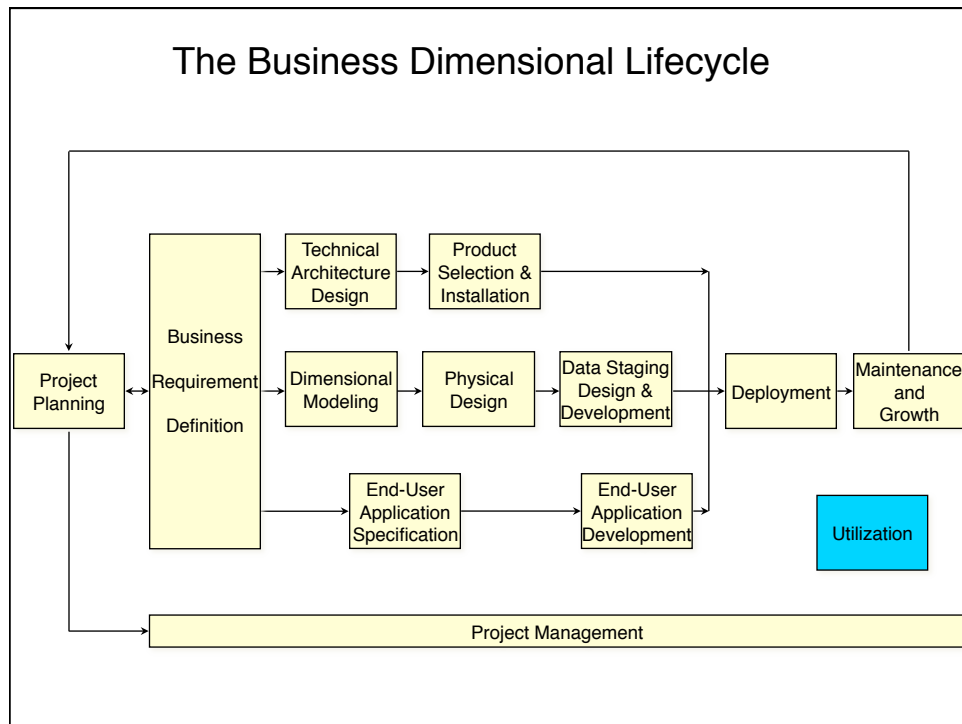# CSPP 53017: Data Warehousing
Winter 2013

Lecture 7
Svetlozar Nestorov

Class News

- Make-up class on Saturday, Mar 9 in Gleacher 203 10:30am – 1:30pm.
- Last 15 minute in-class quiz (6:30pm) on Mar 5.
  - Covers the first seven lectures, project submissions and the Gradiance homework.
 - Open book/notes

The Business Dimensional Lifecycle

OLAP

- **Online transaction processing OLTP -** updating (i.e. inserting, modifying and deleting) querying and presenting data from databases for operational purposes

- **Online analytical processing (OLAP) -** querying and presenting data from data warehouses and/or data marts for analytical purposes

OLAP and Business Intelligence

- Designed for analysis of dimensionally modeled data
- Regardless of which data warehousing approach is chosen, the data that is accessible by the end user is typically structured as a dimensional model
  - OLAP/BI tools can be used on analytical data stores created with different modeling approaches


OLAP/BI tools as an interface to data warehouses modeled using different approaches
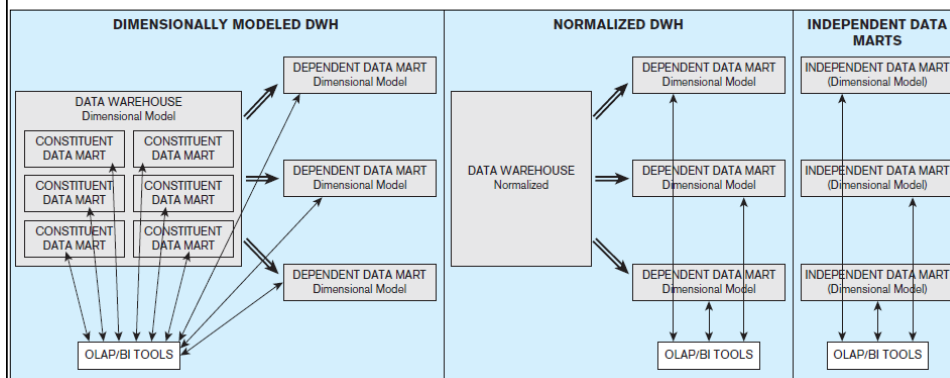
OLAP/BI Tools

- Allow users to query fact and dimension tables by using simple point-and-click query-building applications
- Based on the point-and-click actions by the user of the OLAP/BI tool, the tool writes and executes the code in the language of the data management system (e.g. SQL) that hosts the data warehouse or data mart that is being queried
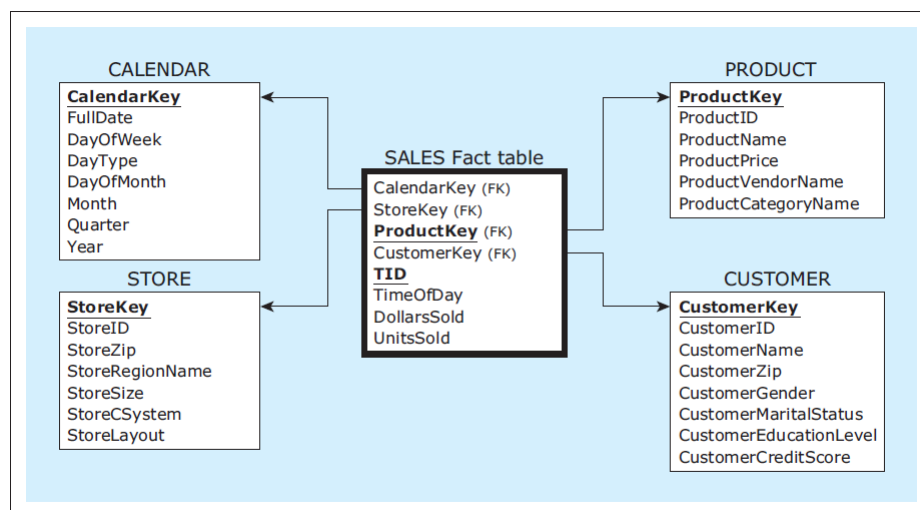
A typical OLAP/BI tool query construction space
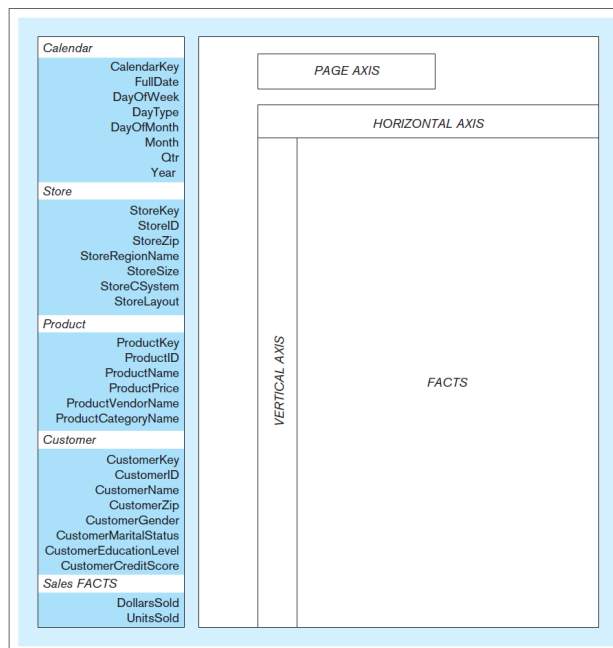
OLAP Query 1

For each individual store, show separately for male and female shoppers the number of product units sold for each product category

Example Star Schema

```
CALENDAR                                      PRODUCT
CalendarKey                                   ProductKey
FullDate                                      ProductID
DayOfWeek                                     ProductName
DayType          SALES Fact table            ProductPrice
DayOfMonth                                    ProductVendorName
Month            CalendarKey (FK)             ProductCategoryName
Quarter          StoreKey (FK)
Year             ProductKey (FK)
                 CustomerKey (FK)
STORE            TID                          CUSTOMER
StoreKey         TimeOfDay                    CustomerKey
StoreID          DollarsSold                  CustomerID
StoreZip         UnitsSold                    CustomerName
StoreRegionName                               CustomerZip
StoreSize                                     CustomerGender
StoreCSystem                                  CustomerMaritalStatus
StoreLayout                                   CustomerEducationLevel
                                              CustomerCreditScore
```

Dimension Table Example Data

CALENDAR Dimension

Calendar Key	FullDate	DayOfWeek	DayType	DayOfMonth	Month	Qtr	Year
1	1/1/2013	Tuesday	Weekend/Holiday	1	January	Q1	2013
2	1/2/2013	Wednesday	Workday	2	January	Q1	2013
3	1/3/2013	Thursday	Workday	3	January	Q1	2013
4	1/4/2013	Friday	Workday	4	January	Q1	2013

PRODUCT Dimension

ProductKey	ProductID	ProductName	ProductPrice	ProductVendor Name	ProductCategory Name
1	1X1	Zzz Bag	$100	Pacifica Gear	Camping
2	2X2	Easy Boot	$70	Mountain King	Footwear
3	3X3	Cosy Sock	$15	Mountain King	Footwear
4	4X4	Dura Boot	$90	Pacifica Gear	Footwear
5	5X5	Tiny Tent	$150	Mountain King	Camping
6	6X6	Biggy Tent	$250	Mountain King	Camping

STORE Dimension

StoreKey	StoreID	StoreZip	StoreRegion Name	Store Size (m²)	Store CSystem	Store Layout
1	S1	60600	Chicagoland	51000	Cashiers	Modern
2	S2	60605	Chicagoland	35000	Self Service	Traditional
3	S3	35400	Tristate	55000	Mixed	Traditional

CUSTOMER Dimension

Customer Key	CustomerID	Customer Name	Customer Zip	Customer Gender	Customer Marital Status	Customer Education Level	Customer Credit Score
1	1-2-333	Tina	60137	Female	Single	College	700
2	2-3-444	Tony	60611	Male	Single	High School	650
3	3-4-555	Pam	35401	Female	Married	College	623

Fact Table Example Data

SALES Fact table

CalendarKey	StoreKey	ProductKey	Customer Key	TID	TimeOfDay	Dollars Sold	UnitsSold
1	1	1	1	T111	8:23:59 AM	$100	1
1	2	2	2	T222	8:24:30 AM	$70	1
2	3	3	1	T333	8:15:08 AM	$75	5
2	3	1	1	T333	8:15:08 AM	$100	1
2	3	4	3	T444	8:20:33 AM	$90	1
2	3	2	3	T444	8:20:33 AM	$140	2
2	3	4	2	T555	8:30:00 AM	$360	4
2	3	5	2	T555	8:30:00 AM	$300	2
2	3	6	2	T555	8:30:00 AM	$250	1
3	1	1	2	T666	8:00:00 AM	$100	1
3	2	2	3	T777	8:10:00 AM	$70	1
3	2	3	3	T777	8:10:00 AM	$30	2
4	3	1	1	T888	8:05:00 AM	$200	2
4	2	1	2	T999	9:07:33 AM	$300	3
4	2	3	2	T999	9:07:33 AM	$60	4
4	2	4	2	T999	9:07:33 AM	$180	2

OLAP Query 1 - OLAP/BI tool query construction actions
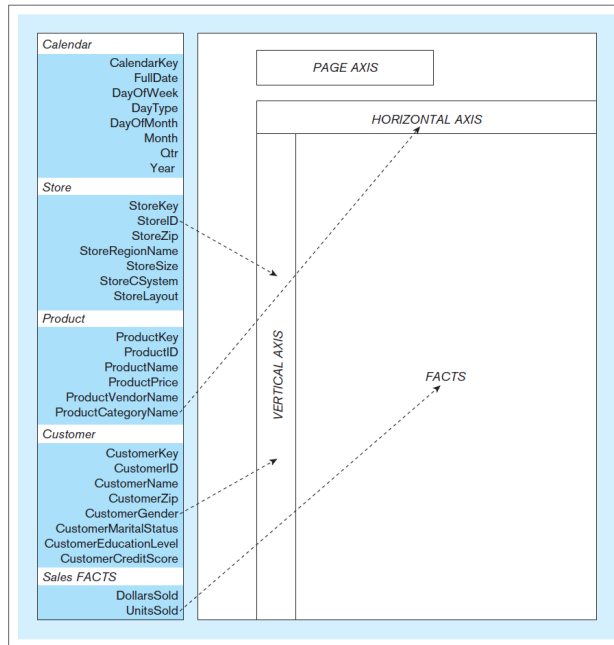
OLAP Query 1: Result

	Sales-Units Sold	
	Camping	Footwear
Store 1		
Female	1	0
Male	1	0
Store 2		
Female	0	3
Male	3	7
Store 3		
Female	3	8
Male	3	4

Basic OLAP Tool Features

- Slice and Dice
- Pivot (Rotate)
- Drill Down / Drill Up

Slice and Dice

- Adds, replaces, or eliminates specified dimension attributes (or particular values of the dimension attributes) from the already displayed result

OLAP Query 1: Result starting point

		Sales-Units Sold	
		Camping	Footwear
Store 1			
	Female	1	0
	Male	1	0
Store 2			
	Female	0	3
	Male	3	7
Store 3			
	Female	3	8
	Male	3	4

OLAP Query 2

~~For each individual store~~, show separately for male and female shoppers the number of product units sold for each product category

For stores 1 and 2, show separately for male and female shoppers the number of product units sold for each product category

OLAP Query 2: Result

| | | Sales-Units Sold | |
		Camping	Footwear
Store 1			
	Female	1	0
	Male	1	0
Store 2			
	Female	0	3
	Male	3	7

OLAP Query 3

For each individual store, show separately for male and female shoppers the number of product units sold ~~for each product category~~

For each individual store, show separately for male and female shoppers the number of product units sold on workdays and on weekends/holidays

OLAP Query 3: Result

	Sales-Units Sold	
	Weekend/ Holiday	Workday
Store 1		
Female	1	0
Male	0	1
Store 2		
Female	0	3
Male	1	9
Store 3		
Female	0	11
Male	0	7

Pivot

- Reorganizes the values displayed in the original query result by moving values of a dimension column from one axis to another

OLAP Query 4

~~For each individual store, show separately for male and female shoppers the number of product units sold for each product category~~
???

OLAP Query 4: Result

		Sales-Units Sold	
		Female	Male
Store 1			
	Camping	1	1
	Footwear	0	0
Store 2			
	Camping	0	3
	Footwear	3	7
Store 3			
	Camping	3	3
	Footwear	8	4

Drill Down and Drill Up

- Drill Down
  - Makes the granularity of the data in the query result *finer*
- Drill Up
  - Makes the granularity of the data in the query result *courser*

Drill Hierarchy

- Set of attributes within a dimension where an attribute is related to one or more attributes at a lower level but only related to one item at a higher level
  - For example:

  StoreRegionName → StoreZip → StoreID
  ProductCategoryName → ProductName → ProductID

- Used for drill down/drill up operations

OLAP Query 5

For each individual store, show separately for male and female shoppers the number of individual product units sold

OLAP Query 5: Result

| | | | Sales-Units Sold | | | | |
| | | | Camping | | | | Footwear | |
		Biggy Tent	Tiny Tent	Zzz Bag	Cosy Sock	Dura Boot	Easy Boot
Store 1							
	Female	0	0	1	0	0	0
	Male	0	0	1	0	0	0
Store 2							
	Female	0	0	0	2	0	1
	Male	0	0	3	4	2	1
Store 3							
	Female	0	0	3	5	1	2
	Male	1	2	0	0	4	0

2/26/13

OLAP Query 5: Result Drill Up

| | *Sales-Units Sold* | |
	Camping	Footwear
Store 1		
Female	1	0
Male	1	0
Store 2		
Female	0	3
Male	3	7
Store 3		
Female	3	8
Male	3	4

OLAP/BI Tools

- Require dimensional organization of underlying data for performing basic OLAP operations (slice, pivot, drill)
- Additional OLAP/BI Tool functionalities:
  - Graphically visualizing the answers
  - Creating and examining calculated data
  - Determining comparative or relative differences
  - Performing exception analysis, trend analysis, forecasting, and regression analysis
  - Number of other analytical functions
- Many OLAP/BI tools are web-based

15

OLAP Query 1: Result

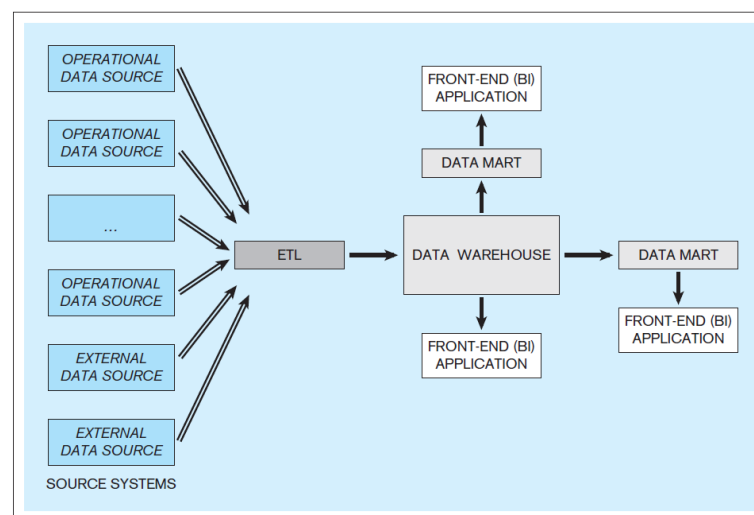|		Camping	Footwear
	Sales-Units Sold		
Store 1			
	Female	1	0
	Male	1	0
Store 2			
	Female	0	3
	Male	3	7
Store 3			
	Female	3	8
	Male	3	4

Result Visualized As Chart
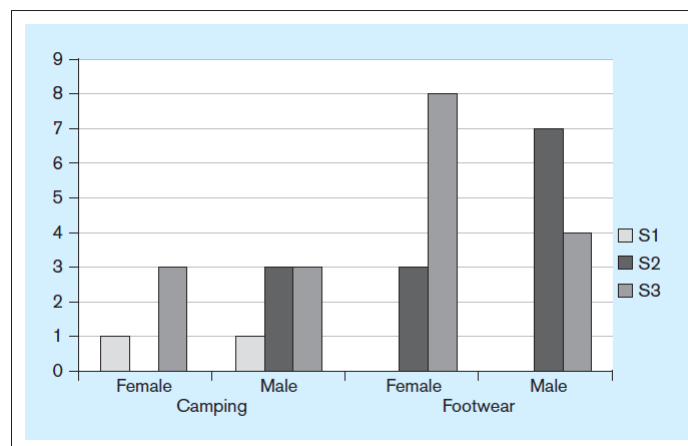
OLAP/BI Tools Usage

- Two purposes:
  - Ad-hoc direct analysis of dimensionally modeled data
  - Creation of front-end (BI) applications
- Data Warehouse/Data Mart Front-End (BI) Applications
  - Provides access to the data warehouse for indirect use

Chapter 9 – Slide

DW system with front-end applications

Front-End Interface Example

WELCOME TO THE ANALYTICAL APPLICATIONS CENTER
(Click on the application you wish to use)

Sales Comparisons Analyzer

Cost Analysis Center

Supply Chain Analytics

Marketing Analysis Reports

Front-End Interface Example 2

SALES COMPARISONS ANALYZER
(Choose from the Options Below)

Quarterly Store Sales Comparisons	Monthly Store Sales Comparisons	Weekly Store Sales Comparisons
Quarterly Product Sales Comparisons	Monthly Product Sales Comparisons	Weekly Product Sales Comparisons
Quarterly Customer Sales Comparisons	Monthly Customer Sales Comparisons	Weekly Customer Sales Comparisons

Template Query Interface

QUARTERLY STORE SALES COMPARISONS
Choose Your Parameters

Choose Year 1: ▼ Choose Year 2: ▼

Choose Quarter in Year 1: ▼ Choose Quarter in Year 2: ▼

Choose Stores (up to 5):

▼ ▼ ▼ ▼ ▼ RUN

Sample Results

QUARTERLY STORE SALES COMPARISONS
Result

	2012 Q4	2013 Q3
Store 1	$3,513,678	$3,713,341
Store 3	$5,453,889	$5,613,111
Store 7	$4,456,109	$4,151,901
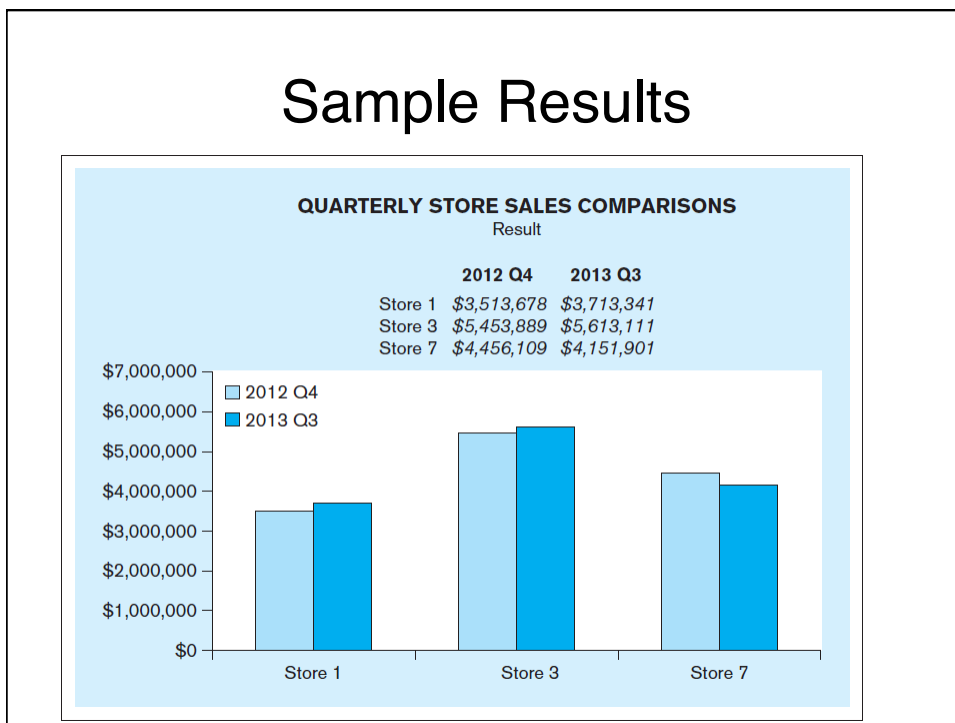
Executive Dashboard

- Intended for use by higher level decision makers within an organization
- Contains an organized easy-to-read display of a number of critically important queries describing the performance of the organization
- In general, the usage of executive dashboards should require little or no effort or training
- Executive dashboards can be web-based

Example – Executive Dashboard
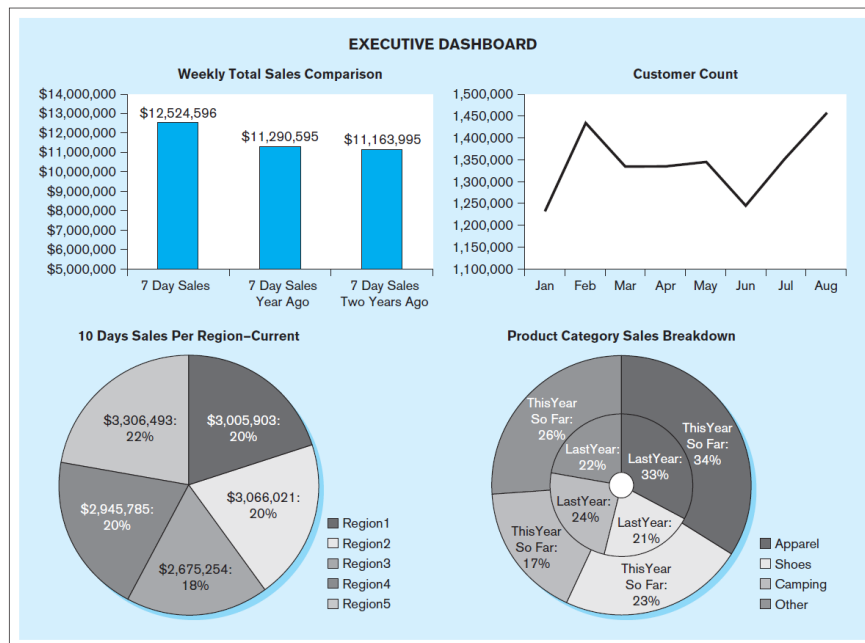
DW Deployment

- Releasing the created and populated data warehouse and its front-end (BI) applications for use by the end-users
- Alpha release
  - Internal deployment of a system to the members of the development team for initial testing of its functionalities
- Beta release
  - Deployment of a system to a selected group of users to test the usability of the system
- Production release
  - The actual deployment of a functioning system

OLAP/BI Tools Database Models

- OLAP/BI tools are designed for access of dimensionally modeled data
- Dimensionally modeled data stores can be implemented as
  - *Relational database model* (database is a collection of tables)
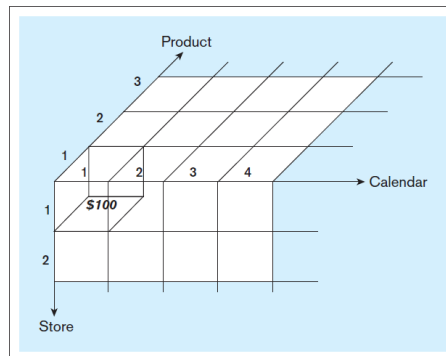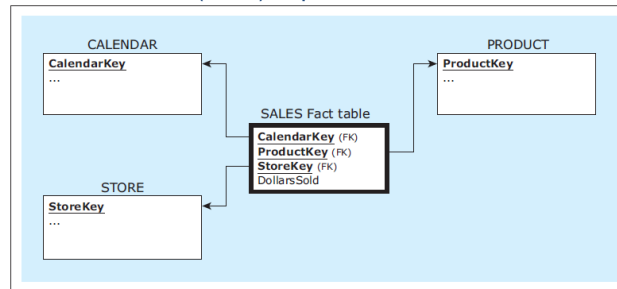 - *Multidimensional database model* (database is a collection of cubes)

Example – Relational implementation of a fact table in a dimensional model
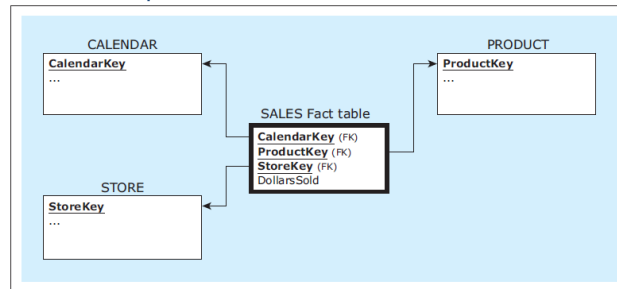
SALES FACT TABLE

| CalendarKey	ProductKey	StoreKey	DollarsSold
1	1	1	$100
1	2	1	$70
1	3	1	$15
.	.	.	.

Example – Multidimensional (cube) implementation of a fact table in a dimensional model
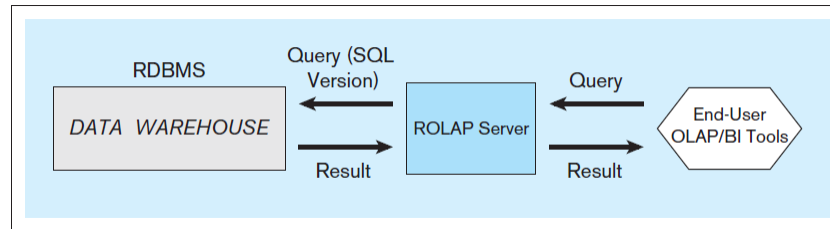
2/26/13

OLAP/BI Tools Data Architecture Options

- Three common categories
  - Multidimensional online analytical processing - **MOLAP**
  - Relational online analytical processing – **ROLAP**
  - Hybrid online analytical processing - **HOLAP**

MOLAP Architecture
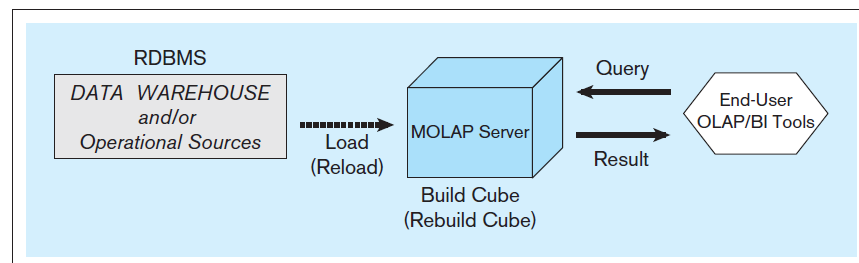
ROLAP Architecture

HOLAP Architecture