**CartDM Data Mart Documentation**

**Purpose**: This document outlines the design and structure of the **CartDM** data mart, which supports analytical queries related to user cart behavior, course popularity, and student engagement for a Udemy-like e-learning platform. The schema follows a **star schema** design optimized for OLAP queries.

**1. Design Principles**

**1.1 Star Schema Foundation**

* **Fact Table**: FactCarts (grain: cart-course interaction).
* **Core Dimensions**: DimStudents, DimCourses, DimDate.
* **Conformed Dimensions**: SDimQuiz, SubDimCrsReq, SubDimSection (reusable across other data marts).

**1.2 SCD Handling**

* **Type 2 SCD**: Used for DimStudents to track historical changes in student profiles (e.g., country, wallet balance).
* **Type 1 SCD**: Used for DimCourses (overwrite changes, no history tracking).

**1.3 Key Metrics**

* Cart activity (add/remove courses).
* Course pricing trends.
* Student demographics and engagement.
* Course content structure (requirements, sections, quizzes).

**2. Data Model Overview**

**2.1 Core Dimension Tables**

DimStudents**(Type 2 SCD)**

* **Purpose**: Track student profiles over time.
* **Key Fields**:
  + StudentId\_SK: Surrogate key (identity).
  + StudentId\_BK: Business key from source (AspNetUsers.Id).
  + Demographics: Combined country/city/state.
  + SocialPresence: Flags for social media links (e.g., 1 if Facebook/X exists).
  + SCD Dates: Start/end dates for validity period.

DimCourses**(Type 1 SCD)**

* **Purpose**: Current state of courses.
* **Key Fields**:
  + CourseId\_SK: Surrogate key.
  + CategoryInfo: Denormalized category/subcategory hierarchy.
  + Popularity: Derived from NoSubscribers.

DimDate

* **Purpose**: Standard date dimension for time-based analysis.
* **Granularity**: Daily.

**2.2 Fact Table**

FactCarts

* **Grain**: One row per cart-course interaction (e.g., course added/removed).
* **Measures**:
  + Amount: Number of courses in the cart.
* **Foreign Keys**:
  + StudentId\_SK: Links to DimStudents.
  + CourseId\_SK: Links to DimCourses.
  + DateKey: Links to DimDate (cart activity date).

**2.3 Supporting Dimensions**

**SDimQuiz**

* **Purpose**: Analyze quiz structure per course.
* **Key Fields**:
  + QuestionTypes: Count of multiple-choice/true-false questions.

**SubDimCrsReq**

* **Purpose**: Track course prerequisites.
* **Key Fields**:
  + Requirement: Text of prerequisite (e.g., "Basic Python knowledge").

**SubDimSection**

* **Purpose**: Analyze course content structure.
* **Key Fields**:
  + ContentMetrics: Aggregated duration/lesson counts.
  + MediaTypes: Video/article ratios.

**3. Key Features**

**3.1 Student-Centric Analysis**

* Track cart behavior by demographics (e.g., "Students aged 20-30 in Germany prefer coding courses").
* Monitor wallet balance trends for premium course purchases.

**3.2 Course Popularity**

* Correlate cart additions with pricing changes (DimCourses.Pricing).
* Identify bestsellers by Popularity and CategoryInfo.

**3.3 Content Optimization**

* Analyze course requirements (SubDimCrsReq) to improve completion rates.
* Compare quiz complexity (SDimQuiz.QuestionTypes) with cart abandonment.

**4. ETL Considerations**

**4.1 Source Tables**

* **Operational Database**:
  + Carts, CartCourse, Courses, AspNetUsers, Students, Subcategories.
  + Sections, Lessons, QuizQuestions.

**4.2 Transformation Logic**

1. **Student Dimension**:
   * Merge AspNetUsers + Students + SocialMedias (social presence flags).
   * Handle Type 2 SCD using CreatedDate/ModifiedDate.
2. **Course Dimension**:
   * Denormalize Categories + Subcategories into CategoryInfo.
3. **FactCarts**:
   * Map CartCourse to fact rows with CartId\_BK as the source key.

**4.3 Incremental Loading**

* Use CreatedDate/ModifiedDate from source tables to identify delta changes.