Database Principles: Fundamentals of Design, Implementation, and Management

Tenth Edition

Chapter 15

Databases for Decision Support

Objectives

In this chapter, you will learn:

- How business intelligence provides a comprehensive business decision support framework
- About business intelligence architecture, its evolution, and reporting styles
- About the relationship and differences between operational data and decision support data
- What a data warehouse is and how to prepare data for one

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Objectives (cont’d.)

- What star schemas are and how they are constructed
- About data analytics, data mining, and predictive analytics
- About online analytical processing (OLAP)
- How SQL extensions are used to support OLAP-type data manipulations

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

The Need for Data Analysis

- Managers track daily transactions to evaluate how the business is performing
- Strategies should be developed to meet organizational goals using operational databases
- Data analysis provides information about short-term tactical evaluations and strategies

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Business Intelligence

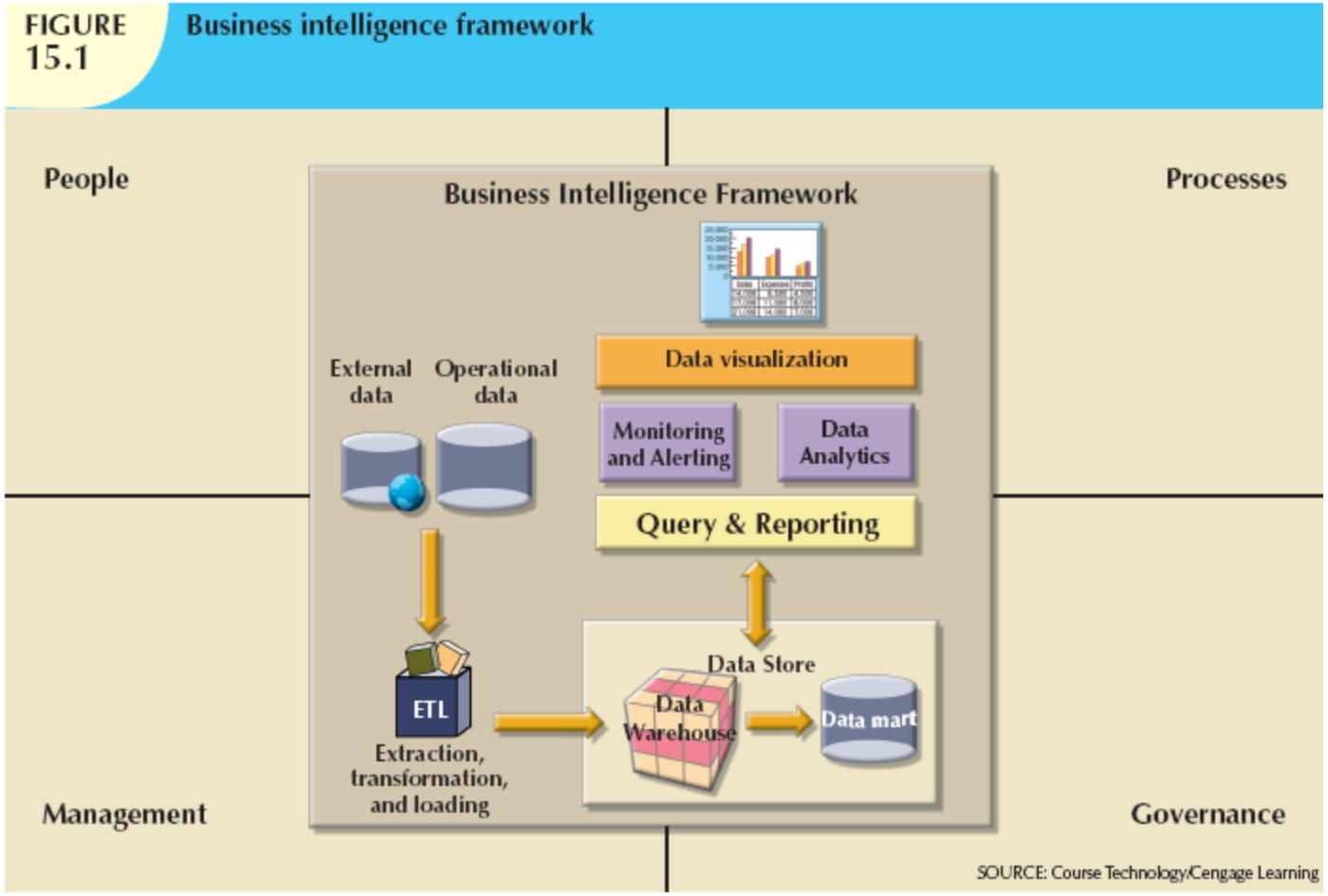
- Comprehensive, cohesive, integrated tools and processes
  - Capture, collect, integrate, store, and analyze data
  - Generate information to support business decision making
- Framework that allows a business to transform:
  - Data into information
  - Information into knowledge
  - Knowledge into wisdom

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Business Intelligence Architecture

- Composed of data, people, processes, technology, and management of components
- Focuses on strategic and tactical use of information
- Key performance indicators (KPI)
  - Measurements that assess company's effectiveness or success in reaching goals
- Multiple tools from different vendors can be integrated into a single BI framework

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.



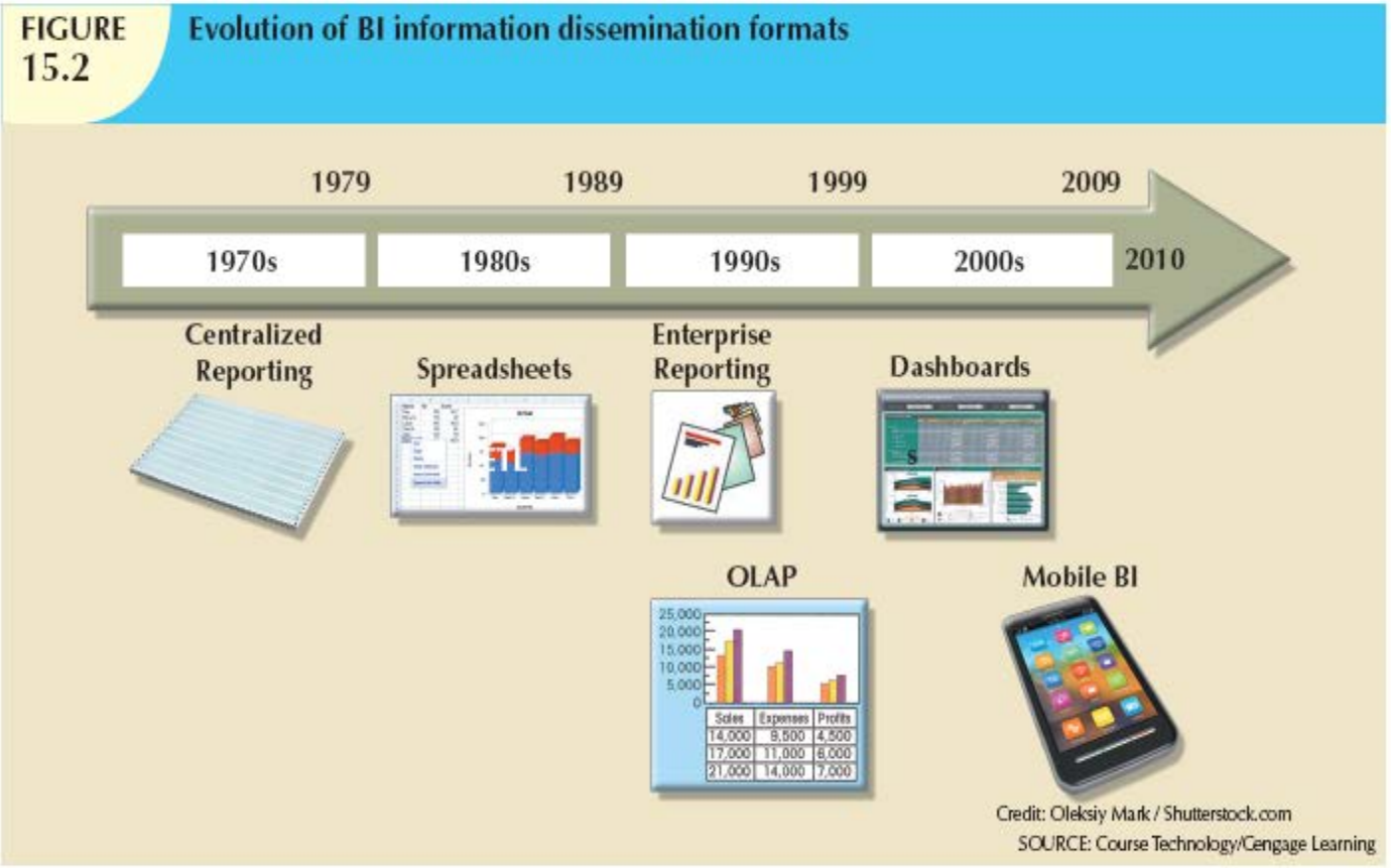
© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Business Intelligence Benefits

- Main goal: improved decision making
- Other benefits
  - Integrating architecture
  - Common user interface for data reporting and analysis
  - Common data repository fosters single version of company data
  - Improved organizational performance

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Business Intelligence Evolution



© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

TABLE 15.4 Business Intelligence Evolution

SYSTEM TYPE	DATA SOURCE	DATA EXTRACTION/ INTEGRATION PROCESS	DATA STORE	END-USER QUERY TOOL	END USER PRESENTATION TOOL
Traditional mainframe-based online transaction processing (OLTP)	Operational data	None Reports read and summarized data directly from operational data	None Temporary files used for report-ing purposes	Very basic Predefined reporting formats Basic sorting, totaling, and averaging	Very basic Menu-driven, predefined reports, text and numbers only
Managerial information system (MIS)	Operational data	Basic extraction and aggregation Read, filter, and summarize operational data into intermediate data store	Lightly aggre-gated data in RDBMS	Same as above, in addition to some ad hoc reporting using SQL	Same as above, in addition to some ad hoc columnar report definitions
First-generation departmental decision support system (DSS)	Operational data External data	Data extraction and integration process popu-lates DSS data store Run periodically	First DSS data-base generation Usually RDBMS	Query tool with some analyti-cal capabilities and predefined reports	Spreadsheet style Advanced pre-sentation tools with plotting and graphics capabilities
First-generation BI	Operational data External data	Advanced data extraction and integration Access diverse data sources, filters, aggrega-tions, classifica-tions, scheduling, and conflict resolution	Data warehouse RDBMS technology Optimized for query purposes Star schema model	Same as above	Same as above, in addition to multidimen-sional presenta-tion tools with drill-down capabilities
Second-generation BI Online analyti-cal processing (OLAP)	Same as above	Same as above	Data warehouse stores data in MDBMS Cubes with multiple dimensions	Adds support for end-user-based data analytics	Same as above, but uses cubes and multidi-mensional matrixes; lim-ited by terms of cube size Dashboards Scorecards Portals
Third-generation Mobile BI and cloud-based	Same as above	Same as above Cloud-based	Same as above Cloud based	Advanced analytics Limited ad hoc interactions	Mobile devices: iPhone, iPad, Blackberry, Android

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Business Intelligence Technology Trends

- Data storage improvements
- Business intelligence appliances
- Business intelligence as a service
- Big Data analytics
- Personal analytics

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Decision Support Data

- BI effectiveness depends on quality of data gathered at operational level
- Operational data seldom well-suited for decision support tasks
- Need reformat data in order to be useful for business intelligence

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Operational Data vs. Decision Support Data

- Operational data
  - Mostly stored in relational database
  - Optimized to support transactions representing daily operations
- Decision support data differs from operational data in three main areas:
  - Time span
  - Granularity
  - Dimensionality

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

TABLE 15.5 **Contrasting Operational and Decision Support Data Characteristics**

CHARACTERISTIC	OPERATIONAL DATA	DECISION SUPPORT DATA
Data currency	Current operations Real-time data	Historic data Snapshot of company data Time component (week/month/year)
Granularity	Atomic-detailed data	Summarized data
Summarization level	Low; some aggregate yields	High; many aggregation levels
Data model	Highly normalized Mostly relational DBMSs	Non-normalized Complex structures Some relational, but mostly multidimensional DBMSs
Transaction type	Mostly updates	Mostly query
Transaction volumes	High update volumes	Periodic loads and summary calculations
Transaction speed	Updates are critical	Retrievals are critical
Query activity	Low to medium	High
Query scope	Narrow range	Broad range
Query complexity	Simple to medium	Very complex
Data volumes	Hundreds of gigabytes	Terabytes to petabytes

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Decision Support Database Requirements

- Specialized DBMS tailored to provide fast answers to complex queries
- Three main requirements
  - Database schema
  - Data extraction and loading
  - Database size

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Decision Support Database Requirements (cont'd.)

- Database schema
  - Complex data representations
  - Aggregated and summarized data
  - Queries extract multidimensional time slices
- Data extraction and filtering
  - Supports different data sources
    - Flat files
    - Hierarchical, network, and relational databases
    - Multiple vendors
  - Checking for inconsistent data

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Decision Support Database Requirements (cont'd.)

- Database size
  - In 2005, Wal-Mart had 260 terabytes of data in its data warehouses
  - DBMS must support very large databases (VLDBs)

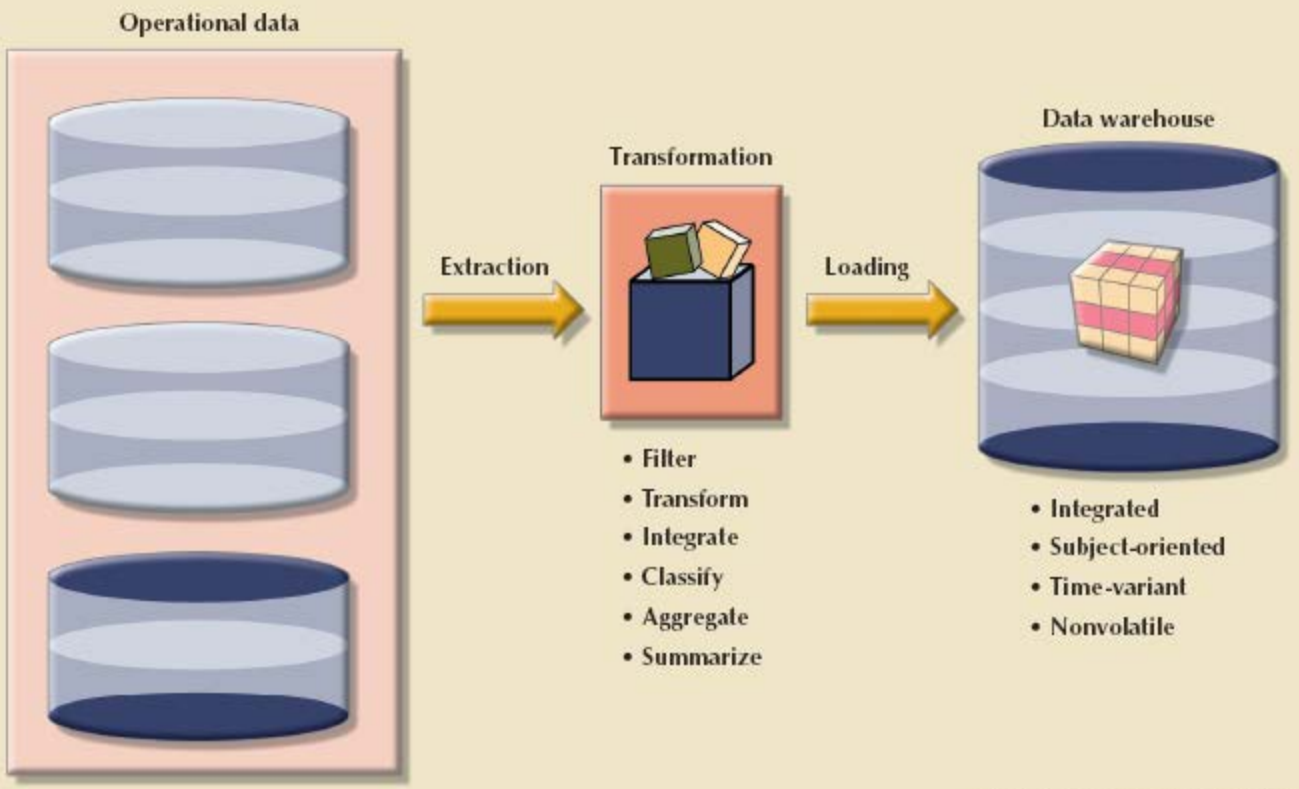
© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

The Data Warehouse

- Integrated, subject-oriented, time-variant, and nonvolatile collection of data
  - Provides support for decision making
- Usually a read-only database optimized for data analysis and query processing
- Requires time, money, and considerable managerial effort to create

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

FIGURE 15.4 The ETL process



SOURCE: Course Technology/Cengage Learning

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Data Marts

- Small, single-subject data warehouse subset
- More manageable data set than data warehouse
- Provides decision support to small group of people
- Typically lower cost and lower implementation time than data warehouse

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Twelve Rules That Define a Data Warehouse

TABLE 15.9 Twelve Rules for a Data Warehouse

RULE NO.	DESCRIPTION
1	The data warehouse and operational environments are separated.
2	The data warehouse data are integrated.
3	The data warehouse contains historical data over a long time.
4	The data warehouse data are snapshot data captured at a given point in time.
5	The data warehouse data are subject oriented.
6	The data warehouse data are mainly read-only with periodic batch updates from operational data. No online updates are allowed.
7	The data warehouse development life cycle differs from classical systems development. Data warehouse development is data-driven; the classical approach is process-driven.
8	The data warehouse contains data with several levels of detail: current detail data, old detail data, lightly summarized data, and highly summarized data.
9	The data warehouse environment is characterized by read-only transactions to very large data sets. The operational environment is characterized by numerous update transactions to a few data entities at a time.
10	The data warehouse environment has a system that traces data sources, transformations, and storage.
11	The data warehouse's metadata are a critical component of this environment. The metadata identify and define all data elements. The metadata provide the source, transformation, integration, storage, usage, relationships, and history of each data element.
12	The data warehouse contains a chargeback mechanism for resource usage that enforces optimal use of the data by end users.

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Star Schemas

- Data-modeling technique
  - Maps multidimensional decision support data into relational database
- Creates near equivalent of multidimensional database schema from relational data
- Easily implemented model for multidimensional data analysis while preserving relational structures
- Four components: facts, dimensions, attributes, and attribute hierarchies

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Facts

- Numeric measurements that represent specific business aspect or activity
  - Normally stored in fact table that is center of star schema
- Fact table contains facts linked through their dimensions
- Metrics are facts computed at run time

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Dimensions

- Qualifying characteristics provide additional perspectives to a given fact
- Decision support data almost always viewed in relation to other data
- Study facts via dimensions
- Dimensions stored in dimension tables

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Attributes

- Use to search, filter, and classify facts
- Dimensions provide descriptions of facts through their attributes
- No mathematical limit to the number of dimensions
- Slice and dice: focus on slices of the data cube for more detailed analysis

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Attribute Hierarchies

- Provide top-down data organization
- Two purposes:
  - Aggregation
  - Drill-down/roll-up data analysis
- Determine how the data are extracted and represented
- Stored in the DBMS's data dictionary
- Used by OLAP tool to access warehouse properly

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Star Schema Representation

- Facts and dimensions represented in physical tables in data warehouse database
- Many fact rows related to each dimension row
  - Primary key of fact table is a composite primary key
  - Fact table primary key formed by combining foreign keys pointing to dimension tables
- Dimension tables are smaller than fact tables
- Each dimension record is related to thousands of fact records

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Performance-Improving Techniques for the Star Schema

- Four techniques to optimize data warehouse design:
  - Normalizing dimensional tables
  - Maintaining multiple fact tables to represent different aggregation levels
  - Denormalizing fact tables
  - Partitioning and replicating tables

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Performance-Improving Techniques for the Star Schema (cont'd.)

- Dimension tables normalized to:
  - Achieve semantic simplicity
  - Facilitate end-user navigation through the dimensions
- Denormalizing fact tables improves data access performance and saves data storage space
- Partitioning splits table into subsets of rows or columns
- Replication makes copy of table and places it in different location

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Data Analytics

- Subset of BI functionality
- Encompasses a wide range of mathematical, statistical, and modeling techniques
  - Purpose of extracting knowledge from data
- Tools can be grouped into two separate areas:
  - Explanatory analytics
  - Predictive analytics

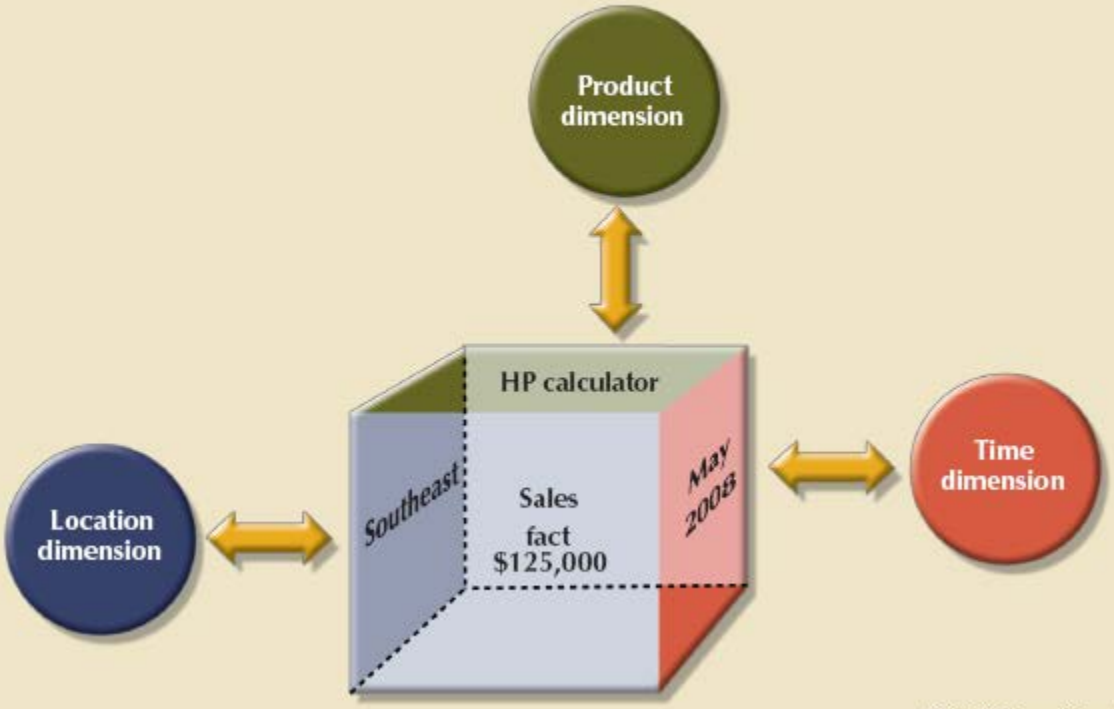
© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Data Mining

- Data-mining tools do the following:
  - Analyze data
  - Uncover problems or opportunities hidden in data relationships
  - Form computer models based on their findings
  - Use models to predict business behavior
- Runs in two modes
  - Guided
  - Automated

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

FIGURE 15.5 Simple star schema



SOURCE: Course Technology/Cengage Learning

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Predictive Analytics

- Employs mathematical and statistical algorithms, neural networks, artificial intelligence, and other advanced modeling tools
- Create actionable predictive models based on available data
- Models are used in areas such as:
  - Customer relationships, customer service, customer retention, fraud detection, targeted marketing, and optimized pricing

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Online Analytical Processing

- Three main characteristics:
  - Multidimensional data analysis techniques
  - Advanced database support
  - Easy-to-use end-user interfaces

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Multidimensional Data Analysis Techniques

- Data are processed and viewed as part of a multidimensional structure
- Augmented by the following functions:
  - Advanced data presentation functions
  - Advanced data aggregation, consolidation, and classification functions
  - Advanced computational functions
  - Advanced data modeling functions

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Advanced Database Support

- Advanced data access features include:
  - Access to many different kinds of DBMSs, flat files, and internal and external data sources
  - Access to aggregated data warehouse data
  - Advanced data navigation
  - Rapid and consistent query response times
  - Maps end-user requests to appropriate data source and to proper data access language
  - Support for very large databases

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Easy-to-Use End-User Interface

- Advanced OLAP features are more useful when access is simple
- Many interface features are “borrowed” from previous generations of data analysis tools
  - Already familiar to end users
  - Makes OLAP easily accepted and readily used

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

OLAP Architecture

- Three main architectural components:
  - Graphical user interface (GUI)
  - Analytical processing logic
  - Data-processing logic

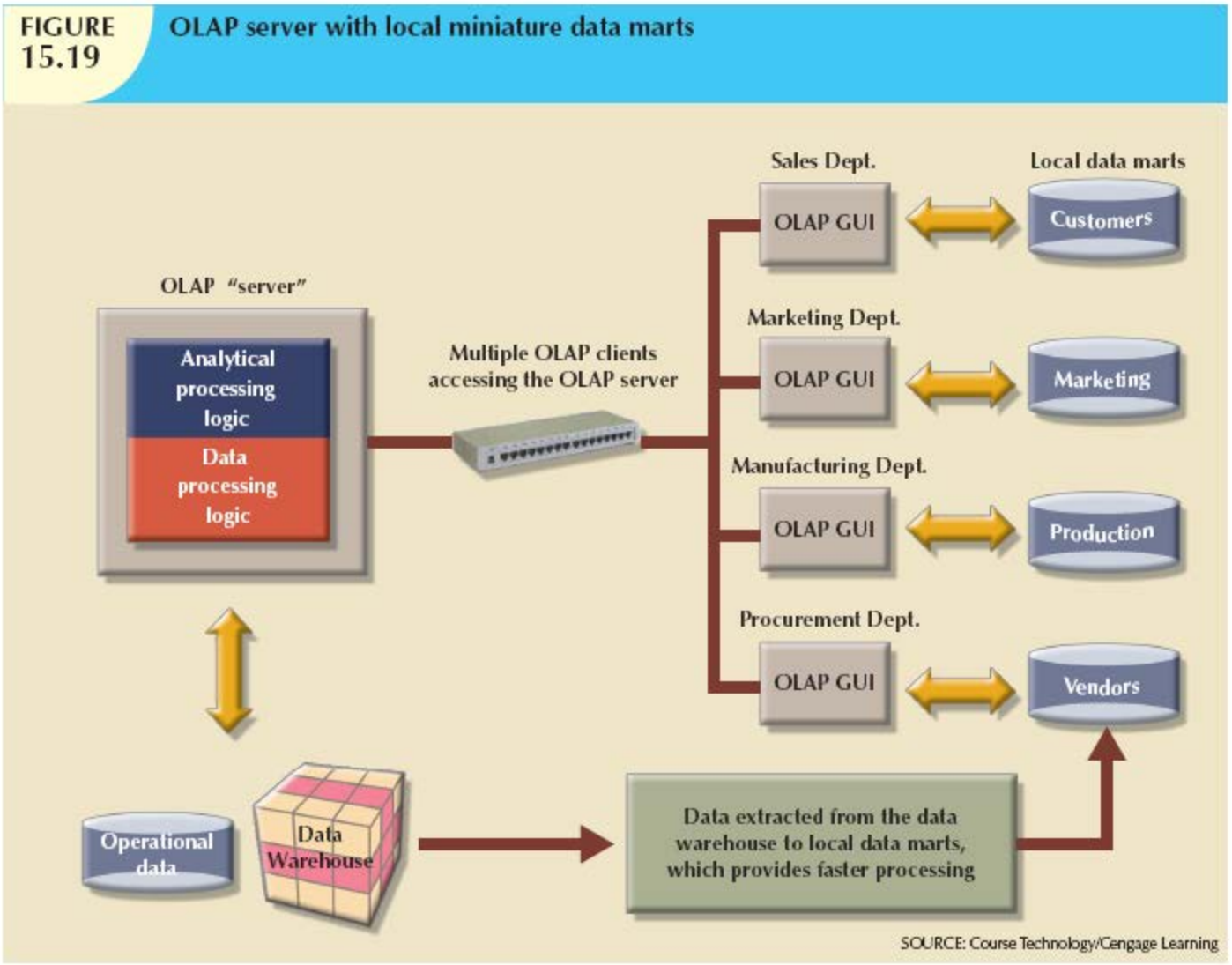
© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

OLAP Architecture (cont'd.)

- Designed to use both operational and data warehouse data
- In most implementations, data warehouse and OLAP are interrelated and complementary
- OLAP systems merge data warehouse and data mart approaches

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

FIGURE 15.19 OLAP server with local miniature data marts



© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Relational OLAP

- Relational online analytical processing (ROLAP) provides the following extensions:
  - Multidimensional data schema support within the RDBMS
  - Data access language and query performance optimized for multidimensional data
  - Support for very large databases (VLDBs)

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Multidimensional OLAP

- Multidimensional online analytical processing (MOLAP) extends OLAP functionality to multidimensional database management systems (MDBMSs)
  - MDBMS end users visualize stored data as a 3D data cube
  - Data cubes can grow to *n* dimensions, becoming hypercubes
  - To speed access, data cubes are held in memory in a cube cache

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Relational vs. Multidimensional OLAP

- Selection of one or the other depends on evaluator's vantage point
- Proper evaluation must include supported hardware, compatibility with DBMS, etc.
- ROLAP and MOLAP vendors working toward integration within unified framework
- Relational databases use star schema design to handle multidimensional data

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

TABLE 15.12 **Relational vs. Multidimensional OLAP**

CHARACTERISTIC	ROLAP	MOLAP
Schema	Uses star schema Additional dimensions can be added dynamically	Uses data cubes Multidimensional arrays, row stores, column stores Additional dimensions require re-creation of the data cube
Database size	Medium to large	Large
Architecture	Client/server Standards-based	Client/server Open or proprietary, depending on vendor
Access	Supports ad hoc requests Unlimited dimensions	Limited to predefined dimensions Proprietary access languages
Speed	Good with small data sets; average for medium-sized to large data sets	Faster for large data sets with predefined dimensions

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

SQL Extensions for OLAP

- Proliferation of OLAP tools fostered development of SQL extensions
- Many innovations have become part of standard SQL
- All SQL commands will work in data warehouse as expected
- Most queries include many data groupings and aggregations over multiple columns

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

The ROLLUP Extension

- Used with GROUP BY clause to generate aggregates by different dimensions
- GROUP BY generates only one aggregate for each new value combination of attributes
- ROLLUP extension enables subtotal for each column listed except for the last one
  - Last column gets grand total
- Order of column list important

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

The CUBE Extension

- CUBE extension used with GROUP BY clause to generate aggregates by listed columns
  - Includes the last column
- Enables subtotal for each column in addition to grand total for last column
  - Useful when you want to compute all possible subtotals within groupings
- Cross-tabulations are good candidates for application of CUBE extension

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Materialized Views

- A dynamic table that contains SQL query command to generate rows
  - Also contains the actual rows
- Created the first time query is run and summary rows are stored in table
- Automatically updated when base tables are updated

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Summary

- Business intelligence generates information used to support decision making
- BI covers a range of technologies, applications, and functionalities
- Decision support systems were the precursor of current generation BI systems
- Operational data not suited for decision support

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Summary (cont'd.)

- Data warehouse provides support for decision making
  - Usually read-only
  - Optimized for data analysis, query processing
- Star schema is a data-modeling technique
  - Maps multidimensional decision support data into a relational database
- Star schema has four components:
  - Facts, dimensions, attributes, and attribute hierarchies

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Summary (cont'd.)

- Data analytics
  - Provides advanced data analysis tools to extract knowledge from business data
- Data mining
  - Automates the analysis of operational data to find previously unknown data characteristics, relationships, dependencies, and trends
- Predictive analytics
  - Uses information generated in the data-mining phase to create advanced predictive models

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.

Summary (cont'd.)

- Online analytical processing (OLAP)
  - Advanced data analysis environment that supports decision making, business modeling, and operations research
- SQL has been enhanced with extensions that support OLAP-type processing and data generation

© 2013 Cengage Learning. All Rights Reserved. This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. May not be scanned, copied, duplicated, or posted to a publicly accessible website, in whole or in part.