Dimensional Modeling and OLAP

Chaudhuri and Dayal paper,
"A Dimensional Modeling Manifesto" by R. Kimbal

Outline

- Multidimensional view of data
- OLAP operations
- Dimensional modeling
- SQL Extensions

OLAP

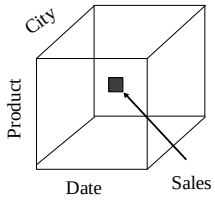
- OLAP applications provide users with *multidimensional view* of data in the DW
  - view data from different perspectives
    - sales by product, sales by region, sales by week
  - view data at different levels of detail
    - sales by region by month → sales by store by week

OLAP (cont.)

- Typical operations include
  - rollup
  - drill-down
  - slice-and-dice
  - pivot

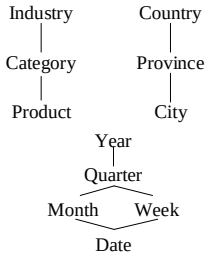
Multidimensional View

- Set of numeric measures
  - Sales
- Set of dimensions
  - Product, Date, City
- Dimensions form a key for the measures



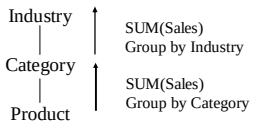
Multidimensional View (cont.)

- Each dimension described by a set of attributes
- Attributes of a dimension may be related by a hierarchy



Multidimensional View (cont.)

- Aggregation of measures along one or more dimensions is a key operation

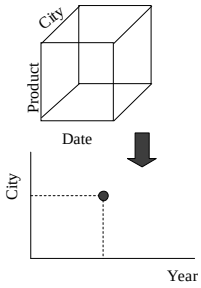


CISC 432/832 Dimensional Modeling 7

OLAP Operations

Pivot

- re-orient multidimensional view of the data
- look at another face of the cube

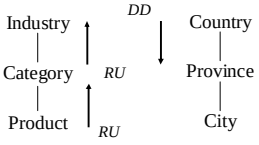


CISC 432/832 Dimensional Modeling 8

OLAP Operations (cont.)

Rollup / Drill-down

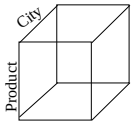
- increase/decrease level of aggregation along one or more dimensions



CISC 432/832 Dimensional Modeling 9

OLAP Operations(cont.)

Slice-and-dice
- do selects and projects on cube



Product	City	Sales
	Ottawa	
	Ottawa	
	Ottawa	
	Ottawa	

CISC 432/832 Dimensional Modeling 10

Dimensional Modeling

- Design technique used for data warehouses
- Maps multidimensional view to relations
- A dimensional model composed of
  - fact table
  - set of dimension tables
- A dimensional model corresponds to one business process

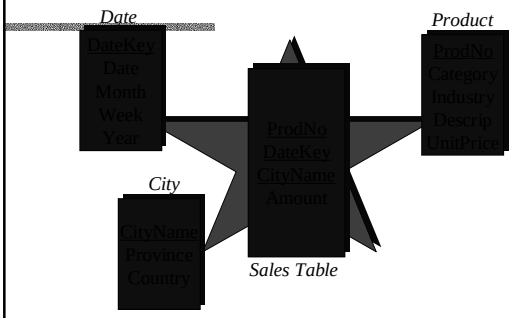
CISC 432/832 Dimensional Modeling 11

Dimensional Modeling (cont.)

- Advantages:
  - predictable, standard framework
  - flexible and easily modified
  - fits business view of the data
  - supports aggregation

CISC 432/832 Dimensional Modeling 12

Star Schema



CISC 432/832 Dimensional Modeling 13

SQL Extensions

- SQL originally designed for OLTP applications
- SQL not suited for OLAP
  - hard to express complex queries
  - inadequate functionality for analysis
  - hard to construct
    - histograms
    - roll-up totals and subtotals
    - cross tabulations

CISC 432/832 Dimensional Modeling 14

SQL Extensions (cont.)

- Aggregate functions
  - support for rank and percentile
  - support for functions from financial analysis like mean, mode and median
- Reporting features
  - moving SUM and AVG
  - computing running totals
  - setting breakpoints to calculate subtotals

CISC 432/832 Dimensional Modeling 15

SQL Extensions (cont.)

- Multiple group-by
  - CUBE operator supports multidimensional data cube

```
select ProdNo, DateKey, CityName, SUM(Amount)
from Sales
where CityName in {'Hong Kong', 'Beijing'}
and DateKey between '040199' and '040299'
group by CUBE (ProdNo, DateKey, CityName)
```

CISC 432/832 Dimensional Modeling 16

SQL Extensions (cont.)

- CUBE generates a table that contains a row for sum over every possible dimension combination

ProdNo	DateKey	CityName	SUM(Sales)
1	040199	Beijing	5
1	040299	Beijing	4
1	ALL	Beijing	9
1	ALL	ALL	20
ALL	ALL	ALL	200

CISC 432/832 Dimensional Modeling 17