


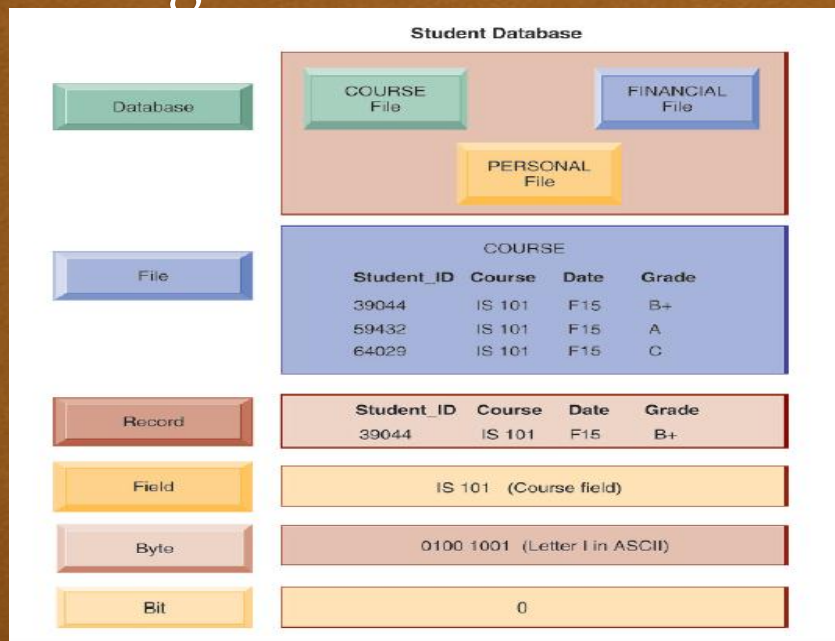
Chapter 6

**Foundations of Business
intelligence: databases and
information Management**



6-1 What is a database and how does a relational database organize data?

Image. 6.1



Entities and Attributes

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)

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

Image 6.2
Image 6.3



Organizing Data in a Relational DataBase

Establishing Relationships

red.

Order Number: 3502
Order Date: 1/15/2016

Supplier Number: 8259
Supplier Name: CBM Inc.
Supplier Address: 74 5th Avenue, Dayton, OH 45220

Order_Number	Part_Number	Part_Quantity	Part_Name	Unit_Price	Extended Price
3502	137	10	Door latch	22.00	\$220.00
3502	152	20	Door lock	31.00	620.00
3502	178	5	Door handle	10.00	50.00
Order Total:					\$890.00

6.4

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

LINE_ITEM

Order_Number	Part_Number	Part_Quantity
3502	137	10
3502	152	20
3502	178	5

ORDER

Order_Number	Order_Date
3502	1/15/2016
3503	1/16/2016
3504	1/17/2016

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	58723
8444	Bryant Corporation	4315 Mill Drive	Rochester	NY	11344

Figure 6.6

The Final Database Design with Sample Records

The final design of the database for suppliers, parts, and orders has four tables. The LINE_ITEM table is a join table that implements the many-to-many relationship between ORDER and PART.

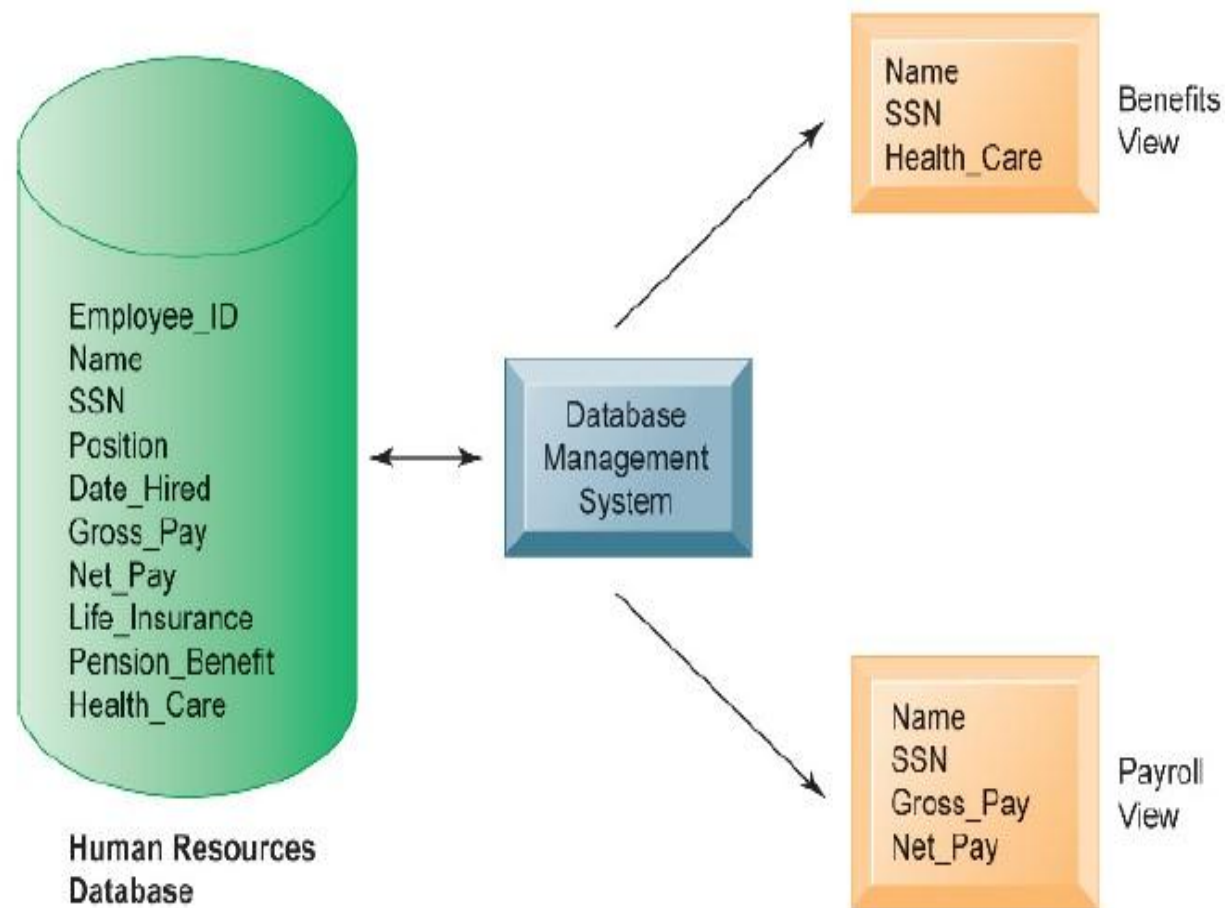
6.5

6-2 What are the principles of a database management system?

- ❧ A database management system (DBMS) is a specific type of software for creating, storing, organizing, and accessing data from a database.
- ❧ The DBMS relieves the end user or programmer from the task of understanding where and how the data are actually stored by separating the logical and physical views of the data.
- ❧ The database management software makes the physical database available for different logical views required by users.

Figure 6.8 Human Resources Database with Multiple Views

A single human resources database provides many 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



Operations of a Relational DBMS



- ❧ In a relational database, tables can be easily combined to deliver data that users require, provided that any two tables share a common data element.
- ❧ In a relational database, three basic operations are used to develop useful sets of data: select, project, and join.

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 5 th 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

Figure 6.9

The Three Basic Operations of a Relational DBMS

The select, join, and project operations enable data from two tables to be combined and only selected attributes to be displayed.

Capabilities of Database Management Systems

- ❧ A DBMS includes capabilities and tools for organizing, managing, and accessing the data in the database.
- ❧ The most important are its data definition capability, data dictionary, and data manipulation language.

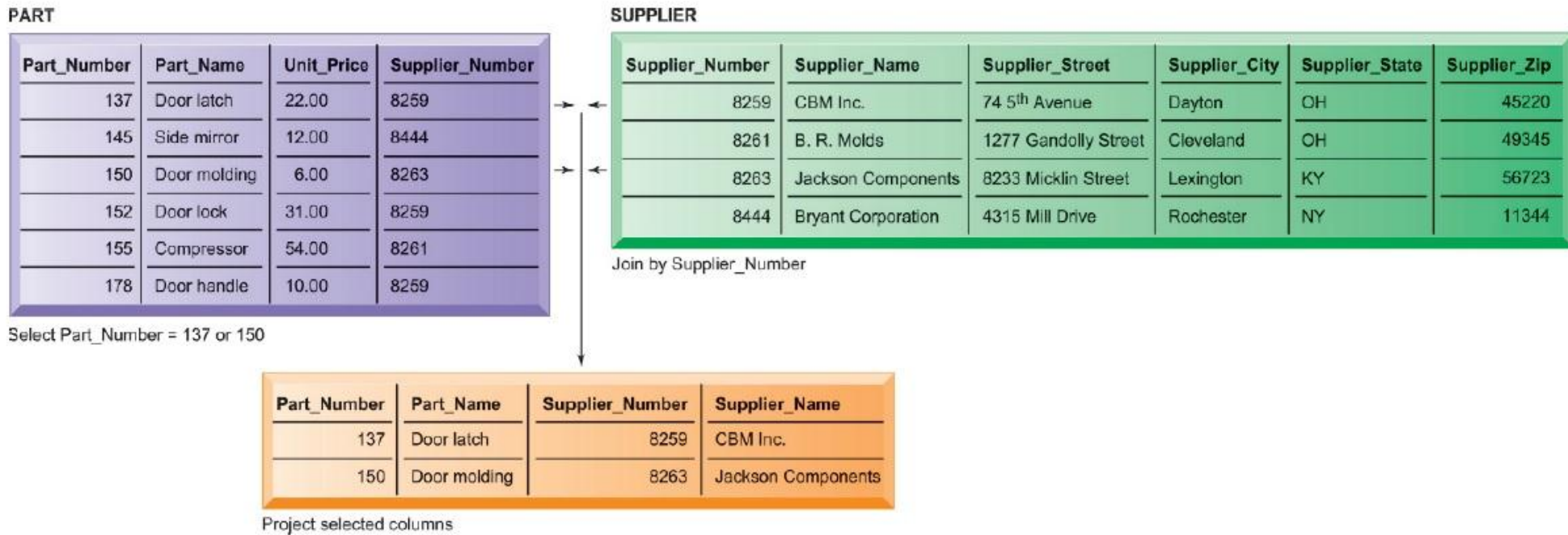


Figure 6.9
 The Three Basic Operations of a Relational DBMS
 The select, join, and project operations enable data from two tables to be combined and only selected attributes to be displayed.

```
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 6.11
 Example of a SQL Query
 Illustrated here are the SQL statements for a query to select suppliers for parts 137 or 150. They produce a list with the same results as Figure 6.9.

Nonrelational Databases and Databases in the Cloud

- ❧ Nonrelational database management systems use a more flexible data model and are designed for managing large data sets across many distributed machines and for easily scaling up or down.
- ❧ Several kinds of NoSQL databases: Oracle NoSQL Database, Amazon's SimpleDB, MetLife 's MongoDB (opensource)
- ❧ In addition to public cloud-based data management services, companies now have the option of using databases in private clouds.

and technologies for accessing
information from databases to
improve business performance
and decision making?



The Challenge of Big Data



- ❧ Businesses use their databases to keep track of basic transactions, such as paying suppliers, processing orders, serving customers, and paying employees, but they also need databases to provide information that will help the company run the business more efficiently and help managers and employees make better decisions. If a company wants to know which product is the most popular or who is its most profitable customer, the answer lies in the data.
- ❧ We now use the term big data to describe these data sets with volumes so huge that they are beyond the ability of typical DBMS to capture, store, and analyze.
- ❧ Big data doesn't designate any specific quantity but usually refers to data in the petabyte and exabyte range – in other words, billions to trillions of records, respectively, from different sources. Big data are produced in much larger quantities and much more rapidly than traditional data.
- ❧ Businesses are interested in big data because they contain more patterns and interesting relationships than smaller data sets, with the potential to provide new insights into customer behavior, weather patterns, financial market activity, or other phenomena.

Business Intelligence Infrastruktur

☞ Data Warehouses and Data Marts

Data warehouse is a database that stores current and historical data of potential interest to decision makers throughout the company. The data originate in many core operational transaction systems, such as systems for sales, customer accounts, and manufacturing, and can include data from website transactions.

A data mart is a subset of a data warehouse, in which a summarized or highly focused portion of the organization's data is placed in a separate database for a specific population of users. For example, a company might develop marketing and sales data marts to deal with customer information.

☞ Hadoop

Hadoop is an open-source software framework the Apache Software Foundation manages that enables distributed parallel processing of huge amounts of data across inexpensive computers.

Hadoop can process large quantities of any kind of data, including structured transactional data, loosely structured data such as Facebook and Twitter feeds, complex data such as web server log files, and unstructured audio and video data.



∞ In-Memory Computing

Another way of facilitating big data analysis is to use in-memory computing, which relies primarily on a computer's main memory (RAM) for data storage.

∞ Analytic Platforms

These analytic platforms, such as IBM PureData System for Analytics, feature preconfigured hardware-software systems that are specifically designed for query processing and analytics.

Analytical Tools: Relationships, Patterns, Trends

☞ Online Analytical Processing (OLAP)

OLAP supports multidimensional data analysis, enabling users to view the same data in different ways using multiple dimensions. Each aspect of information – product, pricing, cost, region, or time period – represents a different dimension.

OLAP enables users to obtain online answers to ad hoc questions such as these in rapid time, even when the data are stored in very large databases, such as sales figures for multiple years.

☞ Data Mining

Data mining is more discovery-driven. Data mining provides insights into corporate data that cannot be obtained with OLAP by finding hidden patterns and relationships in large databases and inferring rules from them to predict future behavior.



☞ Text Mining and Web Mining

Text mining tools are now available to help businesses analyze these data. These tools can extract key elements from unstructured big data sets, discover patterns and relationships, and summarize the information.

The discovery and analysis of useful patterns and information from the World Wide Web is called web mining. Businesses might turn to web mining to help them understand customer behavior, evaluate the effectiveness of a particular website, or quantify the success of a marketing campaign.



THANK YOU ☺



- ❧ Ruri: Fadly belum menjelaskan tentang pentingnya DBMS... beri contoh yang sering kita temui biar gampang aja memahaminya
- ❧ Ricko: menambahkan adit, mengelola data juga perlu ketelatenan dari diri sendiri
- ❧ Barqi Tobroni: Mas Ocid harus pake ppt, jangan Cuma teori aja, biar gampang memahami