

1.

ความรู้เบื้องต้นเกี่ยวกับระบบฐานข้อมูล

DB304 การจัดการฐานข้อมูลทางธุรกิจดิจิทัล

อาจารย์เฉลิมพล ศิริกายน และ ดร.ปาณิตา ฐุสรานนท์

หัวข้อที่สำคัญ

- ความหมายของข้อมูล สารสนเทศ และระบบสารสนเทศ
- ความรู้พื้นฐานเรื่องเขตข้อมูล ระเบียบ และเพิ่มข้อมูล
- ฐานข้อมูล (database)
- การประมวลผลแบบเพิ่มข้อมูล กับ การประมวลผลแบบฐานข้อมูล
- ระบบจัดการฐานข้อมูล (Database Management System)

ความหมายของข้อมูล สารสนเทศ และระบบสารสนเทศ

ข้อมูล (Data)

สิ่งที่อยู่ในรูปแบบของ ตัวเลข ตัวอักษร สัญลักษณ์ รวมถึง ภาพ เสียง

สารสนเทศ (Information)

ข้อมูลเชิงสรุปที่ได้หลังจากการประมวลผล เช่น การคำนวณ การจัดกลุ่ม การเรียงลำดับ หรือสรุปผล เพื่อสร้างเป็นรายงานสรุป หรือจัดให้อยู่ในรูปแบบที่เหมาะสมต่อการนำเสนอสำหรับผู้บริหาร

ระบบสารสนเทศ (Information System)

ชุดขององค์ประกอบที่ทำหน้าที่รวบรวม ประมวลผล จัดเก็บ และแจกจ่าย สารสนเทศ เพื่อช่วยในการตัดสินใจ และการควบคุมในองค์กร การทำงานของระบบ สารสนเทศประกอบไปด้วยกิจกรรม 3 อย่าง คือ การนำข้อมูลเข้าสู่ระบบ (Input) การประมวลผล (Processing) และ การนำเสนอผลลัพธ์ (Output)

Example: Tony's snack company data

Order Date	Customer	Sales Representative	Product	Qty	Unit Price	Total Sales	Unit Cost	Total Cost	Profit
4-Jan	Walmart	PJ Helgoth	Doritos	41	\$24	\$ 984	\$18	\$738	\$246
4-Jan	Walmart	Roberta Cross	Ruffles	90	\$15	\$1,350	\$10	\$900	\$450
5-Jan	Safeway	Craig Schultz	Ruffles	27	\$15	\$ 405	\$10	\$270	\$135
6-Jan	Walmart	Roberta Cross	Ruffles	67	\$15	\$1,005	\$10	\$670	\$335
7-Jan	7-Eleven	Craig Schultz	Pringles	79	\$12	\$ 948	\$ 6	\$474	\$474
7-Jan	Walmart	Roberta Cross	Ruffles	52	\$15	\$ 780	\$10	\$520	\$260
8-Jan	Kroger	Craig Schultz	Ruffles	39	\$15	\$ 585	\$10	\$390	\$195
9-Jan	Walmart	Craig Schultz	Ruffles	66	\$15	\$ 990	\$10	\$660	\$330
10-Jan	Target	Craig Schultz	Ruffles	40	\$15	\$ 600	\$10	\$400	\$200
11-Jan	Walmart	Craig Schultz	Ruffles	71	\$15	\$1,065	\$10	\$710	\$355

- ▷ Who are my best customer?
- ▷ Who are my least-profitable customer?
- ▷ What is my best-selling product?

Example: Tony's data sorted by customer "Walmart" and sales representative "Roberta Cross"

Order Date	Customer	Sales Representative	Product	Quantity	Unit Price	Total Sales	Unit Cost	Total Cost	Profit
26-Apr	Walmart	Roberta Cross	Fritos	86	\$ 19	\$ 1,634	\$ 17	\$ 1,462	\$ 172
29-Aug	Walmart	Roberta Cross	Fritos	76	\$ 19	\$ 1,444	\$ 17	\$ 1,292	\$ 152
7-Sep	Walmart	Roberta Cross	Fritos	20	\$ 19	\$ 380	\$ 17	\$ 340	\$ 40
22-Nov	Walmart	Roberta Cross	Fritos	39	\$ 19	\$ 741	\$ 17	\$ 663	\$ 78
30-Dec	Walmart	Roberta Cross	Fritos	68	\$ 19	\$ 1,292	\$ 17	\$ 1,156	\$ 136
7-Jul	Walmart	Roberta Cross	Pringles	79	\$ 18	\$ 1,422	\$ 8	\$ 632	\$ 790
6-Aug	Walmart	Roberta Cross	Pringles	21	\$ 12	\$ 252	\$ 6	\$ 126	\$ 126
2-Oct	Walmart	Roberta Cross	Pringles	60	\$ 18	\$ 1,080	\$ 8	\$ 480	\$ 600
15-Nov	Walmart	Roberta Cross	Pringles	32	\$ 12	\$ 384	\$ 6	\$ 192	\$ 192
21-Dec	Walmart	Roberta Cross	Pringles	92	\$ 12	\$ 1,104	\$ 6	\$ 552	\$ 552
28-Feb	Walmart	Roberta Cross	Ruffles	67	\$ 15	\$ 1,005	\$ 10	\$ 670	\$ 335
6-Mar	Walmart	Roberta Cross	Ruffles	8	\$ 15	\$ 120	\$ 10	\$ 80	\$ 40
16-Mar	Walmart	Roberta Cross	Ruffles	68	\$ 15	\$ 1,020	\$ 10	\$ 680	\$ 340
28-Oct	Walmart	Roberta Cross	Ruffles	50	\$ 15	\$ 750	\$ 10	\$ 500	\$ 250
21-Nov	Walmart	Roberta Cross	Ruffles	79	\$ 15	\$ 1,185	\$ 10	\$ 790	\$ 395
29-Jan	Walmart	Roberta Cross	Sun Chips	5	\$ 22	\$ 110	\$ 18	\$ 90	\$ 20
12-Apr	Walmart	Roberta Cross	Sun Chips	85	\$ 22	\$ 1,870	\$ 18	\$ 1,530	\$ 340
16-Jun	Walmart	Roberta Cross	Sun Chips	55	\$ 22	\$ 1,210	\$ 18	\$ 990	\$ 220
				1,206	\$383	\$20,243	\$273	\$14,385	\$5,858

Example: Information gained after analyzing Tony's data

Tony's Business Information	Name	Total Profit
Who is Tony's best customer by total sales?	Walmart	\$ 560,789
Who is Tony's least-valuable customer by total sales?	Walgreens	\$45,673
Who is Tony's best customer by profit?	7-Eleven	\$ 324,550
Who is Tony's least-valuable customer by profit?	King Soopers	\$ 23,908
What is Tony's best-selling product by total sales?	Ruffles	\$ 232,500
What is Tony's weakest-selling product by total sales?	Pringles	\$ 54,890
What is Tony's best-selling product by profit?	Tostitos	\$ 13,050
What is Tony's weakest-selling product by profit?	Pringles	\$ 23,000
Who is Tony's best sales representative by profit?	R. Cross	\$1,230,980
Who is Tony's weakest sales representative by profit?	Craig Schultz	\$ 98,980
What is the best sales representative's best-selling product by total profit?	Ruffles	\$ 98,780
Who is the best sales representative's best customer by total profit?	Walmart	\$ 345,900
What is the best sales representative's weakest-selling product by total profit?	Sun Chips	\$ 45,600
Who is the best sales representative's weakest customer by total profit?	Krogers	\$ 56,050

ความสำคัญของข้อมูล

ระดับผู้ปฏิบัติงาน: ช่วยให้การปฏิบัติงานมีประสิทธิภาพมากยิ่งขึ้น เช่น

- การตรวจสอบการชำระเงิน
- รายงานการขายประจำวัน ประจำเดือน หรือปี
- การตรวจสอบจำนวนสินค้าคงเหลือของแต่ละคลังสินค้า เป็นต้น

ระดับผู้บริหาร: สามารถนำข้อมูลมาวิเคราะห์เพื่อกำหนดกลยุทธ์ต่างๆของบริษัท

- วิเคราะห์ข้อมูลการขาย ทำการวิเคราะห์ว่า กลุ่มสินค้าใดบ้างที่ลูกค้ามักจะซื้อไปด้วยกันเสมอ เพื่อแนะนำลูกค้าเมื่อสั่งซื้อสินค้า ก แล้วควรซื้อสินค้า ข หรือ ค ด้วย เป็นต้น >> ซึ่งสามารถกำหนดเป็นกลยุทธ์ในการขายสินค้าเป็นชุด (Package)

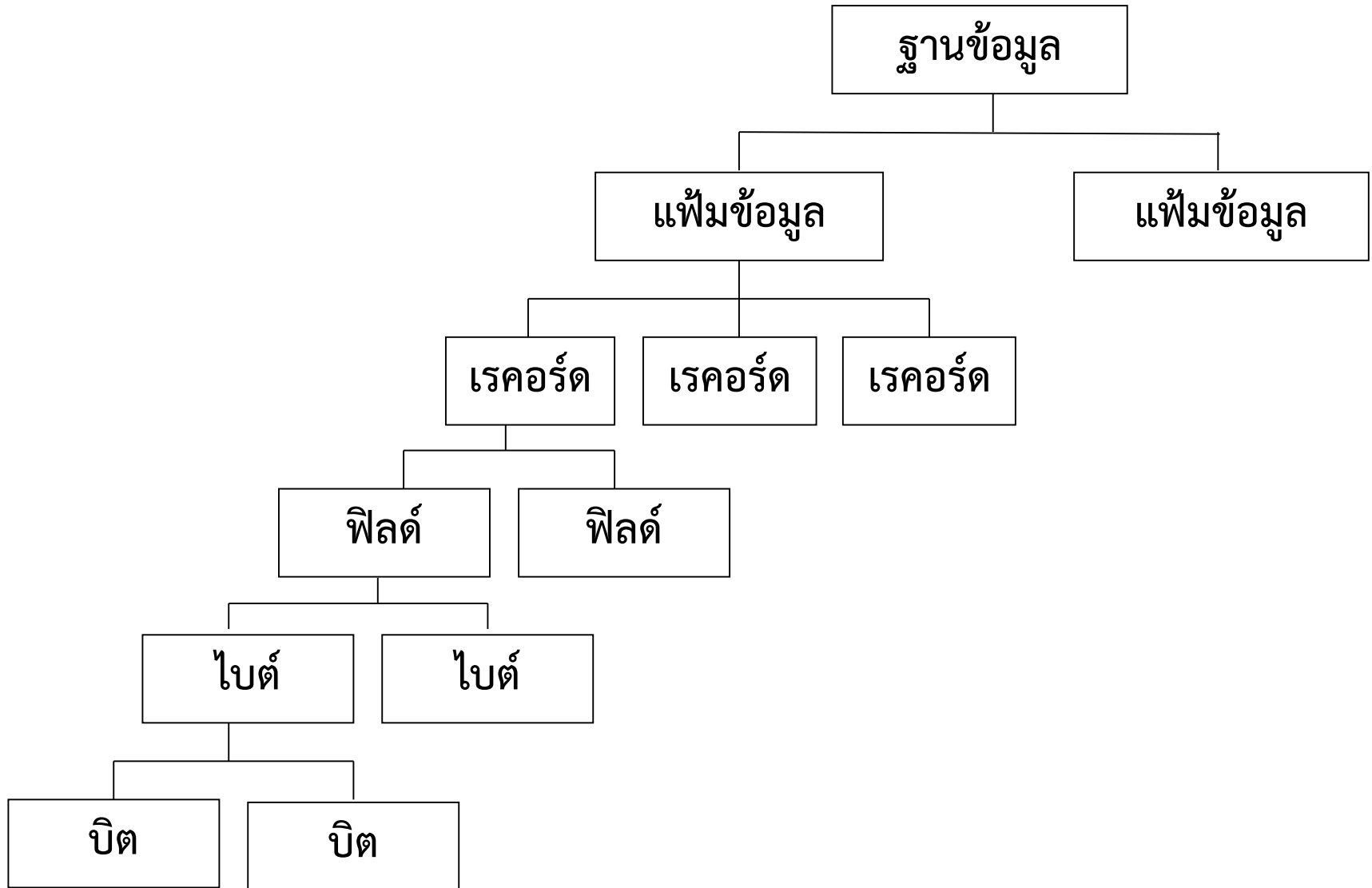
ความรู้พื้นฐานเรื่องเขตข้อมูล ระเบียน และแฟ้มข้อมูล



The hierarchy of data

Data Level	Example
Character	223287695 Doe John 1987
Field	223287695 Doe John 1987
Record	200987845 Jewel Mark 1987 223287695 Doe John 1987 249876587 Smith Justin 1987
File	<div>Student ID</div> <div>Last Name</div> <div>First Name</div> <div>200987845 Jewel Mark 1995</div> <div>223287695 Doe John 1995</div> <div>349876587 Smith Justin 1994</div> <div>410098456 Jones Jose 1993</div> <div>... ..</div> <div>Student File</div>
Database	<div>Year of Birth</div> <div>First Name</div> <div>Last Name</div> <div>Department</div> <div>Campus Phone Number</div> <div>200987845 Jewel Mark 1987</div> <div>223287695 Doe John 1987</div> <div>349876587 Smith Justin 1986</div> <div>410098456 Jones Jose 1985</div> <div>... ..</div> <div>ACC Dor Avi 9-8776</div> <div>MKT Jenings Rich 9-8776</div> <div>FIN Dor Jim 9-8776</div> <div>... ..</div> <div>Student File</div> <div>Professor File</div>

ความรู้พื้นฐานเรื่องเขตข้อมูล ระเบียน และแฟ้มข้อมูล



ความรู้พื้นฐานเรื่องเขตข้อมูล ระเบียน และแฟ้มข้อมูล

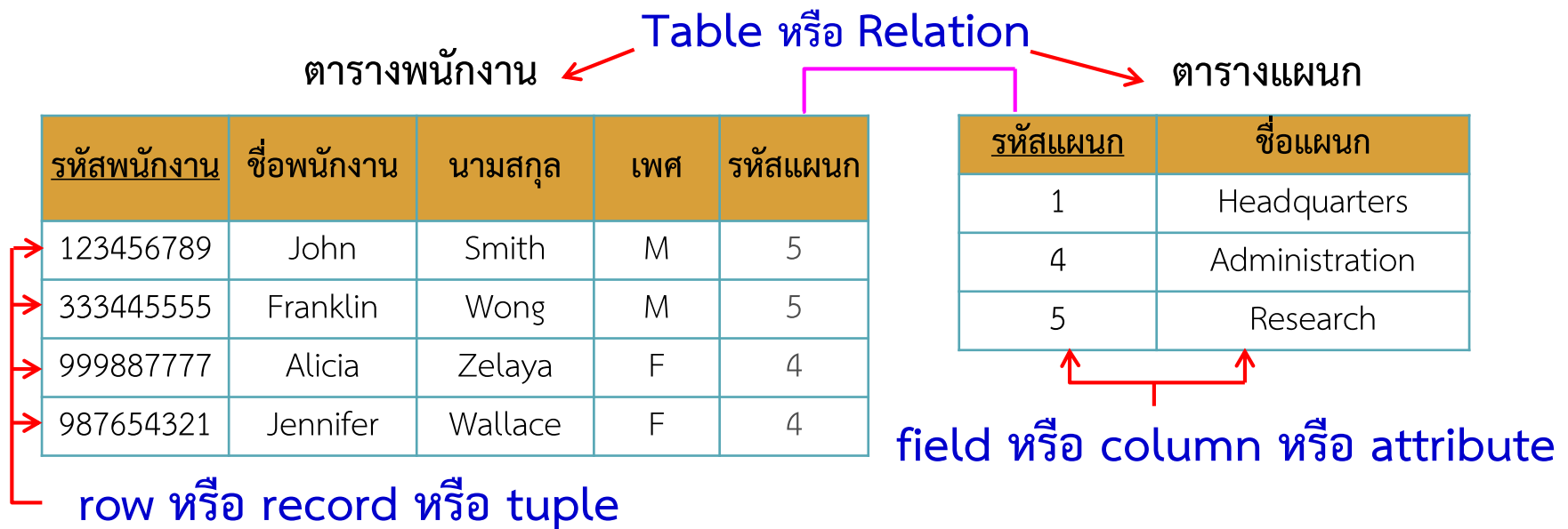
เขตข้อมูล คือ การนำเอาอักขระหลายๆ ตัวรวมกัน โดยมีความหมายอย่างใดอย่างหนึ่ง หรืออาจเรียกอีกอย่างว่า **ฟิลด์ (Field)** หรือ **แอทริบิวต์ (Attribute)** หรือ **คอลัมน์ (Column)** แบ่งได้เป็น 3 ประเภท

- ตัวเลข: จำนวนเต็ม ทศนิยม บวก ลบ
- ตัวอักษร: อักขระที่เป็นตัวอักษร หรือ ช่องว่าง
- อักขระ: อักขระที่เป็นตัวอักษร หรือ ตัวเลข

ระเบียน คือ กลุ่มของฟิลด์ที่สัมพันธ์กันหรืออาจเรียกอีกอย่างหนึ่งว่า **เรคอร์ด (Record)** หรือ **แถว (Row)** หรือ **ทูเพิล (Tuple)**

ความรู้พื้นฐานเรื่องเขตข้อมูล ระเบียน และแฟ้มข้อมูล

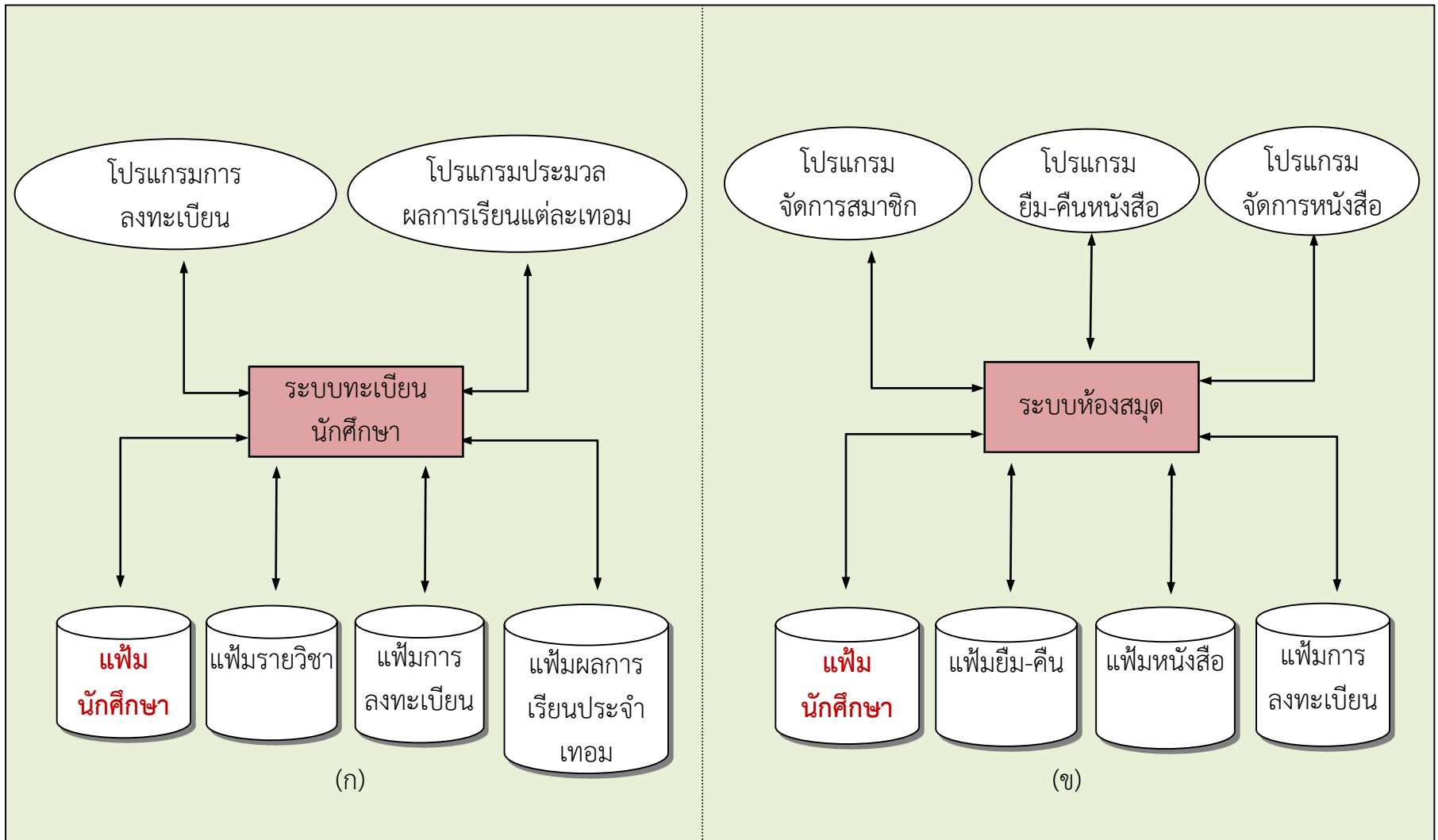
แฟ้มข้อมูล/ตารางข้อมูล คือ การเก็บข้อมูลที่มีการนำระเบียนหลายๆ ระเบียนที่มีความสัมพันธ์กันมารวมกัน เช่น ตารางข้อมูลนักศึกษา, ตารางข้อมูลอาจารย์/เจ้าหน้าที่, ตารางข้อมูลอาคาร/สถานที่, ตารางข้อมูลการจัดตารางการสอน ฯลฯ ขึ้นอยู่กับการออกแบบของผู้ออกแบบระบบ



ฐานข้อมูล (database)

- กลุ่มของข้อมูลที่มีความสัมพันธ์กัน และถูกนำมาจัดเก็บในที่เดียวกัน
- ข้อมูลอาจเก็บไว้ในแฟ้มข้อมูลเดียวกัน หรือแยกเก็บหลายๆ แฟ้มข้อมูล แต่ต้องมีการสร้างความสัมพันธ์ระหว่างข้อมูล เพื่อประสิทธิภาพในการจัดการข้อมูล
- ในการจัดเก็บข้อมูลใน “**ระบบฐานข้อมูล**” มีข้อดีกว่าการจัดเก็บข้อมูลใน “**ระบบแฟ้มข้อมูล**” สรุปประเด็นหลักๆ ได้ดังนี้
 - มีการใช้ข้อมูลร่วมกัน (data sharing)
 - ลดความซ้ำซ้อนของข้อมูล (reduce data redundancy)
 - ข้อมูลมีความถูกต้องมากขึ้น (improved data integrity)
 - เพิ่มความปลอดภัยให้กับข้อมูล (increased security)
 - มีความเป็นอิสระของข้อมูล (data independency)

ตัวอย่างระบบการประมวลผลแบบเพิ่มข้อมูล



ระบบการประมวลผลแบบเพิ่มข้อมูล

ข้อดี

- การประมวลผลข้อมูลทำได้อย่างรวดเร็ว
- ค่าลงทุนในเบื้องต้นจะต่ำ เนื่องจากอาจไม่จำเป็นต้องใช้คอมพิวเตอร์ที่มีความสามารถมากก็สามารถทำการประมวลผลข้อมูลได้
- โปรแกรมประยุกต์แต่ละโปรแกรมสามารถควบคุมการใช้ข้อมูลในเพิ่มข้อมูลของตนเองได้

ข้อเสีย

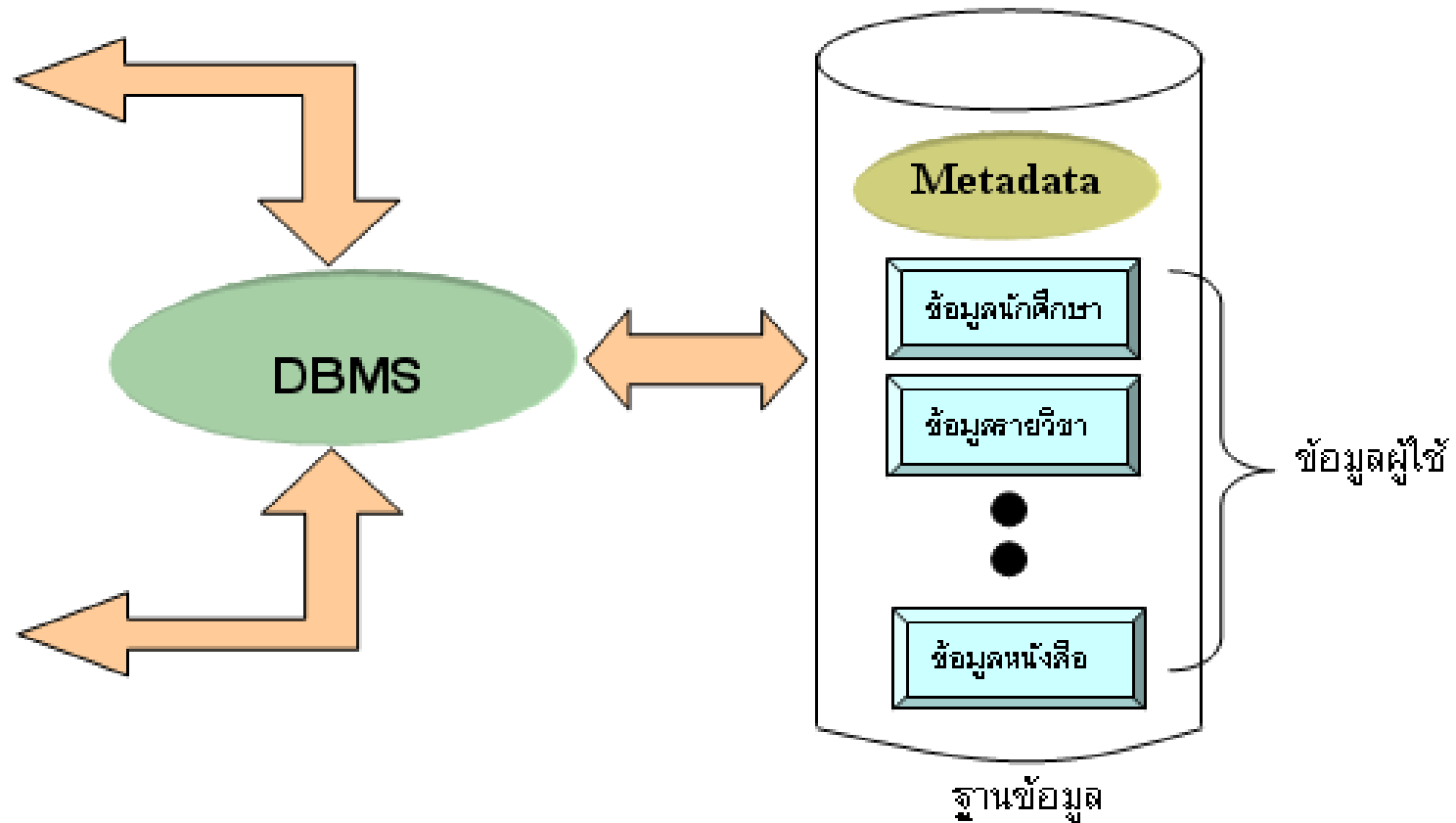
- การขึ้นต่อกันระหว่างโปรแกรมและข้อมูล
- ความซ้ำซ้อนกันของข้อมูล
- ความขัดแย้งกันของข้อมูล
- ค่าใช้จ่ายของแต่ละหน่วยงาน
- ข้อจำกัดในการใช้ข้อมูลร่วมกัน
- ความซับซ้อนของการประมวลผลจากหลายเพิ่มข้อมูล

ตัวอย่างระบบการประมวลผลแบบฐานข้อมูล

โปรแกรมประยุกต์สำหรับ
ระบบทะเบียนนักศึกษา



โปรแกรมประยุกต์สำหรับ
ระบบห้องสมุด



ระบบการประมวลผลแบบฐานข้อมูล

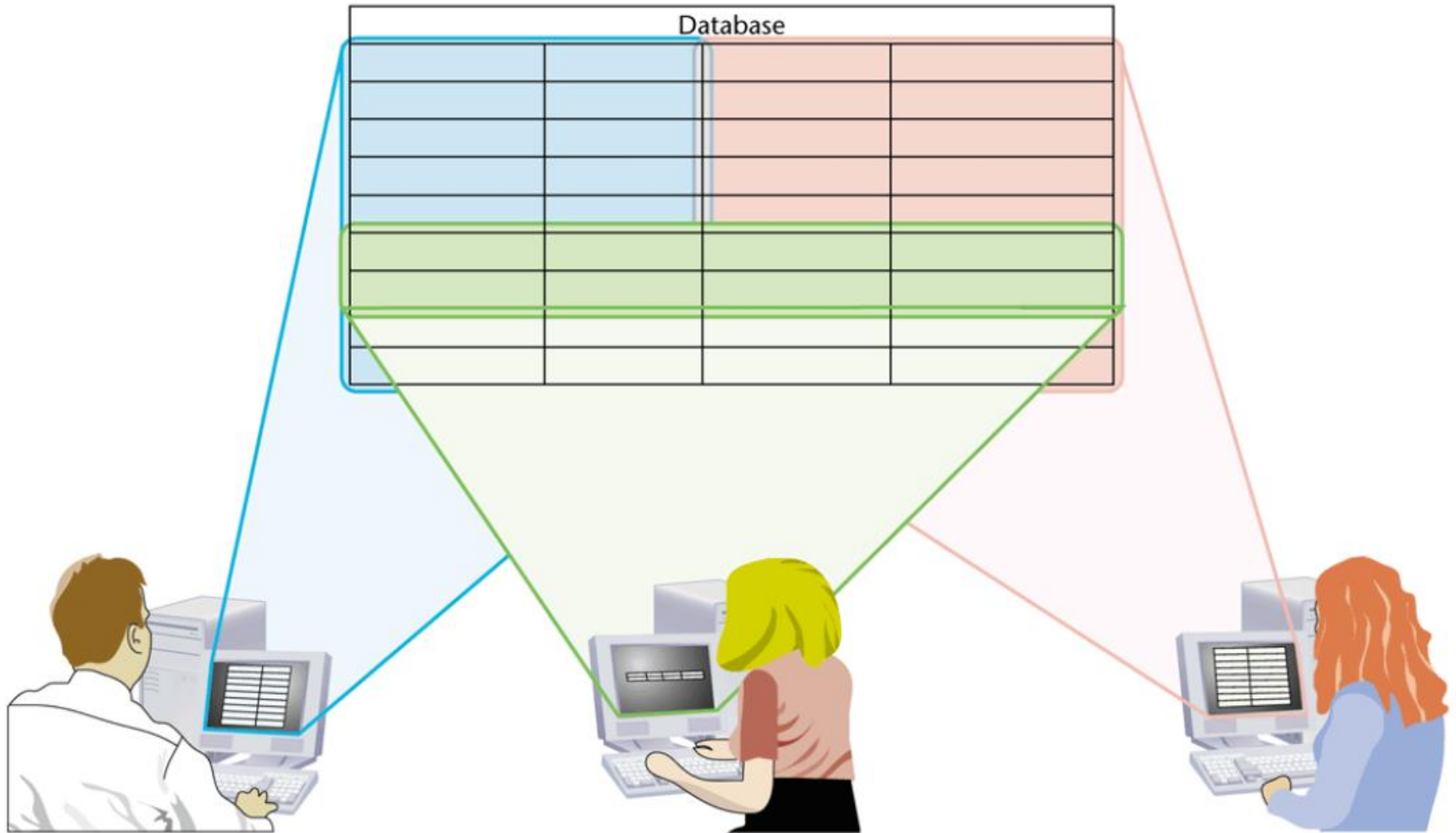
ข้อดี

- มีความเป็นอิสระระหว่างโปรแกรมประยุกต์และข้อมูล
- ลดความซ้ำซ้อนและความขัดแย้งกันของข้อมูล
- ข้อมูลถูกเก็บอยู่รวมกันและสามารถใช้ข้อมูลร่วมกันได้
- สามารถควบคุมความคงสภาพของข้อมูลได้
- การจัดการฐานข้อมูลในฐานข้อมูลจะทำได้ง่าย
- มีการควบคุมการใช้งานฐานข้อมูลของผู้ใช้งานหลายคน

ข้อเสีย

- ค่าใช้จ่ายในการติดตั้งระบบจัดการฐานข้อมูล
- ค่าใช้จ่ายระหว่างดำเนินงาน

Using database views, users may have different structural pictures of a database



Using database views, users may have different structural pictures of a database

View of Human Resource Manager						
Hourly Rate	SSN	Name	D.O.B.	Hire Date	Marital Status	Benefits Code

View of Payroll Personnel			
SSN	Hourly Rate	Benefits Code	Hours Worked

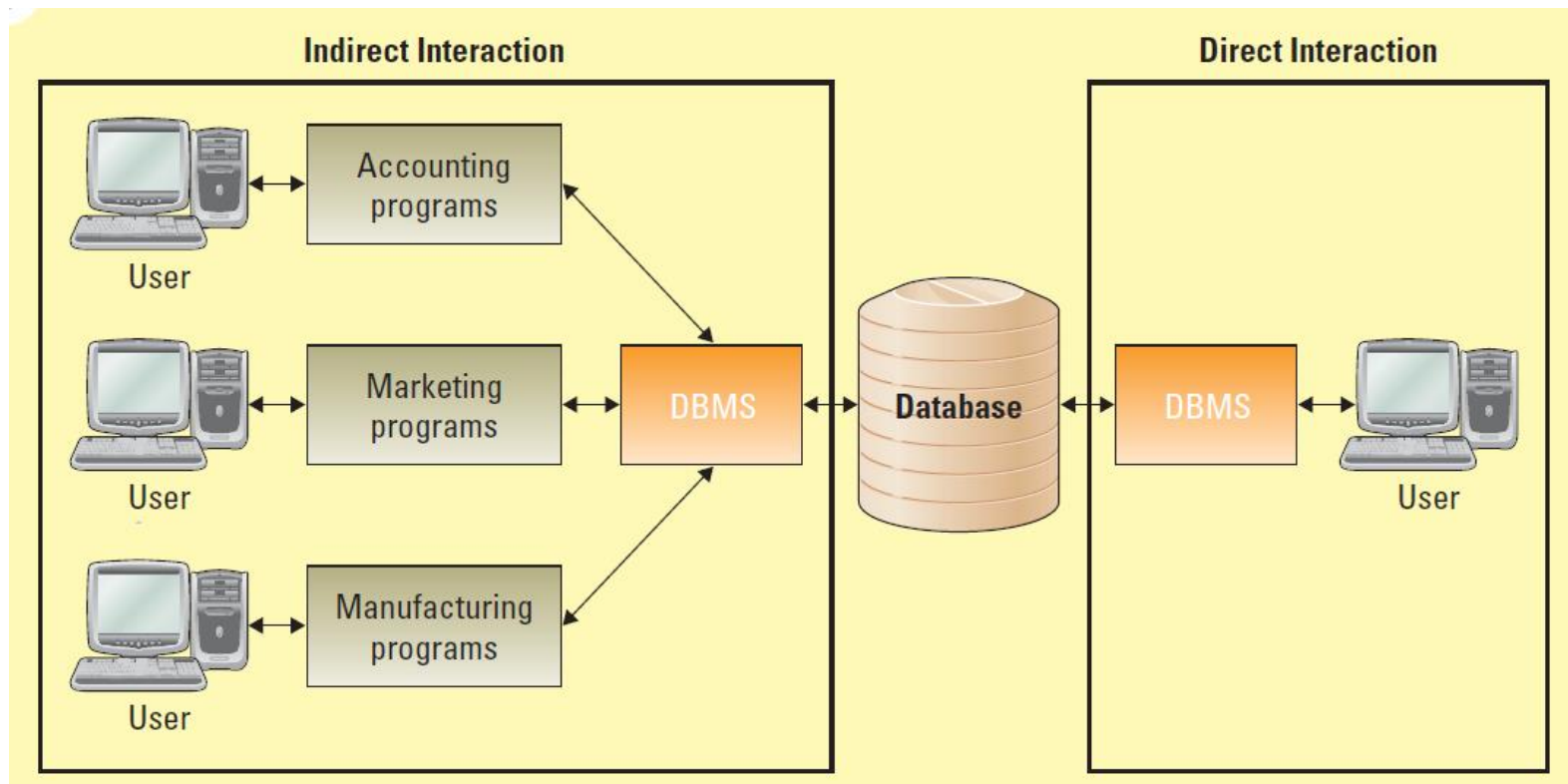
View of Project Manager	
Name	Hours Worked

ระบบจัดการฐานข้อมูล (Database Management System)

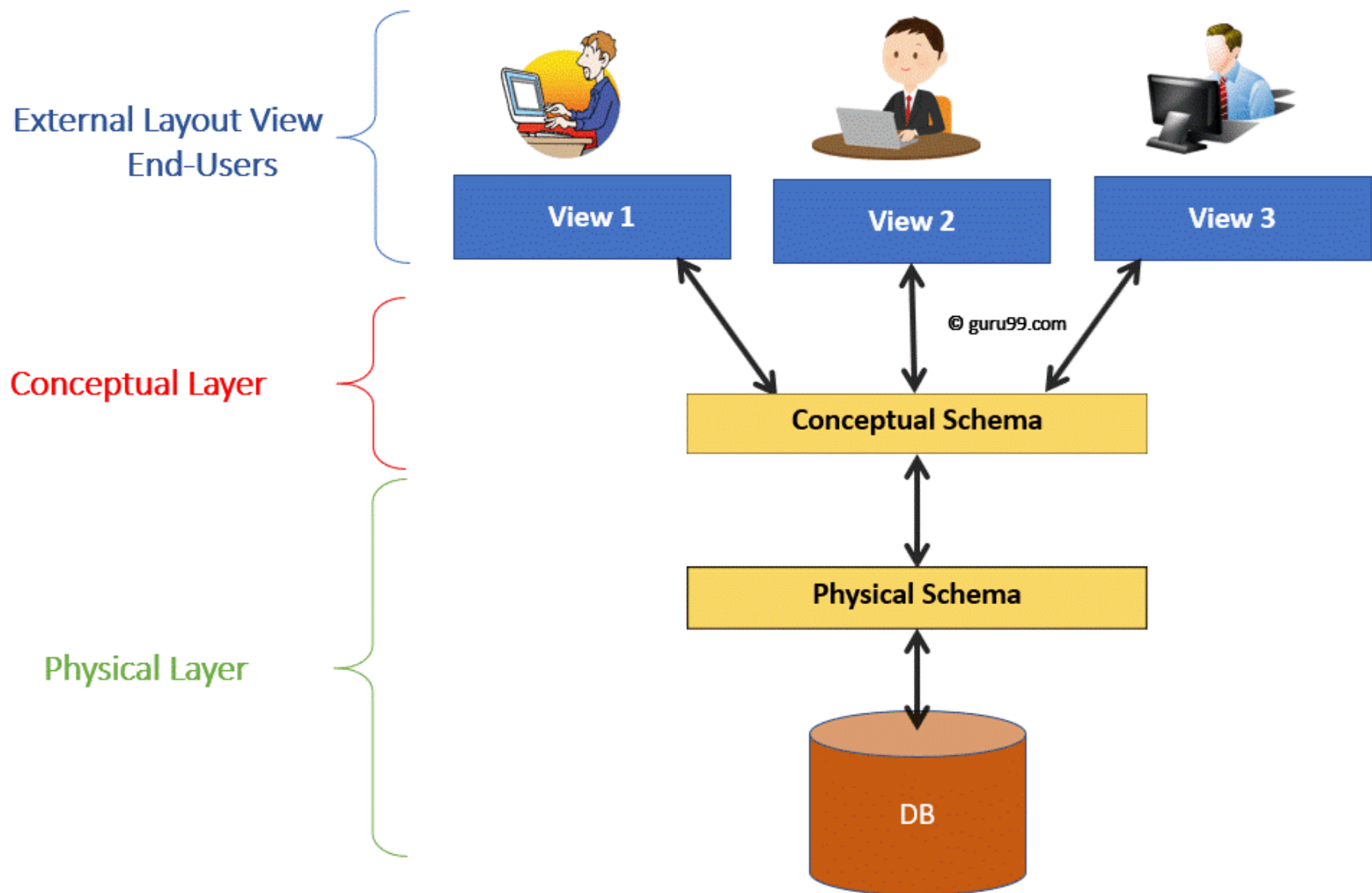
- ระบบจัดการฐานข้อมูล (Database Management System: DBMS)
คือ โปรแกรมที่ทำหน้าที่เป็นตัวกลางในการติดต่อระหว่างผู้ใช้กับฐานข้อมูล เพื่อจัดการและควบคุมความถูกต้อง ความซ้ำซ้อน และความสัมพันธ์ระหว่างข้อมูลต่างๆ ภายในฐานข้อมูล
- ต่างจากระบบแฟ้มข้อมูลที่หน้าที่เหล่านี้จะเป็นหน้าที่ของโปรแกรมเมอร์
- ระบบจัดการฐานข้อมูลที่นิยมใช้กันในปัจจุบัน เช่น Oracle, IBM DB2, Microsoft SQL Server, Sybase, PostgreSQL, MySQL, Interbase, Firebird, SAP DB, Microsoft Office Access, SQLite

ระบบจัดการฐานข้อมูล (Database Management System)

- ▷ Database management system (DBMS) is a program used to:
- Build databases
 - Populate a database with data
 - Manipulate data in a database



DBMS three level architecture



RDBMS ประเภทต่าง ๆ

DB-Engines Ranking of Relational DBMS

