COMP 333 — Week 5 Datacubes and OLAP

Datacubes and OLAP

The lecture will discuss Data Cubes and OLAP

READ Golfarelli and Rizzi Sections 1.5, and 1.7.2

Introduction to Data Warehousing, Chapter 1 of Data Warehouse Design: Modern Principles and Methodologies, 2009, by Matteo Golfarelli and Stefano Rizzi.

Data Cubes

A *data cube* is an n-dimensional cube of values.

The multidimensional model begins with the observation
that the factors affecting decision-making processes
are enterprise-specific *facts*,
such as sales, shipments, hospital admissions, surgeries, and so on.

Instances of a *fact* correspond to *events* that occurred.
For example, every single sale or shipment carried out is an *event*

Each *fact* is described by the *values* of a set of relevant measures
that provide a quantitative description of events.
For example, sales receipts, amounts shipped, hospital admission costs, and surgery time are measures.

FIGURE 1-10
The three-dimensional cube modeling sales in a store chain: 10 packs of Shiny were sold on 4/5/2008 in the EverMore store, totaling $25.

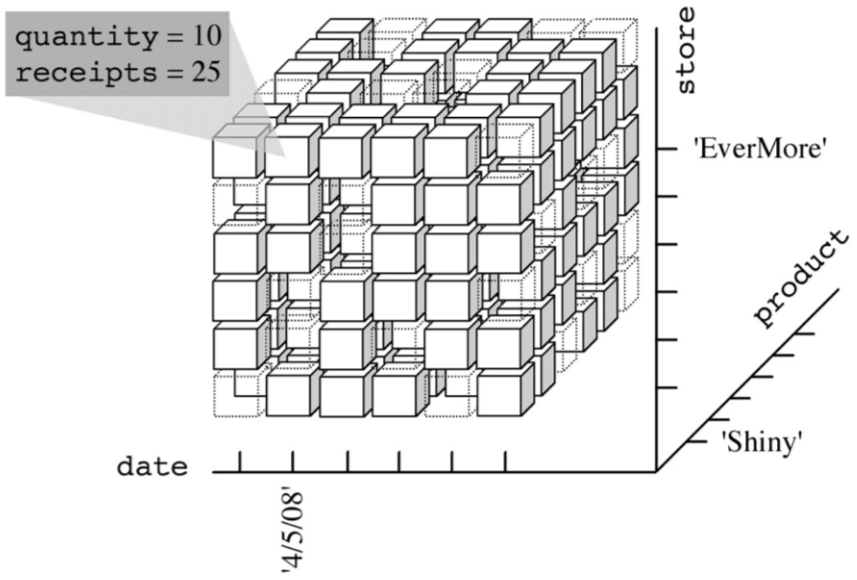
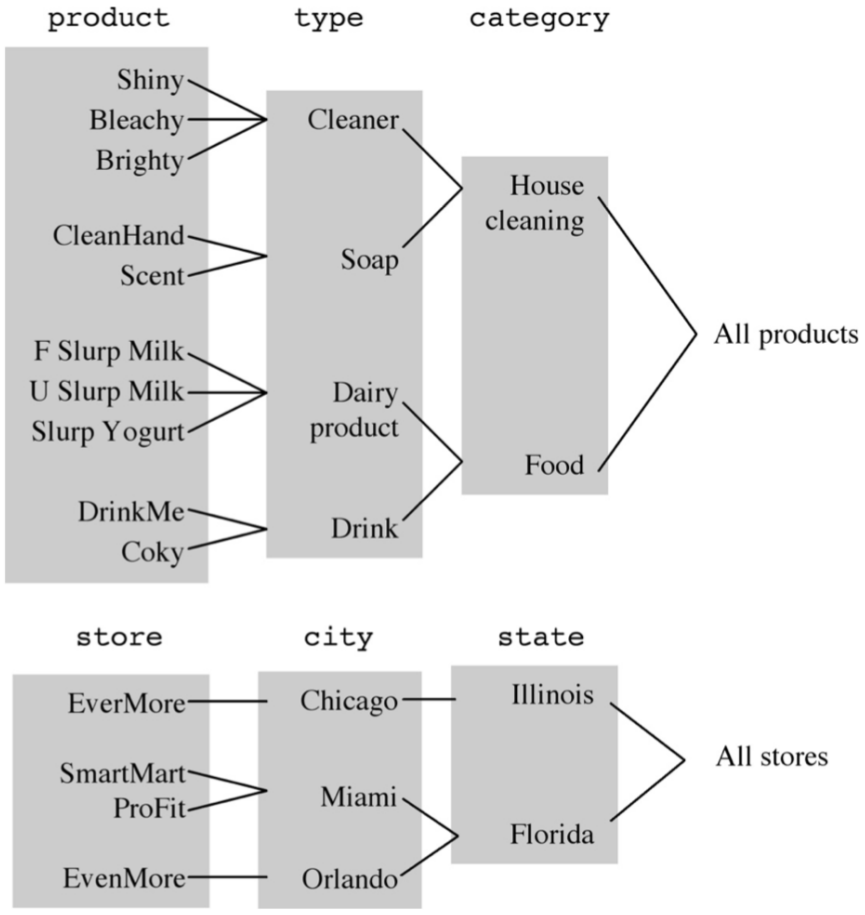


FIGURE 1-11
Aggregation hierarchies built on the `product` and `store` dimensions



OLAP

OLAP stands for Online Analytic Processing
in constrast to OLTP (Online Transactional Processing)
of typical interaction with a database
that reads, creates, deletes, and updates entries.

OLAP operates on a datawarehouse/datamart as a datacube.
These are read-only.
They are updated in bulk when they are re-built periodically.