

3.2 Slowly Changing Dimensions (Langsam veränderliche Dimensionen)

Smart City IoT Projekt

1. Definition und Zielsetzung

In Smart-City-Systemen ändern sich Dimensionsdaten (z.B. Sensor-, Standort-, Wetter- oder Marktdaten) im Laufe der Zeit. Um sowohl den aktuellen Zustand als auch die historische Entwicklung darstellen zu können, wird das Konzept der **Slowly Changing Dimensions (SCD)** verwendet. SCD beschreibt Verfahren, mit denen Veränderungen in den Dimensionsdaten eines Data Warehouse systematisch nachverfolgt werden.

2. Arten von SCD

Tabelle 1: Übersicht der SCD-Typen

Typ	Beschreibung	Vorteil	Nachteil
Typ 0	Keine Änderungen – nur der ursprüngliche Wert bleibt erhalten.	Einfacher Aufbau	Keine Historie
Typ 1	Neue Werte überschreiben alte Werte.	Schnelle Aktualisierung	Verlust historischer Daten
Typ 2	Für jede Änderung wird ein neuer Datensatz mit Start- und Enddatum angelegt.	Vollständige Historie bleibt erhalten	Komplexere ETL-Prozesse
Typ 3	Altdaten werden in zusätzlichen Attributen gespeichert (z.B. <i>Vorheriger_Wert</i>).	Einfacher als Typ 2	Nur begrenzte Historie

In diesem Projekt wird **SCD Typ 2** verwendet, da eine vollständige Historisierung für Trendanalysen essenziell ist.

3. Umsetzung im Smart City IoT Projekt

SCD Typ 2 wird auf folgenden Dimensionen eingesetzt:

- **D_Sensor** – Änderungen an Gerätetyp/Hersteller/Installationsdatum.
- **D_Lokation** – Umbenennungen oder organisatorische Anpassungen.
- **D_MarketInfo** – Änderungen bei Strompreis und prognostizierter Last.
- **D_Wetter** – Kalibrierungen/Anpassungen meteorologischer Werte.

Die Faktentabelle **EnergieAblesung** (zeitpunktgenaue Messungen) erhält *keine* SCD-Logik.

A. Dimension „Sensor“

Tabelle vor der Änderung:

Tabelle 2: Sensor – Zustand vor der Änderung

Sensor_ID	Standort_ID	SensorTyp	Hersteller	Installdatum	StartDate	EndDate	IsCurrent
S-9F42EE91	AST-001	IndustrialMeter	ABB	2024-01-01	2024-01-01	NULL	TRUE

Änderung: Am 30.04.2024 wurde der Sensortyp auf *SmartMeter* aktualisiert.

Tabelle nach Anwendung von SCD Typ 2:

Tabelle 3: Sensor – nach Anwendung von SCD Typ 2

Sensor_ID	Standort_ID	SensorTyp	Hersteller	Installdatum	StartDate	EndDate	IsCurrent
S-9F42EE91	AST-001	IndustrialMeter	ABB	2024-01-01	2024-01-01	2024-04-30	FALSE
S-9F42EE91	AST-001	SmartMeter	ABB	2024-04-30	2024-04-30	NULL	TRUE

Ergebnis: Der alte Datensatz wird geschlossen, ein neuer dokumentiert die Geräteänderung.

B. Dimension „Standort“

Tabelle vor der Änderung:

Tabelle 4: Standort – Zustand vor der Änderung

Standort_ID	Name	City	Region	Country	StartDate	IsCurrent
AST-001	Astana_PowerPlant	Astana	Asia	SolarPark	2024-01-01	TRUE

Änderung: Am 01.06.2024 wurde der Standort in *Astana_East_PowerPlant* umbenannt.

Tabelle nach Anwendung von SCD Typ 2:

Tabelle 5: Standort – nach Anwendung von SCD Typ 2

Standort_ID	Name	City	Region	Country	StartDate	EndDate	IsCurrent
AST-001	Astana_PowerPlant	Astana	Asia	SolarPark	2024-01-01	2024-06-01	FALSE
AST-001	Astana_East_PowerPlant	Astana	Asia	SolarPark	2024-06-01	NULL	TRUE

Ergebnis: Historische Berichte zeigen den alten Standortnamen, neue Analysen den aktuellen.

Tabelle 6: MarketInfo – Zustand vor der Änderung

MarketInfo.ID	Standort.ID	Zeitstempel	Electricity Price (EUR/kWh)	Predicted Load (kW)	StartDate	IsCurrent
MKT-0001	AST-001	2024-04-30	0.447	10.44	2024-04-30	TRUE

C. Dimension „Market Information“

Tabelle vor der Änderung:

Änderung: Am 27.10.2024 änderten sich Strompreis und prognostizierte Last.

Tabelle nach Anwendung von SCD Typ 2:

Tabelle 7: MarketInfo – nach Anwendung von SCD Typ 2

MarketInfo.ID	Standort.ID	Zeitstempel	Electricity Price (EUR/kWh)	Predicted Load (kW)	StartDate	EndDate	IsCurrent
MKT-0001	AST-001	2024-04-30	0.447	10.44	2024-04-30	2024-10-27	FALSE
MKT-0002	AST-001	2024-10-27	0.480	11.10	2024-10-27	NULL	TRUE

Ergebnis: Alte Marktwerde bleiben historisch nachvollziehbar.

D. Dimension „Wetter“

Tabelle vor der Änderung:

Tabelle 8: Wetter – Zustand vor der Änderung

Wetter.ID	Temperatur (°C)	Luftfeuchtigkeit	Sonneneinstrahlung	Windgeschwindigkeit	StartDate	EndDate	IsCurrent
W-1001	27.92	30.24	1200	2.47	2024-04-30	2024-04-30	TRUE

Änderung: Am 15.05.2024 wurde die Wetterstation neu kalibriert.

Tabelle nach Anwendung von SCD Typ 2:

Tabelle 9: Wetter – nach Anwendung von SCD Typ 2

Wetter.ID	Temperatur (°C)	Luftfeuchtigkeit	Sonneneinstrahlung	Windgeschwindigkeit	StartDate	EndDate	IsCurrent
W-1001	27.92	30.24	1200	2.47	2024-04-30	2024-05-15	FALSE
W-1002	27.60	31.10	1180	2.40	2024-05-15	NULL	TRUE

4. Allgemeine ETL-Logik für SCD Typ 2

```

IF incoming_record differs from current_record THEN
    UPDATE current_record SET EndDate = CURRENT_DATE, IsCurrent = FALSE;
    INSERT new_record SET StartDate = CURRENT_DATE, EndDate = NULL, IsCurrent = TRUE
END IF;

```

5. Gründe für die Wahl von SCD Typ 2

1. Vollständige Historisierung der Dimensionen Sensor, Standort, Wetter und Markt.
2. Exakte zeitliche Analysen und Trendverfolgung über Perioden.
3. Nachvollziehbarkeit und Datenintegrität im gesamten Lebenszyklus.
4. Unterstützung für Prognose-, Optimierungs- und KI-Modelle.

6. Fazit

Durch die Anwendung von **SCD Typ 2** bleiben aktuelle und historische Zustände vollständig nachvollziehbar. **Technischer Hinweis:** Da SCD Typ 2 die Datenmenge und die Anzahl der Tabellen erhöht, empfiehlt es sich, umfangreiche Beispieldatensätze in den **Anhang** bzw. separate CSV-Dateien auszulagern, damit die Hauptdokumentation übersichtlich bleibt.