ON-LINE ANALYTICAL PROCESSING

Dr. Yashvardhan Sharma

Department of Computer Science & Information Systems

BITS, Pilani

L11

Dimensional Analysis

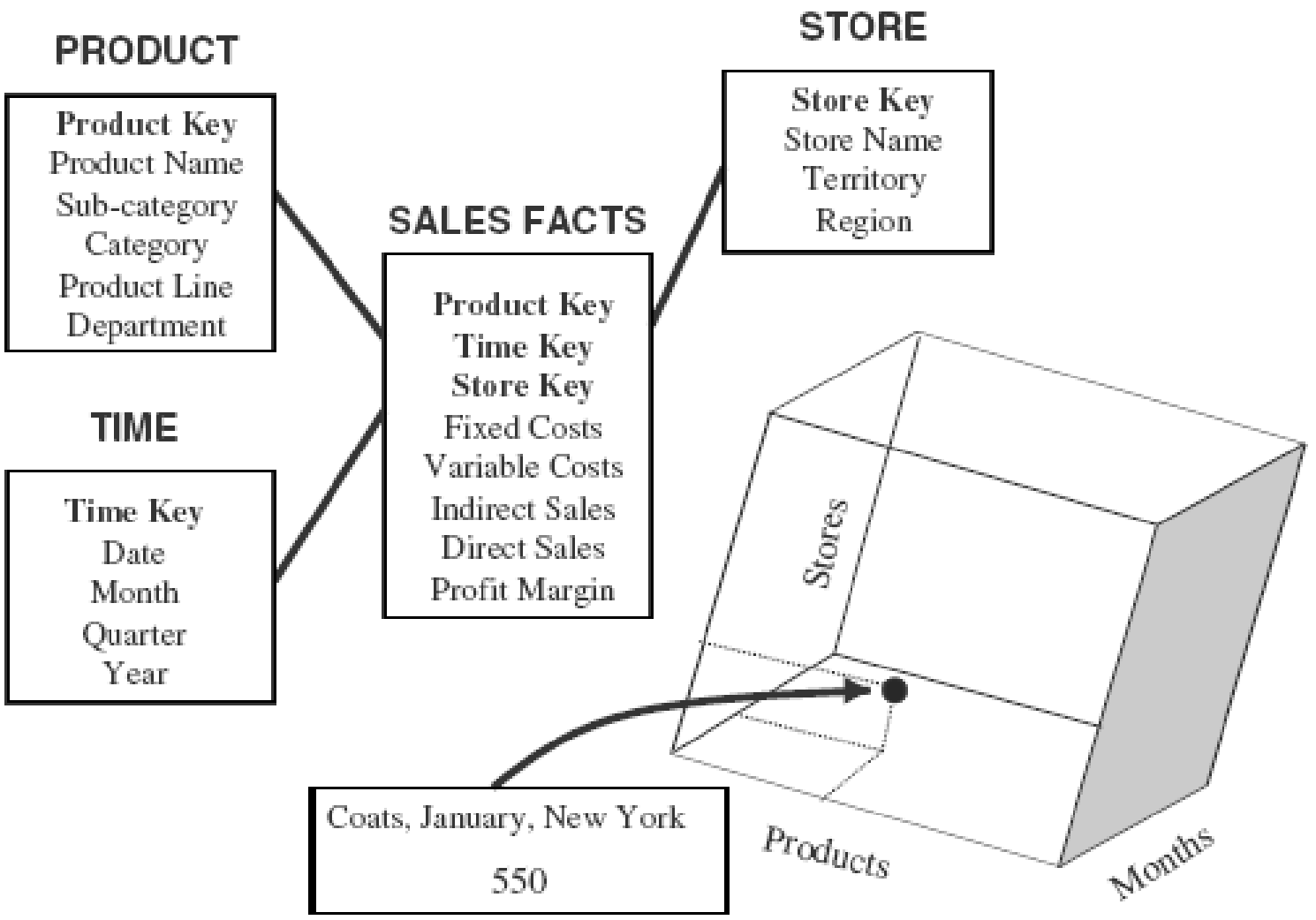


Figure 15-5 Simple STAR schema.

Store: New York

PAGES: STORE dimension

Products

COLUMNS: PRODUCT dimension

ROWS: TIME dimension

Months

	Hats	Coats	Jackets	Dresses	Shirts	Slacks
Jan	200	550	350	500	520	490
Feb	210	480	390	510	530	500
Mar	190	480	380	480	500	470
Apr	190	430	350	490	510	480
May	160	530	320	530	550	520
Jun	150	450	310	540	560	330
Jul	130	480	270	550	570	250
Aug	140	570	250	650	670	230
Sep	160	470	240	630	650	210
Oct	170	480	260	610	630	250
Nov	180	520	280	680	700	260
Dec	200	560	320	750	770	310

Figure 15-6 A Three-dimensional display.

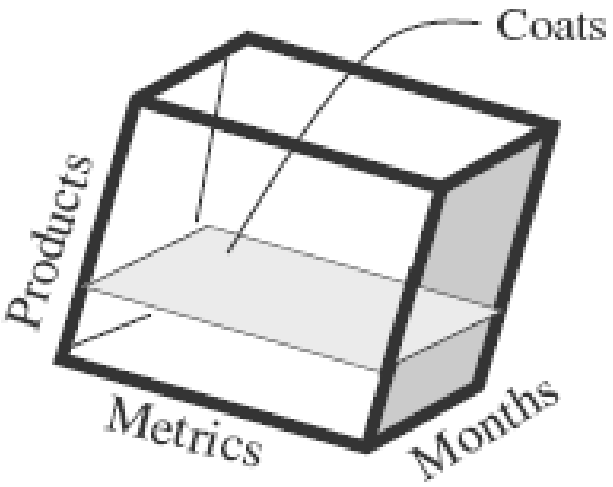
Hypercubes

PRODUCT: Coats

PAGES: PRODUCT dimension COLUMNS: Metrics

ROWS: TIME dimension

	Fixed Cost	Variable Cost	Indirect Sales	Direct Sales	Profit Margin
Jan	340	110	230	320	100
Feb	270	90	200	260	100
Mar	310	100	210	270	70
Apr	340	110	210	320	80
May	330	110	230	300	90
Jun	260	90	150	300	100
Jul	310	100	180	300	70
Aug	380	130	210	360	60
Sep	300	100	180	290	70
Oct	310	100	170	310	70
Nov	330	110	210	310	80
Dec	350	120	200	360	90



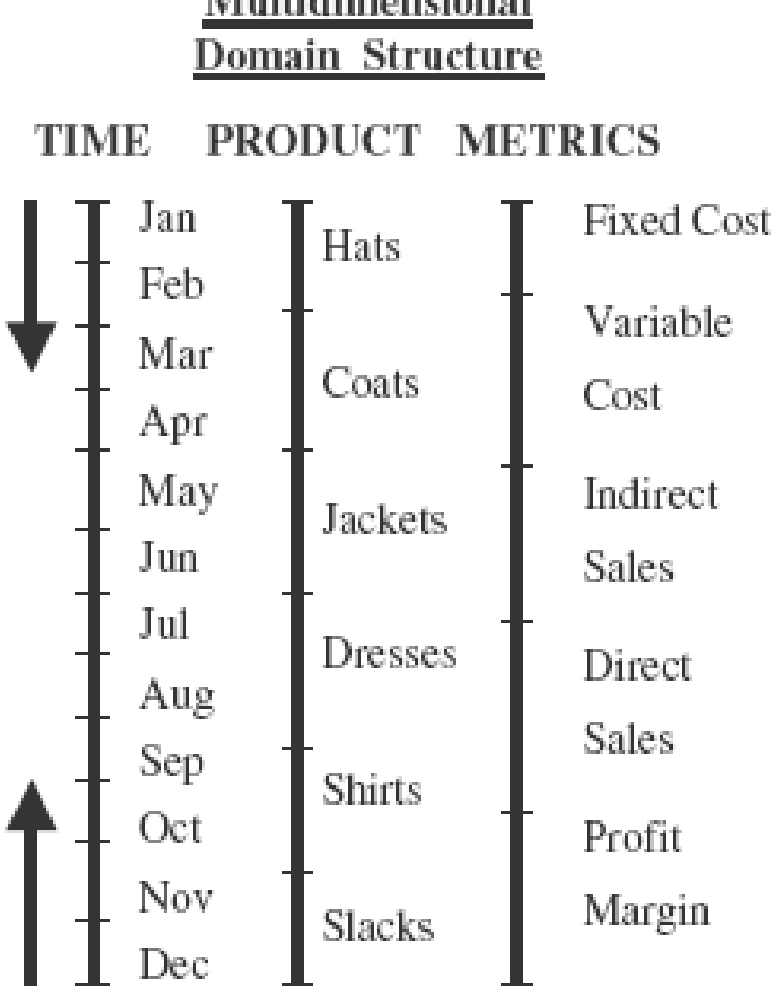


Figure 15-7 Display of columns, rows, and pages.

- In the figure, note the three straight lines, two of which represent the two business dimensions and the third, the metrics. You can independently move up or down along the straight lines.
- Some experts refer to this representation of a multidimension as a **multidimensional domain structure (MDS)**.

Multidimensional Domain Structure

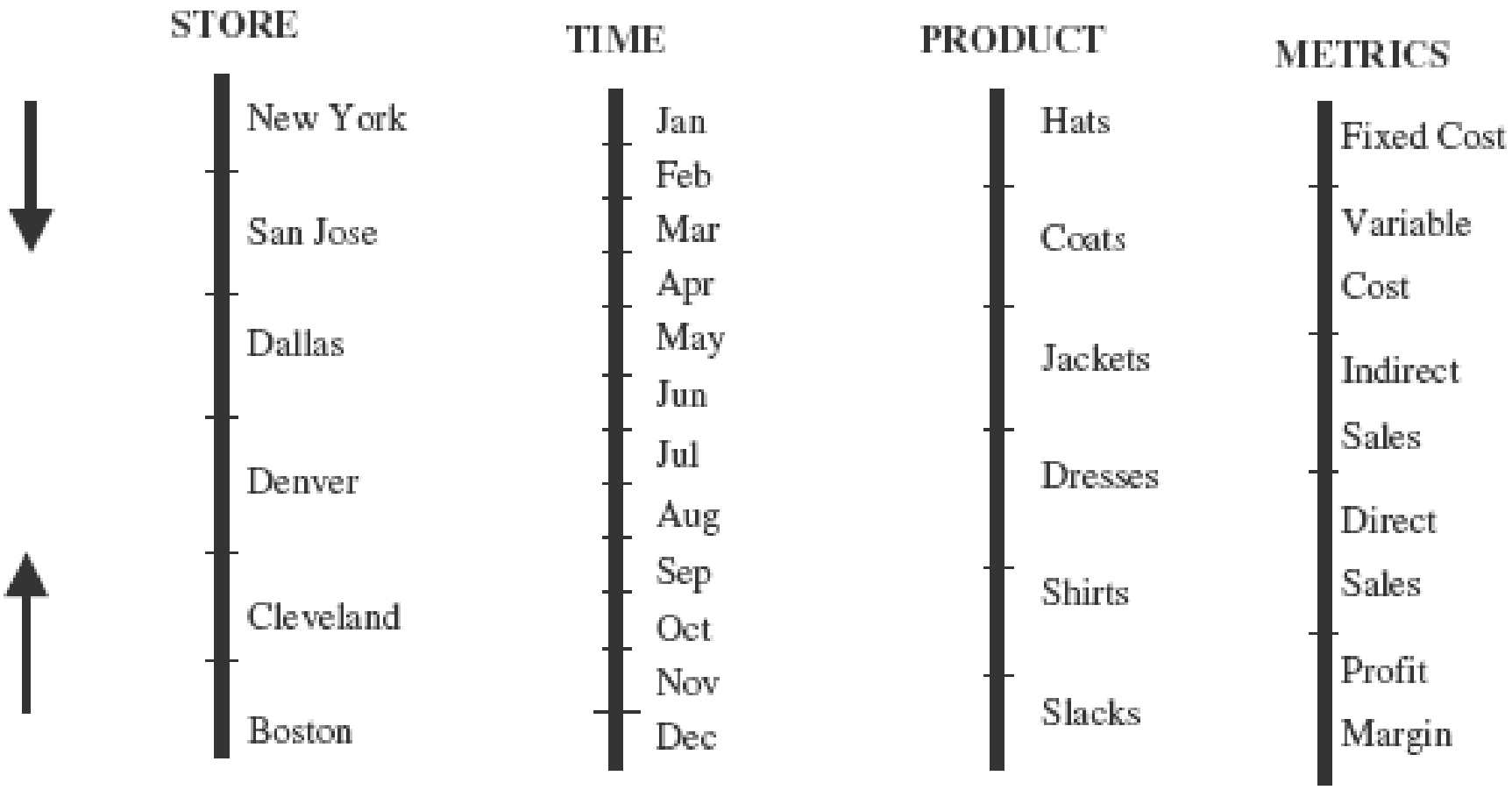
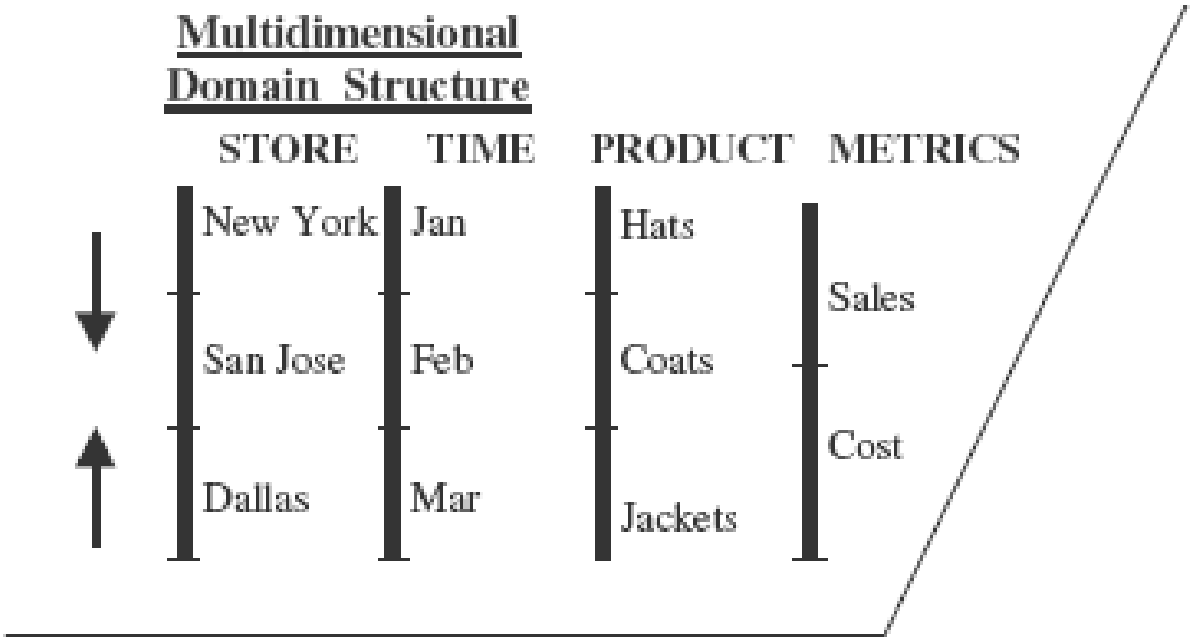


Figure 15-8 MDS for four dimensions.



New York Store

	Hats:Sales	Hats:Cost	Coats:Sales	Coats:Cost	Jackets:Sales	Jackets:Cost
Jan	450	350	550	450	500	400
Feb	380	280	460	360	400	320
Mar	400	310	480	410	450	400

Figure 15-9 Page displays for four-dimensional data.

Multidimensional Domain Structure

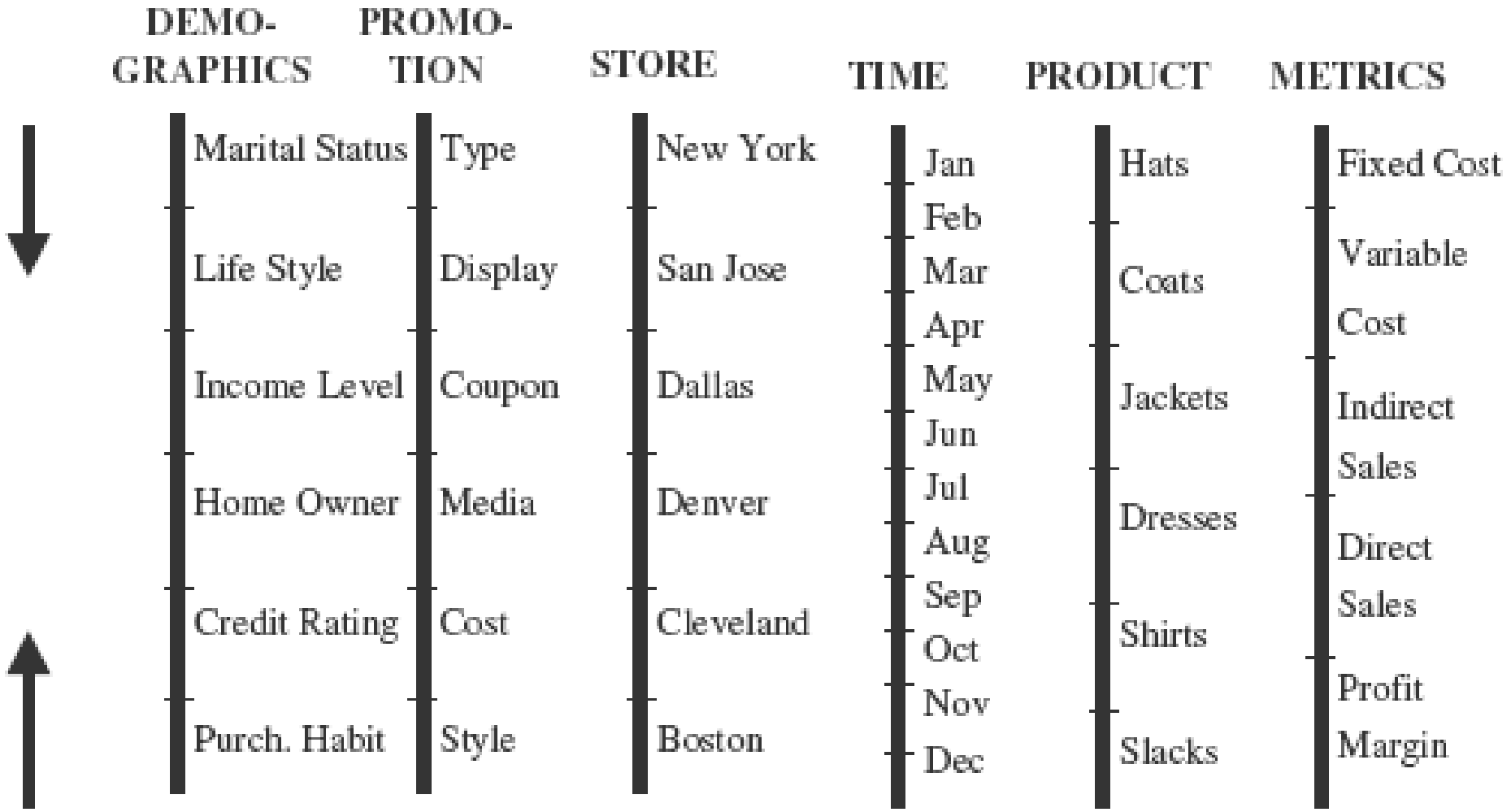
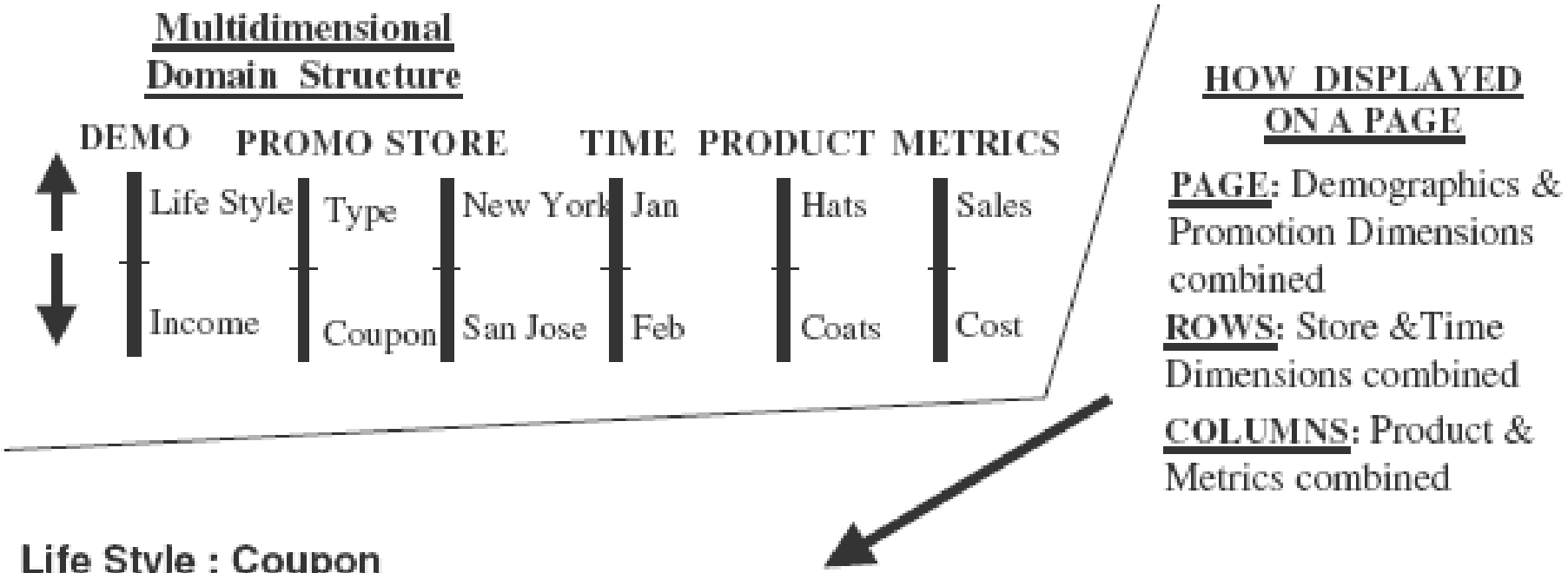


Figure 15-10 Six-dimensional MDS.



Life Style : Coupon

		Hats	Hats	Coats	Coats
		Sales	Cost	Sales	Cost
New York	Jan	220	170	270	220
	Feb	190	140	230	180
Boston	Jan	200	160	240	200
	Feb	180	130	220	170

Figure 15-11 Page displays for six-dimensional data.

Drill down: It refers to the process of viewing data at a level of increased detail

Roll up: It refers to the process of viewing data with decreasing detail

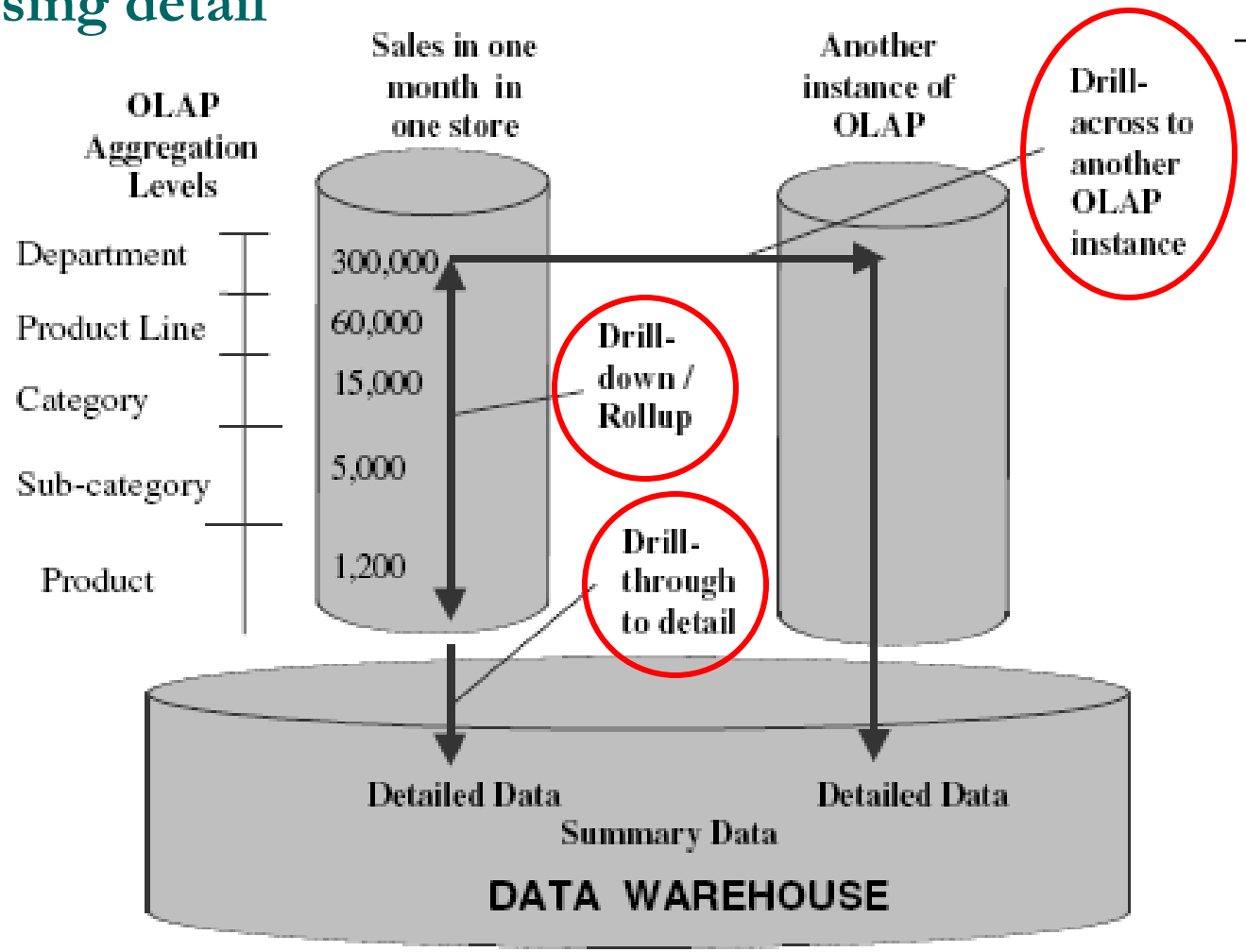


Figure 15-12 Roll-up and drill-down features of OLAP.

Example of roll-up

Store: New York

PAGES: STORE dimension

Products

COLUMNS: PRODUCT dimension

ROWS: TIME dimension

Months

	Hats	Coats	Jackets	Dresses	Shirts	Slacks
Jan	200	550	350	500	520	490
Feb	210	480	390	510	530	500
Mar	190	480	380	480	500	470
Apr	190	430	350	490	510	480
May	160	530	320	530	550	520
Jun	150	450	310	540	560	330
Jul	130	480	270	550	570	250
Aug	140	570	250	650	670	230
Sep	160	470	240	630	650	210
Oct	170	480	260	610	630	250
Nov	180	520	280	680	700	260
Dec	200	560	320	750	770	310

Figure 15-6 A Three-dimensional display.

Store: New York

Sub-categories

PAGES: STORE dimension

COLUMNS: PRODUCT dimension

ROWS: TIME dimension

Months

	Outer	Dress	Casual
Jan	1,100	1,020	490
Feb	1,080	1,040	500
Mar	1,050	980	470
Apr	970	1,000	480
May	1,010	1,080	520
Jun	910	1,100	330
Jul	880	1,120	250
Aug	960	1,320	230
Sep	870	1,280	210
Oct	910	1,240	250
Nov	980	1,380	260
Dec	1,080	1,520	310

Figure 15-13 Three-dimensional display with roll-up.

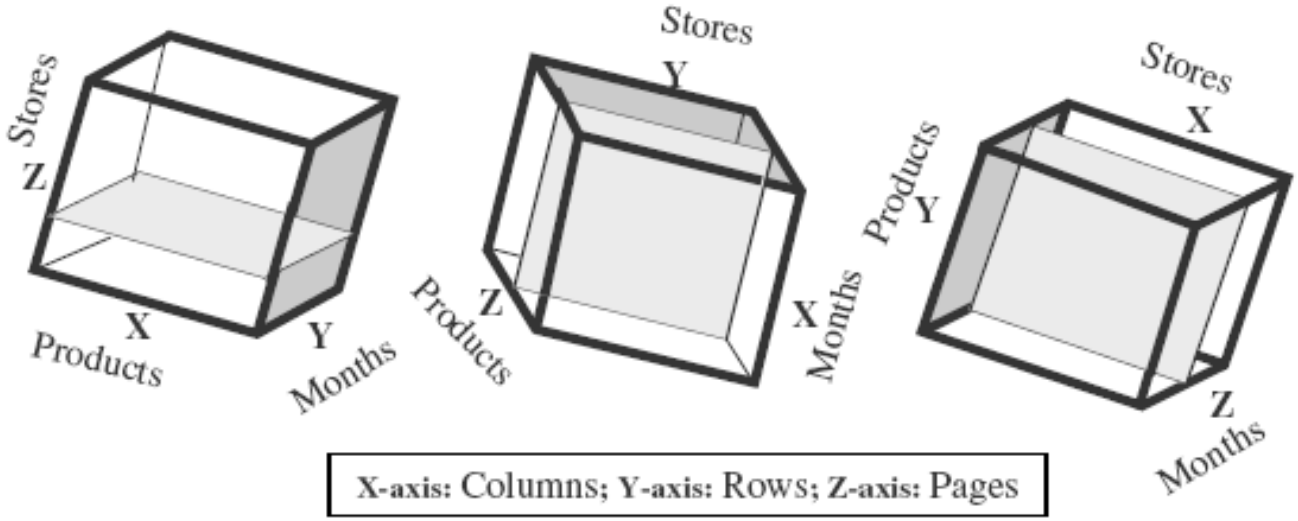
OLAP Operations

- Roll up (drill-up): summarize data
  - *by climbing up hierarchy or by dimension reduction*
- Drill down (roll down): reverse of roll-up
  - *from higher level summary to lower level summary or detailed data, or introducing new dimensions*
- Slice and dice:
  - *project and select*
- Pivot (rotate):
  - *reorient the cube, visualization, 3D to series of 2D planes.*
- Other operations
  - *drill across: involving (across) more than one fact table*
  - *drill through: through the bottom level of the cube to its back-end relational tables (using SQL)*

Rollup & Drill-down

- OLAP permit users to view data at ay desired level of granularity.
- Rollup: moving from finer-granularity data to coarser granularity
- Drill-down: opposite to Rollup

Slice n dice: It is an ability to move between different combinations of dimensions when viewing data with an OLAP browser



Slice-and-Dice or Rotation

Store: New York

	Hats	Coats	Jackets
Jan	200	550	350
Feb	210	480	390
Mar	190	480	380

Product: Hats

	Jan	Feb	Mar
New York	200	210	190
Boston	210	250	240
San Jose	130	90	70

Month: January

	New York	Boston	San Jose
Hats	200	210	130
Coats	550	500	200
Jackets	350	400	100

Figure 15-14 Slicing and dicing.

Slicing & Dicing

- Additional Functionality that can be thought of as viewing a slice of the data cube, particularly when values for multiple dimensions are fixed.
- Slicing/Dicing simply consists of selecting specific values for these attributes, which are then displayed on top of the cross-tab

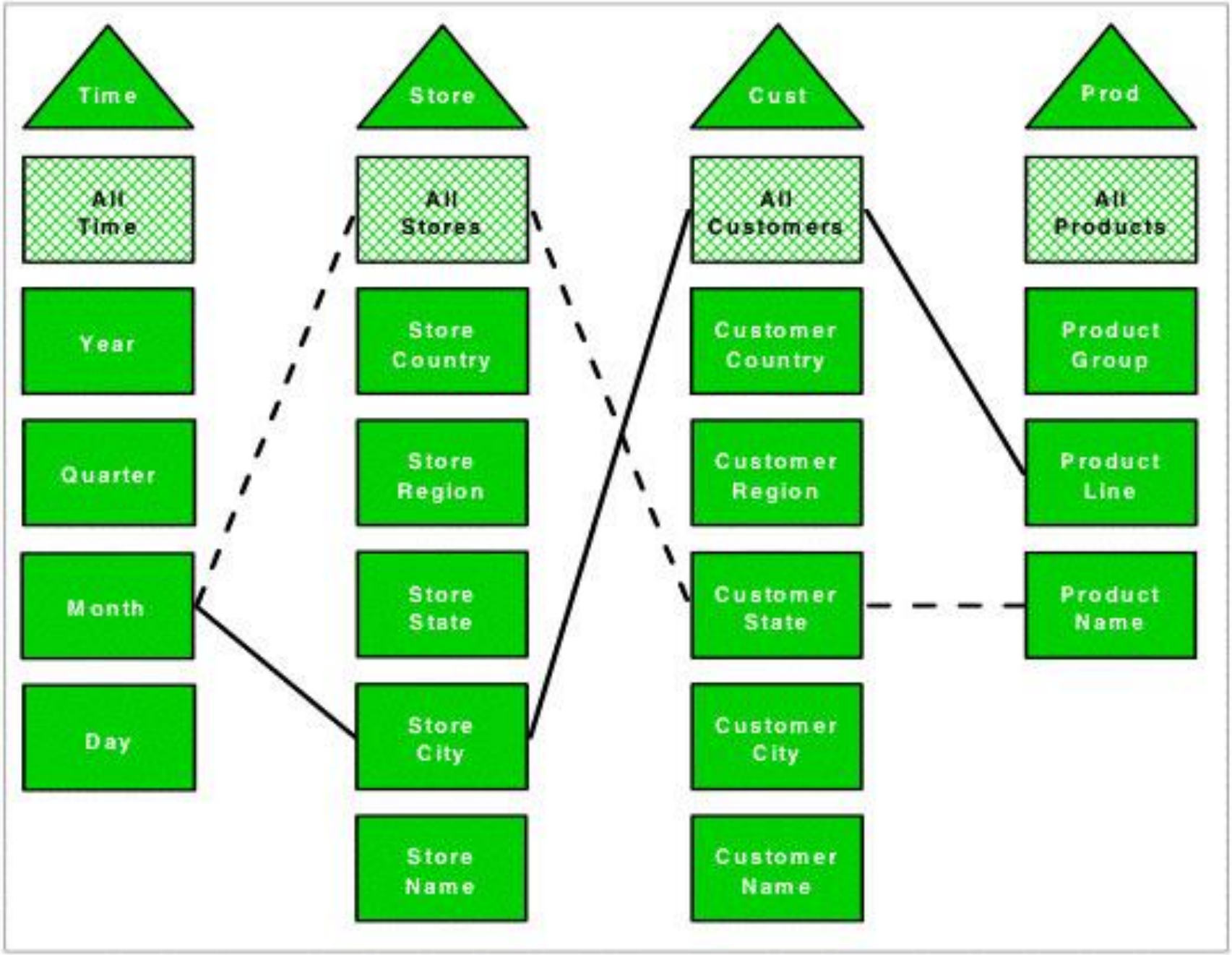
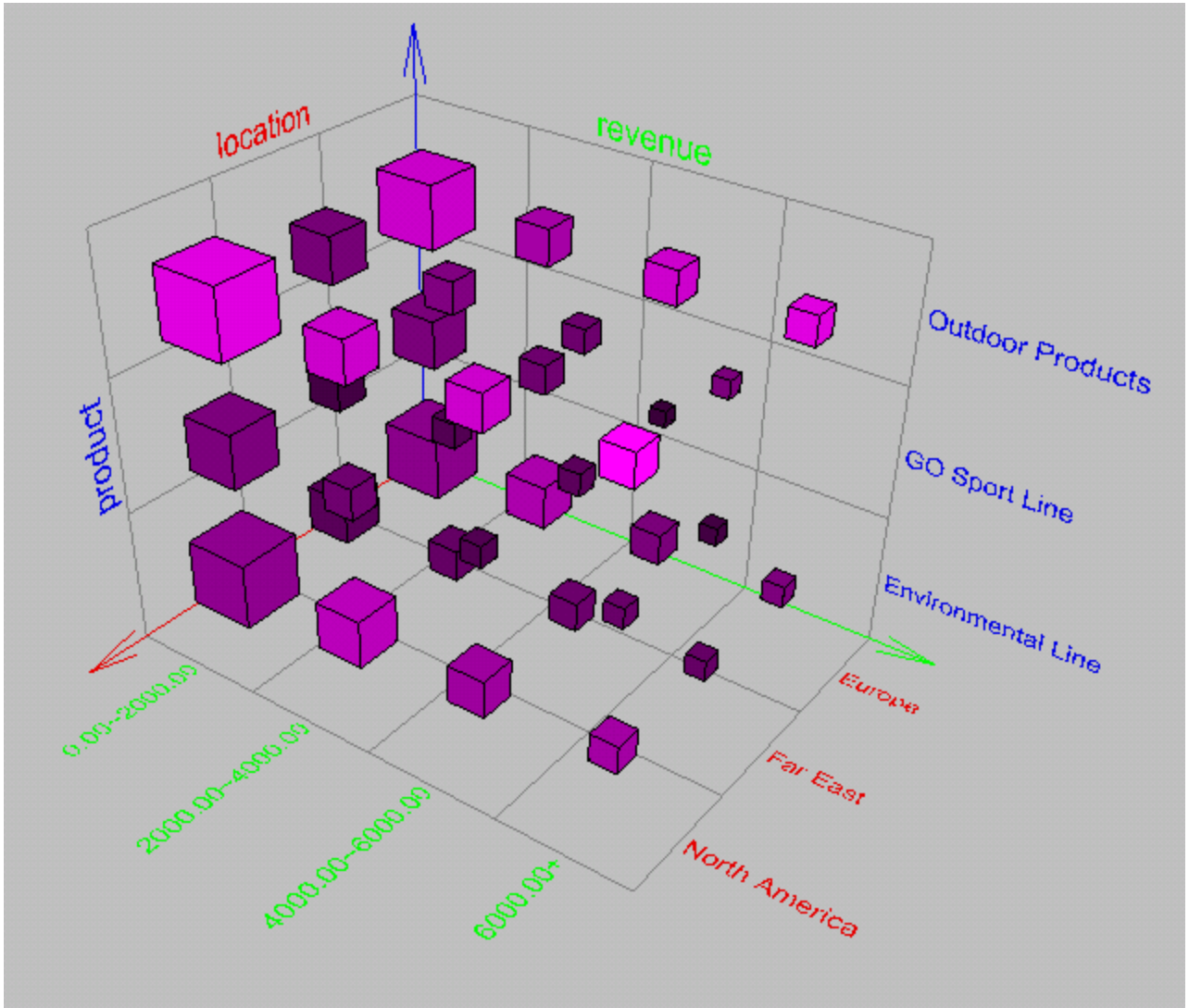


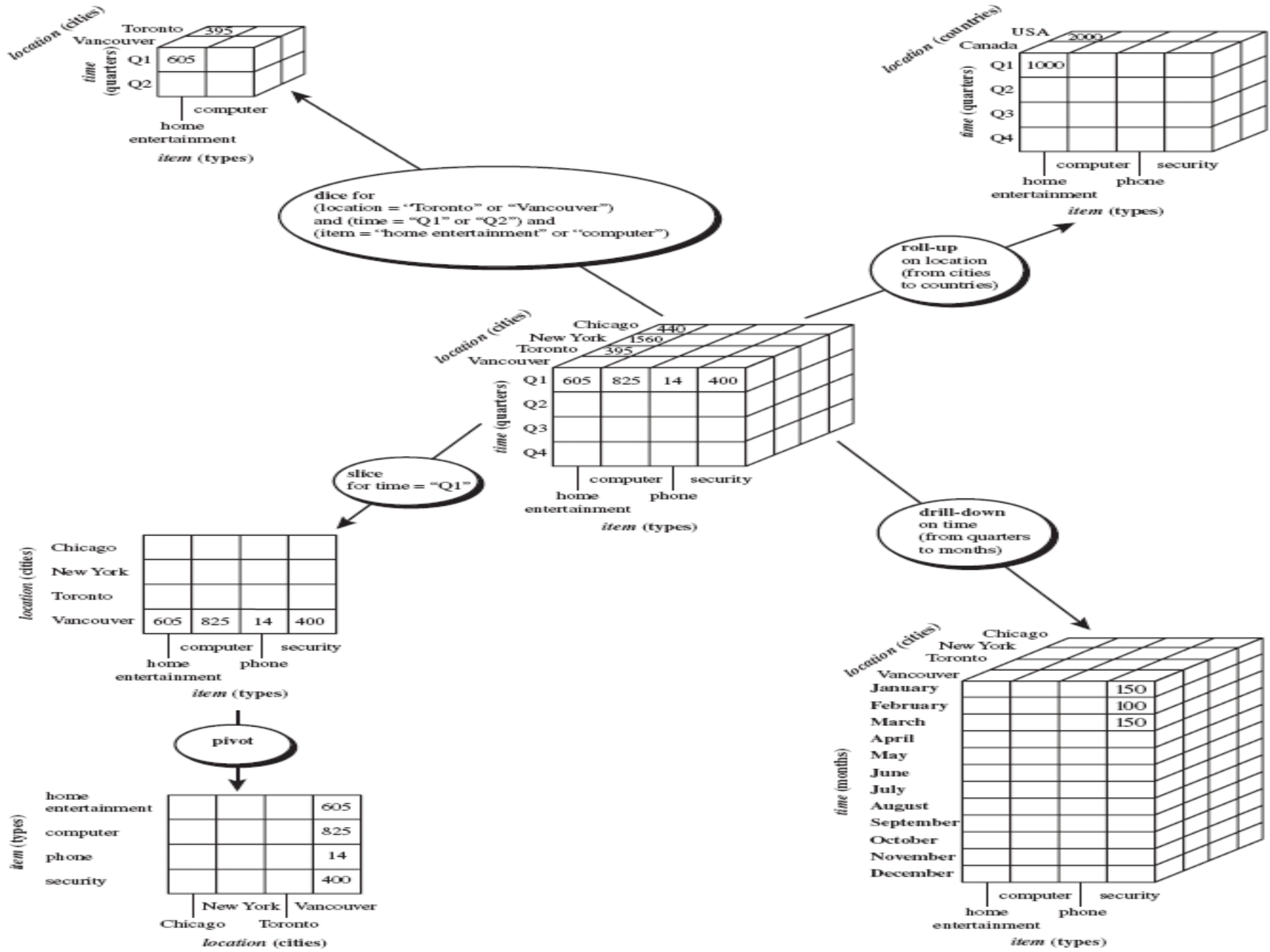
Figure 1-2 Database slices

Browsing a Data Cube



OLAP Operation

- So, how are *concept hierarchies* useful in OLAP?
- In the multidimensional model, data are organized into multiple dimensions,
- And each dimension contains multiple levels of abstraction defined by concept hierarchies


location (cities)
Toronto
Vancouver
395
time (quarters)
Q1
Q2
605
computer
home entertainment
item (types)
location (countries)
USA
Canada
2000
Q1
Q2
Q3
Q4
1000
time (quarters)
computer
security
home entertainment
phone
item (types)
dice for
(location = "Toronto" or "Vancouver")
and (time = "Q1" or "Q2") and
(item = "home entertainment" or "computer")
roll-up
on location
(from cities
to countries)
location (cities)
Chicago
New York
Toronto
Vancouver
440
1560
395
Q1
Q2
Q3
Q4
605
825
14
400
time (quarters)
computer
security
home entertainment
phone
item (types)
slice
for time = "Q1"
drill-down
on time
(from quarters
to months)
location (cities)
Chicago
New York
Toronto
Vancouver
605
825
14
400
computer
security
home entertainment
phone
item (types)
pivot
item (types)
home entertainment
computer
phone
security
605
825
14
400
New York
Vancouver
Chicago
Toronto
location (cities)
location (cities)
Chicago
New York
Toronto
Vancouver
January
February
March
April
May
June
July
August
September
October
November
December
150
100
150
time (months)
computer
security
home entertainment
phone
item (types)

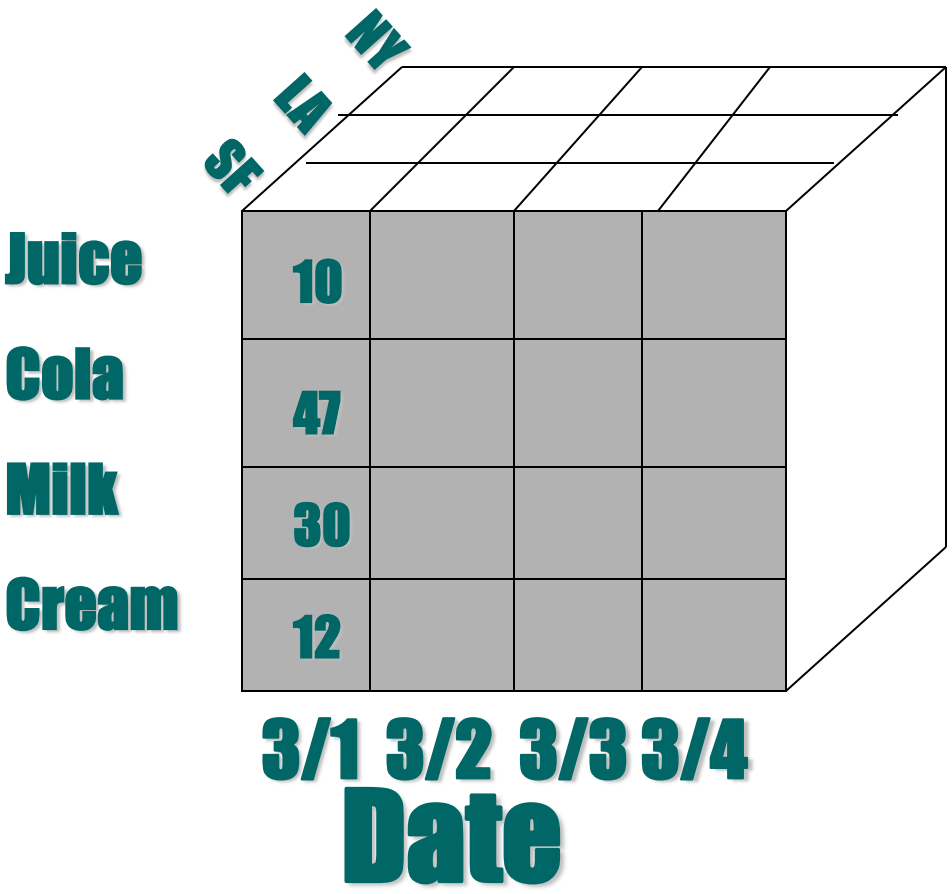
Typical OLAP Operations

- Roll up (drill-up): summarize data
  - *by climbing up hierarchy or by dimension reduction*
- Drill down (roll down): reverse of roll-up
  - *from higher level summary to lower level summary or detailed data, or introducing new dimensions*

Typical OLAP Operations

- Slice and dice: *project and select*
- Pivot (rotate):
  - *reorient the cube, visualization, 3D to series of 2D planes*

Multidimensional Data

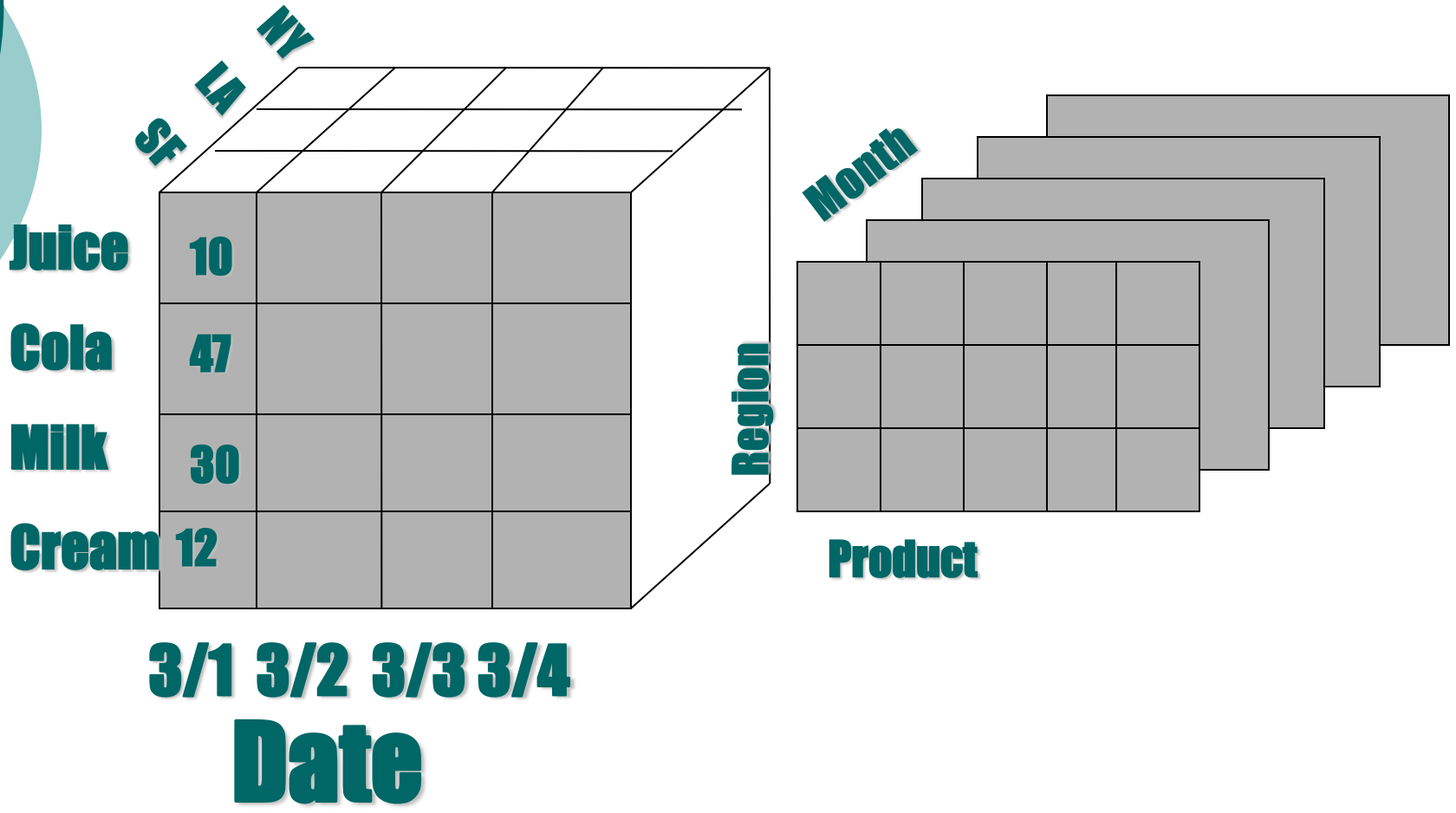


Sales Volume as a function of time, city and product

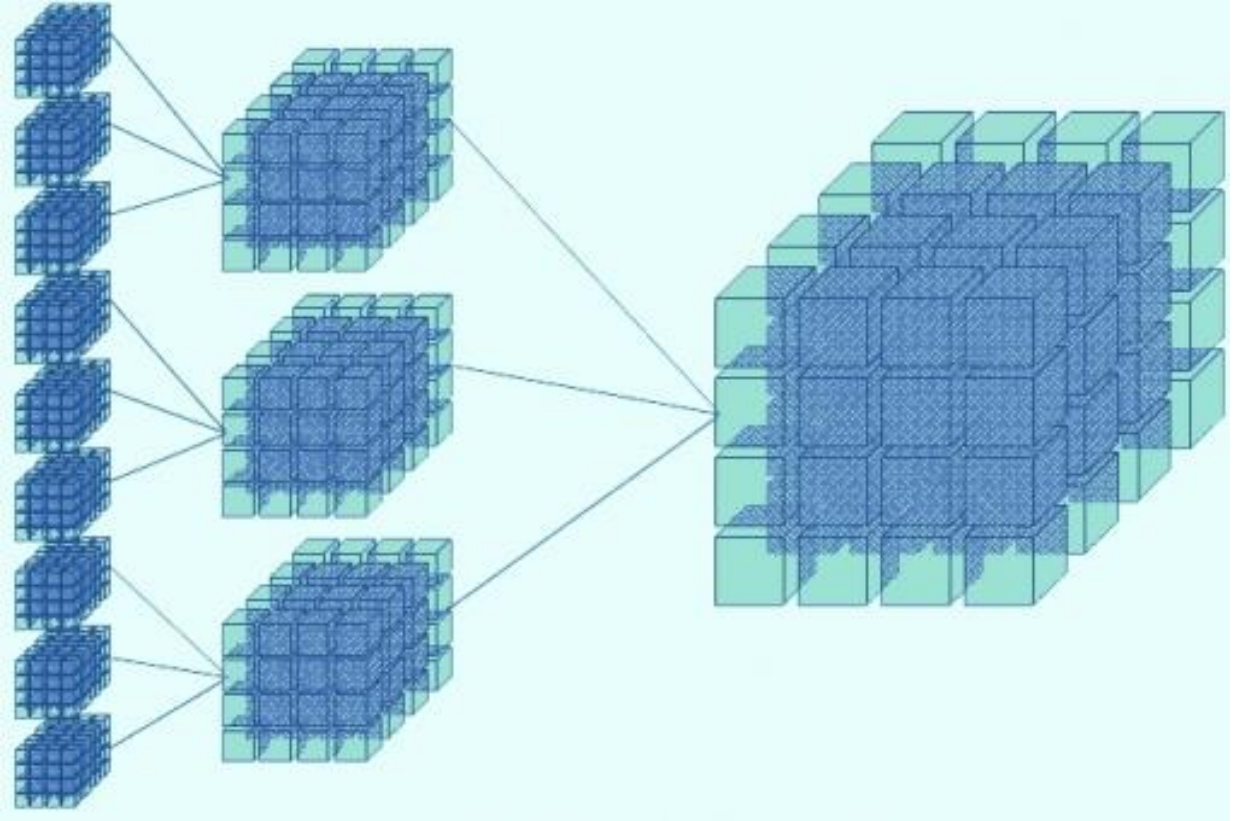
Operations in Multidimensional Data Model

- Aggregation (*roll-up*)
  - dimension reduction: e.g., total sales by city
  - summarization over aggregate hierarchy: e.g., total sales by city and year -> total sales by region and by year
- Selection (*slice*) defines a subcube
  - e.g., sales where city = Palo Alto and date = 1/15/96
- Navigation to detailed data (*drill-down*)
  - e.g., (sales - expense) by city, top 3% of cities by average income
- Visualization Operations (e.g., Pivot)

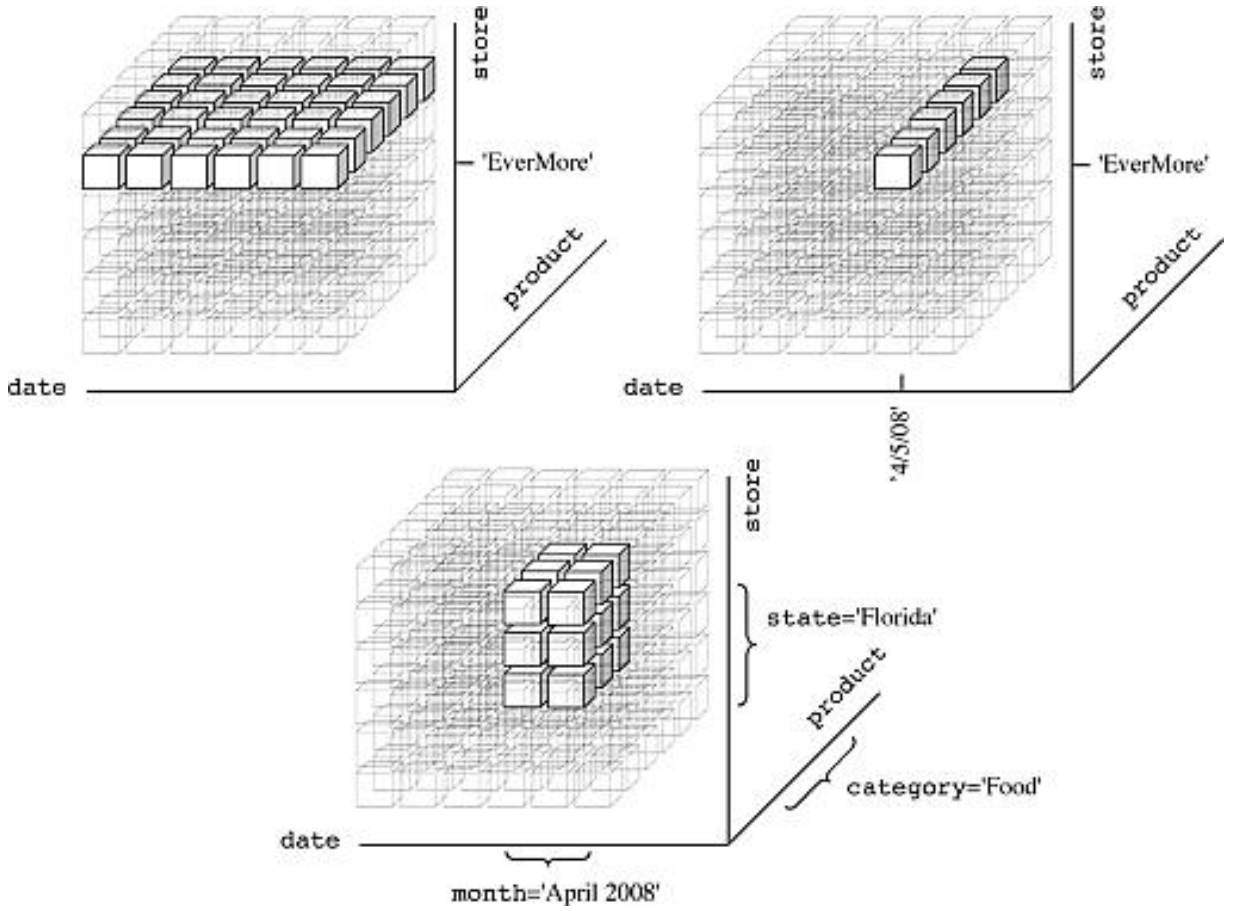
A Visual Operation: Pivot (Rotate)



***Roll-up*:** (Aggregate, Consolidate) A roll-up involves computing all of the data relationships for one or more dimensions. To do this, a computational relationship or formula might be defined.

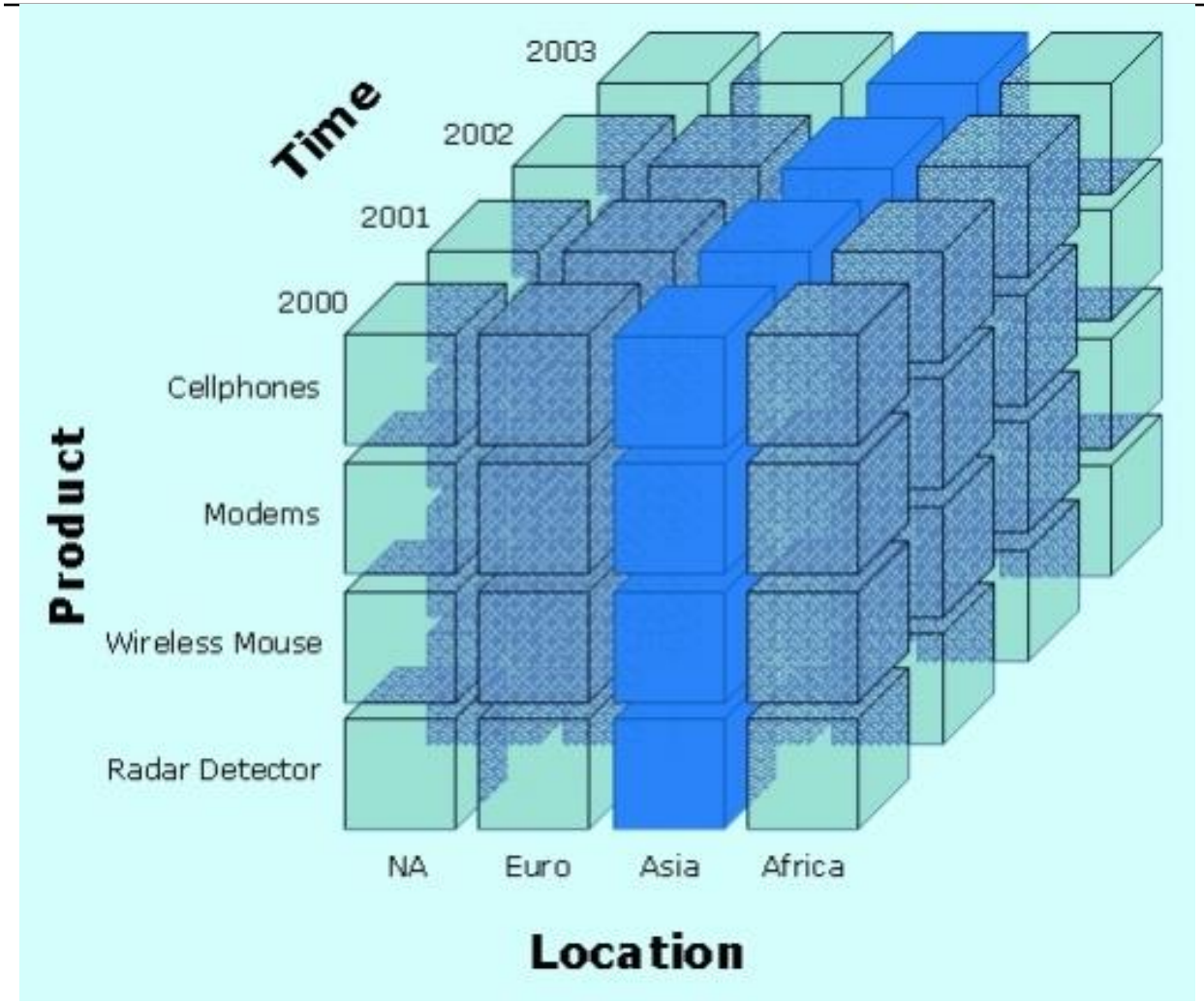


***Drill Down*:** Drilling down is a specific analytical technique whereby the user navigates among levels of data ranging from the most summarized (up) to the most detailed (down).



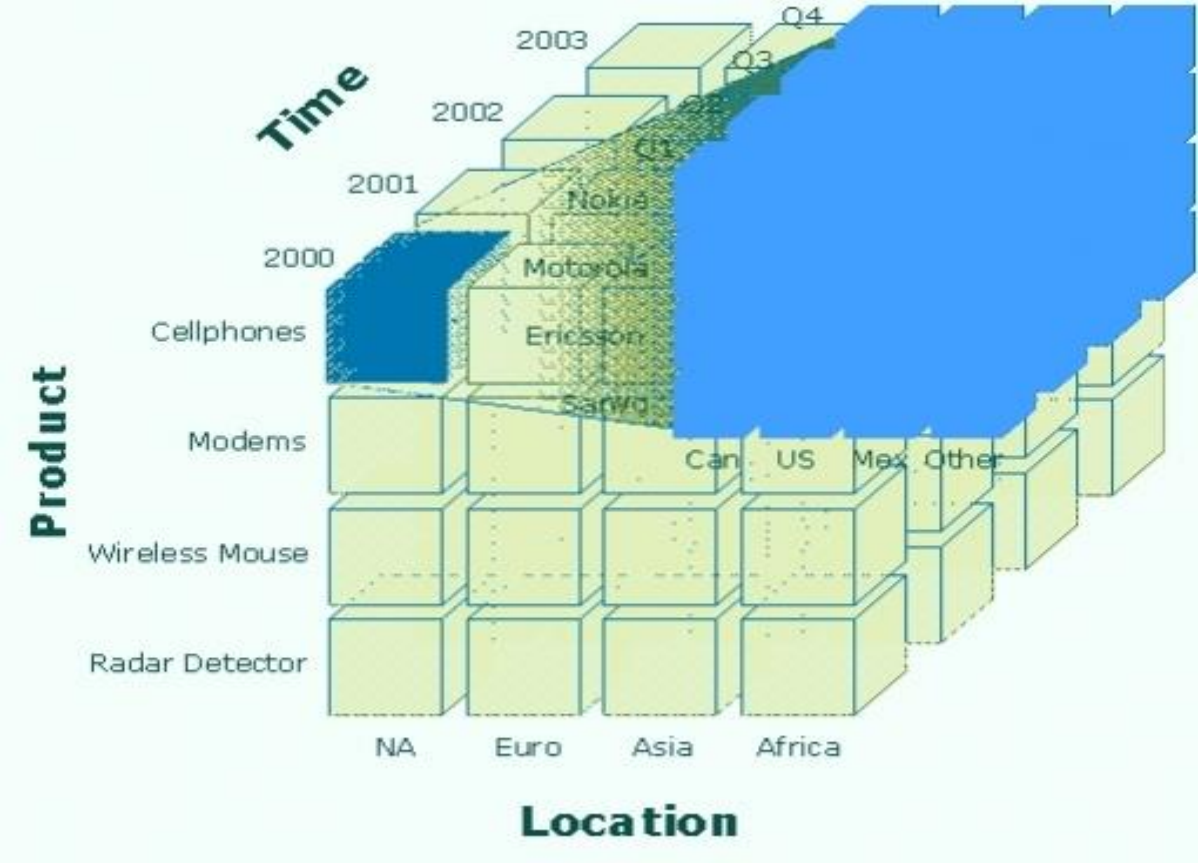
Slice:

A slice is a subset of a multi-dimensional array corresponding to a single value for one or more members of the dimensions not in the subset.

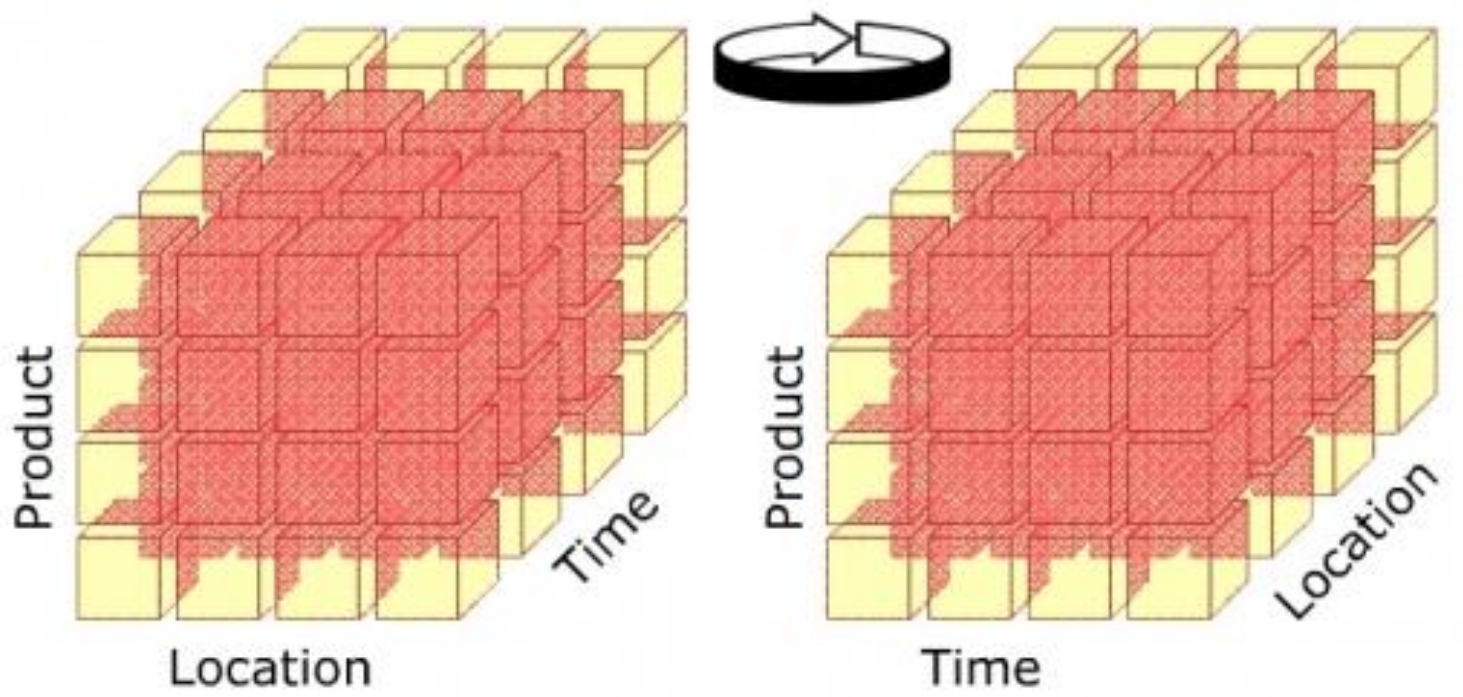


Dice:

The dice operation is a slice on more than two dimensions of a data cube (or more than two consecutive slices).



Pivot:This operation is also called rotate operation. It rotates the data in order to provide an alternative presentation of data – the report or page display takes a different dimensional orientation.



Uses and Benefits

- Increased productivity of business managers, analysts, and executives
- Faster delivery of applications by IT people
- Self sufficiency of users, resulting in reduction in backlog
- More efficient operation through reducing time on query executions and in net work traffic
- Ability to model real world challenges with business metrics and dimensions