

Dimensional Model (DDL)

SQL - OLAP star schema with SCD2 & naming conventions

Sanitized work sample authored by Bernard Millet. All client-specific data, names and identifiers have been removed; it is representative of approach and craft, not a client deliverable.

```

/* =====
DIMENSIONAL MODEL: star schema with SCD2, surrogate keys and conventions
Author: Bernard Millet. DAMA-DMBOK: Data Modeling and Design
Portfolio work sample. Sanitized and representative of delivery at a global
bank (anonymized). Contains no proprietary data.
-----
A reference star schema for a finance/regulatory data mart. Demonstrates the
design standards I apply: surrogate keys, conformed dimensions, slowly
changing dimension (Type 2) history, explicit grain, and naming conventions
that make models self-documenting and BCBS 239-traceable.
Dialect: ANSI / Snowflake-compatible DDL.
-----
NAMING CONVENTIONS
dim_* dimension      fct_* fact          br_* bridge
*_sk surrogate key (system)          *_id natural/business key
*_eff_from / *_eff_to SCD2 validity    is_current current-row flag
===== */

-- ----- DATE DIM
CREATE TABLE dim_date (
    date_sk          INTEGER      NOT NULL PRIMARY KEY,    -- yyyyymmdd surrogate
    calendar_date    DATE        NOT NULL,
    day_of_week      SMALLINT    NOT NULL,
    month_number     SMALLINT    NOT NULL,
    month_name       VARCHAR(12) NOT NULL,
    quarter          SMALLINT    NOT NULL,
    year             SMALLINT    NOT NULL,
    is_month_end     BOOLEAN     NOT NULL,
    is_business_day  BOOLEAN     NOT NULL
);

-- ----- CUSTOMER DIM (SCD2)
-- Grain: one row per customer per version. History preserved via eff_from/to.
CREATE TABLE dim_customer (
    customer_sk      BIGINT      NOT NULL PRIMARY KEY,    -- surrogate (stable join key)
    customer_id      VARCHAR(20) NOT NULL,                -- natural/business key
    legal_name       VARCHAR(200) NOT NULL,
    country_code     CHAR(2)     NOT NULL,                -- ISO-3166 alpha-2
    risk_rating      VARCHAR(10),                          -- conformed reference value
    classification   VARCHAR(20) NOT NULL DEFAULT 'Confidential',
    -- SCD Type 2 control columns
    eff_from         TIMESTAMP   NOT NULL,
    eff_to           TIMESTAMP   NOT NULL DEFAULT TIMESTAMP '9999-12-31 00:00:00',
    is_current       BOOLEAN     NOT NULL DEFAULT TRUE,
    record_source    VARCHAR(30) NOT NULL,                -- lineage: system of record
    load_ts          TIMESTAMP   NOT NULL
);

-- ----- ACCOUNT DIM
CREATE TABLE dim_account (
    account_sk       BIGINT      NOT NULL PRIMARY KEY,
    account_id       VARCHAR(34) NOT NULL,                -- IBAN-length safe
    account_type     VARCHAR(20) NOT NULL,
    currency_code    CHAR(3)     NOT NULL,                -- ISO-4217
    booking_entity   VARCHAR(40) NOT NULL
);

-- ----- TRANSACTION FACT
-- Grain: one row per posted transaction. Additive measure = amount.

```

```

CREATE TABLE fct_transaction (
    transaction_sk      BIGINT      NOT NULL PRIMARY KEY,
    -- dimension foreign keys (star joins)
    date_sk             INTEGER     NOT NULL REFERENCES dim_date(date_sk),
    customer_sk         BIGINT      NOT NULL REFERENCES dim_customer(customer_sk),
    account_sk          BIGINT      NOT NULL REFERENCES dim_account(account_sk),
    -- degenerate dimension (carried on the fact, no separate table)
    transaction_ref     VARCHAR(40) NOT NULL,
    -- measures
    amount              DECIMAL(18,2) NOT NULL,                -- additive, in account ccy
    amount_chf          DECIMAL(18,2) NOT NULL,                -- conformed reporting ccy
    -- lineage / audit (BCBS 239 traceability)
    record_source       VARCHAR(30) NOT NULL,
    load_ts             TIMESTAMP   NOT NULL
);

-- Recommended access pattern: filter on dim_date, join customer/account on _sk.
CREATE INDEX ix_fct_txn_date ON fct_transaction(date_sk);
CREATE INDEX ix_fct_txn_cust ON fct_transaction(customer_sk);

/* -----
   SCD2 UPSERT PATTERN (close the old version, insert the new one)
   ----- */
-- 1) expire the current row when a tracked attribute changes
UPDATE dim_customer d
SET   eff_to = CURRENT_TIMESTAMP, is_current = FALSE
FROM   stg_customer s
WHERE  d.customer_id = s.customer_id
      AND d.is_current = TRUE
      AND (d.legal_name <> s.legal_name
          OR d.country_code <> s.country_code
          OR d.risk_rating IS DISTINCT FROM s.risk_rating);

-- 2) insert the new current version (surrogate key from a sequence/identity)
INSERT INTO dim_customer
(customer_sk, customer_id, legal_name, country_code, risk_rating,
 classification, eff_from, eff_to, is_current, record_source, load_ts)
SELECT
    nextval('seq_customer_sk'), s.customer_id, s.legal_name, s.country_code,
    s.risk_rating, 'Confidential', CURRENT_TIMESTAMP,
    TIMESTAMP '9999-12-31 00:00:00', TRUE, s.record_source, CURRENT_TIMESTAMP
FROM stg_customer s
LEFT JOIN dim_customer d
    ON d.customer_id = s.customer_id AND d.is_current = TRUE
WHERE d.customer_id IS NULL      -- brand new customer
    OR d.is_current = FALSE;    -- or a freshly expired version

```