IntelliKidz | Strategy | Education | Not Applicable |

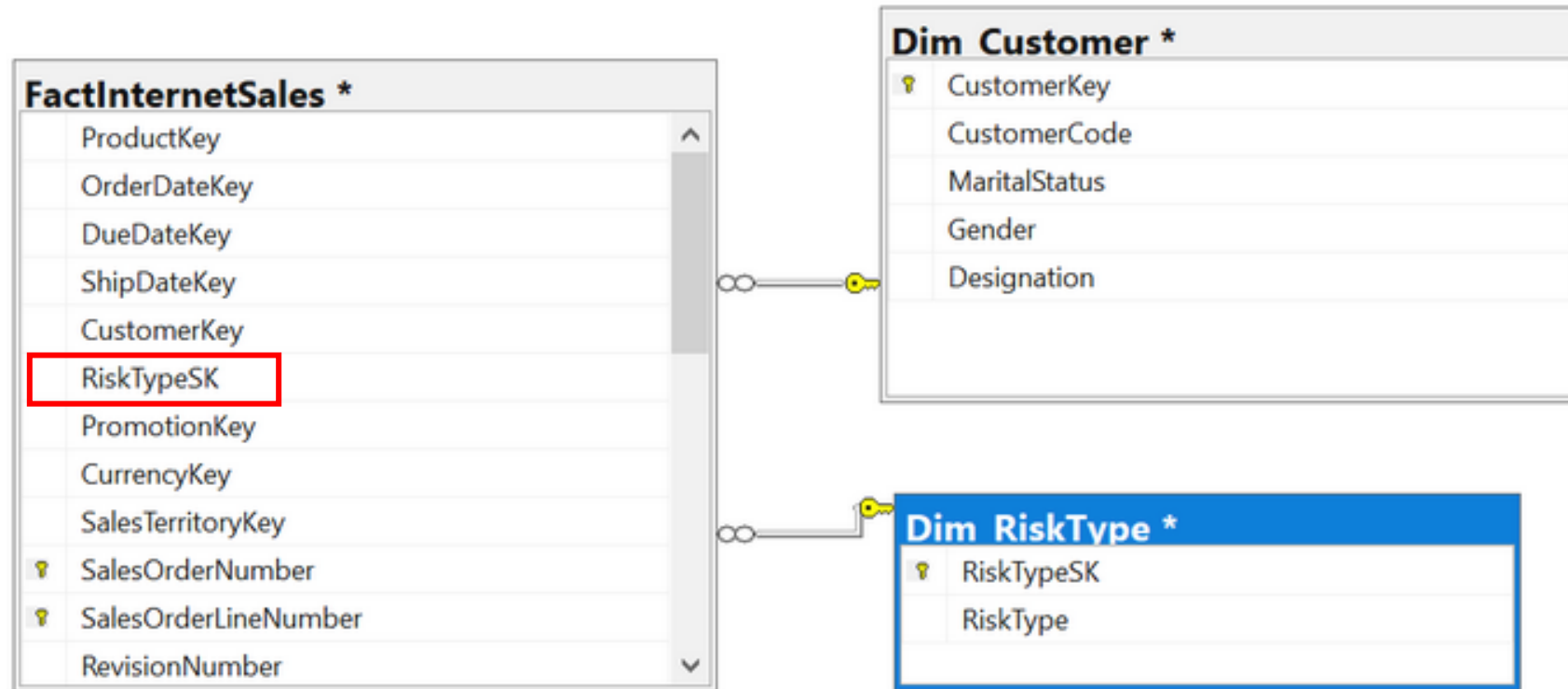
SCD Type 4 – Add Mini Dimension

- Digunakan jika terdapat sekumpulan atribut pada dimension table yang sering berubah nilainya
- Sekumpulan atribut tersebut dipindah ke dimensi baru → sering disebut dengan **mini dimension** atau rapidly changing monster dimension
- Mini dimension membutuhkan primary key nya sendiri
- Baik primary key base dimension dan mini-dimension disimpan pada fact table sebagai foreign key

SCD Type 4 – Add Mini Dimension



SCD Type 4 – Add Mini Dimension

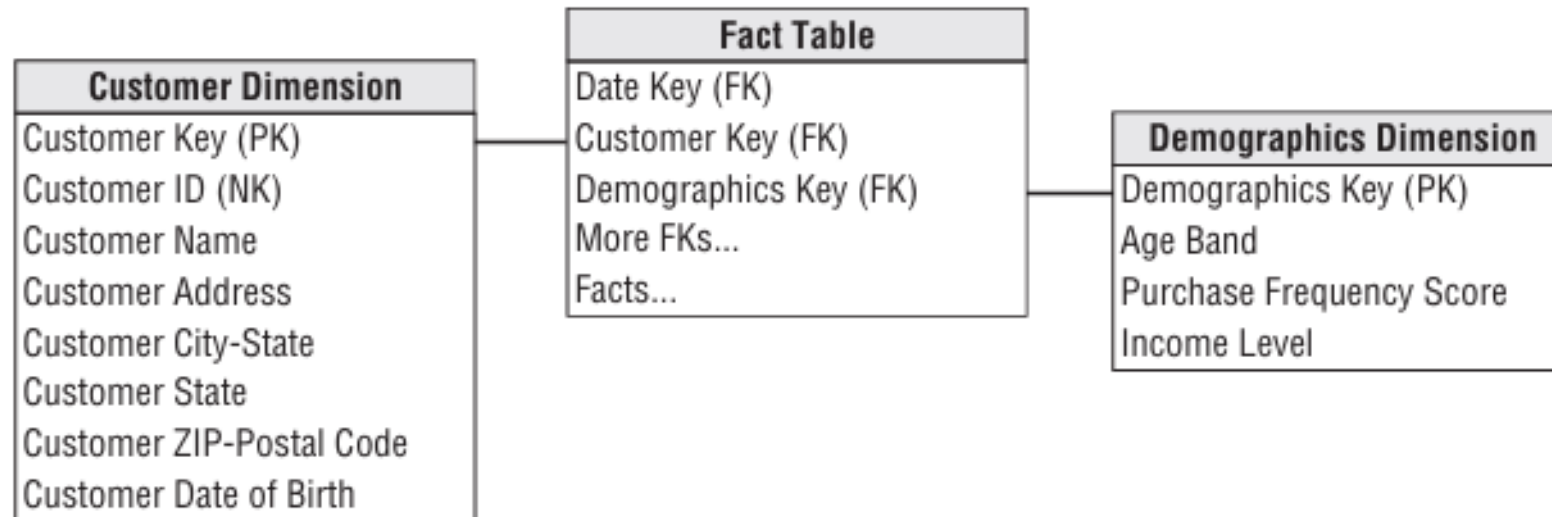


SCD Type 5 – Combine Type 1+4

- Gabungan Type 1 dan 4 → Overwrite dan base dimension + mini dimension
- Data terbaru mini dimension juga tersedia di base dimension dengan konsep overwrite (Type 1)

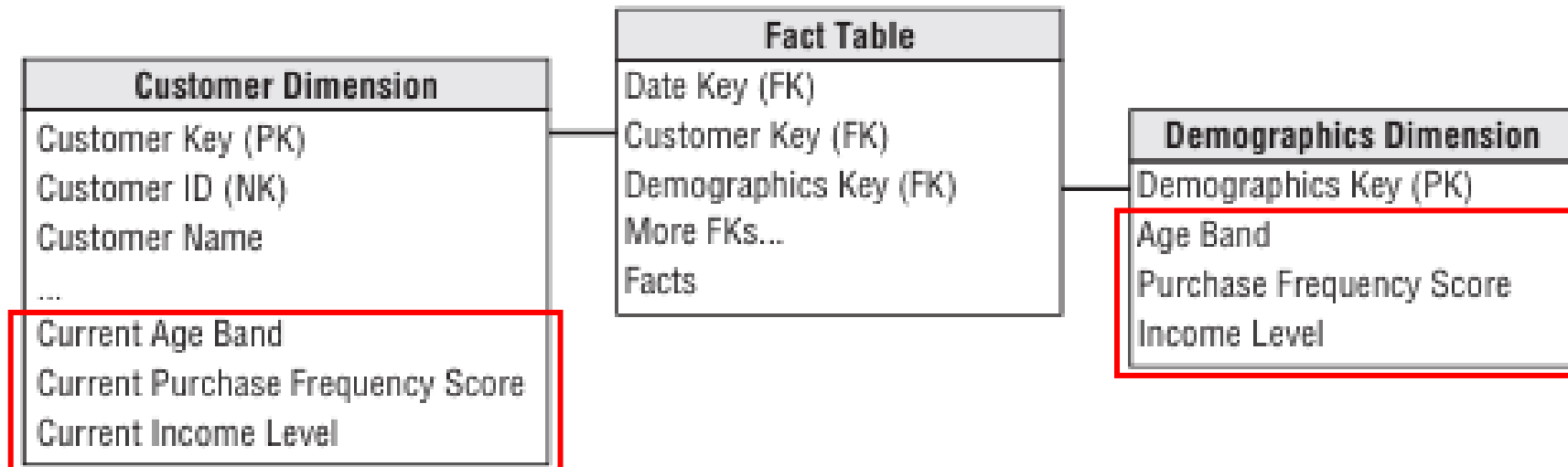
SCD Type 5 – Combine Type 1+4

- Misalnya, 3 column pada table Customer Dimension yaitu *Age Band*, *Purchase Frequency Score*, dan *Income Level* paling sering berubah. Oleh karena itu, 3 column tersebut dibuatkan mini dimension baru yaitu Demographic Dimension → [SCD Type 4](#)



SCD Type 5 – Combine Type 1+4

- Jika menggunakan Type 5, maka ketiga kolom tersebut tetap disimpan juga di Customer Dimension, tapi konsepnya overwrite ([Type 1](#)), artinya, jika ada perubahan Age Band/PFS/Income Level, maka nilainya akan di-overwrite



SCD Type 6 – Combine Type 1+2+3

- Buat baris baru saat ada perubahan data (Type 2)
- Buat kolom tambahan untuk menandakan current data (Type 3)
- Update kolom untuk current data (Type 1)

SCD Type 6 – Combine Type 1+2+3

Original row in Product dimension:

| Product Key | SKU (NK) | Product Description | Historic Department Name | Current Department Name | ... | Row Effective Date | Row Expiration Date | Current Row Indicator |
|-------------|----------|---------------------|--------------------------|-------------------------|-----|--------------------|---------------------|-----------------------|
| 12345 | ABC922-Z | IntelliKidz | Education | Education | ... | 2012-01-01 | 9999-12-31 | Current |

Rows in Product dimension following first department reassignment:

| Product Key | SKU (NK) | Product Description | Historic Department Name | Current Department Name | ... | Row Effective Date | Row Expiration Date | Current Row Indicator |
|-------------|----------|---------------------|--------------------------|-------------------------|-----|--------------------|---------------------|-----------------------|
| 12345 | ABC922-Z | IntelliKidz | Education | Strategy | ... | 2012-01-01 | 2013-01-31 | Expired |
| 25984 | ABC922-Z | IntelliKidz | Strategy | Strategy | ... | 2013-02-01 | 9999-12-31 | Current |

Rows in Product dimension following second department reassignment:

| Product Key | SKU (NK) | Product Description | Historic Department Name | Current Department Name | ... | Row Effective Date | Row Expiration Date | Current Row Indicator |
|-------------|----------|---------------------|--------------------------|-------------------------|-----|--------------------|---------------------|-----------------------|
| 12345 | ABC922-Z | IntelliKidz | Education | Critical Thinking | ... | 2012-01-01 | 2013-01-31 | Expired |
| 25984 | ABC922-Z | IntelliKidz | Strategy | Critical Thinking | ... | 2013-02-01 | 2013-06-30 | Expired |
| 31726 | ABC922-Z | IntelliKidz | Critical Thinking | Critical Thinking | ... | 2013-07-01 | 9999-12-31 | Current |

Notes

- Jangan menyimpulkan bahwa perubahan atribut pada dimension table tidak perlu dihandle karena tidak disebutkan pada saat requirement gathering
- Tetapi jangan juga men-tracking perubahan untuk setiap atribut
- Jika dimension table punya kolom dengan **SCD Type 2 atau 6** → **wajib** punya primary key baru yang berbeda dengan primary key pada OLTP. Jika tidak, maka sebaiknya tetap punya primary key yang berbeda, tapi tidak wajib.
- Primary key baru **wajib** dibuat jika terdapat **multiple data source**, misalnya data warehouse digunakan untuk gudang data beberapa toko cabang



Reference

- The Data Warehouse Toolkit – The Definitive Guide to Dimensional Modelling 3rd edition by Ralph Kimball & Margy Ross



Terima kasih