Machine Learning and Data Mining

DFM – Example

Claudio Sartori

DISI
Department of Computer Science and Engineering – University of Bologna, Italy
claudio.sartori@unibo.it

Example of Conceptual Modeling Using the Dimensional Fact Model (DFM)

Retail Sales Data Warehouse

- **Business Scenario**
  - Build a data warehouse for a retail store chain to analyze sales by product, store, and time.

Identify the Fact: Sales Transaction

- Each row represents a sale of a product in a store at a particular time.
- **Measures:**
  - Revenue
  - Quantity Sold
  - Profit

Fact Table: Sales Fact

- **Measures:**
  - Revenue (continuous measure)
  - Quantity Sold (discrete measure)
  - Profit (calculated measure: Revenue - Cost)

Identify the Dimensions: Primary Dimensions I

- **Product**
  - Attributes: Product ID, Product Name, Category, Brand
- **Store**
  - Attributes: Store ID, Store Name, Location, Store Type
- **Time**
  - Attributes: Date, Month, Quarter, Year

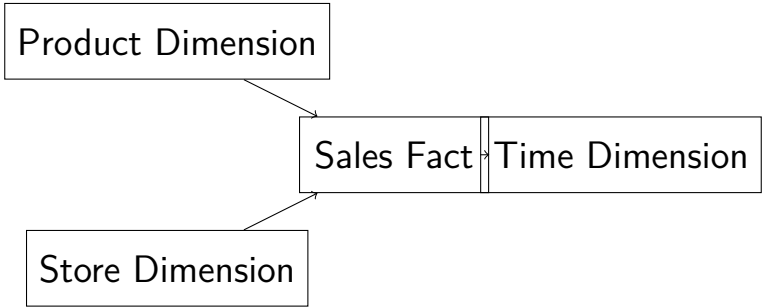
Dimension Tables I

- **Dimension Table 2: Product Dimension**
  - Product ID (Primary Key)
  - Product Name
  - Category
  - Brand
- **Dimension Table 2: Store Dimension**
  - Store ID (Primary Key)
  - Store Name
  - Location
  - Store Type

Dimension Tables II

- **Dimension Table 3: Time Dimension**
  - Date (Primary Key)
  - Month
  - Quarter
  - Year

Dimensional Fact Model Diagram



- Sales Fact table at the center
- Product, Store, and Time dimensions surrounding the fact table
- Measures stored in the fact table

Define the Granularity

- The granularity of the fact table is the daily sales of a product in a store.
- Each row in the fact table corresponds to the sale of a specific product in a particular store on a specific date.

Example Queries

- What is the total revenue for Product A in Store X during Q1 2024?
- How many units of Product B were sold across all stores in January 2024?
- What was the profit margin for each product category last year?