Chapter 6
FOUNDATIONS OF BUSINESS INTELLIGENCE: DATABASES AND INFORMATION MANAGEMENT

Dr. N. Abdolvand

Management Information System

Sources

- Management Information Systems, Ken Laudon & Jane Laudon, Prentice Hall

Cases

- Case 1a: City of Dubuque Uses Cloud Computing and Sensors to Build a Smarter, Sustainable City
- Case 1b: IBM Smarter City: Portland, Oregon
- Case 2: Data Warehousing at REI: Understanding the Customer
- Case 3: Maruti Suzuki Business Intelligence and Enterprise Databases

Learning Objectives

- Describe how the problems of managing data resources in a traditional file environment are solved by a database management system
- Describe the capabilities and value of a database management system
- Apply important database design principles
- Evaluate tools and technologies for accessing information from databases to improve business performance and decision making
- Assess the role of information policy, data administration, and data quality assurance in the management of a firm's data resources

RR Donnelley Tries to Master Its Data

- Problem: Explosive growth created information management challenges.
- Solutions: Use MDM to create an enterprise-wide set of data, preventing unnecessary data duplication.
- Master data management (MDM) enables companies like R.R. Donnelley to eliminate outdated, incomplete or incorrectly formatted data.
- Demonstrates IT's role in successful data management.
- Illustrates digital technology's role in storing and organizing data.

5

Organizing Data in a Traditional File Environment

- File organization concepts
  - Database: Group of related files
  - File: Group of records of same type
  - Record: Group of related fields
  - Field: Group of characters as word(s) or number
    - Describes an entity (person, place, thing on which we store information)
    - Attribute: Each characteristic, or quality, describing entity
      - E.g., Attributes Date or Grade belong to entity COURSE

6

THE DATA HIERARCHY

Organizing Data in a Traditional File Environment

A computer system organizes data in a hierarchy that starts with the bit, which represents either a 0 or a 1. Bits can be grouped to form a byte to represent one character, number, or symbol. Bytes can be grouped to form a field, and related fields can be grouped to form a record. Related records can be collected to form a file, and related files can be organized into a database.

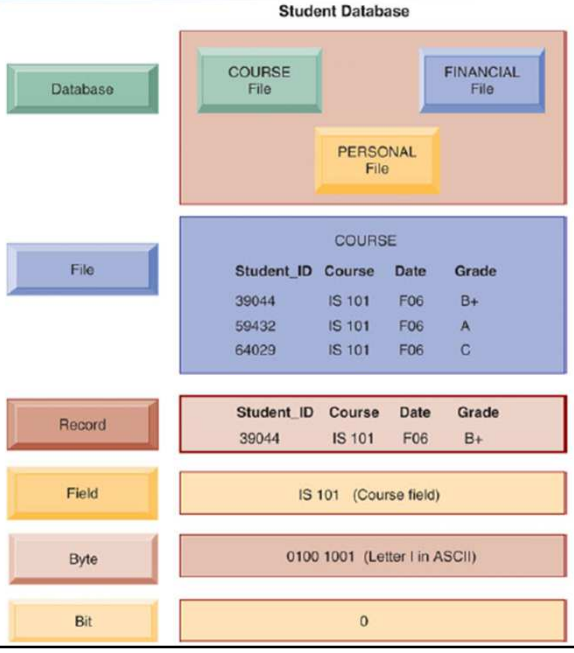


FIGURE 6-1

Organizing Data in a Traditional File Environment

- Problems with the traditional file environment (files maintained separately by different departments)
  - Data redundancy:
    - Presence of duplicate data in multiple files
  - Data inconsistency:
    - Same attribute has different values
  - Program-data dependence:
    - When changes in program requires changes to data accessed by program
  - Lack of flexibility
  - Poor security
  - Lack of data sharing and availability

8

The Database Approach to Data Management

- Database
  - Serves many applications by centralizing data and controlling redundant data
- Database management system (DBMS)
  - Interfaces between applications and physical data files
  - Separates logical and physical views of data
  - Solves problems of traditional file environment
    - Controls redundancy
    - Eliminates inconsistency
    - Uncouples programs and data
    - Enables organization to centrally manage data and data security

9

HUMAN RESOURCES DATABASE WITH MULTIPLE VIEWS

The Database Approach to Data Management

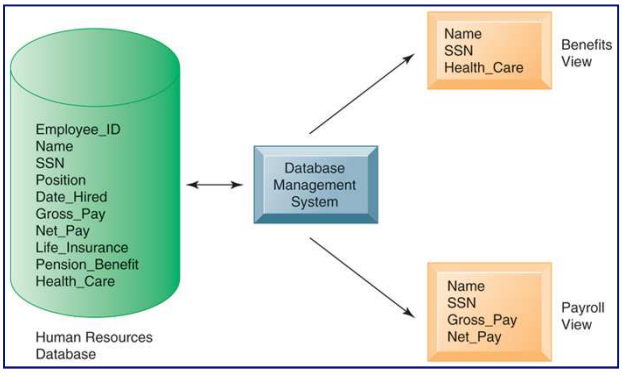


FIGURE 6-3 A single human resources database provides many different views of data, depending on the information requirements of the user. Illustrated here are two possible views, one of interest to a benefits specialist and one of interest to a member of the company's payroll department.

10

The Database Approach to Data Management

- Relational DBMS
  - Represent data as two-dimensional tables called relations or files
  - Each table contains data on entity and attributes
- Table: grid of columns and rows
  - Rows (tuples): Records for different entities
  - Fields (columns): Represents attribute for entity
  - Key field: Field used to uniquely identify each record
  - Primary key: Field in table used for key fields
  - Foreign key: Primary key used in second table as look-up field to identify records from original table

11

RELATIONAL DATABASE TABLES

The Database Approach to Data Management

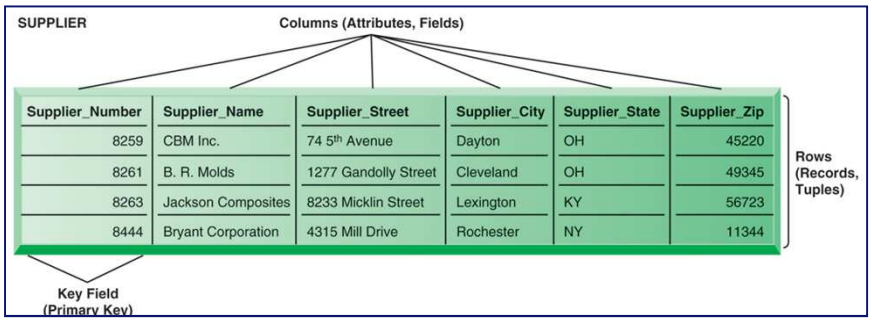


FIGURE 6-4 A relational database organizes data in the form of two-dimensional tables. Illustrated here are tables for the entities SUPPLIER and PART showing how they represent each entity and its attributes. Supplier Number is a primary key for the SUPPLIER table and a foreign key for the PART table.

12

Using Databases to Improve Business Performance and Decision Making

- Big data
  - Massive sets of unstructured/semi-structured data from Web traffic, social media, sensors, and so on
  - Petabytes, exabytes of data
    - Volumes too great for typical DBMS
  - Can reveal more patterns and anomalies

13

Using Databases to Improve Business Performance and Decision Making

- Business intelligence infrastructure
  - Today includes an array of tools for separate systems, and big data
- Contemporary tools:
  - Data warehouses
  - Data marts
  - Analytical platforms

14

Using Databases to Improve Business Performance and Decision Making

- Data warehouse:
  - Stores current and historical data from many core operational transaction systems
  - Consolidates and standardizes information for use across enterprise, but data cannot be altered
  - Data warehouse system will provide query, analysis, and reporting tools
- Data marts:
  - Subset of data warehouse
  - Summarized or highly focused portion of firm's data for use by specific population of users
  - Typically focuses on single subject or line of business
- Analytic platforms
  - High-speed platforms using both relational and non-relational tools optimized for large datasets

15

COMPONENTS OF A DATA WAREHOUSE

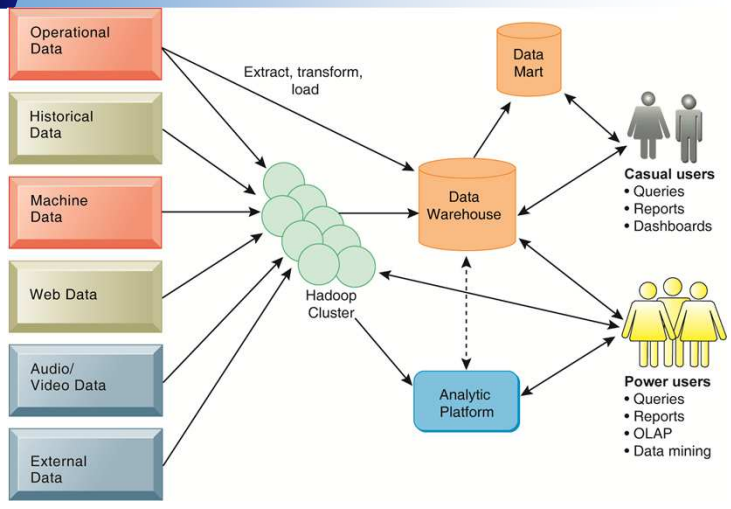


FIGURE 6-12 The data warehouse extracts current and historical data from multiple operational systems inside the organization. These data are combined with data from external sources and reorganized into a central database designed for management reporting and analysis. The information directory provides users with information about the data available in the warehouse.

16

Decision Making and Information Systems

- Business value of improved decision making
  - Improving hundreds of thousands of "small" decisions adds up to large annual value for the business
- Types of decisions:
  - Unstructured: Decision maker must provide judgment, evaluation, and insight to solve problem
  - Structured: Repetitive and routine; involve definite procedure for handling so they do not have to be treated each time as new
  - Semistructured: Only part of problem has clear-cut answer provided by accepted procedure

17

Decision Making and Information Systems

- Senior managers:
  - Make many unstructured decisions
  - E.g. Should we enter a new market?
- Middle managers:
  - Make more structured decisions but these may include unstructured components
  - E.g. Why is order fulfillment report showing decline in Minneapolis?
- Operational managers, rank and file employees
  - Make more structured decisions
  - E.g. Does customer meet criteria for credit?

18

STAGES IN DECISION MAKING

The decision-making process is broken down into four stages.

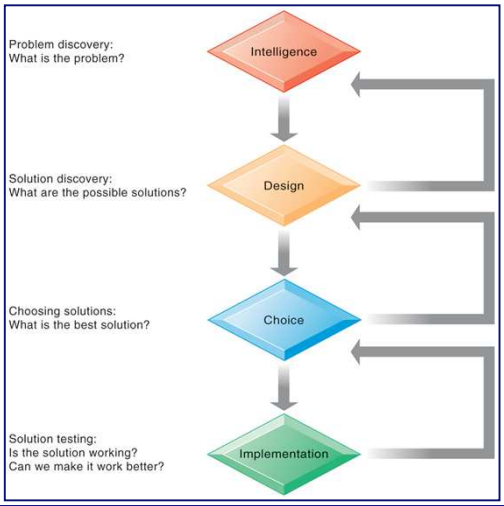


FIGURE 12-2

19

Business Intelligence in the Enterprise

- Business intelligence
  - Infrastructure for collecting, storing, analyzing data produced by business
  - Databases, data warehouses, data marts
- Business analytics
  - Tools and techniques for analyzing data
  - OLAP, statistics, models, data mining

20

Using Databases to Improve Business Performance and Decision Making

- Business Intelligence:
  - Tools for consolidating, analyzing, and providing access to vast amounts of data to help users make better business decisions
  - E.g., Harrah's Entertainment analyzes customers to develop gambling profiles and identify most profitable customers
  - Principle tools include:
    - Software for database query and reporting
    - Online analytical processing (OLAP)
    - Data mining

21

Business Intelligence in the Enterprise

- Six elements in the business intelligence environment
  1. Data from the business environment
  2. Business intelligence infrastructure
  3. Business analytics toolset
  4. Managerial users and methods
  5. Delivery platform – MIS, DSS, ESS
  6. User interface

22

BUSINESS INTELLIGENCE AND ANALYTICS FOR DECISION SUPPORT

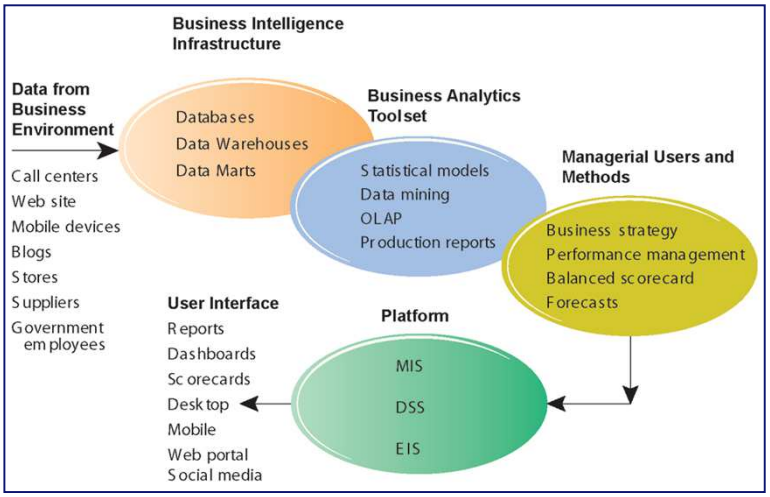


FIGURE 12-3 Business intelligence and analytics requires a strong database foundation, a set of analytic tools, and an involved management team that can ask intelligent questions and analyze data.

Business Intelligence in the Enterprise

- Business intelligence and analytics capabilities
  - Goal is to deliver accurate real-time information to decision-makers
  - Main functionalities of BI systems
    1. Production reports
    2. Parameterized reports
    3. Dashboards/scorecards
    4. Ad hoc query/search/report creation
    5. Drill down
    6. Forecasts, scenarios, models

Using Databases to Improve Business Performance and Decision Making

- Online analytical processing (OLAP)
  - Supports multidimensional data analysis
    - Viewing data using multiple dimensions
    - Each aspect of information (product, pricing, cost, region, time period) is different dimension
    - E.g., how many washers sold in the East in June compared with other regions?
  - OLAP enables rapid, online answers to ad hoc queries

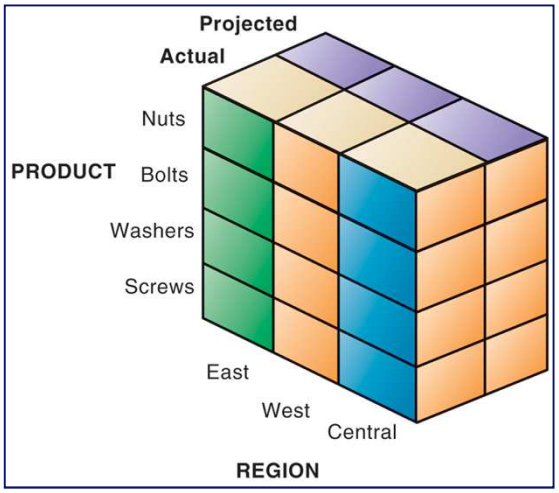
25

MULTIDIMENSIONAL DATA MODEL

- The Database Approach to Data Management

The view that is showing is product versus region. If you rotate the cube 90 degrees, the face that will show is product versus actual and projected sales. If you rotate the cube 90 degrees again, you will see region versus actual and projected sales. Other views are possible.

FIGURE 6-13



26

Using Databases to Improve Business Performance and Decision Making

- Data mining:
  - More discovery driven than OLAP
  - Finds hidden patterns, relationships in large databases and infers rules to predict future behavior
  - E.g., Finding patterns in customer data for one-to-one marketing campaigns or to identify profitable customers.
  - Types of information obtainable from data mining
    - Associations
    - Sequences
    - Classification
    - Clustering
    - Forecasting

27

Using Databases to Improve Business Performance and Decision Making

- Predictive analysis
  - Uses data mining techniques, historical data, and assumptions about future conditions to predict outcomes of events
  - E.g., Probability a customer will respond to an offer
- Text mining
  - Extracts key elements from large unstructured data sets
    - Stored e-mails
    - Call center transcripts
    - Legal cases
    - Patent descriptions
    - Service reports, and so on
  - Sentiment analysis software
    - Mines e-mails, blogs, social media to detect opinions

28

Using Databases to Improve Business Performance and Decision Making

- Web mining
  - Discovery and analysis of useful patterns and information from WWW
    - E.g., to understand customer behavior, evaluate effectiveness of Web site, etc.
  - Web content mining
    - Knowledge extracted from content of Web pages
  - Web structure mining
    - E.g., links to and from Web page
  - Web usage mining
    - User interaction data recorded by Web server

29

Business Intelligence in the Enterprise

- Business intelligence users
  - 80% are casual users relying on production reports
  - Senior executives
    - Use monitoring functionalities
  - Middle managers and analysts
    - Ad-hoc analysis
  - Operational employees
    - Prepackaged reports
    - E.g. sales forecasts, customer satisfaction, loyalty and attrition, supply chain backlog, employee productivity

30

BUSINESS INTELLIGENCE USERS

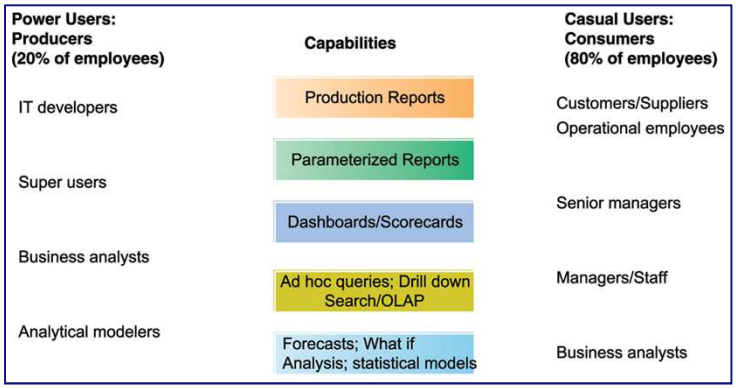


FIGURE 12-4 Casual users are consumers of BI output, while intense power users are the producers of reports, new analyses, models, and forecasts.

31

Decision Making and Information Systems

- Three main reasons why investments in information technology do not always produce positive results
  1. Information quality
     - High-quality decisions require high-quality information
  2. Management filters
     - Managers have selective attention and have variety of biases that reject information that does not conform to prior conceptions
  3. Organizational inertia and politics
     - Strong forces within organizations resist making decisions calling for major change

32

Managing Data Resources

- Establishing an information policy
  - Firm's rules, procedures, roles for sharing, managing, standardizing data
  - Data administration:
    - Firm function responsible for specific policies and procedures to manage data
  - Data governance:
    - Policies and processes for managing availability, usability, integrity, and security of enterprise data, especially as it relates to government regulations
  - Database administration:
    - Defining, organizing, implementing, maintaining database; performed by database design and management group

33

Managing Data Resources

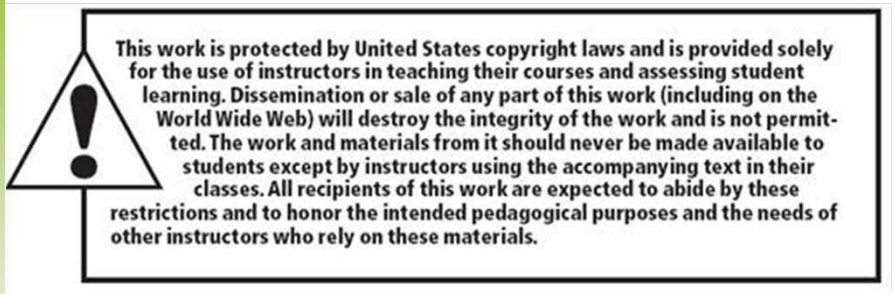
- Ensuring data quality
  - More than 25% of critical data in Fortune 1000 company databases are inaccurate or incomplete
  - Most data quality problems stem from faulty input
  - Before new database in place, need to:
    - Identify and correct faulty data
    - Establish better routines for editing data once database in operation

34

Managing Data Resources

- Data quality audit:
  - Structured survey of the accuracy and level of completeness of the data in an information system
    - Survey samples from data files, or
    - Survey end users for perceptions of quality
- Data cleansing
  - Software to detect and correct data that are incorrect, incomplete, improperly formatted, or redundant
  - Enforces consistency among different sets of data from separate information systems

35




Questions?