บทที่ 10
การออกแบบฐานข้อมูล

หลังจากการวิเคราะห์ข้อมูลภายในระบบจะทำให้ได้รู้ว่าในระบบมีข้อมูลอะไรบ้างและข้อมูลนั้นมีความสัมพันธ์กันอย่างไร ซึ่งแสดงโดยใช้แผนภาพแสดงความสัมพันธ์ของข้อมูล ในแผนภาพอี-อาร์ที่แสดงนั้นยังไม่ได้ระบุว่าข้อมูลเหล่านั้นได้มาจากแหล่งใด มีการเก็บในฐานข้อมูลลักษณะอย่างไร และยังไม่คำนึงถึงระบบการจัดการฐานข้อมูลที่จะใช้ ซึ่งจะเรียกแผนภาพอี-อาร์ที่ได้ในขั้นตอนการวิเคราะห์ระบบนี้ว่าเป็นการออกแบบฐานข้อมูลในระดับแนวคิด (Conceptual Database Design) การออกแบบฐานข้อมูลต้องอาศัยแผนภาพอี-อาร์เพื่อช่วยในการออกแบบตารางที่ใช้เก็บข้อมูลเหล่านั้นลงในฐานข้อมูลได้อย่างเหมาะสม โดยต้องผ่านกระบวนการลดความซ้ำซ้อนของข้อมูลที่เรียกว่า การนอร์มัลไลท์เซชัน (Normalization) เพื่อให้อยู่ในรูปของฐานข้อมูลที่ถูกต้องต่อไป

10.1 ระดับของการออกแบบฐานข้อมูล

ฐานข้อมูลเป็นส่วนที่สำคัญสำหรับการสร้างระบบสารสนเทศที่ใช้คอมพิวเตอร์ในการประมวลผล เนื่องจากเป็นส่วนที่ใช้เก็บข้อมูลต่าง ๆ ที่ใช้เป็นข้อมูลในการทำงานของทุกระบบ การออกแบบระบบจะต้องให้ความสำคัญกับการออกแบบฐานข้อมูล การออกแบบฐานข้อมูลแบ่งออกเป็น 3 ระดับคือ การออกแบบฐานข้อมูลระดับแนวคิด (Conceptual Database Design) การออกแบบฐานข้อมูลระดับตรรก (Logical Database Design) และการออกแบบฐานข้อมูลระดับกายภาพ (Physical Database Design)

10.1.1 การออกแบบฐานข้อมูลระดับแนวคิด

การออกแบบฐานข้อมูลระดับแนวคิดเป็นขั้นตอนที่ได้ดำเนินการมาแล้วในขั้นตอนของการวิเคราะห์ระบบซึ่งสิ่งที่ได้คือแผนภาพอี-อาร์ที่แสดงให้เห็นถึงข้อมูลที่เกิดขึ้นในระบบว่ามีอะไรบ้าง ประกอบด้วยรายละเอียดข้อมูลอะไร และข้อมูลเหล่านั้นมีความสัมพันธ์กันอย่างไร โดยยังไม่คำนึงถึงความซ้ำซ้อนกันของข้อมูล หรือเทคโนโลยีฐานข้อมูลที่จะใช้

10.1.2 การออกแบบฐานข้อมูลในระดับตรรก

การออกแบบฐานข้อมูลในระดับตรรก คือการออกแบบฐานข้อมูลให้อยู่ในแบบจำลองฐานข้อมูลที่เหมาะสมกับระบบการจัดการฐานข้อมูลที่จะเลือกใช้ เช่น แบบจำลองฐานข้อมูลเชิงสัมพันธ์ แบบจำลองฐานข้อมูลเชิงลดหลั่น แบบจำลองฐานข้อมูลเชิงโครงข่าย หรือแบบจำลองฐานข้อมูลเชิงวัตถุ การออกแบบฐานข้อมูลในระดับตรรกเป็นขั้นตอนต่อจากการออกแบบฐานข้อมูลในระดับแนวคิด คือการแปลงจากเค้าโครงร่างในระดับแนวคิดให้เป็นเค้าโครงร่างในระดับตรรก (Logical Schema) ตามรูปแบบของระบบฐานข้อมูล ในกรณีที่ใช้แบบจำลองฐานข้อมูลเชิงสัมพันธ์ เมื่อได้แผนภาพอี-อาร์แล้วต้องทำการแปลงให้อยู่ในรูปแบบรีเลชันจากนั้นจึงทำการนอร์มอลไลซ์รีเลชันเพื่อลดความซ้ำซ้อนของข้อมูลต่อไป

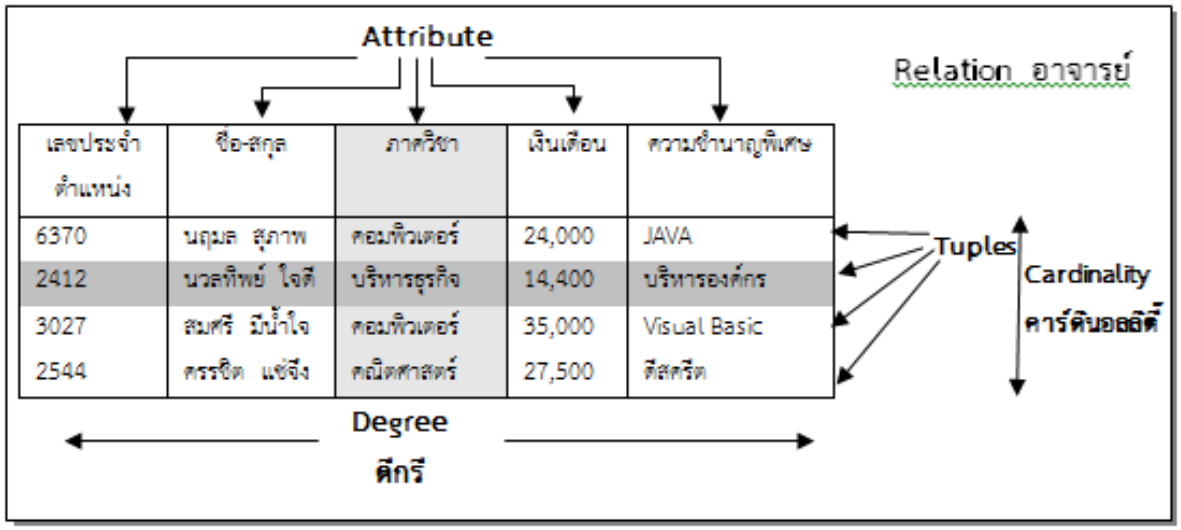
10.1.3 การออกแบบฐานข้อมูลในระดับกายภาพ

การออกแบบฐานข้อมูลในระดับกายภาพ เป็นกระบวนการในการแปลงจากโครงสร้างฐานข้อมูลในระดับตรรก (Logical Database Structure) ไปเป็นโครงสร้างฐานข้อมูลในระดับกายภาพ (Physical Database Structure) คือ ฟิลด์ (Field) เรคอร์ด (Record) ไฟล์ (File) และฐานข้อมูล

(Database) โดยพิจารณาการจัดเก็บข้อมูลจริงในสื่อ มีจุดประสงค์หลักของการออกแบบก็เพื่อประสิทธิภาพของระบบ ซึ่งสามารถพิจารณาได้จากความเร็วในการทำงานของระบบ ความสะดวกในการใช้งานและการประหยัดเนื้อที่ในการจัดเก็บข้อมูล

10.2 โครงสร้างแบบจำลองฐานข้อมูลเชิงสัมพันธ์

แบบจำลองฐานข้อมูลเชิงสัมพันธ์นี้ข้อมูลในฐานข้อมูลจะถูกจัดเก็บอยู่ในลักษณะของตาราง(Table) หลาย ๆ ตารางที่มีความสัมพันธ์กันซึ่งเรียกแต่ละตารางว่า รีเลชัน (Relation) รีเลชันมีองค์ประกอบดังภาพที่ 10.1



ภาพที่ 10.1 องค์ประกอบของรีเลชัน

รีเลชันหรือตาราง 2 มิติซึ่งเปรียบเสมือนแฟ้มข้อมูลจะประกอบด้วยแถวในแนวนอนซึ่งเรียกว่าทัพเปิลและคอลัมน์ในแนวตั้งซึ่งเรียกว่าแอททริบิวท์แต่ละแถวหรือทัพเปิล คือ ข้อมูล 1 เรคอร์ด แต่ละคอลัมน์หรือแอททริบิวท์คือข้อมูล 1 ฟิลด์ ซึ่งจะแสดงคุณสมบัติของรีเลชันนั้น โดยต้องมีการกำหนดชื่อให้กับแต่ละคอลัมน์

กลุ่มของข้อมูลที่เป็นไปได้ของแต่ละแอททริบิวท์ คือโดเมน (Domain) ของแอททริบิวท์นั้น โดเมนเป็นกรอบของค่าที่เป็นไปได้ เช่น โดเมนของค่าของเดือนคือ 1-12 โดเมนของชื่อคือข้อความที่มีความยาวไม่เกิน 40 ตัวอักษร โดเมนของอายุนักศึกษา มีค่าระหว่าง 15 – 30 ปี เป็นต้น

จำนวนของทัพเปิลในหนึ่งรีเลชัน คือจำนวนของแถวในหนึ่งตารางซึ่งเรียกว่า คาร์ดินอลลิตี้ ของรีเลชันนั้น จำนวนของแอททริบิวท์ในหนึ่งรีเลชันก็คือจำนวนคอลัมน์ในตาราง เรียกว่า ดีกรี (Degree) ของรีเลชัน เช่น จากภาพที่ 10.1 รีเลชันอาจารย์จะมีคาร์ดินอลลิตี้เป็น 4 และมีดีกรีเป็น 5 เป็นต้น

การแสดงโครงสร้างของรีเลชันด้วยข้อความประกอบด้วยชื่อของรีเลชันตามด้วยชื่อของแอททริบิวท์ทั้งหมดของรีเลชันนั้นอยู่ในวงเล็บ โดยใช้เครื่องหมายจุลภาค (,) คั่นระหว่างแอททริบิวท์แต่ละตัว ซึ่งมีรูปแบบดังนี้

ชื่อรีเลชัน(ชื่อแอททริบิวท์ตัวที่1, ชื่อแอททริบิวท์ตัวที่2, ..., ชื่อแอททริบิวท์ตัวสุดท้าย)

เช่น

อาจารย์ (เลขประจำตำแหน่ง, ชื่อ-สกุล, ภาควิชา, เงินเดือน, ความชำนาญพิเศษ)

10.3 คุณสมบัติที่สำคัญของรีเลชัน

รีเลชันในแบบจำลองฐานข้อมูลเชิงสัมพันธ์ จะมีคุณสมบัติ ดังนี้

1) แต่ละช่องในตารางจะบรรจุข้อมูลเพียง 1 ค่าเท่านั้น

2) แต่ละคอลัมน์ในตารางหนึ่ง ๆ ต้องมีชื่อเรียกที่แตกต่างกัน โดยชื่อของแต่ละคอลัมน์เป็นชื่อของแต่ละแอททริบิวท์

3) ค่าของข้อมูลที่อยู่ในแต่ละคอลัมน์คือค่าของแอททริบิวท์ และต้องอยู่ในขอบเขตของโดเมนที่กำหนดไว้สำหรับคอลัมน์นั้น

4) ข้อมูลในแต่ละแถวของตารางต้องแตกต่างกัน

5) การเรียงลำดับก่อนหลังจากซ้ายไปขวาของคอลัมน์ในตารางไม่มีความสำคัญ

6) การเรียงลำดับก่อนหลังจากบนลงล่างของแถวในตารางไม่มีความสำคัญ

10.4 ประเภทของคีย์

ในฐานข้อมูลเชิงสัมพันธ์ประกอบด้วยรีเลชันที่มีความสัมพันธ์กันซึ่งแต่ละรีเลชันประกอบด้วยคีย์ประเภทต่าง ๆ ที่มีลักษณะที่แตกต่างกันไป

10.4.1 คีย์ (Key)

หมายถึง แอททริบิวท์ที่มีค่าไม่ซ้ำ (Unique) ที่สามารถระบุค่าของทัพเพิลในรีเลชันได้ในการกำหนดคีย์ อาจประกอบด้วยแอททริบิวท์มากกว่าหนึ่งแอททริบิวท์ประกอบกันเพื่อให้ค่าที่ไม่ซ้ำกัน คีย์ที่ประกอบด้วยหลายแอททริบิวท์เรียกว่า คีย์ผสม (Composite Key) ดังนั้น คีย์จะต้องไม่เป็นค่าว่าง (Not Null) เพื่อระบุค่าของทัพเพิลได้

10.4.2 คีย์หลัก (Primary Key : PK)

คีย์หลัก หมายถึง แอททริบิวท์หรือกลุ่มของแอททริบิวท์ที่สามารถใช้เพื่อชี้เฉพาะ เจาะจงว่าต้องการระบุถึงข้อมูลแถวใดในรีเลชันนั้น คุณสมบัติที่สำคัญของคีย์หลักมี 3 ประการคือ

1) ค่าของข้อมูลในแต่ละแถวของรีเลชัน จะไม่ซ้ำกัน

2) ในทุกรีเลชันจะต้องมีคีย์หลัก โดยที่คีย์หลักจะต้องประกอบด้วยแอททริบิวท์จำนวนที่น้อยที่สุดที่จะใช้ระบุถึงข้อมูลแถวหนึ่งแถวใดได้ (Minimality) คือจะต้องไม่สามารถดึงเอาแอททริบิวท์ใดออกไปแล้วส่วนที่เหลือยังคงสามารถเป็นคีย์หลักได้

3) ค่าแอททริบิวท์คีย์หลักจะต้องไม่เป็นค่านัล (Not Null)

4) การแสดงโครงสร้างของรีเลชันด้วยข้อความนั้นเพื่อแสดงถึงแอททริบิวท์ที่เป็นคีย์หลักจะใช้การขีดเส้นใต้ เช่น ในรีเลชันอาจารย์คีย์หลักคือแอททริบิวท์เลขประจำตำแหน่ง เนื่องจากอาจารย์แต่ละคนจะมีเลขประจำตำแหน่งไม่ซ้ำกัน โครงสร้างของรีเลชันอาจารย์จึงแสดง ได้ดังนี้

อาจารย์ (<u>เลขประจำตำแหน่ง</u>, ชื่อ-สกุล, ภาควิชา, เงินเดือน, ความชำนาญพิเศษ)

10.4.3 คีย์นอก (Foreign Key : FK)

คีย์นอกหรือคีย์อ้างอิง คือ แอททริบิวท์หรือกลุ่มของแอททริบิวท์ในรีเลชันหนึ่งซึ่งจะเป็นคีย์หลักในอีกรีเลชันหนึ่ง ดังนั้นค่าคีย์นอกจะมีค่าเท่ากับค่าคีย์หลักในแถวใดแถวหนึ่งของอีกรีเลชัน หรือมีค่าเป็น

ค่านัล (Null Value) คีย์นอกและคีย์หลักของอีกรีเลชันที่มีความสัมพันธ์กันไม่จำเป็นต้องมีชื่อเดียวกัน แต่ต้องอยู่ในโดเมนเดียวกัน คีย์นอกใช้ในการกำหนดความสัมพันธ์ระหว่างรีเลชัน และใช้ควบคุมความคงสภาพ (Integrity Rule) โดยคีย์นอกนี้จะอ้างอิงไปยังค่าที่มีอยู่จริงในอีกรีเลชันหนึ่ง

การแสดงโครงสร้างของรีเลชันด้วยข้อความแอททริบิวท์ที่เป็นคีย์นอกจะใช้การขีดเส้นประเพื่อแสดง เช่น รีเลชันเกี่ยวกับสังกัดของอาจารย์ซึ่งประกอบด้วยสองรีเลชัน คือ รีเลชันอาจารย์และรีเลชันคณะ แอททริบิวท์รหัสคณะจะเป็นคีย์นอกของรีเลชันอาจารย์และเป็นคีย์หลักของรีเลชันคณะ ดังนี้

อาจารย์ (<u>รหัสอาจารย์</u>, ชื่อ-สกุล,ห้องพัก , <u>รหัสคณะ</u>)

คณะ (<u>รหัสคณะ</u>, ชื่อคณะ, ที่ตั้งสำนักงานคณะ)

ตารางที่ 10.1 รีเลชันอาจารย์

<u>รหัสอาจารย์</u>	ชื่อ-สกุล	ห้องพัก	<u>รหัสคณะ</u>
6001	กีรติ แช่มรัมย์	521	A001
6002	จิรพันธ์ สุบินนาม	511	A001
6013	บัณฑิต สาระปัญญา	151004	A004
6034	นุศรา วิเศษ	523	A001
6045	เสาวลักษณ์ กันสุมาโส	1311	A002

ตารางที่ 10.2 รีเลชันคณะ

<u>รหัสคณะ</u>	ชื่อคณะ	ที่ตั้งสำนักงานคณะ
A001	วิทยาศาสตร์	1211
A002	วิทยาการจัดการ	1311
A003	มนุษยศาสตร์	611
A004	เกษตรศาสตร์	อาคารเกษตร
A005	ครุศาสตร์	1011

จากรีเลชันอาจารย์มีแอททริบิวท์รหัสคณะเป็นตัวเชื่อมความสัมพันธ์กับรีเลชันคณะหากเราต้องการรู้ว่าอาจารย์ที่ชื่อว่า กีรติ แช่มรัมย์ สังกัดคณะอะไรเราก็สามารถใช้รหัสคณะเชื่อมมาเพื่อดูรายละเอียดของรีเลชันคณะได้

10.4.4 คีย์คู่แข่ง (Candidate Key)

คีย์คู่แข่งคือ คีย์ที่เป็นคู่แข่งสำหรับคีย์หลักในรีเลชันหนึ่งเนื่องจากเป็นแอททริบิวท์ หรือกลุ่มของแอททริบิวท์ที่สามารถเป็นคีย์หลักได้ คีย์คู่แข่งที่ไม่ได้รับเลือกเป็นคีย์หลักจะเรียกว่าคีย์สำรอง (Alternate Key) จากรีเลชันพนักงานเราจะเห็นคีย์คู่แข่ง ดังนี้

พนักงาน(<u>รหัสพนักงาน</u>, ชื่อ-สกุล, แผนก, เงินเดือน, หมายเลขบัตรประชาชน)

ตารางที่ 10.3 รีเลชันพนักงาน

รหัสพนักงาน	ชื่อ-สกุล	แผนก	เงินเดือน	หมายเลขบัตรประชาชน
1001	ครรชิต สุขดี	IT	23,400	33001 01197 598
1002	ณัฐพล โพธิ์จันทา	IT	18,900	41014 04025 387
2015	นันทนา ฤดี	บัญชี	14,500	33201 11658 458
1003	พร้อมวิทย์ รัตนากุล	IT	19,300	53247 21101 235
3012	จันทรา มานนท์เวช	ประชาสัมพันธ์	12,500	42121 37544 215

แอททริบิวท์ที่เป็นคีย์คู่แข่งในรีเลชันพนักงานคือ แอททริบิวท์รหัสพนักงานกับแอททริบิวท์หมายเลขบัตรประชาชน เนื่องจาก รหัสพนักงานของพนักงานแต่ละคนจะไม่ซ้ำกันและ หมายเลขบัตรประชาชนของพนักงานแต่ละคนก็จะไม่ซ้ำกันเช่นกัน ดังนั้นแอททริบิวท์ รหัสพนักงานหรือแอททริบิวท์หมายเลขบัตรประชาชน จึงสามารถเป็นคีย์หลักในรีเลชันพนักงานได้ ดังนี้

พนักงาน (<u>รหัสพนักงาน</u>, ชื่อ-สกุล, แผนก, เงินเดือน, หมายเลขบัตรประชาชน) หรือ
พนักงาน (รหัสพนักงาน, ชื่อ-สกุล, แผนก, เงินเดือน, <u>หมายเลขบัตรประชาชน</u>)

ดังนั้นคีย์หลักก็คือคีย์ที่เลือกขึ้นมาจากบรรดาคีย์คู่แข่งในรีเลชันหนึ่งและหากคีย์ใดไม่ได้รับเลือกก็จะเป็นคีย์สำรองต่อไป

10.4.5 คีย์รวม (Composite Key)

คีย์รวมคือ คีย์ที่ประกอบด้วยแอททริบิวท์มากกว่าหนึ่งแอททริบิวท์ เช่น รีเลชันผลการเรียน ดังนี้

ผลการเรียน(<u>รหัสนักศึกษา</u>, <u>ภาคการศึกษา</u>, เกรดเฉลี่ย)

ตารางที่ 10.4 แสดงรีเลชันผลการเรียน

<u>รหัสนักศึกษา</u>	<u>ภาคเรียน</u>	เกรดเฉลี่ย
565146072	1/2557	2.98
565146073	2/2557	3.22
565146074	3/2557	3.14
565146072	2/2557	2.15
565146047	3/2557	1.98

คีย์หลักของรีเลชันผลการเรียนเป็นคีย์รวมที่ประกอบด้วยแอททริบิวท์รหัสนักศึกษา และแอททริบิวท์ภาคเรียน ด้วยในตารางมีข้อมูลของเกรดเฉลี่ยมากกว่าหนึ่งภาคเรียนสำหรับนักศึกษาแต่ละคน

10.4.6 นอลคีย์แอททริบิวท์ (Nonkey Attribute)

นอลคีย์แอททริบิวท์ คือ แอททริบิวท์อื่นในรีเลชันที่ไม่ใช่ส่วนใดส่วนหนึ่งของคีย์หลักในรีเลชันนั้น จากรีเลชันผลการเรียนจะมีคีย์หลักเป็นคีย์รวมที่ประกอบด้วยแอททริบิวท์รหัสนักศึกษาและแอททริบิวท์ภาคเรียน ส่วนนอลคีย์แอททริบิวท์ คือ แอททริบิวท์เกรดเฉลี่ย

10.5 การแปลงแผนภาพอี-อาร์ เป็นแบบจำลองฐานข้อมูลเชิงสัมพันธ์

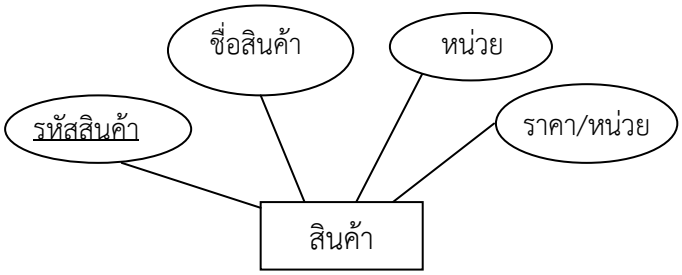
ปัจจุบันแบบจำลองฐานข้อมูลที่เป็นที่นิยมคือ แบบจำลองฐานข้อมูลเชิงสัมพันธ์ เพื่อให้สามารถออกแบบฐานข้อมูลได้อย่างถูกต้อง ขั้นตอนการแปลงจากแผนภาพอี-อาร์เป็นแบบจำลองฐานข้อมูลเชิงสัมพันธ์จึงถือว่าสำคัญซึ่งมีขั้นตอนคือ การแปลงจากแผนภาพอี-อาร์มาเป็นรีเลชัน และนำรีเลชันที่ได้มาผ่านกระบวนการนอร์มัลไลซ์ มีรายละเอียดดังนี้

10.5.1 การเปลี่ยนจากแผนภาพอี-อาร์มาเป็นรีเลชัน

การแปลงจากแผนภาพอี-อาร์มาเป็นรีเลชันประกอบด้วย 2 ขั้นตอนคือ การแปลงเอนทิตี้ และ การแปลงความสัมพันธ์ ดังนี้

1) การแปลงเอนทิตี้นั้นทำโดยการแปลงเอนทิตี้แต่ละชนิดในแผนภาพอี-อาร์ให้มาเป็นรีเลชันโดยให้แต่ละรีเลชันมีคอลัมน์ซึ่งมาจากแอทริบิวท์ของเอนทิตี้ และให้แอททริบิวท์ที่เป็นคีย์หลักของเอนทิตี้แต่ละชนิดนั้นมาเป็นคีย์หลักของรีเลชัน

ตัวอย่าง จากเอนทิตี้สินค้าประกอบด้วยแอทริบิวท์ดังภาพที่ 10.2



ภาพที่ 10.2 เอนทิตี้สินค้า

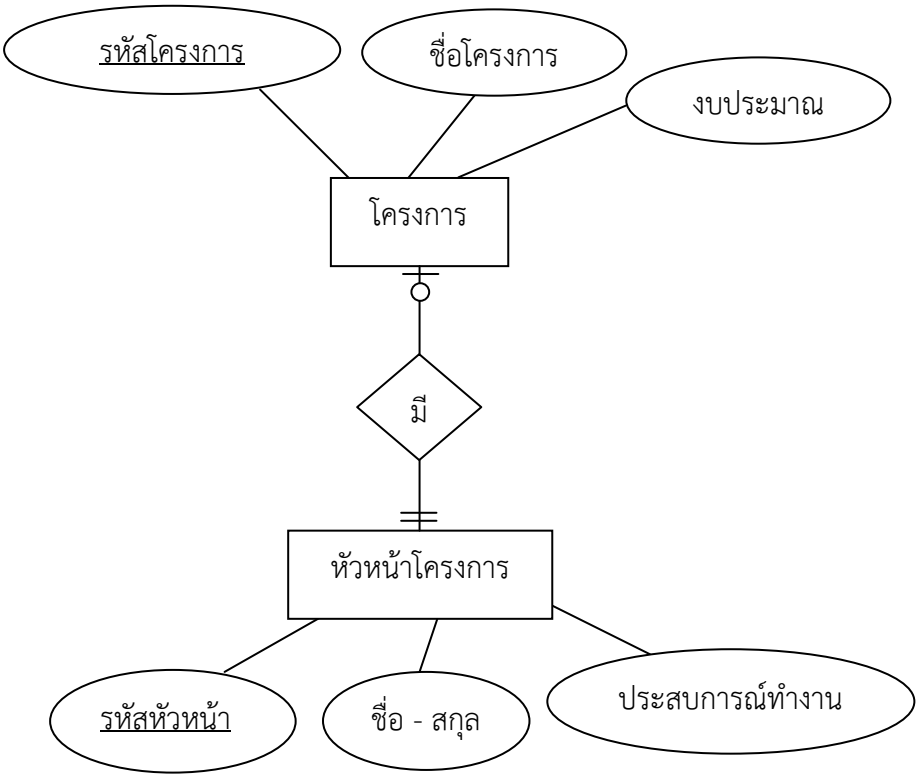
จากเอนทิตี้สินค้าทำการแปลงเป็นรีเลชันโดยใช้ชื่อเดียวกัน ซึ่งประกอบด้วยแอทริบิวท์ที่มาจากเอนทิตี้สินค้าและให้ รหัสสินค้าซึ่งเป็นคีย์หลักของเอนทิตี้สินค้ามาเป็นคีย์หลักของรีเลชันด้วย ดังนี้

สินค้า(<u>รหัสสินค้า</u>, ชื่อสินค้า, หน่วย, ราคา/หน่วย)

2) การแปลงความสัมพันธ์ระหว่างเอนทิตี้ ลักษณะความสัมพันธ์ระหว่างเอนทิตี้มี 3 แบบ ได้แก่ ความสัมพันธ์แบบหนึ่งต่อหนึ่ง (1:1) ความสัมพันธ์แบบหนึ่งต่อกลุ่ม (1:M) และความสัมพันธ์แบบกลุ่มต่อกลุ่ม (M:N) แต่ละแบบจะมีลักษณะของการแปลง ดังนี้

(1) การแปลงความสัมพันธ์แบบหนึ่งต่อหนึ่ง สามารถแทนความสัมพันธ์ระหว่างเอนทิตี้ A และเอนทิตี้ B โดยให้คีย์หลักของรีเลชันซึ่งแทนเอนทิตี้ A ไปเป็นคีย์นอกของรีเลชันซึ่งแทนเอนทิตี้ B หรือนำเอาคีย์หลักของรีเลชันซึ่งแทนเอนทิตี้ B ไปเป็นคีย์นอกของรีเลชันซึ่งแทนเอนทิตี้ A

ตัวอย่าง พิจารณาความสัมพันธ์แบบหนึ่งต่อหนึ่งระหว่างเอนทิตี้โครงการและเอนทิตี้หัวหน้าโครงการ ซึ่งมีแอททริบิวท์ดังภาพที่ 10.3



ภาพที่ 10.3 แผนภาพอี-อาร์แสดงความสัมพันธ์ระหว่างเอนทิตี้โครงการกับเอนทิตี้หัวหน้าโครงการ

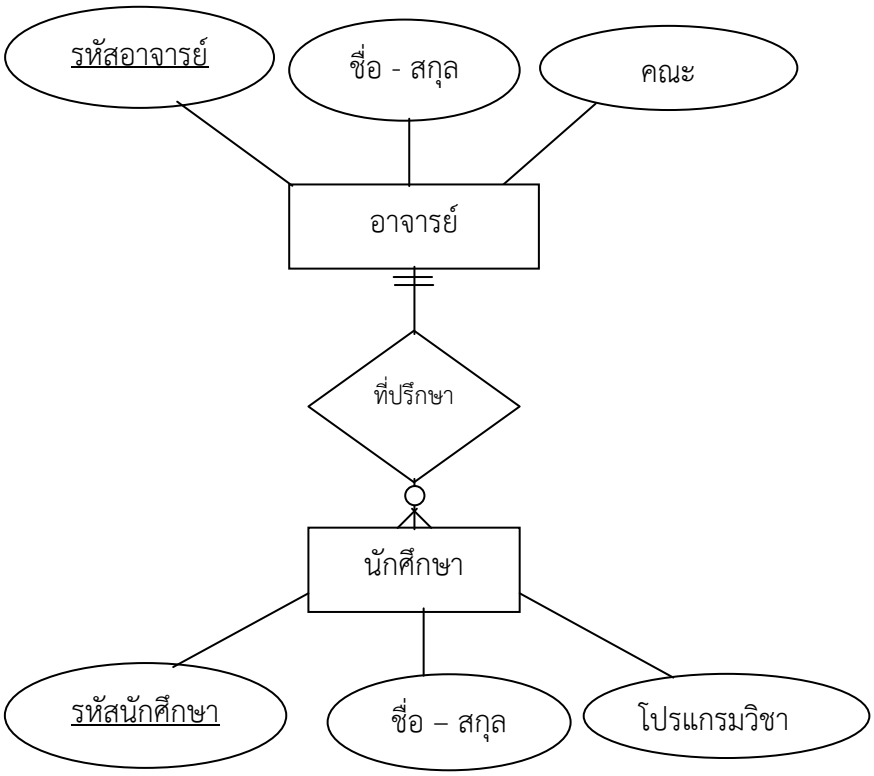
จากแผนภาพอีอาร์แสดงความสัมพันธ์แบบหนึ่งต่อหนึ่งระหว่างเอนทิตี้โครงการกับเอนทิตี้หัวหน้าโครงการจะถูกแปลงไปเป็นรีเลชันได้โดยกำหนดให้รหัสหัวหน้า เป็นคีย์นอกของรีเลชันโครงการ เนื่องจากรหัสหัวหน้าเป็นคีย์หลักของรีเลชันหัวหน้าโครงการดังนี้

โครงการ(รหัสโครงการ, ชื่อโครงการ, งบประมาณ, รหัสหัวหน้า)
หัวหน้าโครงการ(รหัสหัวหน้า, ชื่อ-สกุล, ประสบการณ์ทำงาน)

สำหรับการแปลงความสัมพันธ์แบบ 1:1 นี้ จะได้ตารางที่ไม่มีความซ้ำซ้อนจึงไม่ต้องผ่านกระบวนการนอร์มัลไลท์เซชัน

(2) การแปลงความสัมพันธ์แบบหนึ่งต่อกลุ่ม ให้นำคีย์หลักของเอนทิตี้ทางด้านความสัมพันธ์เป็นหนึ่ง (One) ไปเป็นคีย์นอกของรีเลชันทางด้านความสัมพันธ์เป็นกลุ่ม (Many)

ตัวอย่าง พิจารณาความสัมพันธ์แบบหนึ่งต่อกลุ่มระหว่างเอนทิตี้อาจารย์และเอนทิตี้นักศึกษา ซึ่งมีแอททริบิวท์ดังภาพที่ 10.4 โดยกำหนดให้นักศึกษาแต่ละคนมีอาจารย์ที่ปรึกษาเพียงหนึ่งคน และอาจารย์แต่ละคนจะเป็นที่ปรึกษาให้นักศึกษาได้หลายคน



ภาพที่ 10.4 แผนภาพอี-อาร์ แสดงความสัมพันธ์ระหว่างเอนทิตี้อาจารย์และเอนทิตี้นักศึกษา

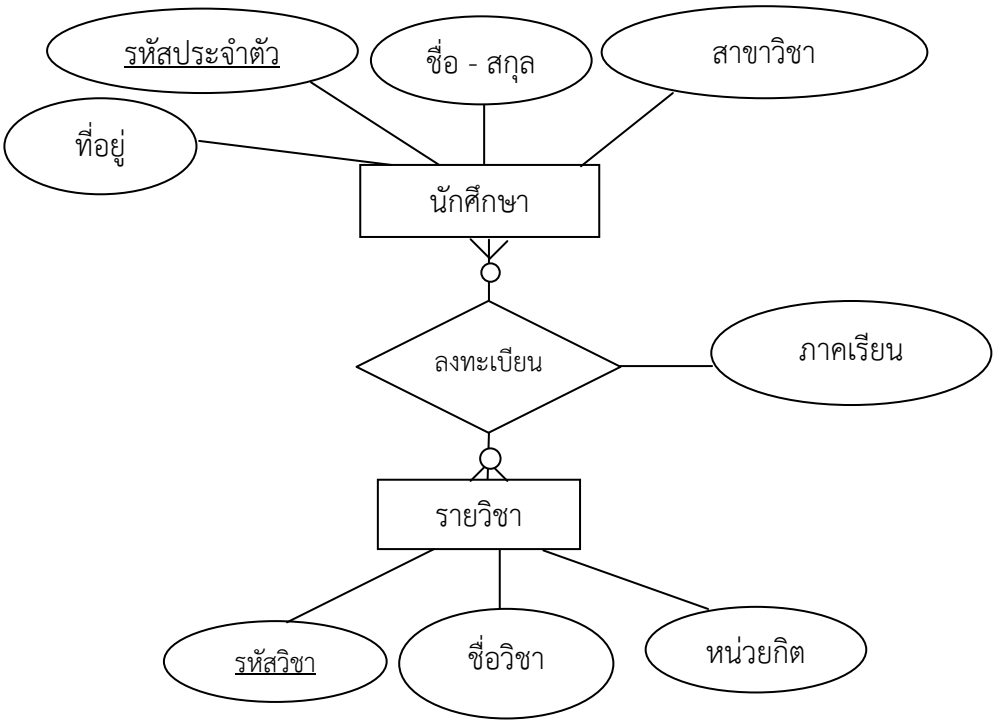
จากแผนภาพอี-อาร์แสดงความสัมพันธ์แบบหนึ่งต่อกลุ่มระหว่างเอนทิตี้อาจารย์และเอนทิตี้นักศึกษา จะถูกแปลงไปเป็นรีเลชันได้โดยกำหนดให้ “รหัสอาจารย์” ซึ่งเป็นคีย์หลักของเอนทิตี้ทางด้านความสัมพันธ์เป็นหนึ่งให้เป็นคีย์นอกของรีเลชันนักศึกษาทางด้านความสัมพันธ์เป็นกลุ่ม ดังนี้

อาจารย์(รหัสอาจารย์ , ชื่อ – สกุล, คณะ)

นักศึกษา(รหัสนักศึกษา , ชื่อ – สกุล, โปรแกรมวิชา, รหัสอาจารย์)

(3) การแปลงความสัมพันธ์แบบกลุ่มต่อกลุ่มระหว่างเอนทิตี้ 2 เอนทิตี้คือ A และ B ทำโดยสร้างรีเลชันขึ้นอีก 1 รีเลชันคือ C โดยให้รีเลชัน C มีคีย์หลักเป็นคีย์รวมที่สร้างขึ้นจากคีย์หลักของรีเลชัน A และ B และนอลคีย์แอททริบิวท์ของความสัมพันธ์แบบกลุ่มต่อกลุ่มนั้นจะเป็นแอททริบิวท์ของรีเลชัน C ด้วย

ตัวอย่าง พิจารณาความสัมพันธ์แบบกลุ่มต่อกลุ่มระหว่างเอนทิตี้นักศึกษาและเอนทิตี้รายวิชา ซึ่งมีแอททริบิวท์ดังภาพที่ 10.5 กำหนดให้นักศึกษาแต่ละคนลงทะเบียนเรียนได้หลายวิชาและแต่ละรายวิชาจะมีนักศึกษามาลงทะเบียนได้หลายคน



ภาพที่ 10.5 แผนภาพอี-อาร์แสดงความสัมพันธ์ระหว่างนักศึกษากับรายวิชา

จากแผนภาพอี-อาร์แสดงความสัมพันธ์แบบกลุ่มต่อกลุ่มระหว่างเอนทิตี้นักศึกษากับเอนทิตี้รายวิชาจะถูกแปลงไปเป็นรีเลชันได้โดยสร้างเป็นรีเลชันการลงทะเบียนเพิ่มขึ้นอีกรีเลชันโดยให้รีเลชันการลงทะเบียนมีคีย์หลักเป็นคีย์รวมที่สร้างขึ้นจากคีย์หลักของรีเลชันนักศึกษาและรีเลชันรายวิชาแล้วนำแอททริบิวท์ภาคเรียนของความสัมพันธ์ลงทะเบียนมาเป็นแอททริบิวท์ของรีเลชันการลงทะเบียนด้วยจึงเกิดเป็น 3 รีเลชันดังนี้

นักศึกษา(<u>รหัสประจำตัว</u>, ชื่อ-สกุล, ที่อยู่, โปรแกรมวิชา)

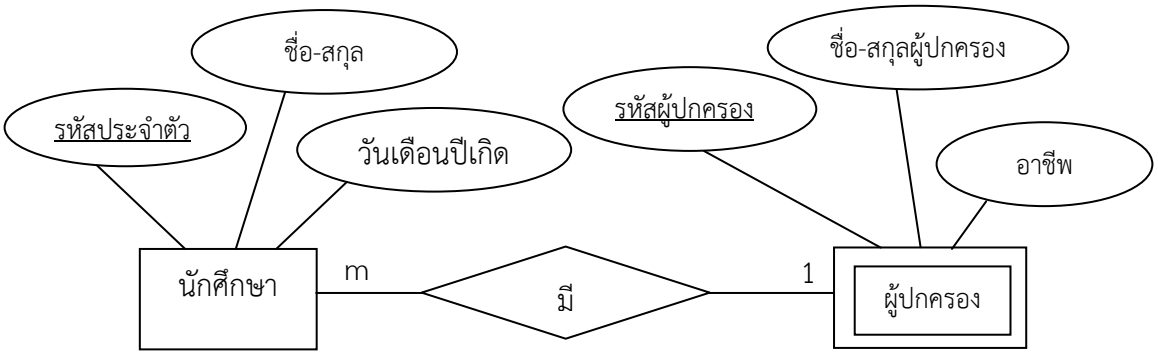
รายวิชา(<u>รหัสวิชา</u>, ชื่อวิชา, หน่วยกิต)

การลงทะเบียน(<u>รหัสประจำตัว, รหัสวิชา</u>, ภาคเรียน)

10.5.2 การแปลงเอนทิตี้แบบอ่อน (Weak Entity)

การแปลงเอนทิตี้แบบอ่อน ซึ่งเป็นเอนทิตี้ที่จะปรากฏในฐานข้อมูลได้ต่อเมื่อมีอีกเอนทิตี้หนึ่งอยู่ในฐานข้อมูลด้วยเท่านั้นสามารถแปลงเอนทิตี้แบบอ่อนให้เป็นรีเลชันทำโดยการสร้างรีเลชันของทั้งเอนทิตี้แบบอ่อน และเอนทิตี้ที่เอนทิตี้แบบอ่อนนั้นมีการขึ้นต่อกันเชิงระบุ โดยให้รีเลชันของเอนทิตี้แบบอ่อนประกอบด้วยแอททริบิวท์ทั้งหมดของเอนทิตี้นั้น และจะต้องนำแอททริบิวท์ที่เป็นคีย์ของเอนทิตี้ที่เอนทิตี้แบบอ่อนนั้นมีการขึ้นต่อกันเชิงระบุมาอยู่ในรีเลชันของเอนทิตี้แบบอ่อนและให้เป็นส่วนหนึ่งของคีย์หลักของรีเลชันนั้น

ตัวอย่าง พิจารณาความสัมพันธ์ระหว่างเอนทิตี้นักศึกษากับเอนทิตี้แบบอ่อนคือเอนทิตี้ผู้ปกครอง ซึ่งมีแอททริบิวท์ดังภาพที่ 10.6



ภาพที่ 10.6 ความสัมพันธ์ระหว่างเอนทิตี้นักศึกษากับเอนทิตี้ผู้ปกครอง

จากแผนภาพอี-อาร์แสดงความสัมพันธ์ระหว่างเอนทิตี้นักศึกษากับเอนทิตี้ผู้ปกครองจะเปลี่ยนไปเป็นรีเลชัน 2 รีเลชัน ดังนี้

นักศึกษา (รหัสประจำตัว, ชื่อ-นามสกุล, วันเดือนปีเกิด)

ผู้ปกครอง (รหัสประจำตัว, รหัสผู้ปกครอง, ชื่อ-สกุลผู้ปกครอง, อาชีพ)

10.6 การรวมรีเลชัน

เมื่อเราทำการแปลงรีเลชันมาจากแผนภาพอี-อาร์แล้วบางรีเลชันที่ได้นั้นเป็นรีเลชันของเอนทิตี้ชนิดเดียวกัน ดังนั้นเราควรจะทำการรวมรีเลชันเหล่านั้นเพื่อจะลดความซ้ำซ้อนของข้อมูลลง และจะต้องพิจารณาถึงความหมายของแอทริบิวท์ของแต่ละรีเลชันด้วยซึ่งบางครั้งชื่อแอทริบิวท์เดียวกันอาจแทนข้อมูลไม่เหมือนกันหรือชื่อของแอทริบิวท์ต่างกันแต่แทนข้อมูลเดียวกันซึ่งต้องอาศัยการทำความเข้าใจของนักวิเคราะห์ระบบ

เช่น จากการแปลงแผนภาพอี-อาร์ในระบบเดียวกันจะได้รีเลชันลักษณะเดียวกัน 2 รีเลชัน คือ รีเลชันนักศึกษา1 และรีเลชันนักศึกษา2 ดังนี้

นักศึกษา1(รหัสประจำตัว, ชื่อ-สกุล, โปแกรรมวิชา)

นักศึกษา2(รหัสประจำตัว, ชื่อ-สกุล, ที่อยู่)

เราสามารถจะรวมให้เป็นรีเลชันเดียวกันได้ดังนี้

นักศึกษา(รหัสประจำตัว, ชื่อ-สกุล, ที่อยู่, โปรแกรมวิชา)

10.7 การนอร์มอลไลซ์เซชัน

การนอร์มอลไลซ์เซชัน (Normalization) คือ การแปลงข้อมูลที่มีโครงสร้างแบบรีเลชันนอลจากรูปแบบที่มีความซ้ำซ้อน (Redundancy) ให้อยู่ในรูปแบบที่ใช้ได้ง่ายเพื่อทำให้การเพิ่มข้อมูล ลบข้อมูล หรือแก้ไขข้อมูล ที่อยู่ในรีเลชันได้โดยไม่เกิดความผิดพลาด หรือเกิดความไม่คงที่ ความไม่แน่นอนและความขัดแย้ง (Inconsistency) ของข้อมูลซึ่งเรียกว่าความผิดปกติ (Anomaly)

ความซ้ำซ้อนของข้อมูลในรีเลชันอาจจะทำให้เกิดความผิดปกติ 3 ลักษณะคือ ความผิดปกติจากการเพิ่มข้อมูล (Insertion Anomaly) ความผิดปกติจากการลบข้อมูล (Deletion Anomaly) และความผิดปกติจากการแก้ไขข้อมูล (Update Anomaly)

รูปแบบบรรทัดฐานข้อมูล (Normal Form : NF) เป็นการจัดแบ่งระดับรีเลชันตามคุณสมบัติของรีเลชันนั้น รูปแบบบรรทัดฐานที่สำคัญมีดังนี้

1) รูปแบบบรรทัดฐานระดับที่หนึ่ง (First Normal Form :1NF)

2) รูปแบบบรรทัดฐานระดับที่สอง (Second Normal Form :2NF)

3) รูปแบบบรรทัดฐานระดับที่สาม (Third Normal Form :3NF)

โดยทั่วไปการทำนอร์มอลไลซ์จะพิจารณาจากโครงสร้างของข้อมูลว่าจัดอยู่ในบรรทัดฐานระดับใดแล้วจึงทำนอร์มอลไลซ์จากระดับนั้นต่อไปเรื่อย ๆ ส่วนการพิจารณาว่าจะสิ้นสุดการนอร์มอลไลซ์ที่ระดับใดนั้นขึ้นอยู่กับว่าโครงสร้างของข้อมูลที่ได้จากการนอร์มอลไลซ์นั้นมีความถูกต้องตามความหมายที่ได้กำหนดไว้แล้วหรือไม่ ถ้าผลที่ได้จากการนอร์มอลไลซ์ในระดับใดส่งผลให้โครงสร้างของข้อมูลมีความหมายตามที่กำหนดไว้ การทำนอร์มอลไลซ์ก็จะสิ้นสุดในระดับนั้น

10.7.1 ฟังก์ชันการขึ้นต่อกัน (Functional Dependency : FD)

เมื่อพิจารณาความสัมพันธ์ระหว่างข้อมูลของแอททริบิวท์สองแอททริบิวท์ คือ แอททริบิวท์ A และ B ในรีเลชัน จะกล่าวได้ว่า B มีฟังก์ชันการขึ้นต่อกันอยู่กับ A ก็ต่อเมื่อค่าของ A สามารถใช้ในการกำหนดค่าของ B ได้เพียง 1 ค่าเสมอ การที่ B มีฟังก์ชันการขึ้นอยู่กับ A หรือกล่าวสั้นๆว่า B "ขึ้นกับ" A จะเขียนแทนได้ด้วยสัญลักษณ์ A→ B

การที่ข้อมูลแต่ละค่าของแอททริบิวท์ A จะใช้ในการกำหนดค่า(Determines) ข้อมูลของแอททริบิวท์ B เพียง 1 ค่า จะเรียก A ว่าเป็นตัวกำหนดค่า(Determinant) เรียก B ว่าเป็นตัวถูกกำหนดค่า (Dependent) สัญลักษณ์ A→ B จึงใช้เพื่อแสดงว่า A กำหนดค่า B

เช่น ชื่อของนักศึกษา "ขึ้นกับ" รหัสของนักศึกษา เขียนแทนด้วยสัญลักษณ์ รหัสนักศึกษา → ชื่อของนักศึกษา เพราะเมื่ออ้างถึง รหัสนักศึกษา 1 ค่า จะได้ชื่อนักศึกษามา 1 ชื่อเสมอ แต่รหัสนักศึกษาจะไม่ขึ้นกับชื่อนักศึกษาเพราะอาจมีนักศึกษาที่มีชื่อซ้ำกันในรีเลชัน นอลคีย์แอททริบิวท์จะมีฟังก์ชันการขึ้นอยู่กับคีย์หลัก

10.7.2 รูปแบบบรรทัดฐานระดับที่ 1

รูปแบบบรรทัดฐานระดับที่หนึ่งคือรูปแบบรีเลชันที่ไม่มีกลุ่มข้อมูลซ้ำ(Repeating Group) คือทุกแอททริบิวท์ของแต่ละแถวจะมีข้อมูลเพียงหนึ่งค่าเท่านั้น

การทำให้รีเลชันอยู่ในรูปแบบบรรทัดฐานระดับที่หนึ่ง ทำได้โดย

1) แยกคอลัมน์ที่มีมากกว่า 1 ค่าออกเป็นแถวใหม่

2) เพิ่มข้อมูลที่เหมาะสมเข้าไปในคอลัมน์ที่ว่างอยู่ของแถวที่เกิดขึ้นใหม่

3) กำหนดคีย์หลักให้กับรีเลชัน

ตัวอย่าง การทำเป็นรูปแบบบรรทัดฐานระดับที่ 1 ในรีเลชันอาจารย์ ดังตารางที่ 10.5

อาจารย์(<u>รหัสอาจารย์</u>, ชื่อ-สกุล, แผนก, เงินเดือน, <u>การอบรม</u>, วันที่สำเร็จการอบรม)

ตารางที่ 10.5 ข้อมูลในรีเลชันอาจารย์

	รหัสอาจารย์	ชื่อ-สกุล	คณะ	เงินเดือน	การอบรม	วันที่สำเร็จการอบรม
แถวที่ 1	4102	ดวงฤดี ซื่อสัตย์	IT	24,500	ภาษา C VB.NET	12/2/57 27/10/57
แถวที่ 2	2519	นฤมล สุภาพ	บริหารธุรกิจ	35,000	นโยบายการเงิน	22/9/55
แถวที่ 3	5214	สมคิด ดวงไธสง	IT	18,000	Unix ภาษา C	20/5/56 25/7/58
แถวที่ 4	4095	สุดา มหาดำรงค์	วิทยาศาสตร์	34,000	การเพาะเนื้อเยื่อ	1/11/57

จากตารางข้อมูลในรีเลชันอาจารย์นั้นจะมีกลุ่มข้อมูลซ้ำ คือ แอททริบิวท์การอบรมและแอททริบิวท์วันที่สำเร็จการอบรมในแถวที่ 1 และ 3

การแปลงให้รีเลชันอยู่ในรูปแบบบรรทัดฐานระดับที่หนึ่ง ทำโดยแยกแถวข้อมูลที่มีกลุ่มข้อมูลซ้ำคือแถวที่ 1 และ 3 ออกโดยการเติมข้อมูลให้ครบทุกแอททริบิวท์ และกำหนดคีย์หลักของรีเลชันอาจารย์ ซึ่งเป็นคีย์รวมที่ประกอบด้วยแอททริบิวท์รหัสอาจารย์และแอททริบิวส์การอบรม ดังตารางที่ 10.6

ตารางที่ 10.6 รีเลชันอาจารย์ในรูปแบบบรรทัดฐานระดับที่หนึ่ง

รหัสอาจารย์	ชื่อ-สกุล	คณะ	เงินเดือน	การอบรม	วันที่สำเร็จการอบรม
4102	ดวงฤดี ซื่อสัตย์	IT	24,500	ภาษา C	12/2/57
4102	ดวงฤดี ซื่อสัตย์	IT	24,500	VB.NET	27/10/57
2519	นฤมล สุภาพ	บริหารธุรกิจ	35,000	นโยบายการเงิน	22/9/55
5214	สมคิด ดวงไธสง	IT	18,000	Unix	20/5/56
5214	สมคิด ดวงไธสง	IT	18,000	JAVA	25/7/58
4095	สุดา มหาดำรงค์	วิทยาศาสตร์	34,000	การเพาะเนื้อเยื่อ	1/11/57

อาจารย์(รหัสอาจารย์, ชื่อ-สกุล, คณะ, เงินเดือน, การอบรม, วันที่สำเร็จการอบรม)

10.7.3 รูปแบบบรรทัดฐานระดับที่สอง

รีเลชันที่อยู่ในรูปแบบบรรทัดฐานระดับที่สองคือรีเลชันที่อยู่ในรูปแบบบรรทัดฐานระดับที่หนึ่งและนอลคีย์แอททริบิวท์ทุกตัวจะต้องขึ้นกับคีย์หลักอย่างสมบูรณ์ นั่นก็คือไม่มีนอลคีย์แอททริบิวท์ใดที่ขึ้นอยู่กับเฉพาะส่วนใดส่วนหนึ่งของคีย์หลัก กรณีที่คีย์หลักเป็นคีย์รวมที่ประกอบด้วยแอททริบิวท์มากกว่าหนึ่งแอททริบิวท์ การที่มีนอลคีย์แอททริบิวท์ใดขึ้นอยู่กับเฉพาะส่วนใดส่วนหนึ่งของคีย์หลักเรียกว่า การขึ้นต่อกันเพียงบางส่วน (Partial Functional Dependency)

การแปลงรีเลชันในรูปแบบบรรทัดฐานระดับที่หนึ่ง ให้เป็นรีเลชันอยู่ในรูปแบบบรรทัดฐานระดับที่สอง ทำโดย

1) การแยกแอททริบิวท์กลุ่มที่ขึ้นอยู่กับบางส่วนของคีย์หลักออกมาเป็นรีเลชันใหม่

2) โดยให้แอททริบิวท์ที่เป็นตัวกำหนดค่าในกลุ่มนี้เป็นคีย์หลักของรีเลชันใหม่

จากรีเลชันอาจารย์ในตารางที่ 10.6 รีเลชันอาจารย์อยู่ในรูปแบบบรรทัดฐานระดับที่หนึ่งแล้ว มีคีย์หลักเป็นคีย์รวมที่ประกอบด้วยแอททริบิวท์รหัสอาจารย์ และการอบรม การขึ้นต่อกันของแอททริบิวท์ต่างๆในรีเลชันเป็นดังนี้

รหัสอาจารย์, การอบรม → ชื่อ-สกุล, คณะ ,เงินเดือน, วันที่สำเร็จการอบรม

นั่นคือ “รหัสอาจารย์” และ “การอบรม” จะเป็นตัวกำหนดค่าให้กับแอททริบิวท์ “ชื่อ-สกุล” “คณะ” “เงินเดือน” และ “วันที่สำเร็จการอบรม” และมีการขึ้นต่อกันเพียงบางส่วนของคีย์หลักคือเฉพาะแอททริบิวท์“รหัสอาจารย์” จะเป็นตัวกำหนดค่าให้กับแอททริบิวท์“ชื่อ-สกุล” “คณะ” “ เงินเดือน” ดังนี้

รหัสอาจารย์ → ชื่อ-สกุล, คณะ, เงินเดือน

นั่นคือนอลคีย์แอททริบิวท์“ชื่อ-สกุล”,“คณะ”และ“เงินเดือน”ขึ้นอยู่กับแอททริบิวท์ “รหัสอาจารย์” เท่านั้น ซึ่ง “รหัสอาจารย์” เป็นเพียงส่วนหนึ่งของคีย์หลัก

ปัญหาที่อาจเกิดขึ้นกับรีเลชันที่มีรูปแบบบรรทัดฐานระดับที่ 1 รีเลชันที่อยู่ในรูปแบบบรรทัดฐานระดับที่หนึ่งแล้วแต่มีการเก็บข้อมูลซ้ำซ้อนจะทำให้เกิดปัญหาขึ้นได้ พิจารณาจากรีเลชันอาจารย์ซึ่งในรูปแบบบรรทัดฐานระดับที่หนึ่ง แล้วแต่ยังมีปัญหาในลักษณะของความผิดปกติเกิดขึ้นได้ดังนี้

1) ความผิดปกติจากการเพิ่มข้อมูล การเพิ่มข้อมูลแต่ละแถวจะทำได้เมื่อได้ข้อมูลของทั้ง 2 แอททริบิวท์คือ แอททริบิวท์รหัสอาจารย์และแอททริบิวท์การอบรมเท่านั้น เนื่องจากคีย์หลักของรีเลชันอาจารย์ คือคีย์รวมที่ประกอบด้วยแอททริบิวท์รหัสอาจารย์และแอททริบิวท์การอบรม ดังนั้นข้อมูลในคอลัมน์รหัสอาจารย์หรือการอบรมของแถวใด ๆ จะไม่สามารถเป็นค่านัลได้ กรณีที่มีอาจารย์เข้ามาใหม่ซึ่งยังไม่ได้ผ่านการอบรมหลักสูตรใด ๆ เลย จะยังไม่สามารถเพิ่มข้อมูลอาจารย์ผู้นั้นลงในรีเลชันอาจารย์ได้

2) ความผิดปกติจากการลบข้อมูล ถ้ามีการลบข้อมูลของอาจารย์รหัส 4102 ออกไปจากรีเลชันอาจารย์ ในตารางที่ 10.6 จะทำให้รายละเอียดที่แสดงว่าอาจารย์ผู้นี้สำเร็จการอบรมภาษา C เมื่อ 12 กุมภาพันธ์ 2557 และสำเร็จการอบรม VB.NET เมื่อ 27 ตุลาคม 2557 หายไป และส่งผลให้ข้อมูลที่แสดงว่ามีการเปิดอบรมหลักสูตร VB.NET ซึ่งสิ้นสุดในวันที่ 27 ตุลาคม 2557 หายไปด้วย นั่นคือการลบข้อมูลส่วนหนึ่งอาจจะทำให้สูญเสียข้อมูลส่วนอื่นที่ยังคงต้องเก็บไว้ไปด้วย

ตารางที่ 10.7 รีเลชันอาจารย์หลังจากการลบข้อมูลอาจารย์รหัส 4102 ออกไป

รหัสอาจารย์	ชื่อ-สกุล	แผนก	เงินเดือน	การอบรม	วันที่สำเร็จการอบรม
2519	นฤมล สุภาพ	การเงิน	35,000	นโยบายการเงิน	22/9/55
5214	สมคิด ดวงไธสง	IT	18,000	Unix	20/5/56
5214	สมคิด ดวงไธสง	IT	18,000	JAVA	25/7/58
4095	สุดา มหาดำรงค์	บุคคล	34,000	กฎหมายแรงงาน	1/11/57

3) ความผิดปกติจากการแก้ไขข้อมูล ถ้าต้องการแก้ไขข้อมูลเงินเดือนของอาจารย์คนใดคนหนึ่งจะต้องทำให้ครบทุกแถวที่เป็นข้อมูลของอาจารย์คนนั้น เช่น กรณีที่อาจารย์รหัส 5214 ได้รับการขึ้นเงินเดือน เป็น 19,500 บาท จะต้องทำการแก้ไขเงินเดือนอาจารย์คนนี้ให้ครบทุก ๆ แถว (จากตัวอย่างในตารางที่ 6.2 คือแถวที่ 4 และแถวที่ 5) มิฉะนั้นแล้วเงินเดือนของอาจารย์คนนี้จะไม่ถูกต้องตรงกัน

จากตัวอย่างรีเลชันอาจารย์จะแปลงรีเลชันอาจารย์ให้เป็นรีเลชันในรูปแบบบรรทัดฐานระดับที่สอง โดยการแยกแอททริบิวท์กลุ่มที่ขึ้นอยู่กับบางส่วนของคีย์หลักออกมาเป็นรีเลชันใหม่ชื่อรีเลชัน “รายละเอียดอาจารย์” โดยให้แอททริบิวท์ “รหัสอาจารย์” ซึ่งเป็นตัวกำหนดค่าเป็นคีย์หลักของรีเลชันใหม่นี้ดังตารางที่ 10.8 และตารางที่ 10.9 จากรีเลชันอาจารย์ จึงแยกออกเป็น 2 รีเลชันคือ “รายละเอียดอาจารย์” และ “การอบรมอาจารย์” ดังนี้

รายละเอียดอาจารย์(รหัสอาจารย์, ชื่อ-สกุล,แผนก, เงินเดือน)

การอบรมอาจารย์(รหัสอาจารย์, การอบรม, วันที่สำเร็จการอบรม)

ตารางที่ 10.8 รีเลชันรายละเอียดอาจารย์

รหัสอาจารย์	ชื่อ-สกุล	คณะ	เงินเดือน
4102	ดวงฤดี ซื่อสัตย์	IT	24,500
2519	นฤมล สุภาพ	บริหารธุรกิจ	35,000
5214	สมคิด ดวงไธสง	IT	18,000
4095	สุดา มหาดำรงค์	วิทยาศาสตร์	34,000

ตารางที่ 10.9 รีเลชันการอบรมอาจารย์

รหัสอาจารย์	การอบรม	วันที่สำเร็จการอบรม
4102	ภาษา C	12/2/57
4102	VB.NET	27/10/57
2519	กฎหมายแรงงาน	22/9/55
5214	Unix	20/5/56
5214	ภาษา C	25/7/58
4095	เพาะเลี้ยงเนื้อเยื่อ	1/11/57

ในกรณีที่รีเลชันอยู่ในรูปแบบบรรทัดฐานระดับที่หนึ่งอยู่แล้ว รีเลชันนั้นจะอยู่ในรูปแบบบรรทัดฐานระดับที่สองเมื่อมีลักษณะดังนี้

(1) คีย์หลักประกอบด้วยแอททริบิวท์เพียงหนึ่งแอททริบิวท์เท่านั้น หรือ

(2) ในรีเลชันนั้นไม่มีนอลคีย์แอททริบิวท์ เนื่องจากคีย์หลักประกอบด้วยทุกๆแอททริบิวท์ หรือ

(3) นอลคีย์แอททริบิวท์ทุกตัวจะต้องขึ้นกับคีย์หลักอย่างสมบูรณ์

10.7.4 รูปแบบบรรทัดฐานระดับที่สาม

รีเลชันที่อยู่ในรูปแบบบรรทัดฐานระดับที่สาม คือรีเลชันที่อยู่ในรูปแบบบรรทัดฐานระดับที่สอง และไม่มีแอททริบิวท์ใดขึ้นอยู่กับแอททริบิวท์อื่นๆ ที่ไม่ใช่คีย์หลัก นั่นคือนอลคีย์แอททริบิวท์ทุกตัวจะต้องขึ้นกับคีย์หลักเท่านั้น การขึ้นต่อกันระหว่างนอลคีย์แอททริบิวท์ด้วยกันเองเรียกว่าการขึ้นต่อกันแบบทรานซิทีพ (Transitive Dependency) ซึ่งเป็นผลมาจากยังมีความซ้ำซ้อนในการเก็บข้อมูล

ตัวอย่าง พิจารณารีเลชัน “การสังกัดชมรม” ซึ่งมีข้อกำหนดดังนี้

- นักศึกษาแต่ละคนสังกัดเพียงชมรมเดียวเท่านั้น
- แต่ละชมรมจะมีห้องประจำชมรมแตกต่างกัน
- นักศึกษาที่อยู่ชมรมเดียวกันต้องใช้ห้องประจำชมรมเดียวกัน

การสังกัดชมรม (<u>รหัสนักศึกษา</u>, ชมรม, ห้องประจำชมรม)

ตารางที่ 10.10 รีเลชันการสังกัดชมรม

<u>รหัสนักศึกษา</u>	ชมรม	ห้องประจำชมรม
565146052	คณิตศาสตร์	511
565146079	คอมพิวเตอร์	522
575146034	คณิตศาสตร์	511
575146043	ห้องสมุด	15011
555146001	ดนตรี	1402

“รหัสนักศึกษา” ซึ่งเป็นคีย์หลักจะเป็นตัวกำหนดค่าให้แอททริบิวท์ “ชมรม” และ“ห้องประจำชมรม” และมีการขึ้นต่อกันระหว่างนอลคีย์แอททริบิวท์ “ชมรม” และ “ห้องประจำชมรม” ดังนี้

ชมรม → ห้องประจำชมรม

ปัญหาที่อาจเกิดขึ้นกับรีเลชันที่มีรูปแบบบรรทัดฐานระดับที่ 2 การขึ้นต่อกันระหว่างนอลคีย์แอททริบิวท์ในรีเลชัน “การสังกัดชมรม” จะทำให้เกิดปัญหาเมื่อมีการเปลี่ยนแปลงแก้ไขข้อมูลดังนี้

1) ความผิดปกติจากการเพิ่มข้อมูล ถ้ามีชมรมใหม่ตั้งขึ้นมาชื่อ “ภูมิสารสนเทศ” โดยกำหนดห้องประจำชมรมให้เป็น 216 เราจะยังไม่สามารถเพิ่มข้อมูลนี้ลงไปในตารางได้จนกว่าจะมีนักศึกษาสังกัดชมรมนั้นเนื่องจากคีย์หลักคือรหัสนักศึกษาจึงจะต้องไม่เป็นค่านัล

2) ความผิดปกติจากการลบข้อมูล ถ้าลบข้อมูลแถวที่ 2 ออกเนื่องจากนักศึกษารหัส 565146079 ลาออกไปจะทำให้สูญเสียข้อมูลที่แสดงว่าชมรมคอมพิวเตอร์ที่มีห้องประจำชมรมคือ 522 จะหายออกไปด้วย

3) ความผิดปกติจากการแก้ไขข้อมูล ถ้ามีการเปลี่ยนแปลงห้องชมรมคณิตศาสตร์เป็น 535 จะต้องทำการเปลี่ยน แปลงทั้งแถวที่ 1 และ 3 มิฉะนั้นแล้วข้อมูลจะไม่ตรงกัน

การแปลงรีเลชันในรูปแบบบรรทัดฐานระดับที่สอง ให้เป็นรีเลชันอยู่ในรูปแบบบรรทัดฐานระดับที่สาม ทำโดย

(1) ย้ายนอลคีย์แอททริบิวท์กลุ่มที่ขึ้นต่อกันออกมาเป็นรีเลชันใหม่ โดยให้แอททริบิวท์ที่เป็นตัวกำหนดค่า (Determinant) ในกลุ่มนี้มาเป็นคีย์หลักของรีเลชันใหม่

(2) และให้คีย์หลักของรีเลชันใหม่นี้เป็นคีย์นอก (Foreign Key) ในอีกรีเลชันซึ่งประกอบด้วยแอททริบิวท์ที่เหลือจากรีเลชันเดิม

จากตัวอย่างจะแปลงรีเลชันการสังกัดชมรมให้เป็นรีเลชันในรูปแบบบรรทัดฐานระดับที่สาม โดยย้ายนอลคีย์แอททริบิวท์กลุ่มที่ขึ้นต่อกันคือ “ชมรม” และ “ห้องประจำชมรม” ออกมาเป็นรีเลชันใหม่ชื่อรีเลชันห้องชมรม โดยให้แอททริบิวท์“ชมรม” ซึ่งเป็นตัวกำหนดค่าเป็นคีย์หลักของรีเลชันใหม่นี้และให้ “ชมรม”นี้เป็นคีย์นอกของอีกรีเลชันคือรีเลชัน “ชมรมนักศึกษา” จากรีเลชันการสังกัดชมรมของนักศึกษา จึงแตกออกเป็นรีเลชันใหม่ 2 รีเลชันคือรีเลชันชมรมนักศึกษาและรีเลชันห้องชมรม ดังตารางที่ 10.11 และตารางที่ 10.12 เส้นประใต้แอททริบิวท์“ชมรม” ในรีเลชันชมรมนักศึกษา แสดงว่าเป็นคีย์นอกของรีเลชันชมรมนักศึกษา

ชมรมนักศึกษา(<u>รหัสนักศึกษา</u>, <u>ชมรม</u>)

ห้องชมรม(<u>ชมรม</u>, หมายเลขห้อง)

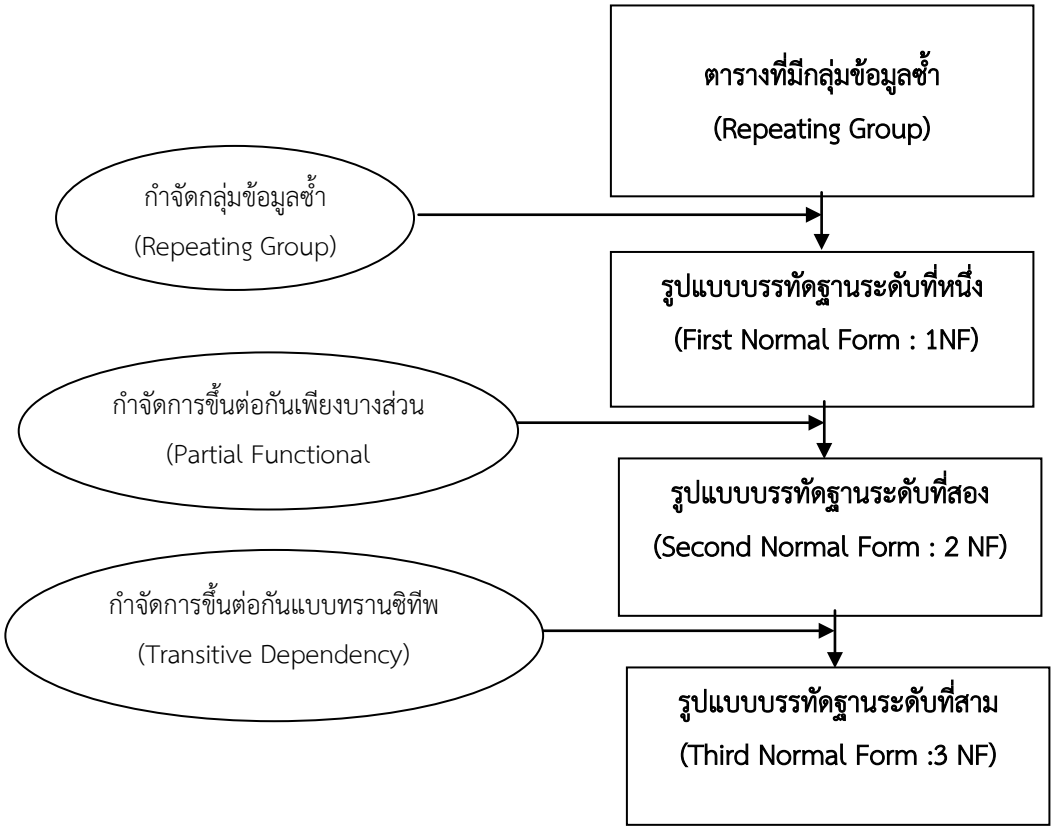
ตารางที่ 10.11 รีเลชันชมรมนักศึกษา

<u>รหัสนักศึกษา</u>	<u>ชมรม</u>
565146052	คณิตศาสตร์
565146079	คอมพิวเตอร์
575146034	คณิตศาสตร์
575146043	ห้องสมุด
555146001	ดนตรี

ตารางที่ 10.12 รีเลชันห้องชมรม

<u>ชมรม</u>	ประจำชมรม
คณิตศาสตร์	511
คอมพิวเตอร์	522
ห้องสมุด	15011
ดนตรี	1402

โดยสรุปการทำนอร์มอลไลซ์เซชันจะมีขั้นตอนดังภาพที่ 10.7



ภาพที่ 10.7 ขั้นตอนการนอร์มอลไลซ์เซชัน

10.8 การออกแบบฐานข้อมูลระดับกายภาพ

การออกแบบฐานข้อมูลระดับกายภาพเป็นการออกแบบเพื่อกำหนดโครงสร้างทางกายภาพของฐานข้อมูลจากรีเลชันที่ผ่านกระบวนการนอร์มัลไลท์มาแล้ว มาสร้างเป็นตาราง (Table) และทำการกำหนดโครงสร้างทางกายภาพ กำหนดกฎความถูกต้องและกำหนดโครงสร้างของแฟ้มข้อมูลตลอดจนวิธีการรักษาความปลอดภัยให้กับฐานข้อมูล ปัจจัยที่จะต้องคำนึงถึงในการออกแบบฐานข้อมูลในระดับกายภาพ ได้แก่

1) อุปกรณ์ที่ใช้เพื่อเก็บข้อมูล วิธีการเข้าถึงข้อมูล วิธีการบำรุงรักษาและเพิ่มประสิทธิภาพ
2) จำนวนของเอนทิตี้แต่ละชนิดในฐานข้อมูล เช่นประมาณว่าในฐานข้อมูลจะเก็บข้อมูลทั้งหมดกี่เอนทิตี้
3) จำนวนผลลัพธ์ (Output) ที่ต้องการจากฐานข้อมูล เช่น จำนวนรายงาน ขนาดของแต่ละรายงาน และความถี่ในการใช้รายงาน
4) กฎเกณฑ์ในการควบคุมความถูกต้องและความปลอดภัยของข้อมูล

10.9 บทสรุป

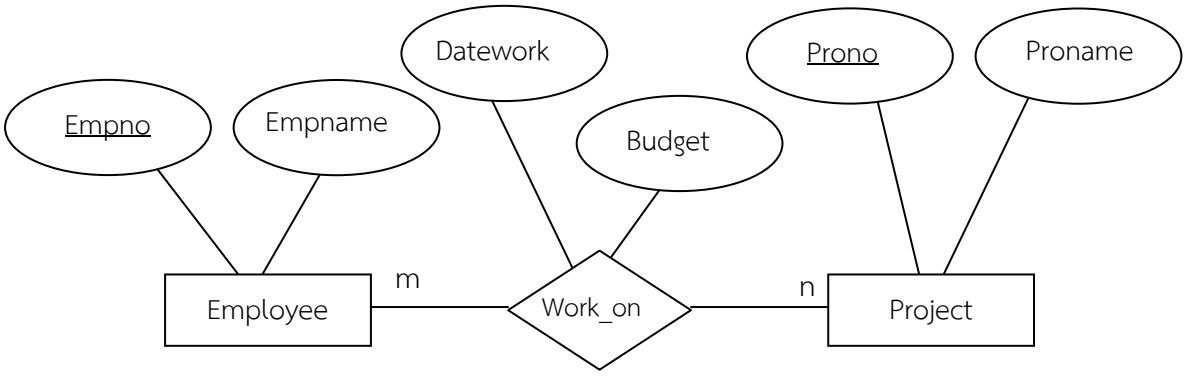
การออกแบบฐานข้อมูลจะเป็นการนำข้อมูลจากขั้นตอนการวิเคราะห์ระบบมาออกแบบรายละเอียดเพิ่มเติม การออกแบบฐานข้อมูลจะแบ่งเป็น 3 ระดับ คือ การออกแบบฐานข้อมูลระดับแนวคิด การออกแบบฐานข้อมูลระดับตรรก และการออกแบบฐานข้อมูลระดับกายภาพ จากขั้นตอนการวิเคราะห์ระบบแบบจำลองข้อมูลที่ได้คือ แผนภาพอี-อาร์ซึ่งจะอยู่ในขั้นตอนการออกแบบฐานข้อมูลในระดับตรรกะซึ่งข้อมูลที่ได้ยังมีความซ้ำซ้อน กระบวนการนอร์มัลไลท์เซชันเป็นกระบวนการที่ช่วยขจัดปัญหานี้ทำให้ฐานข้อมูลที่ได้

ไม่มีความผิดปกติ ในการเพิ่มข้อมูล การลบข้อมูล และการแก้ไขข้อมูล การทำนอร์มัลไลท์เซชันระดับที่ 1(1NF) จะเป็นการกำจัดกลุ่มข้อมูลที่มีความซ้ำซ้อน การทำนอร์มัลไลท์เซชันระดับที่ 2 (2NF) จะต้องอยู่ในรูปแบบ 1NF และต้องกำจัดการขึ้นต่อกันเพียงบางส่วน การทำนอร์มัลไลท์เซชันระดับที่ 3 (3NF) จะต้องมีคุณสมบัติ 2NF และกำจัดกำจัดการขึ้นต่อกันแบบทรานซิทีพ

การออกแบบฐานข้อมูลในระดับกายภาพเป็นการออกแบบโดยพิจารณาถึงการใช้งานจริง การจัดเก็บข้อมูล สื่อที่ใช้บันทึกข้อมูลตลอดจนโครงสร้างของฐานข้อมูลที่ทำการออกแบบ

คำถามทบทวน

1. อะไรคือสิ่งที่ได้จากการออกแบบฐานข้อมูลในระดับแนวคิด
2. จงบอกคุณสมบัติของรีเลชัน
3. จงอธิบายฟังก์ชันการขึ้นต่อกันพร้อมยกตัวอย่าง
4. สิ่งที่ควรคำนึงถึงในการออกแบบฐานข้อมูลในระดับกายภาพมีอะไรบ้าง
5. การทำนอร์มัลไลท์เซชันมีประโยชน์ต่อฐานข้อมูลอย่างไร
6. จากแผนภาพอี-อาร์ต่อไปนี้ จงแปลงเป็นรีเลชัน



7. จากรีเลชันต่อไปนี้อยู่ในรูปแบบบรรทัดฐานที่สมบูรณ์หรือไม่ หากไม่เกิดความปกติใด และทำให้อยู่ในรูปแบบปกติ

STUDENT-NUMBER	STUDENT-NAME	TOTAL-CREDITS	GPA	ADVISOR-NUMBER	ADVISOR-NAME	COURSE-NUMBER	COURSE-DESC	NUM-CREDITS	GRADE
1035	Linda	47	3.647	49	Smith	CSC151	Computer Science I	4	B
						MKT212	Marketing Management	3	A
						ENG101	English Composition	3	B
						CHM112	General Chemistry I	4	A
						BUS105	Introduction to Business	2	A
3397	Sam	29	3.000	49	Smith	ENG101	English Composition	3	A
						MKT212	Marketing Management	3	C
						CSC151	Computer Science I	4	B
4070	Kelly	14	2.214	23	Jones	CSC151	Computer Science I	4	B
						CHM112	General Chemistry I	4	C
						ENG101	English Composition	3	C
						BUS105	Introduction to Business	2	C