

บทที่ 10 การประมวลผลข้อมูล

โดย
ร.ศ. พร摊ิภา ไพบูลย์นิมิตร

วิชา 204100 Information Technology and Modern Life

1

ข้อมูลดิบ

- ข้อมูลที่รวบรวมมาจากแหล่งข้อมูลต่าง ๆ
- ตัวเลข ตัวอักษร รูปภาพ เสียง อุณหภูมิ ความเร็ว
- คะแนนสอบของนักศึกษา 1 กลุ่ม

วิชา 204100 Information Technology and Modern Life

3

การประมวลผล (Processing)

- การดำเนินการกับข้อมูลดิบ (Raw Data) เพื่อให้ได้ ข้อมูลสารสนเทศ (Information)



วิชา 204100 Information Technology and Modern Life

2

ข้อมูลสารสนเทศ

- ข้อมูลที่ผ่านการประมวลผลแล้ว สามารถนำไปใช้ประโยชน์ได้
- ค่าเฉลี่ยคะแนนสอบของนักศึกษา 1 กลุ่ม

วิชา 204100 Information Technology and Modern Life

4

ประเภทของการประมวลผล

- การประมวลผลด้วยมือ (Manual Data Processing)
- การประมวลผลอัตโนมัติ (Automatic Data Processing)
 - Online Processing
 - Batch Processing

Online Processing

- Real-time Processing
- ระบบสำรองที่นั่งในเครื่องบิน
- Time Sharing
- เป็นการประมวลผลแบบโดยตรงหรือสุ่ม (Random)
- สื่อที่มีการเข้าถึงแบบโดยตรง เช่น ฮาร์ดดิสก์

Online Processing (ต่อ)

- ข้อมูลทันต่อเหตุการณ์
- สามารถทำรายการได้ทันที
- I/O Device ต้องเชื่อมต่อกับ CPU ตลอดเวลา

Batch Processing

- ไฟล์หลัก (Master Files)
 - ไฟล์รายการข้อมูลสินค้า
- ไฟล์รายการเปลี่ยนแปลง (Transaction Files)
 - รายการขายสินค้า
 - รายการซื้อสินค้า

Batch Processing (ต่อ)

- ประมวลผลตามรายการเปลี่ยนแปลงทีละรายการ
- ประมวลผลทุกวันหรือทุกเดือน
- เป็นการประมวลผลแบบลำดับ (Sequential)
- สื้อที่มีการเข้าถึงแบบลำดับ เช่น เทปแม่เหล็ก
- ค่าใช้จ่ายต่ำ

การประมวลผลด้วยคอมพิวเตอร์

- การเก็บข้อมูล (Data Collection)
- การประมวลผลข้อมูล (Data Processing)
- ผลลัพธ์ (Output / Information)



Batch Processing (ต่อ)

- ข้อมูลทันไม่ทันต่อเหตุการณ์
- รายงานล่าช้าไม่ทันต่อเหตุการณ์
- I/O Device ไม่จำเป็นต้องเชื่อมต่อกับ CPU ตลอดเวลา

การเก็บข้อมูล

- วิธีการเก็บข้อมูล
- ลักษณะของข้อมูลที่ได้
- การเข้ารหัสข้อมูล
- การตรวจสอบความถูกต้องของข้อมูล
- การเก็บข้อมูลในสื่อบันทึกข้อมูล

วิธีการเก็บข้อมูล

- สังเกต (Observation)
- วัดค่า (Measurement)
- เอกสาร (Document)
- สัมภาษณ์ (Interview)

แบบสอบถาม

ชื่อ	นาย	นามสกุล	
อายุ	พี่สาว 20	ระดับการศึกษา	ทักษะ เช่น...
ตำแหน่ง		ชื่อเมือง	
หน่วยงาน		ประเภทหน่วยงาน	หน่วยงานราชการ
ที่อยู่บ้านที่		ถนน	
แขวง/ตำบล		เขต/อำเภอ	
Province	กรุงเทพมหานคร	รหัสไปรษณีย์	
โทรศัพท์		โทรศัพท์	

วิธีการเก็บข้อมูล

- แบบสอบถาม (Questionnaire)
 - Manual
 - Internet

ลักษณะของข้อมูลที่ดี

- ถูกต้อง (Accuracy)
- ปัจจุบัน (Recency)
- สมบูรณ์ (Completeness)

การเข้ารหัสข้อมูล

- ประยุกต์ใช้ในการเก็บข้อมูล
- มีความเร็วในการค้นหาสูง
- ใช้ตัวเลขหรือรหัสแทนข้อความ
- การปรับอัตราส่วนของจำนวน

1 กรุงเทพ
2 เชียงใหม่
3 ลำพูน
....

หน่วย = *1000 บาท

การตรวจสอบความถูกต้องของข้อมูล

- Edit การแก้ไข
- Verify การตรวจสอบ
- Validate ความสมเหตุสมผล
 - Range Check เช่น อายุ เป็น -10 ไม่ได้
 - Relation Check เช่น เป็นผู้หญิง คำนำหน้านามเป็นนาย ไม่ได้

การเข้ารหัสข้อมูล

ข้อมูลดิบ

นาย	สมชาย	คนไทย	เชียงใหม่	12,000 บาท
-----	-------	-------	-----------	------------

เข้ารหัส

ข้อมูลผ่านการเข้ารหัส

M	สมชาย	คนไทย	2	12
---	-------	-------	---	----

M แทน นาย
N แทน นางสาว
O แทน นาง

1 กรุงเทพ
2 เชียงใหม่
3 ลำพูน
....

หน่วย = *1000 บาท

การเก็บข้อมูลในสื่อบันทึกข้อมูล

- การเปลี่ยนสื่อบันทึกข้อมูล
- ข้อมูลจริง
- ข้อมูลในสื่อบันทึกข้อมูล

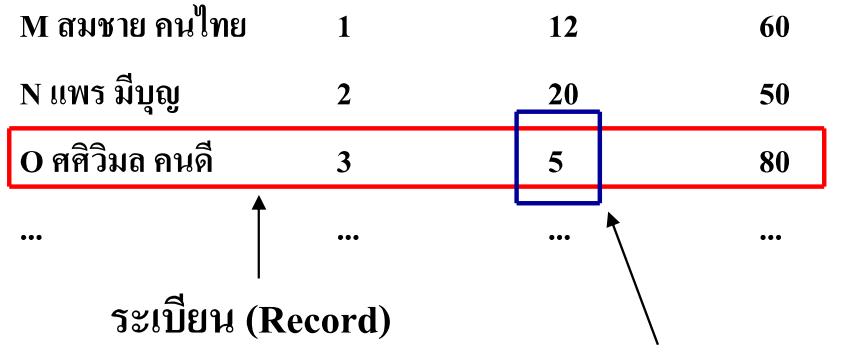
ข้อมูลจริง

ชื่อ	แผนก	ชั่วโมง	อัตรา
นายสมชาย คนไทย	คอมพิวเตอร์	12	60
นางสาวแพร มีบุญ	บัญชี	20	50
นางศศิวิมล คงดี	ประชาสัมพันธ์	5	80
...

M แทน นาย
N แทน นางสาว
O แทน นาง

1 คอมพิวเตอร์
2 บัญชี
3 ประชาสัมพันธ์
...

ข้อมูลในสื่อบันทึกข้อมูล



การประมวลผลข้อมูล

- คำนวณ (Calculating)
- ปรับปรุง (Updating)
- รายงาน (Reporting)
- สรุป (Summarizing)
- เรียงลำดับ (Sorting)
- ค้นหา (Searching)
- แบ่งประเภท (Classification)

คำนวณ

- นำกลุ่มของแอทริบิวต์ที่มีอยู่มาคำนวณหาค่า
- ค่าเฉลี่ย (Mean)
- มัธยฐาน (Median)
- ฐานนิยม (Mode)

ปรับปรุง

- นำแอพทริบิวต์ที่มีอยู่มาคำนวณ แล้วแทนที่
- การปรับค่าอัตราเงินเดือนของพนักงาน
- การปรับค่าเงินสุทธิของบัญชีธนาคาร

วิชา 204100 Information Technology and Modern Life

25

สรุป- คำนวณและแสดงผลลัพธ์แบบรวมทุกรายการ

ชื่อ	แผนก	อัตรา	ชั่วโมง	ค่าจ้าง
นายสมชาย คนไทย	คอมพิวเตอร์	60	12	720
นางสาวแพร มีบุญ	บัญชี	50	20	1,000
นางศศิวิมล คงดี	ประชาสัมพันธ์	80	5	400
...

รวมรายจ่ายทั้งหมด
ข้อมูลเป็นประโยชน์ต่อเจ้าของกิจการ (คนจ่ายเงิน)

72 4,170

วิชา 204100 Information Technology and Modern Life

27

รายงาน - แสดงรายการทั้งหมด

ชื่อ	แผนก	อัตรา	ชั่วโมง	ค่าจ้าง
นายสมชาย คนไทย	คอมพิวเตอร์	60	12	720
นางสาวแพร มีบุญ	บัญชี	50	20	1,000
นางศศิวิมล คงดี	ประชาสัมพันธ์	80	5	400
...

Data หรือ ข้อมูลเดิม

Information หรือ สารสนเทศ
ได้จาก อัตรา x ชั่วโมง

วิชา 204100 Information Technology and Modern Life

26

สรุป - แสดงข้อมูลสรุปแบบกลุ่ม

แผนก	อัตรา	ชั่วโมง	ค่าจ้าง
คอมพิวเตอร์	60	22	1,320
บัญชี	50	33	1,650
ประชาสัมพันธ์	80	15	1,200

70

4,170

สรุปค่าจ้างแยกตามแผนก
เป็นประโยชน์ในการลงทุนขยายกิจการ

วิชา 204100 Information Technology and Modern Life

28

เรียงลำดับ

- ตัวเลข เช่น คะแนนสอบ
- ข้อความ เช่น ชื่อ
- รหัส เช่น รหัสสินค้า

ค้นหา

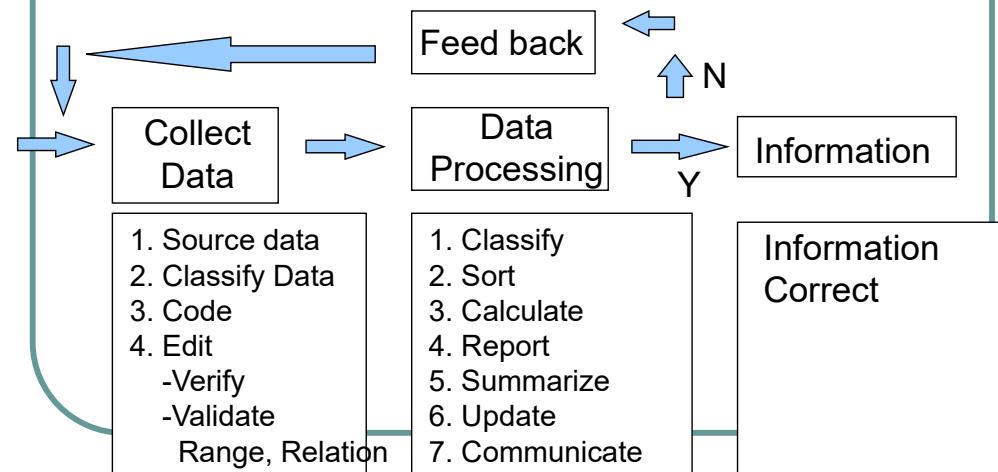
- ค้นหาระเบียนที่ต้องการ จากเงื่อนไขที่กำหนด
- ค้นหานักเรียนที่ $GPA < 1.75$
- ค้นหาพนักงานที่ เงินเดือน $\leq 5,000$ หรือ อยู่ฝ่ายบัญชี
- ค้นหาสินค้าที่ ราคา = 100 และ เป็นอาหาร

แบ่งประเภท หรือจัดกลุ่ม

- แบ่งระเบียนทั้งหมดออกเป็นกลุ่ม โดยที่ระเบียนในกลุ่มเดียวกันต้องเป็นประเภทเดียวกัน
- แบ่งพนักงานออกเป็นแผนก
- แบ่งนักศึกษาออกเป็นคณะหรือสาขา
- แบ่งสินค้าออกเป็นสินค้าราคาแพงหรือราคาถูก

ภาพวงจรการประมวลผลด้วยคอมพิวเตอร์

● Processing



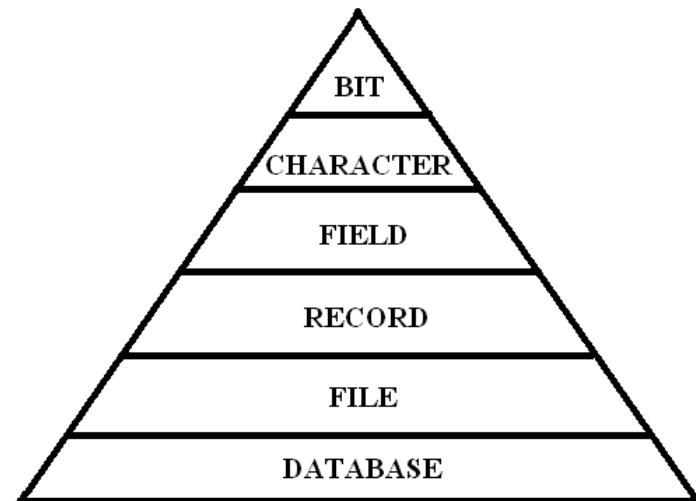
องค์ประกอบข้อมูล (Data Organization)

- บิต (Bit)
- อักขระ (Character)
- ขอบเขตข้อมูล (Field)
- ระเบียน (Record)
- แฟ้มข้อมูล (File)
- ฐานข้อมูล (Database)

บิต

- ตัวเลข 0 และ 1
- หน่วยพื้นฐานที่จัดเก็บในคอมพิวเตอร์

องค์ประกอบข้อมูล (ต่อ)



อักขระ

- กลุ่มของบิต เช่น 011101011
- 1 อักขระ นับเป็น 1 ไบต์
- โดยทั่วไป 1 ไบต์ (อักขระ) มี 8 บิต
- เช่น 01101011 แทน F เป็นต้น

ขอบเขตข้อมูล

- กลุ่มของอักขระ
- จำนวน 1,024
- ข้อความ Cat
- อักขระเดี่ยว A

ระเบียน

- กลุ่มของขอบเขตข้อมูลที่เกี่ยวข้องกัน
- นายพร มีบุญ แผนกบัญชี 20 ชั่วโมง 50 บาท

แฟ้มข้อมูล

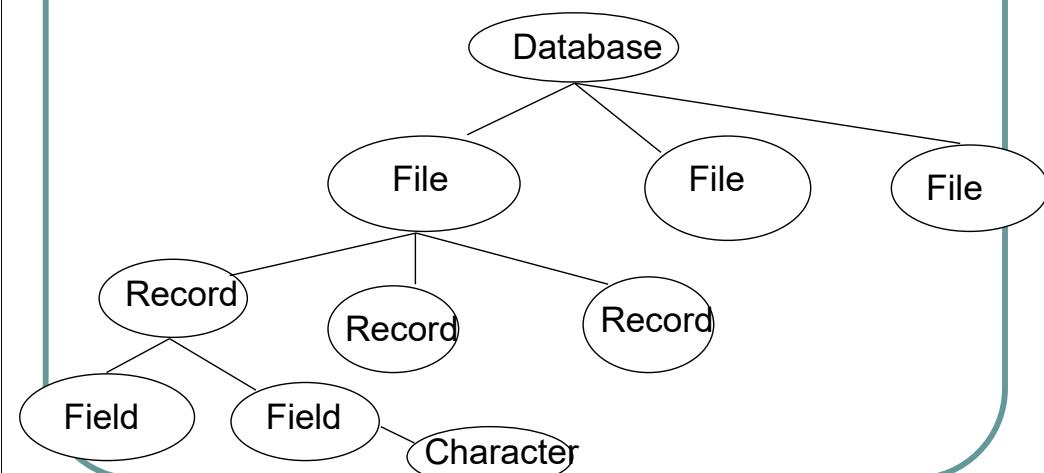
กลุ่มของระเบียนที่มีรูปแบบเดียวกัน

แฟ้ม/ File

ชื่อ	แผนก	ชั่วโมง	อัตรา
นายพร มีบุญ Record	คอมพิวเตอร์	20	เรคคอร์ด 50
นางสาวศศิวิมล คนดี	บัญชี	5	80
นาง ปนัดดา ดวงคำ	ประชาสัมพันธ์ Field	17	90
...

ภาพแสดงจุดข้อมูล

DATABASE



ฐานข้อมูล

กลุ่มของแฟ้มข้อมูลที่สัมพันธ์กัน

ชื่อ	แผนก	ชั่วโมง	อัตรา
นายพร มีนุญ	คอมพิวเตอร์	20	50 File A
นางสาวศศิวิมล คงดี	บัญชี	5	80
...
แผนก	เวลาเริ่มงาน	จำนวนพนักงาน	
คอมพิวเตอร์	12:00	52	File B
บัญชี	9:00	23	
...	

วิชา 204100 Information Technology and Modern Life

41

ลักษณะฐานข้อมูล

- ข้อมูลมีโครงสร้างที่แน่นอน
- มีชุดเดียวใช้ได้ทั้งหน่วยงาน
- เป็นอิสระจากโปรแกรม
- มีภาษาสอบถาม (Query Language)

วิชา 204100 Information Technology and Modern Life

42