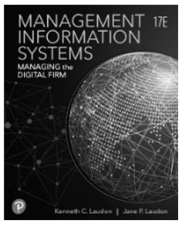
Management Information Systems: Managing the Digital Firm

Seventeenth Edition



Chapter 6

Foundations of Business Intelligence: Databases and Information Management

Learning Objectives

6.1 What are the problems of managing data resources in a traditional file environment?

6.2 What are the major capabilities of database management systems (DBMS), and why is a relational DBMS so powerful?

6.3 What are the principal tools and technologies for accessing information from databases to improve business performance and decision making?

6.4 Why are data governance and data quality assurance essential for managing the firm's data resources?

6.5 How will MIS help my career?

Video Cases

- Case 1: Brooks Brothers Closes In on Omnichannel Retail
- Case 2: Maruti Suzuki Business Intelligence and Enterprise Databases

Domino's Masters Data One Pizza at a Time (Slide 1 of 2)

- Problem
  - Very large volumes of data
  - Data fragmented in multiple systems and files
- Solutions
  - Develop enterprise data strategy
  - Consolidate, standardize, and cleanse data
  - Revise data access rules and procedures
  - Deploy Talend, MDM software, and Hadoop

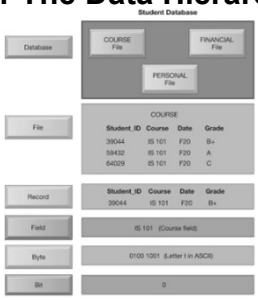
Domino's Masters Data One Pizza at a Time (Slide 2 of 2)

- Enterprise Management Framework makes data more accurate and consistent enterprise-wide, accelerates decision making and improves customer analysis
- Illustrates the importance of data management for better decision making and customer analysis

File Organization Terms and 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(s)
- Entity: Person, place, thing on which we store information
- Attribute: Each characteristic, or quality, describing entity

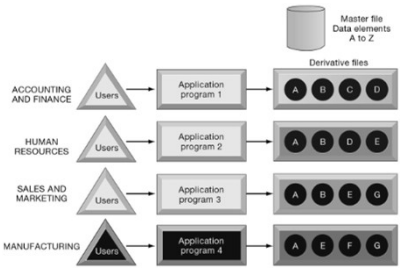
Figure 6.1 The Data Hierarchy



Problems with the Traditional File Environment

- Files maintained separately by different departments
- Data redundancy
- Data inconsistency
- Program-data dependence
- Lack of flexibility
- Poor security
- Lack of data sharing and availability

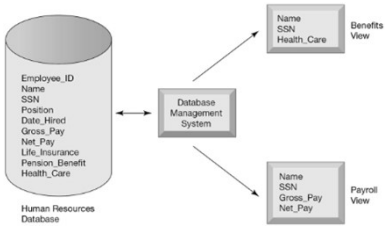
Figure 6.2 Traditional File Processing



Database Management Systems

- 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

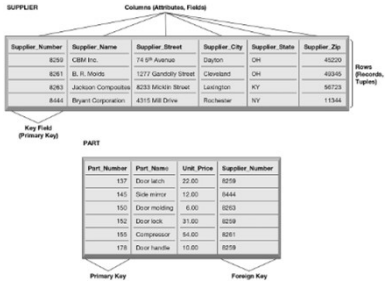
Figure 6.3 Human Resources Database with Multiple Views



Relational DBMS

- Represent data as two-dimensional tables
- 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

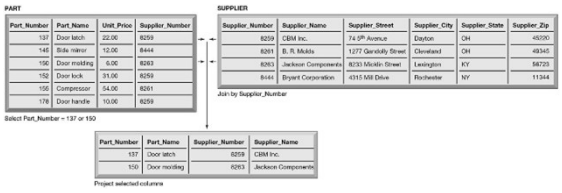
Figure 6.4 Relational Database Tables



Operations of a Relational DBMS

- Three basic operations used to develop useful sets of data
  - SELECT
    - Creates subset of data of all records that meet stated criteria
  - JOIN
    - Combines relational tables to provide user with more information than available in individual tables
  - PROJECT
    - Creates subset of columns in table, creating tables with only the information specified

Figure 6.5 The Three Basic Operations of a Relational DBMS



Capabilities of Database Management Systems

- Data definition
- Data dictionary
- Querying and reporting
  - Data manipulation language
    - Structured Query Language (SQL)
- Many DBMS have report generation capabilities for creating polished reports (Microsoft Access)

Figure 6.6 Access Data Dictionary Features

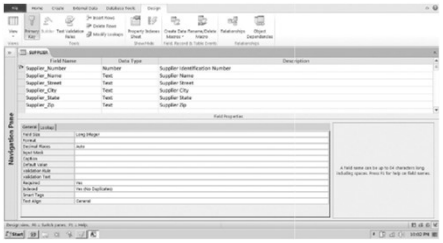
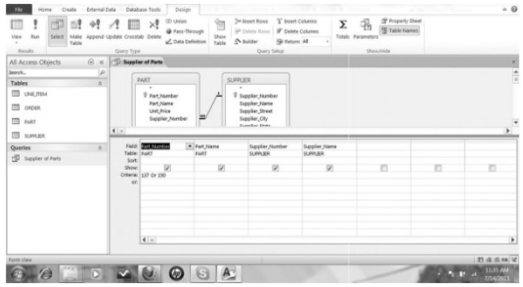


Figure 6.7 Example of an SQL Query

```
SELECT PART.Part_Number, PART.Part_Name, SUPPLIER.Supplier_Number,
SUPPLIER.Supplier_Name
FROM PART, SUPPLIER
WHERE PART.Supplier_Number = SUPPLIER.Supplier_Number AND
Part_Number = 137 OR Part_Number = 150;
```

Figure 6.8 An Access Query



Designing Databases

- Conceptual design vs. physical design
- Normalization
  - Streamlining complex groupings of data to minimize redundant data elements and awkward many-to-many relationships
- Referential integrity
  - Rules used by RDBMS to ensure relationships between tables remain consistent
- Entity-relationship diagram
- A correct data model is essential for a system serving the business well

Figure 6.9 An Unnormalized Relation for Order

Figure 6.10 Normalized Tables Created from Order

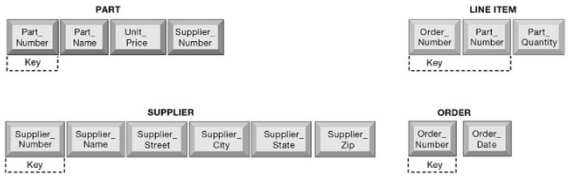


Figure 6.11 An Entity-Relationship Diagram

Non-Relational Databases, Cloud Databases and Blockchain (Slide 1 of 3)

- Non-relational databases: “No SQL”
  - More flexible data model
  - Data sets stored across distributed machines
  - Easier to scale
  - Handle large volumes of unstructured and structured data

Non-Relational Databases, Cloud Databases and Blockchain (Slide 2 of 3)

- Cloud databases
  - Appeal to start-ups, smaller businesses
  - Amazon Relational Database Service, Microsoft SQL Azure
  - Private clouds
- Distributed databases
  - Stored in multiple physical locations
  - Example: Google Spanner

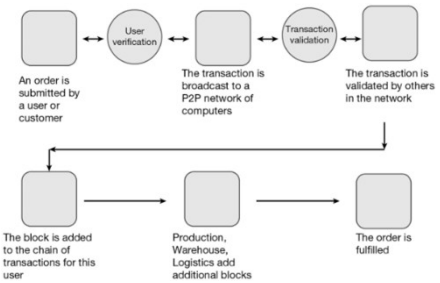
Interactive Session: Technology: New Cloud Database Tools Help Vodafone Fiji Make Better Decisions

- Class discussion
  - Define the problem faced by Vodafone Fiji. What management, organization, and technology factors contributed to the problem?
  - Evaluate Oracle Autonomous Data Warehouse and Oracle Analytics Cloud as a solution for Vodafone Fiji?
  - How did the new Oracle tools change decision making at Vodafone Fiji?
  - Was using cloud services advantageous for Vodafone Fiji? Explain your answer.

Non-relational Databases, Cloud Databases, and Blockchain (Slide 3 of 3)

- Blockchain
  - Distributed ledgers in a peer-to-peer distributed database
  - Maintains a growing list of records and transactions shared by all
  - Encryption used to identify participants and transactions
  - Used for financial transactions, supply chain, and medical records
  - Foundation of Bitcoin, and other crypto currencies

Figure 6.12 How Blockchain Works



The Challenge of Big Data

- Big data
  - Massive sets of unstructured/semi-structured data from web traffic, social media, sensors, and so on
- Volumes too great for typical DBMS
  - Petabytes, exabytes of data
- Can reveal more patterns, relationships and anomalies
- Requires new tools and technologies to manage and analyze

Interactive Session: Management: Big Data Baseball

- Class discussion
  - How did information technology change the game of baseball? Explain.
  - How did information technology affect decision making at MLB teams? What kind of decisions changed as the result of using big data?
  - How much should baseball rely on big data and analytics?

Business Intelligence Infrastructure (1 of 4)

- Array of tools for obtaining information from separate systems and from big data
  - Data warehouse
  - Data mart
  - Hadoop
  - In-memory computing
  - Analytical platforms

Business Intelligence Infrastructure (2 of 4)

- 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
  - Provides analysis and reporting tools
- Data marts
  - Subset of data warehouse
  - Typically focus on single subject or line of business

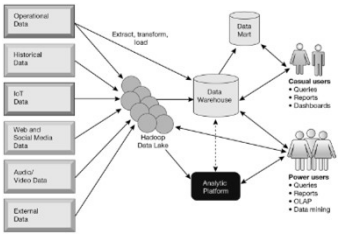
Business Intelligence Infrastructure (3 of 4)

- Hadoop
  - Enables distributed parallel processing of big data across inexpensive computers
  - Key services
    - Hadoop Distributed File System (HDFS): data storage
    - MapReduce: breaks data into clusters for work
    - Hbase: No SQL database
  - Used by Yahoo, NextBio

Business Intelligence Infrastructure (4 of 4)

- In-memory computing
  - Used in big data analysis
  - Uses computers main memory (RAM) for data storage to avoid delays in retrieving data from disk storage
  - Can reduce hours/days of processing to seconds
  - Requires optimized hardware
- Analytic platforms
  - High-speed platforms using both relational and non-relational tools optimized for large datasets

Figure 6.13 Contemporary Business Intelligence Infrastructure



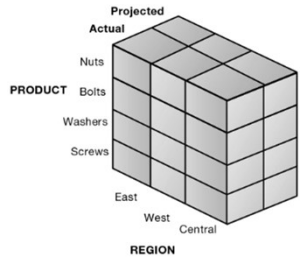
Analytical Tools: Relationships, Patterns, Trends

- Tools for consolidating, analyzing, and providing access to vast amounts of data to help users make better business decisions
  - Multidimensional data analysis (OLAP)
  - Data mining
  - Text mining
  - Web mining

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
  - Example: How many washers sold in the East in June compared to the sales forecast?
- OLAP enables rapid, online answers to ad hoc queries

Figure 6.14 Multidimensional Data Model



Data Mining

- Finds hidden patterns, relationships in datasets
  - Example: customer buying patterns
- Infers rules to predict future behavior
- Types of information obtainable from data mining:
  - Associations
  - Sequences
  - Classification
  - Clustering
  - Forecasting

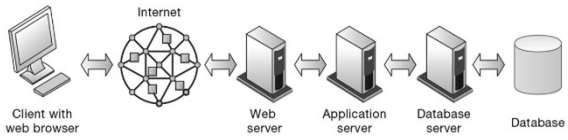
Text Mining and Web Mining

- Text mining
  - Extracts key elements from large unstructured text data sets
  - Sentiment analysis software
- Web mining
  - Discovery and analysis of useful patterns and information from web
  - Web content mining
  - Web structure mining
  - Web usage mining

Databases and the Web

- Many companies use the web to make some internal databases available to customers or partners
- Typical configuration includes:
  - Web server
  - Application server/middleware/scripts
  - Database server (hosting DBMS)
- Advantages of using the web for database access:
  - Ease of use of browser software
  - Web interface requires few or no changes to database
  - Inexpensive to add web interface to system

Figure 6.15 Linking Internal Databases to the Web



Data Governance

- Data governance
  - Encompasses policies and procedures through which data can be managed as an organizational resource.
  - Establishes rules for sharing, disseminating, acquiring, standardizing, classifying and inventorying information
    - Example: Firm information policy that specifies that only selected members of a particular department can view certain information

Data Quality Assurance

- More than 25 percent of critical data in Fortune 1000 company databases are inaccurate or incomplete
- Before new database is in place, a firm must:
  - Identify and correct faulty data
  - Establish better routines for editing data once database in operation
- Data quality audit
- Data cleansing

How Will MIS Help My Career?

- The Company: Mega Midwest Power
- Position Description: Entry-level data analyst
- Job Requirements
- Interview Questions
- Author Tips

Copyright