

ชื่อหน่วย ความรู้เบื้องต้นเกี่ยวกับ ฐานข้อมูล	ใบความรู้ที่ 1.2	รายวิชา ง 43202 การจัดการฐานข้อมูล
ชื่อเรื่อง ฐานข้อมูลเชิงสัมพันธ์		ระดับชั้น มัธยมศึกษาปีที่ 6 เวลาเรียน 2 ชั่วโมง

ฐานข้อมูลเชิงสัมพันธ์ (Relational Database)

ฐานข้อมูลที่ได้รับความนิยมอย่างสูงในปัจจุบัน ผู้ใช้จะมองเห็นกลุ่มของข้อมูลที่จัดเก็บไว้ในรูปแบบของตาราง แบบ 2 มิติ ที่ประกอบด้วยข้อมูลแต่ละแถวในแนวนอน ทำให้สามารถเชื่อมโยงหรือสร้างความสัมพันธ์ระหว่างกลุ่มข้อมูลหรือ ตารางที่เกี่ยวข้องในฐานข้อมูลเดียวกันได้ง่าย

สำหรับโครงสร้างของฐานข้อมูลเชิงสัมพันธ์ที่สามารถเรียกใช้ข้อมูลได้ในรูปของตารางนี้ ผู้ใช้ไม่จำเป็นต้องทราบว่าจะข้อมูลจะถูกเก็บจริงในลักษณะใด แต่สามารถกำหนดความสัมพันธ์ระหว่างตารางต่าง ๆ ในฐานข้อมูลด้วยตัวเองได้

ลูกค้า

รหัสลูกค้า	ชื่อบริษัท	ที่อยู่	...
1001	สมบัติฟูดแวร์	33 หมู่ 2 ต.ทุ่งมะขาม อ.แปลงยาว จ.ปทุมธานี	...
1002	ปิยะพลาสติก	9 ถ.วงแหวนรอบนอก อ.ลำลูกกา จ.ปทุมธานี	...
1003	เอ็มพีแพค	12 หมู่ 8 ต.คูน้ำ อ.เมือง จ.นนทบุรี	...

การสั่งซื้อ

รหัสสั่งซื้อ	รหัสลูกค้า	รหัสสินค้า	...
12001	1001	M-100	...
12002	1003	PE-96	...
12003	1002	PE-15	...

สินค้า

รหัสสินค้า	ชื่อสินค้า	ปริมาณ	ราคา/หน่วย	...
M-100	พลาสติกแท่งอัดเม็ด	25	520	...
PK-5	น้ำยาทำความสะอาด	5	600	...
PE-15	ท่อพีวีซี 15 นิ้ว	200	80	...
PE-30	ท่อพีวีซี 30 นิ้ว	600	200	...

ตารางที่ 1.2.1 แสดงตัวอย่างฐานข้อมูลเชิงสัมพันธ์

จากตารางจะเห็นตารางลูกค้าและตารางสั่งซื้อ มีเขตข้อมูลรหัสลูกค้าเป็นตัวเชื่อมความสัมพันธ์ระหว่าง 2 ตารางนี้ ถ้าต้องการทราบชื่อและที่อยู่ของลูกค้ารหัส 1002 ในตารางการสั่งซื้อ เราจะนำรหัส 1002 ไปค้นในตารางลูกค้าที่มีรหัสตรงกัน

คำศัพท์พื้นฐานที่ควรทราบ

คำศัพท์พื้นฐานที่เกี่ยวข้องกับฐานข้อมูลเชิงสัมพันธ์ที่ควรทราบเช่น Entity , Attribute , Relationship และ Key

เอนทิตี (Entity) และ แอตทริบิวต์ (Attribute)

เอนทิตี หมายถึง สิ่งต่าง ๆ ในฐานข้อมูลที่อ้างอิงถึง สิ่งเหล่านี้อาจเป็นได้ทั้งรูปธรรม คือ คน สัตว์ สิ่งของ เช่น เอนทิตีพนักงาน เอนทิตีนักศึกษา เอนทิตีสินค้า หรือเป็นนามธรรม เช่น เอนทิตีความชำนาญ เป็นต้น ดังนั้นเอนทิตีก็คือกลุ่มของข้อมูลประเภทเดียวกัน เช่น ถ้ากล่าวถึงเอนทิตีพนักงานจะหมายถึงกลุ่มของคนทั้งหมดที่เป็นพนักงาน

แอตทริบิวต์ คือสิ่งที่ใช้บอกองค์ประกอบหรือเนื้อหา (subject) ของเอนทิตี เช่น เอนทิตีพนักงาน จะประกอบด้วยแอตทริบิวต์ต่าง ๆ คือ รหัสประจำตัว ชื่อ นามสกุล วุฒิการศึกษา วันเข้าทำงาน ที่อยู่ เป็นต้น

ตาราง (เอนทิตี) : พนักงาน

รหัสประจำตัว	ชื่อ	นามสกุล	ตำแหน่ง	วุฒิการศึกษา	วันเข้าทำงาน	ที่อยู่
101	นายสมรักษ์	คำหอม	พนักงานชาย	ปวส.	15 ก.พ.2545
102	น.ส.มาลัย	มะลิวัลย์	พนักงานบัญชี	ปริญญาตรี	1 มี.ค. 2546
103	น.ส.ยุพิน	สมศรีคำ	พนักงานชาย	ปวส.	22 ธ.ค. 2548

เรคคอร์ด

ฟิลด์ (แอตทริบิวต์)

ตารางที่ 1.2.2 แสดงตัวอย่างเอนทิตีพนักงาน และคำศัพท์ที่ควรทราบ

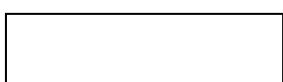
ภาพแสดงกลุ่มข้อมูลในฐานข้อมูลเชิงสัมพันธ์ จะเห็นว่าเอนทิตี หรือกลุ่มข้อมูลพนักงานจะถูกนำมาเก็บในตารางเดียวกันทั้งหมด โดยข้อมูลของพนักงานแต่ละคนก็จะได้เรคคอร์ด ส่วนแอตทริบิวต์จะหมายถึงเขตข้อมูลแต่ละเขตข้อมูลที่ประกอบกันขึ้นเป็นระเบียบ

ผังมโนภาพ (Conceptual Design)

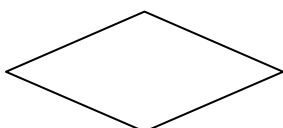
เป็นวิธีออกแบบฐานข้อมูลและความสัมพันธ์ต่าง ๆ ในระดับความคิดโดยใช้แผนภาพแสดงความสัมพันธ์ของเอนทิตี (Entity) และรีเลชัน (Relation) เรียกว่าแผนภาพแบบ E-R สัญลักษณ์ที่ใช้ในแผนภาพ ได้แก่



รูปวงรี แทนแอตทริบิวต์ หรือ เขตข้อมูลสัญลักษณ์และหนึ่งเขตข้อมูล

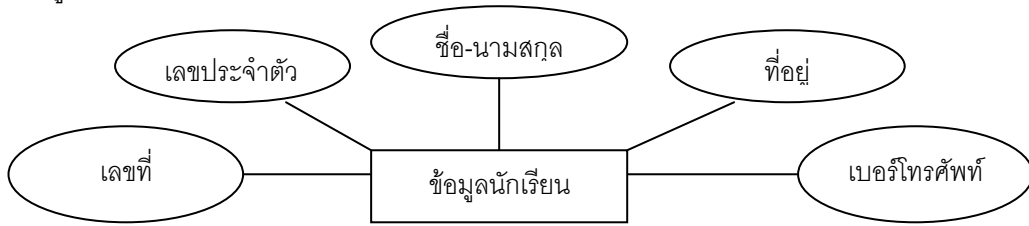


รูปสี่เหลี่ยมผืนผ้า แทนเอนทิตีหรือตารางข้อมูลหนึ่งตาราง



รูปสี่เหลี่ยมข้าวหลามตัด แทนรีเลชันหรือความสัมพันธ์ระหว่างเอนทิตีแต่ละเอนทิตี

ภายในสัญลักษณ์แต่ละอันจะเขียนชื่อหรือความสัมพันธ์กำกับไว้ภายในและเชื่อมโดยถึงกันด้วยเส้นและหัวลูกศร เช่น ตารางข้อมูลของนักเรียน มีแอททริบิวต์เป็น เลขที่ เลขประจำตัว ชื่อ และนามสกุล ที่อยู่ หมายเลขโทรศัพท์ เป็นต้น



แผนภาพที่ 1.2.1 แสดงตัวอย่างแผนภาพแบบ E-R ของเอนทิตี ข้อมูลนักเรียน

ความสัมพันธ์ (Relationship)

การกำหนดความสัมพันธ์ระหว่างเอนทิตีหรือ ตาราง ในฐานข้อมูลเชิงสัมพันธ์ จะมี 3 ลักษณะคือ

ความสัมพันธ์แบบ 1 : 1 (One-to-One)

เป็นความสัมพันธ์ที่เรคคอร์ดๆ ในตารางหนึ่งสามารถจับคู่กับเรคคอร์ดในอีกตารางหนึ่งได้เพียง 1 เรคคอร์ดเท่านั้น จึงเป็นการจับคู่กันแบบตัวต่อตัว ความสัมพันธ์ลักษณะนี้จะพบเห็นได้น้อย เช่น กรณีที่จำนวนฟิลด์ในตารางนั้นมีมากเกินไป จนทำให้ตารางมีขนาดใหญ่ไม่สะดวกในการประมวลผลและบำรุงรักษา จึงต้องแยกข้อมูลบางส่วนที่สำคัญออกมาเก็บไว้อีกตารางหนึ่ง โดยเลือกฟิลด์ที่มีคุณสมบัติเหมาะสมเป็นตัวเชื่อมความสัมพันธ์ระหว่างตารางทั้งสอง ในตัวอย่างนี้จะใช้ฟิลด์รหัสพนักงานเป็นตัวเชื่อมความสัมพันธ์ระหว่างตารางพนักงานและตารางข้อมูลส่วนตัว เนื่องจากฟิลด์รหัสพนักงานเป็นฟิลด์ร่วม (Common field) ของทั้ง 2 ตาราง



พนักงาน

รหัสพนักงาน	ชื่อ	นามสกุล	ที่อยู่	...
1001	สมพร	ความดี	52 หมู่ 12 ต.ลาดยาว อ.เมือง ปทุมธานี	...
1002	สมหมาย	รักดี	22 หมู่ 4 ถ.อารักษ์ อ.เมือง จ.นนทบุรี	...
1003	สุเมธ	คมขำ	122/125 แพลตตินแดง 3 เขตดินแดง กรุงเทพฯ	...

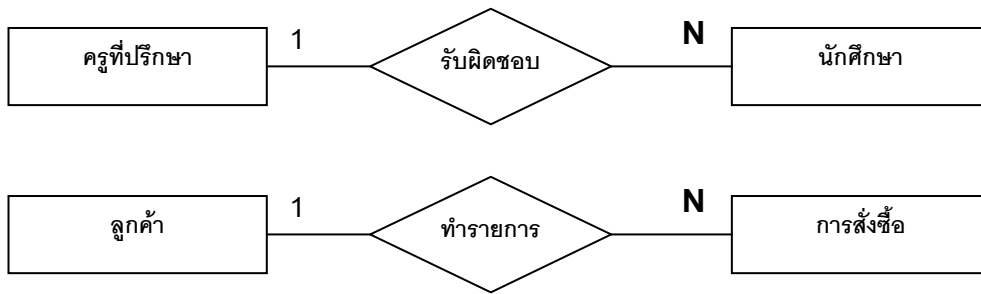
ข้อมูลส่วนตัว

รหัสพนักงาน	เงินเดือน	เลขประกันสังคม
1001	15,000	662-55-2552
1002	13,550	252-66-5259
1003	24,000	141-36-4585

ตารางที่ 1.2.3 แสดงความสัมพันธ์แบบ หนึ่งต่อหนึ่ง

ความสัมพันธ์แบบ 1: N (One-to-Many)

เป็นความสัมพันธ์ที่เรคคอร์ดใดๆ ในตารางหนึ่งสามารถจับคู่กับเรคคอร์ดในอีกตารางหนึ่งได้หลายเรคคอร์ด ตัวอย่างเช่น ความสัมพันธ์ระหว่างตารางลูกค้าและตารางการสั่งซื้อ ลูกค้าหนึ่งคนจะมีใบสั่งซื้อสินค้าได้หลายใบ ในขณะที่ใบสั่งซื้อแต่ละใบจะต้องมาจากลูกค้าเพียงคนเดียวเท่านั้น หรือตัวอย่างของอาจารย์ที่ปรึกษาที่มีนักศึกษาในความรับผิดชอบได้หลายคน แต่นักศึกษาแต่ละคนจะมีอาจารย์ที่ปรึกษาได้เพียงคนเดียวเช่นกัน



ลูกค้า

รหัสลูกค้า	ชื่อบริษัท	ที่อยู่	...
1001	สมบัติฟูดแวร์	33 หมู่ 2 ต.ทุ่งมะขาม อ.แปลงยาว จ.ปทุมธานี	...
1002	ปิยะพลาสติก	9 ถ.วงแหวนรอบนอก อ.ลำลูกกา จ.ปทุมธานี	...
1003	เอ็มพีแพค	12 หมู่ 8 ต.คูน้ำ อ.เมือง จ.นนทบุรี	...

การสั่งซื้อ

รหัสสั่งซื้อ	รหัสลูกค้า	รหัสสินค้า	...
12001	1001	M-100	...
12002	1003	PE-96	...
12003	1002	PE-15	...

ตารางที่ 1.2.4 แสดงตัวอย่างความสัมพันธ์แบบหนึ่งต่อกลุ่ม

ความสัมพันธ์แบบ M : N (Many-to-Many)

เป็นความสัมพันธ์ที่เรคคอร์ดหลาย ๆ เรคคอร์ดในตารางหนึ่ง มีความสัมพันธ์กับอีกหลายๆ เรคคอร์ดในอีกตารางหนึ่งพร้อมกัน เช่น ความสัมพันธ์ระหว่างตารางลูกค้าและตารางสินค้า ลูกค้าหนึ่งคนสามารถซื้อสินค้าได้หลายชนิด ในขณะที่สินค้าแต่ละชนิดก็จะถูกซื้อโดยลูกค้าหลายๆ คนได้ด้วย ถ้าเรานำตารางทั้งสองนี้มาเชื่อมโยงกันโดยตรงจะเห็นว่าทำไม่ได้ เนื่องจากไม่มีฟิลด์ที่เป็นฟิลด์ร่วมของทั้งสองตาราง การสร้างความสัมพันธ์แบบ M:N จึงต้องใช้ตารางอื่นมาช่วยเป็นสะพานในการเชื่อมโยง ซึ่งในที่นี้จะใช้ตาราง การสั่งซื้อมาช่วย



ลูกค้า

รหัสลูกค้า	ชื่อบริษัท	ที่อยู่	...
1001	สมบัติฟูดแวร์	33 หมู่ 2 ต.ทุ่งมะขาม อ.แปลงยาว จ.ปทุมธานี	...
1002	ปิยะพลาสติก	9 ถ.วงแหวนรอบนอก อ.ลำลูกกา จ.ปทุมธานี	...
1003	เอ็มพีแพค	12 หมู่ 8 ต.คูน้ำ อ.เมือง จ.นนทบุรี	...

การสั่งซื้อ

รหัสสั่งซื้อ	รหัสลูกค้า	รหัสสินค้า	...
12001	1002	M-100	...
12002	1003	PE-96	...
12003	1002	PE-15	...
12004	1001	PE-15	...

สินค้า

รหัสสินค้า	ชื่อสินค้า	ปริมาณ	ราคา/หน่วย	...
M-100	พลาสติกแท่งอัดเม็ด	25	520	...
PK-5	น้ำยาทำความสะอาด	5	600	...
PE-15	ท่อพีวีซี 15 นิ้ว	200	80	...
PE-30	ท่อพีวีซี 30 นิ้ว	600	200	...

ตารางที่ 1.2.5 แสดงตัวอย่างตารางที่มีความสัมพันธ์แบบกลุ่มต่อกกลุ่ม

จากภาพจะเกิดความสัมพันธ์แบบ 1 : N สองความสัมพันธ์ ความสัมพันธ์แรกคือความสัมพันธ์ระหว่างตารางลูกค้าและตารางการสั่งซื้อ ความสัมพันธ์ที่สองคือความสัมพันธ์ระหว่างตารางสินค้าและตารางการสั่งซื้อ โดยฟิลด์รหัสลูกค้าและฟิลด์รหัสสินค้าจะถูกนำมาใช้เป็นฟิลด์ร่วมเพื่อสร้างความสัมพันธ์ระหว่าง ตารางลูกค้าและตารางสินค้าให้เป็นแบบ M:N จากตัวอย่างจะเห็นลูกค้ารหัส 1002 ซื้อสินค้านาม M-100 และ PE-15 ในขณะที่สินค้า PE-15 ก็ถูกสั่งซื้อโดยลูกค้ารหัส 1001 และ 1002

คีย์ (Key)

เพื่อเพิ่มประสิทธิภาพในการจัดการข้อมูลควรกำหนด **คีย์ (key)** ให้กับตารางเพื่อใช้จำแนกเรคคอร์ดและกำหนดความสัมพันธ์ระหว่างตาราง เช่น กำหนดรหัสประจำตัวให้กับพนักงานทุกคนในตารางพนักงาน คีย์ที่ใช้ในระบบฐานข้อมูลมีหลายอย่างดังนี้

คีย์หลัก (Primary Key)

เป็นคีย์ที่กำหนดจากฟิลด์ที่จะต้องไม่มีข้อมูลซ้ำกันโดยเด็ดขาดในตารางนั้น เช่น ฟิลด์รหัสพนักงานในตารางข้อมูลพนักงาน หรือฟิลด์รหัสสินค้าในตารางข้อมูลสินค้านำไปใช้จัดเรียงและแยกแยะข้อมูลในแต่ละเรคคอร์ดออกจากกัน ฟิลด์ที่กำหนดให้เป็นคีย์หลักต้องมีค่าเสมอจะเป็นค่าว่างไม่ได้

พนักงาน

รหัสพนักงาน	ชื่อ	นามสกุล	ที่อยู่	...
1001	สมพร	ความดี	52 หมู่ 12 ต.ลาดยาว อ.เมือง ปทุมธานี	...
1002	สมหมาย	รักดี	22 หมู่ 4 ถ.อารักษ์ อ.เมือง จ.นนทบุรี	...
1003	สุเมธ	คมขำ	122/125 แพลตตินแดง 3 เขตดินแดง กรุงเทพฯ	...

↑
ใช้ฟิลด์นี้เป็นคีย์หลัก

คีย์คู่แข่ง (Candidate Key)

ถ้าในตารางหนึ่งมีฟิลด์ที่มีคุณสมบัติที่สามารถใช้เป็นคีย์หลักแทนกันได้ จะเรียกฟิลด์เหล่านั้นว่า คีย์คู่แข่ง เช่น ในตารางพนักงาน ถ้าไม่มีชื่อพนักงานซ้ำกันเลย ก็สามารถใช้ฟิลด์รหัสพนักงานหรือฟิลด์ชื่อพนักงานเป็นคีย์หลักได้ ทำให้ทั้งสองฟิลด์นี้กลายเป็นฟิลด์คู่แข่งกัน ถ้าใช้รหัสพนักงานเป็นคีย์หลัก ก็จะทำให้ฟิลด์ชื่อพนักงานเป็น คีย์สำรอง (Alternate key)

พนักงาน

รหัสพนักงาน	ชื่อ	นามสกุล	ที่อยู่	...
1001	สมพร	ความดี	52 หมู่ 12 ต.ลาดยาว อ.เมือง ปทุมธานี	...
1002	สมหมาย	รักดี	22 หมู่ 4 ถ.อารักษ์ อ.เมือง จ.นนทบุรี	...
1003	สุเมธ	คมขำ	122/125 แพลตตินแดง 3 เขตดินแดง กรุงเทพฯ	...



2 ฟิลด์นี้เป็นฟิลด์คู่แข่ง ในกรณีที่ ฟิลด์รหัสพนักงาน หรือ ฟิลด์ชื่อ ไม่ซ้ำกัน

คีย์รวม (Compound Key)

เกิดจากการนำฟิลด์หลายๆฟิลด์มารวมกัน เพื่อให้มีคุณสมบัติเป็นคีย์หลัก คือไม่มีข้อมูลซ้ำและไม่มีค่าว่าง เนื่องจากในบางครั้งการสร้างคีย์หลักจากฟิลด์เดียวอาจมีโอกาสที่จะเกิดข้อมูลซ้ำกันได้ เช่น ในตารางพนักงานด้านล่าง ตารางพนักงานหากไม่ได้กำหนดฟิลด์รหัสพนักงาน เราอาจใช้ฟิลด์ชื่อ และนามสกุล ประกอบกันเป็นคีย์หลักของตารางก็ได้

พนักงาน

ชื่อ	นามสกุล	ที่อยู่	เบอร์โทรศัพท์	...
สมพร	ความดี	52 หมู่ 12 ต.ลาดยาว อ.เมือง ปทุมธานี	02-953-2222	...
สมหมาย	รักดี	22 หมู่ 4 ถ.อารักษ์ อ.เมือง จ.นนทบุรี	02-253-2552	...
สุเมธ	คมขำ	122/125 แพลตตินแดง 3 เขตดินแดง กรุงเทพฯ	02-225-3987	...



หากไม่กำหนดฟิลด์รหัสพนักงาน อาจใช้ฟิลด์ชื่อและนามสกุล เป็นคีย์รวมกัน

คีย์นอก (Foreign key)

เป็นคีย์ที่ใช้เชื่อมโยงตารางที่มีข้อมูลเกี่ยวข้องกัน เช่น ในตารางลูกค้าจะมีฟิลด์รหัสลูกค้าเป็นคีย์หลัก เราจะให้รหัสลูกค้าในตารางลูกค้าเชื่อมโยงกับรหัสลูกค้าในตารางการสั่งซื้อ เพื่อที่จะได้ทราบชื่อและที่อยู่ของลูกค้าที่สั่งซื้อสินค้านั้น

ลูกค้า

รหัสลูกค้า	ชื่อบริษัท	ที่อยู่	...
1001	สมมติฟุตแวร์	33 หมู่ 2 ต.ทุ่งมะขาม อ.แปลงยาว จ.ปทุมธานี	...
1002	ปิยะพลาสติก	9 ถ.วงแหวนรอบนอก อ.ลำลูกกา จ.ปทุมธานี	...
1003	เอ็มพีแพค	12 หมู่ 8 ต.คูน้ำ อ.เมือง จ.นนทบุรี	...

การสั่งซื้อ

รหัสสั่งซื้อ	รหัสลูกค้า	รหัสสินค้า	...
12001	1002	M-100	...
12002	1003	PE-96	...
12003	1002	PE-15	...
12004	1001	PE-15	...

ในตัวอย่างนี้ฟิลด์รหัสลูกค้าในตารางการสั่งซื้อ (อยู่ด้าน N ของความสัมพันธ์) จะเป็นคีย์นอกที่เชื่อมโยงกับฟิลด์รหัสลูกค้าซึ่งเป็นคีย์หลักของตารางลูกค้า

กฎที่ใช้ควบคุมการคงสภาพความถูกต้องของข้อมูล (Integrity Constraint)

เนื่องจากฐานข้อมูลเชิงสัมพันธ์ประกอบด้วยข้อมูลหลาย ๆ ตารางที่กำหนดให้มีความสัมพันธ์ในลักษณะต่าง ๆ ไม่ว่าจะเป็นแบบ One-to-One , One-to-Many หรือ Many-to-Many การลบเรคคอร์ดหรือแก้ไขเปลี่ยนแปลงข้อมูลในตารางหนึ่งย่อมมีผลกระทบต่อข้อมูลในตารางอื่นที่สัมพันธ์กัน เพื่อให้ข้อมูลทั้งหมดที่เกี่ยวข้องมีการคงสภาพความถูกต้องของข้อมูลที่สอดคล้องตามเงื่อนไขที่กำหนดไว้ตลอดเวลา จะต้องมีการกำหนดกฎที่ใช้ควบคุมความถูกต้องซึ่งมี 2 ลักษณะคือ

- กฎการคงสภาพของเอนทิตี (Entity Integrity Constraint) ควบคุมการคงสภาพความถูกต้องของข้อมูลในตารางเดียวกัน ซึ่งจะต้องไม่มีค่าว่างในฟิลด์ใด ๆ ที่ประกอบกันขึ้นเป็นคีย์หลักของตารางนั้น

- กฎการคงสภาพการอ้างอิง (Referential Integrity Constraint) ควบคุมการคงสภาพความถูกต้องของความสัมพันธ์ระหว่างตาราง ซึ่งถ้ามีคีย์นอกอยู่ในตารางใด ข้อมูลที่อยู่ในคีย์นอกนั้นจะต้องเป็นข้อมูลที่อยู่ในคีย์หลักของอีกตารางหนึ่งด้วย ถ้าไม่เช่นนั้นแล้วข้อมูลในคีย์นอกจะต้องเป็นค่าว่าง

ขั้นตอนการออกแบบฐานข้อมูล

การที่จะนำข้อมูลไปใช้งานได้อย่างมีประสิทธิภาพหรือไม่นั้น หัวใจสำคัญอยู่ที่การออกแบบฐานข้อมูลโดยผู้ออกแบบจะต้องวิเคราะห์หาเอนทิตีและแอตทริบิวต์ทั้งหมดที่จะนำมาใช้ในฐานข้อมูลได้อย่างครบถ้วน รวมทั้งกำหนดความสัมพันธ์ระหว่างเอนทิตีได้อย่างถูกต้องและเหมาะสม การออกแบบฐานข้อมูลใน Access จะมีลำดับขั้นตอนดังนี้

ขั้นที่ 1 เก็บรวบรวมข้อมูลรายละเอียดทั้งหมด

เป็นขั้นตอนของการเก็บรวบรวมข้อมูลและรายละเอียดต่าง ๆ ของงาน รวมทั้งความต้องการของผู้ใช้ เช่น

- ข้อมูลใดบ้างที่เป็นเรื่องเดียวกัน ให้จัดกลุ่มข้อมูลนั้นเป็นเอนทิตี
- ชนิดของข้อมูลเป็นแบบใด ตัวอักษรหรือตัวเลข มีเงื่อนไขหรือข้อกำหนดอย่างไร เช่น รหัสพนักงานจะต้องเป็นเลข 4 หลัก อายุพนักงานต้องไม่เกิน 55 ปี วุฒิการศึกษาของพนักงานต้องไม่ต่ำกว่าระดับ ปวส.
- มีข้อมูลอะไรบ้างที่จะต้องนำมาค้นหาหรือประมวลผล ผลที่ได้ต้องส่งออกระบบภายนอกหรือไม่
- ใครบ้างที่เป็นผู้ใช้งานข้อมูลนี้ ใช้บ่อยแค่ไหน มีความสำคัญอย่างไร
- ลักษณะของรายงานประกอบด้วยรายงานอะไรบ้าง ระยะเวลาในการออกรายงาน
- ข้อมูลอื่น ๆ ที่สามารถรวบรวมได้ โดยพยายามเก็บรายละเอียดให้มากที่สุด

ขั้นที่ 2 กำหนดโครงสร้างของตาราง (Table)

จากกลุ่มข้อมูลหรือเอนทิตีที่รวบรวมได้จากเอกสารต่างๆในขั้นที่ 1 นำมากำหนดแอตทริบิวต์หรือเนื้อหา (subject) ของข้อมูล จะได้ทราบว่าในเอนทิตีนั้นจะต้องนำข้อมูลอะไรมาใช้บ้าง หลังจากนั้นนำเอนทิตีและแอตทริบิวต์มากำหนดโครงสร้างของตาราง โดยแปลงแอตทริบิวต์เป็นฟิลด์ พร้อมกำหนดชนิดและขนาดข้อมูลในแต่ละฟิลด์รวมทั้งเงื่อนไขหรือกฎเกณฑ์ที่ใช้กำหนดลักษณะข้อมูล

ขั้นที่ 3 กำหนดคีย์

ขั้นตอนนี้จะพิจารณาว่าฟิลด์ใดบ้างในตารางที่มีคุณสมบัติเหมาะสมที่จะนำมาเป็นคีย์ ถ้าไม่มีฟิลด์ใดเลยที่เหมาะสม ต้องกำหนดฟิลด์ใหม่เพื่อใช้เป็นคีย์โดยเฉพาะ เช่น ในตารางลูกค้าจะกำหนดฟิลด์รหัสลูกค้าขึ้นใหม่เพื่อใช้เป็นคีย์หลักแทนฟิลด์ชื่อลูกค้า เนื่องจากชื่อลูกค้าอาจซ้ำกันได้

ขั้นที่ 4 การทำ Normalization

ถ้าตารางที่สร้างขึ้นในขั้นที่ 2 ยังมีความซ้ำซ้อนกันของข้อมูล หรือข้อมูลบางฟิลด์ไม่เกี่ยวข้องโดยตรงกับเนื้อหาในตารางนั้น จะต้องนำมาปรับแก้ให้มีโครงสร้างหรือรูปแบบที่เหมาะสมก่อนนำไปประมวลผล ถ้านำโครงสร้างไปใช้เลยโดยไม่ทำ Normalization ก่อนอาจเกิดปัญหาได้ เช่น ปัญหาการสิ้นเปลืองเนื้อที่เนื่องจากเก็บข้อมูลซ้ำซ้อน ปัญหาความผิดปกติ (Anomaly) ของข้อมูลเมื่อมีการเพิ่ม/ลบ/แก้ไขเรคคอร์ด รวมทั้งปัญหาการกำหนดความสัมพันธ์ในขั้นที่ 5 อาจทำได้ยาก

ขั้นที่ 5 กำหนดความสัมพันธ์

นำตารางทั้งหมดหลังจากทำ Normalization มาสร้างความสัมพันธ์โดยใช้คีย์ที่กำหนดในขั้นที่ 3 หรือคีย์ที่เกิดขึ้นใหม่จากการทำ Normalization เป็นตัวเชื่อม ซึ่งอาจเป็นแบบ One-to-One, One-to-Many หรือ Many-to-Many ขึ้นกับลักษณะของข้อมูล การกำหนดความสัมพันธ์ระหว่างตารางนี้มีความสำคัญมาก ผู้ออกแบบจะต้องวิเคราะห์ให้ได้ว่าข้อมูลในตารางต่าง ๆ นั้นมีความสัมพันธ์ในลักษณะใด