Management Information Systems

Database and Information Management

By: Binit Patel

Assistant Professor, I²IM - CHARUSAT

File organization concepts

- 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

Student Database

Database	COURSE File	FINANCIAL File
	PERSONAL File	

File

COURSE

Student_ID	Course	Date	Grade
39044	IS 101	F06	B+
59432	IS 101	F06	A
64029	IS 101	F06	C

Record

Student_ID	Course	Date	Grade
39044	IS 101	F06	B+

Field

IS 101   (Course field)

Byte

0100 1001   (Letter I in ASCII)

Bit

0

Problems with the traditional file environment

- 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
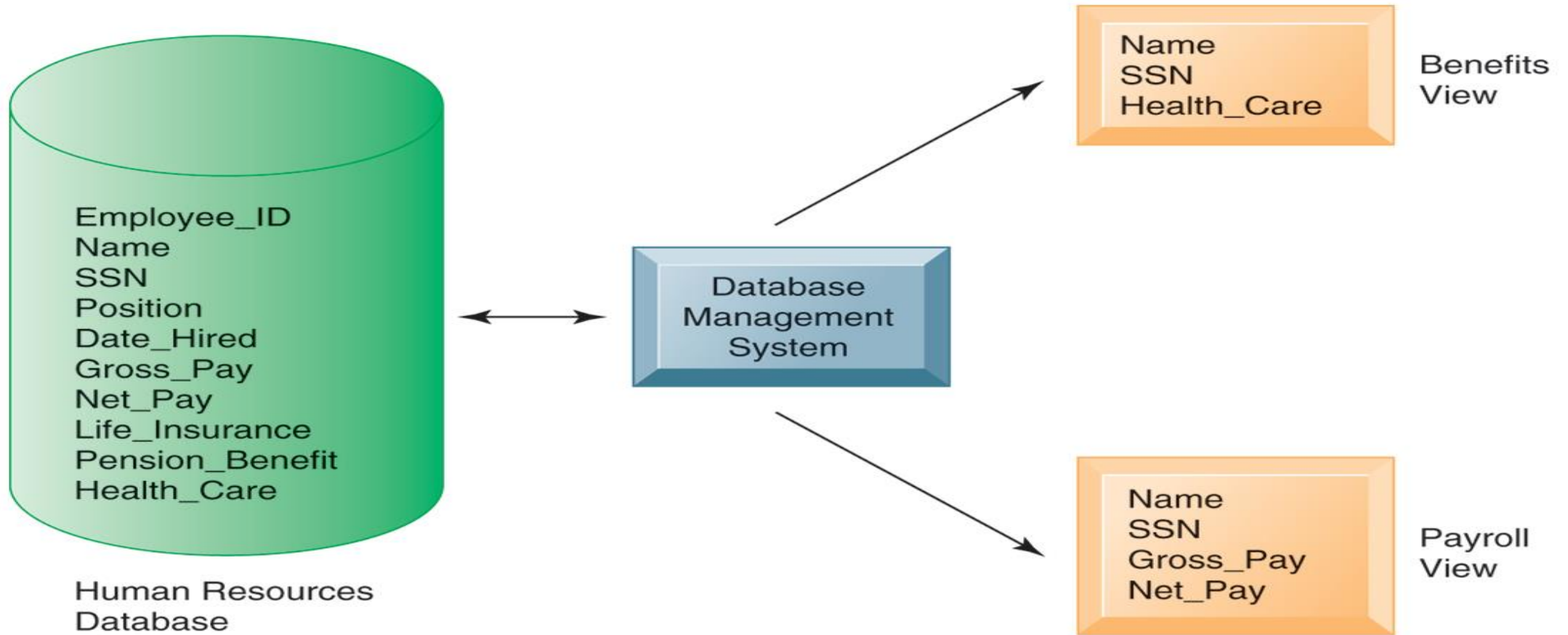- Lack of data sharing and availability

Traditional File Processing
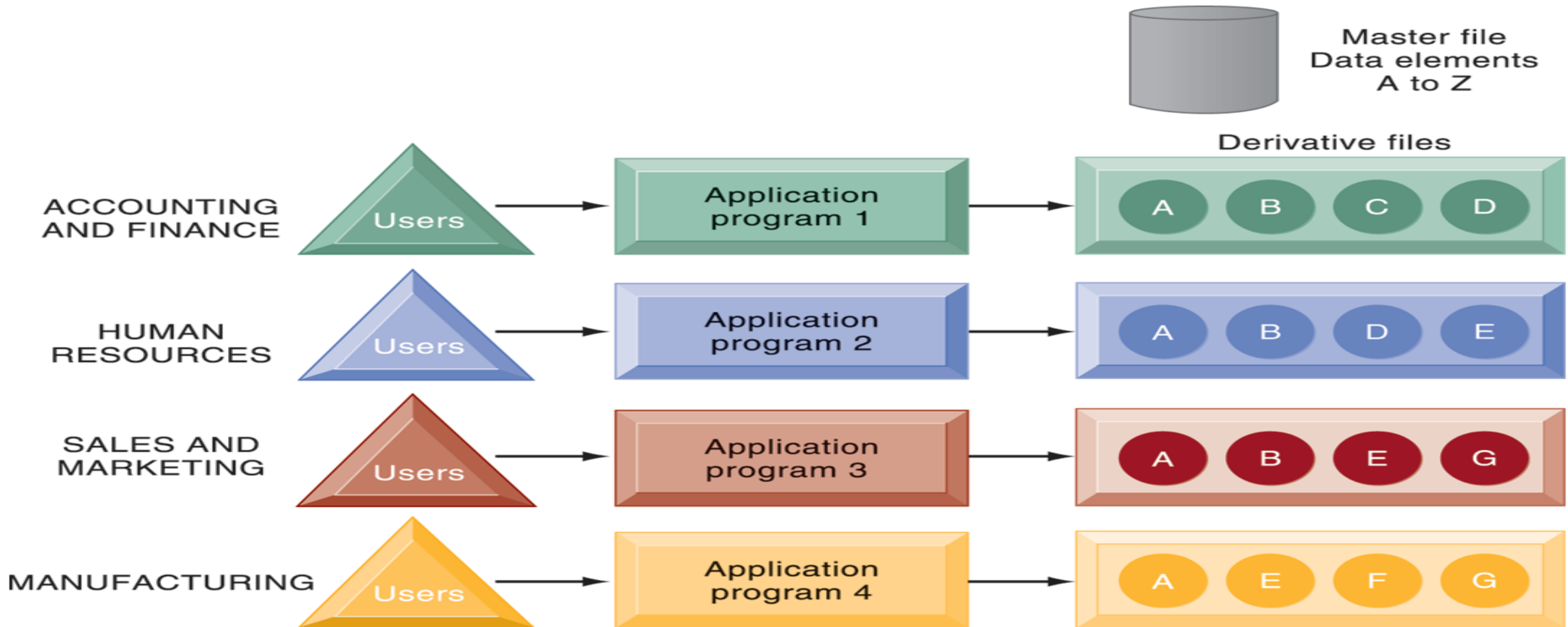
Database Approach

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

HR DATABASE WITH MULTIPLE VIEWS

Human Resources Database:
Employee_ID
Name
SSN
Position
Date_Hired
Gross_Pay
Net_Pay
Life_Insurance
Pension_Benefit
Health_Care

Database Management System

Benefits View:
Name
SSN
Health_Care

Payroll View:
Name
SSN
Gross_Pay
Net_Pay

Database Approach

- 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

Relational Database Table

SUPPLIER

Columns (Attributes, Fields)

Supplier_Number	Supplier_Name	Supplier_Street	Supplier_City	Supplier_State	Supplier_Zip
8259	CBM Inc.	74 5th Avenue	Dayton	OH	45220
8261	B. R. Molds	1277 Gandolly Street	Cleveland	OH	49345
8263	Jackson Composites	8233 Micklin Street	Lexington	KY	56723
8444	Bryant Corporation	4315 Mill Drive	Rochester	NY	11344

Rows (Records, Tuples)

Key Field (Primary Key)

Relational Database Table (Cont…)

PART

Part_Number	Part_Name	Unit_Price	Supplier_Number
137	Door latch	22.00	8259
145	Side mirror	12.00	8444
150	Door molding	6.00	8263
152	Door lock	31.00	8259
155	Compressor	54.00	8261
178	Door handle	10.00	8259

Primary Key   Foreign Key

Operations of a Relational DBMS

- 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

THE THREE BASIC OPERATIONS OF A RDBMS

PART

Part_Number	Part_Name	Unit_Price	Supplier_Number
137	Door latch	22.00	8259
145	Side mirror	12.00	8444
150	Door molding	6.00	8263
152	Door lock	31.00	8259
155	Compressor	54.00	8261
178	Door handle	10.00	8259

Select Part_Number = 137 or 150

SUPPLIER

Supplier_Number	Supplier_Name	Supplier_Street	Supplier_City	Supplier_State	Supplier_Zip
8259	CBM Inc.	74 5th Avenue	Dayton	OH	45220
8261	B. R. Molds	1277 Gandolly Street	Cleveland	OH	49345
8263	Jackson Components	8233 Micklin Street	Lexington	KY	56723
8444	Bryant Corporation	4315 Mill Drive	Rochester	NY	11344

Join by Supplier_Number
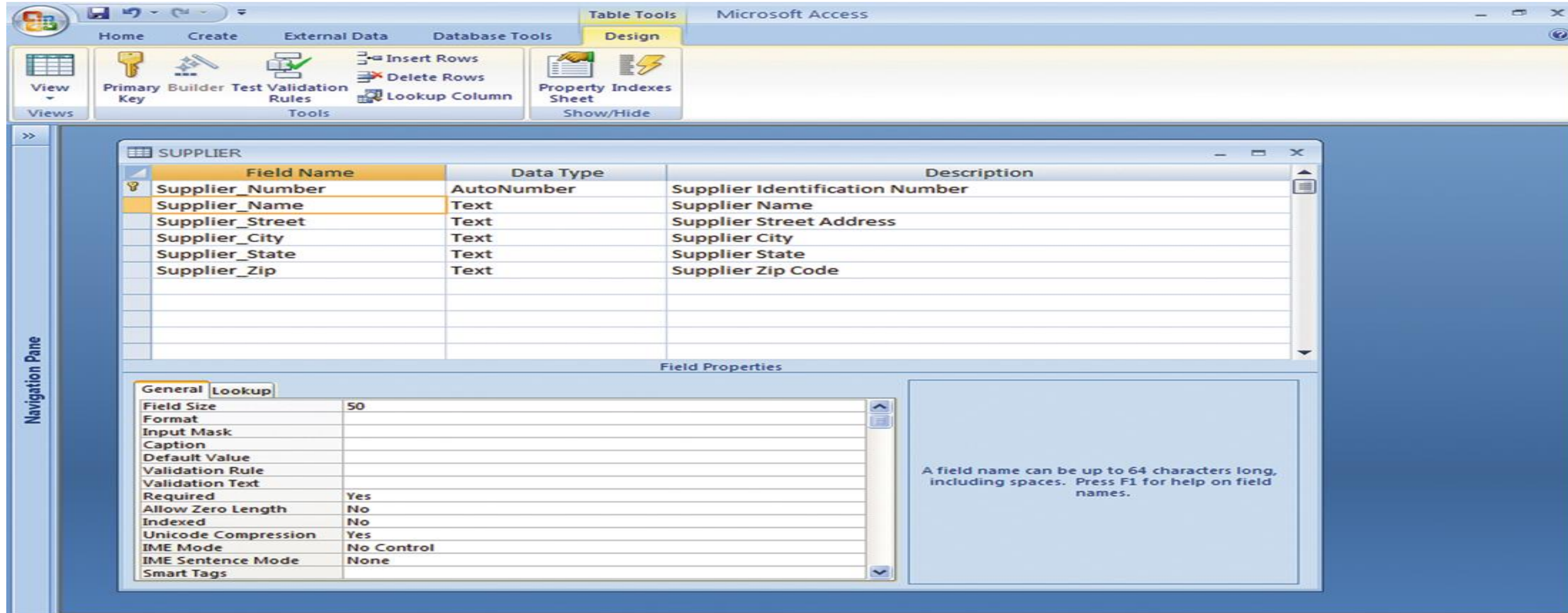
Part_Number	Part_Name	Supplier_Number	Supplier_Name
137	Door latch	8259	CBM Inc.
150	Door molding	8263	Jackson Components

Project selected columns

Capabilities of Database Management Systems

- Capabilities of Database Management Systems
  - Data definition capability: Specifies structure of database content, used to create tables and define characteristics of fields
  - Data dictionary: Automated or manual file storing definitions of data elements and their characteristics
  - Data manipulation language: Used to add, change, delete, retrieve data from database
    - Structured Query Language (SQL)
    - Microsoft Access user tools for generation SQL
  - Many DBMS have report generation capabilities for creating polished reports (Crystal Reports)

Database Approach

Three Key Techniques for Decision Making

1. Data warehousing
2. Data mining
3. Tools for accessing internal databases through the Web

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

COMPONENTS OF A DATA WAREHOUSE
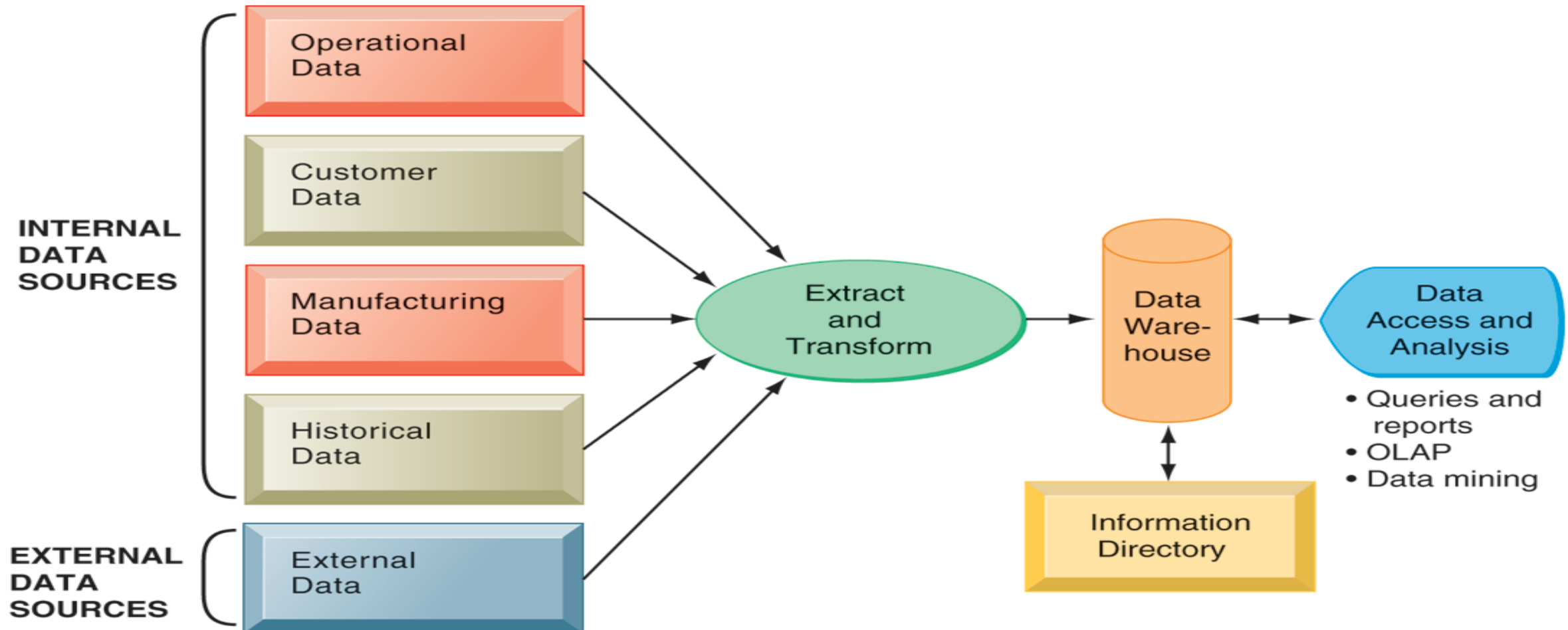
Data Mining

- – 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.

Online analytical processing (OLAP)

- 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

MULTIDIMENSIONAL DATA MODEL
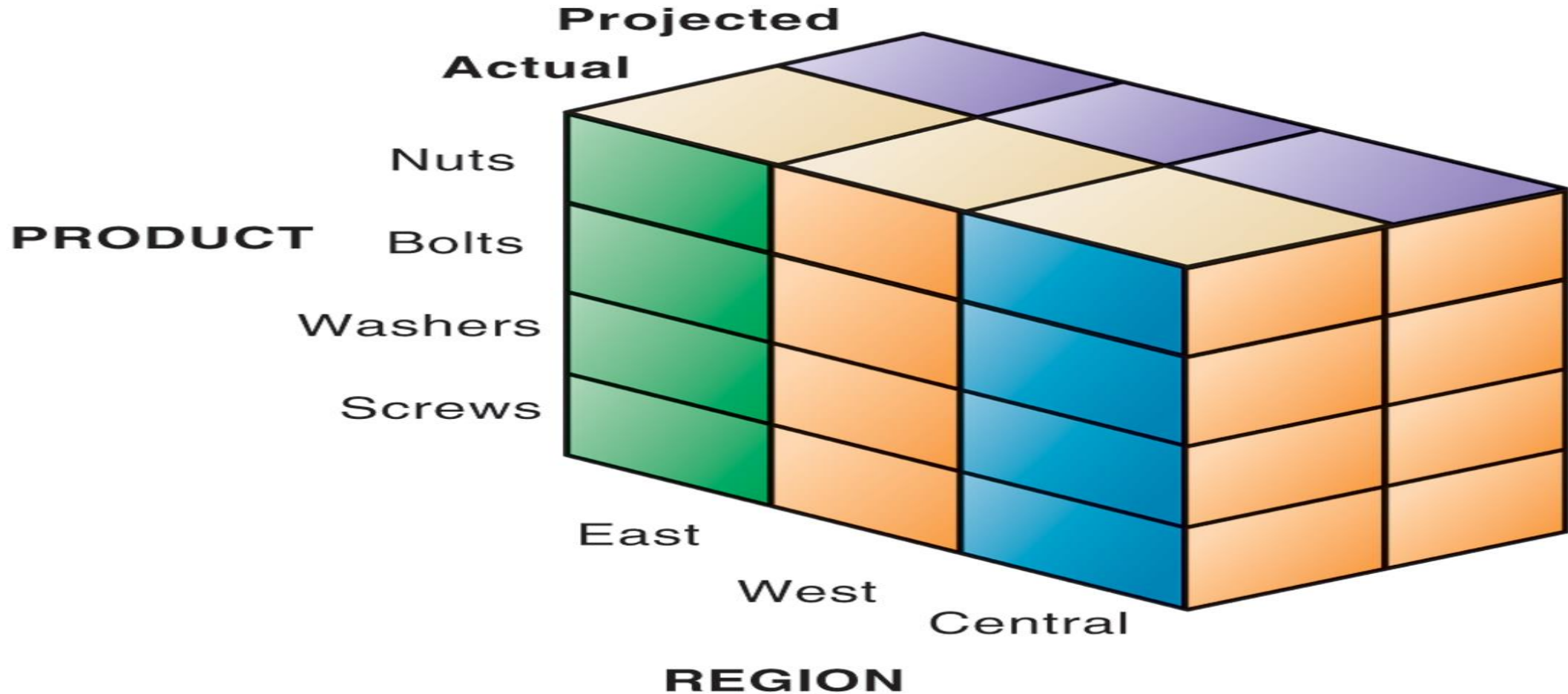
Thank You

Scan this QR Code to access HoneyBee – The Learning Platform