

## บทที่ 2

### กรอบแนวคิดและทฤษฎีเกี่ยวกับฐานข้อมูล

กรอบแนวคิดและทฤษฎีเกี่ยวกับฐานข้อมูลนำไปสู่การพัฒนาวิธีการออกแบบฐานข้อมูล ให้เหมาะสมกับหน่วยงานและองค์กรต่างๆ บุคลากรที่เกี่ยวข้องกับการจัดการฐานข้อมูลแต่ละคน ต้องเข้าใจบทบาทและหน้าที่ความรับผิดชอบของตนเอง โดยเฉพาะอย่างยิ่งผู้ออกแบบฐานข้อมูล ที่จะต้องทำให้การติดต่อระหว่างฐานข้อมูลกับผู้ใช้เกิดความสะดวกในการใช้งาน ข้อมูลที่จัดเก็บ มีความถูกต้องและปลอดภัย ดังนั้นจึงต้องเข้าใจถึงสถาปัตยกรรมของฐานข้อมูลในแต่ละระดับและรูปแบบของฐานข้อมูลแบบต่างๆ

#### 1. สถาปัตยกรรมฐานข้อมูล

ระบบฐานข้อมูลจะมีประโยชน์ก็ต่อเมื่อผู้ใช้สามารถเข้าถึงข้อมูลหรือจัดการกับข้อมูลได้ อย่างมีประสิทธิภาพ นอกจากนั้นการที่ผู้ใช้ฐานข้อมูลส่วนใหญ่ไม่ใช่ผู้เชี่ยวชาญทางด้านคอมพิวเตอร์ ในการออกแบบฐานข้อมูลจึงได้มีการซ่อนรายละเอียดที่ซับซ้อนต่างๆ ไว้ภายใน โดยผู้ใช้จะเห็นข้อมูลในเชิงนามธรรมเท่านั้น และสามารถมองเห็นได้ในมุมมองที่แตกต่างกัน ในปี ค.ศ. 1975 American National Standard Institute ได้นำเสนอสถาปัตยกรรมฐานข้อมูลที่แบ่งเป็น 3 ระดับ (วิเชียร เปรมชัยสวัสดิ์, 2546, หน้า 14) ดังต่อไปนี้

##### 1.1 ระดับภายใน (Internal Level)

เรียกอีกอย่างหนึ่งว่า ระดับกายภาพ (physical level) เป็นข้อมูลเชิงนามธรรมในระดับล่างสุด ซึ่งอยู่ในความรับผิดชอบของผู้ออกแบบฐานข้อมูล โดยเป็นผู้พิจารณาว่าจะเก็บข้อมูลไว้ในอุปกรณ์ใด การจัดเก็บข้อมูลจริงๆ ควรทำอย่างไร วิธีการเข้าถึงข้อมูลเพื่อค้นหาหรือปรับปรุงข้อมูลจะใช้วิธีการใด รวมถึงวิธีการบำรุงรักษา และการเพิ่มประสิทธิภาพของฐานข้อมูล ผู้ใช้งานฐานข้อมูลทั่วไปจะไม่เกี่ยวข้องกับการจัดการข้อมูลในระดับนี้

##### 1.2 ระดับแนวคิด (Conceptual Level)

เรียกอีกอย่างหนึ่งว่า ระดับตรรกะ (logical level) เป็นข้อมูลเชิงนามธรรมในระดับที่สูงขึ้นมาอีกระดับหนึ่ง ใช้ในการอธิบายว่ามีข้อมูลอะไรบ้างที่ถูกจัดเก็บไว้ในฐานข้อมูลและข้อมูลเหล่านี้มีความสัมพันธ์กันอย่างไร ผู้ใช้ในระดับตรรกะไม่มีความจำเป็นต้องไปคำนึงถึงความยุ่งยากต่างๆ ในระดับกายภาพเลย ข้อมูลในระดับตรรกะนี้จะถูกกำหนดโดยผู้ดูแลฐานข้อมูล ซึ่งต้องตัดสินใจว่าข้อมูลใดบ้างที่จะถูกจัดเก็บลงในฐานข้อมูล ดังนั้นจึงต้องมีการวิเคราะห์ความต้องการของผู้ใช้ เพื่อจะนำข้อมูลที่ได้มาออกแบบฐานข้อมูล

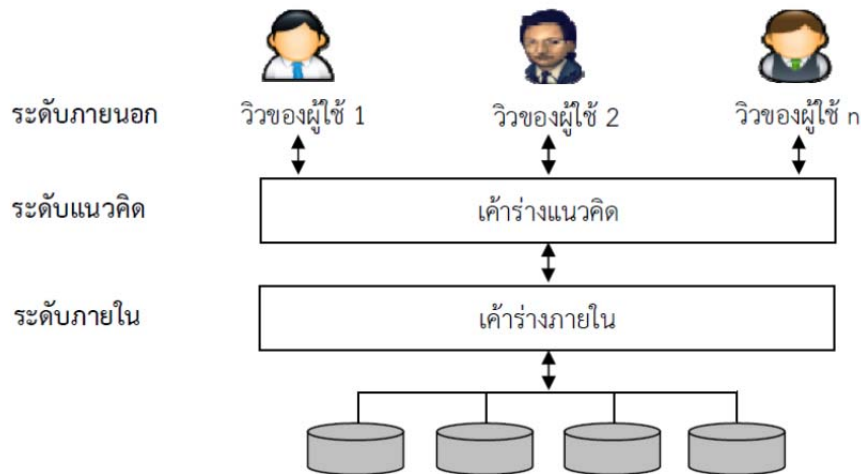
##### 1.3 ระดับภายนอก (External Level)

เรียกอีกอย่างหนึ่งว่า ระดับวิว (view level) เป็นข้อมูลเชิงนามธรรมระดับสูงสุด ใช้อธิบายเกี่ยวกับบางส่วนของฐานข้อมูล เนื่องจากผู้ใช้ฐานข้อมูลส่วนใหญ่ไม่มีความเกี่ยวข้องกับข้อมูลทั้งหมดที่มีอยู่ในฐานข้อมูล แต่มีความจำเป็นต้องเข้าถึงฐานข้อมูลเท่านั้น ดังนั้นเพื่อทำให้การติดต่อกับฐานข้อมูลสามารถกระทำได้ง่ายขึ้น จึงมีการกำหนดข้อมูลเชิงนามธรรมในระดับวิวขึ้น โดยที่ในฐานข้อมูลหนึ่งอาจมีการกำหนด

วิวได้หลายๆ วิวที่แตกต่างกันได้ นอกจากนั้นการกำหนดข้อมูลเชิงนามธรรมในระดับวิว ยังเป็นการช่วยรักษาความปลอดภัยและความมั่นคงให้กับฐานข้อมูลอีกด้วย เนื่องจากผู้ใช้แต่ละคนจะมองเห็นเฉพาะข้อมูลของตนเองได้รับอนุญาตเท่านั้น ทั่วๆ ที่ในฐานข้อมูลมีข้อมูลอื่นๆ ที่เป็นความลับที่อนุญาตให้เฉพาะบุคคลที่มีหน้าที่รับผิดชอบในเรื่องนั้นๆ สามารถมองเห็นและใช้ข้อมูลได้รวมอยู่ด้วย ซึ่งในโปรแกรมการจัดการฐานข้อมูลต่างๆ จะมีส่วนของการสร้างแบบฟอร์มโต้ตอบกับผู้ใช้ ที่เรียกว่า หน้าจอ Interface ดังภาพที่ 2.1

ภาพที่ 2.1 ตัวอย่างหน้าจอโต้ตอบกับผู้ใช้ (interface)

เมื่อเข้าใจแล้วว่าฐานข้อมูลระดับต่างๆ มีความสำคัญอย่างไร ก็ต้องมีการพัฒนาระบบ การจัดการฐานข้อมูลให้สอดคล้องกับข้อมูลที่มีอยู่ในปัจจุบัน และคำนึงถึงสถาปัตยกรรมของฐานข้อมูลด้วย ซึ่งมีแนวคิดในการพัฒนาหลายรูปแบบโดยมีการถ่ายทอดออกมาในลักษณะของแบบจำลองต่างๆ ดังจะกล่าวในรายละเอียดต่อไป



รูปที่ 2.2 แสดงความสัมพันธ์ของเค้าร่างความสัมพันธ์ 3 ระดับ

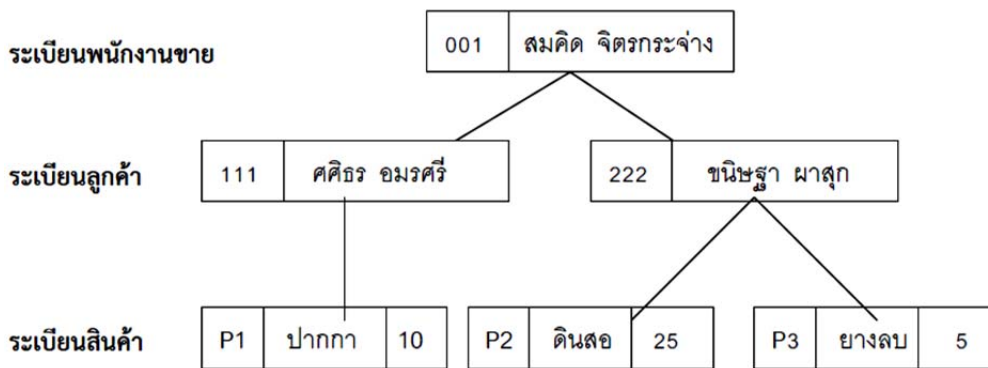
จากรูปที่ 2.2 แสดงให้เห็นว่าฐานข้อมูลจะประกอบด้วยเค้าร่างภายใน 1 ตัว เค้าร่างแนวคิด 1 ตัว และเค้าร่างภายนอกได้หลายๆ ตัว

## 2. รูปแบบของฐานข้อมูล

ฐานข้อมูลเป็นรากฐานสำคัญในการพัฒนาการจัดการข้อมูล จึงมีการสร้างแบบจำลองฐานข้อมูล (database model) ขึ้นมา เพื่อเก็บข้อมูลจริงๆ ตามแนวคิดต่างๆ เพื่อนำเสนอข้อมูลและความสัมพันธ์ระหว่างข้อมูลในรูปแบบที่เข้าใจได้ง่าย โดยแบบจำลองข้อมูลที่ได้รับการยอมรับก็จะถูกนำมาสร้างเป็นระบบฐานข้อมูลยี่ห้อต่างๆ ให้เลือกใช้กัน ในการพัฒนาเป็นฐานข้อมูลมีหลายรูปแบบดังรายละเอียดต่อไปนี้

### 2.1 ฐานข้อมูลแบบลำดับชั้น (The Hierarchical Database Model)

มีลักษณะเป็นแผนภูมิต้นไม้ (tree) ความสัมพันธ์เป็นแบบพ่อกับลูก (parent/child relation) คือระเบียบที่อยู่ด้านบนของโครงสร้างหรือพ่อนั้น สามารถมีลูกได้มากกว่าหนึ่งคน แต่ลูกจะไม่สามารถมีพ่อมากกว่าหนึ่งคนได้ ดังนั้นความสัมพันธ์แบบนี้จึงช่วยลดปัญหาจากการเก็บข้อมูลที่ซ้ำซ้อนกันได้ แต่การที่จะใช้งานข้อมูลได้นั้น ผู้ใช้จะต้องรู้โครงสร้างการเก็บข้อมูลเช่นเดียวกับการเก็บข้อมูลในไฟล์ โดยจะต้องรู้ว่าแผนภูมิต้นไม้นั้นๆ มีข้อมูลอะไรอยู่บ้าง และต้องรู้ว่าด้วยว่าแต่ละลำดับชั้นนั้นเก็บข้อมูลอะไรอยู่



ภาพที่ 2.2 ตัวอย่างข้อมูลในฐานข้อมูลแบบลำดับชั้น

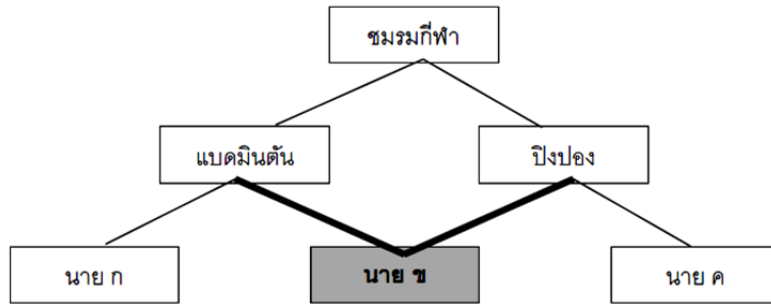
ที่มา (สมจิตร อัจฉรินทร์ และงามนิจ อัจฉรินทร์, 2543, หน้า 24 )

จากตัวอย่างข้างต้นจะเห็นว่าลูกค้าแต่ละคนจะไม่สามารถได้รับบริการจากพนักงานมากกว่าหนึ่งคนได้ เนื่องจากลูกค้าแต่ละคนถือว่าเป็นระเบียบลูก และพนักงานถือว่าเป็นระเบียบพ่อของลูกค้า สินค้าแต่ละชนิดก็จะถูกซื้อโดยลูกค้าเพียงคนเดียวเท่านั้น เนื่องจากสินค้าแต่ละชนิดเป็นระเบียบลูกของระเบียบลูกค้า เป็นต้น ซึ่งในความเป็นจริง ลูกค้าคนหนึ่งๆ สามารถได้รับบริการจากพนักงานมากกว่าหนึ่งคนได้ และสินค้าแต่ละชนิดก็จะถูกซื้อโดยลูกค้าหลายคนได้ แสดงว่าลักษณะของฐานข้อมูลแบบลำดับชั้นมีความสัมพันธ์แบบหนึ่งต่อหนึ่ง (one-to-one) และหนึ่งต่อกลุ่ม (one-to-many) แต่ไม่มีความสัมพันธ์แบบกลุ่มต่อกลุ่ม (many-to-many)

### 2.2 ฐานข้อมูลแบบเครือข่าย (The Network Database Model)

ความเป็นมาของฐานข้อมูลเครือข่าย ได้เริ่มต้นจากความพยายามกำหนดมาตรฐานในการสร้างการจัดการ และการจัดเก็บข้อมูลในฐานข้อมูล โดยกลุ่มผู้พัฒนาภาษา COBOL ได้มีการจัดการประชุมกันโดยชื่อว่า Conference on Data Systems Languages (CODASYL) และได้ตั้งชื่อทีมงานว่า Database Task Group (DBTG) ซึ่งเป็นหน่วยงานที่คิดค้นแบบจำลองฐานข้อมูลแบบเครือข่าย ขึ้นมาใช้แทนแบบจำลองฐานข้อมูล

แบบลำดับชั้น โดยแก้ปัญหาที่ไม่สามารถรองรับความสัมพันธ์แบบกลุ่มต่อกลุ่มได้ ด้วยการนำทฤษฎีเซตทางคณิตศาสตร์มาใช้ ในแบบจำลองข้อมูลนี้ นั่นคือ สมาชิกของเซตหนึ่งๆ สามารถเป็นสมาชิกของเซตอื่นได้อีกด้วย ตัวอย่างเช่น คนที่ชอบเล่นกีฬาแบดมินตันอาจจะชอบเล่นปิงปองด้วยก็ได้ โดยโครงสร้างของแบบจำลองแบบเครือข่ายก็เป็น tree เช่นเดียวกับแบบจำลองแบบลำดับชั้น แต่จะเป็นแผนภูมิต้นไม้ที่ดูซับซ้อนมากขึ้นเพื่อรองรับความสัมพันธ์แบบกลุ่มต่อกลุ่มนั่นเอง

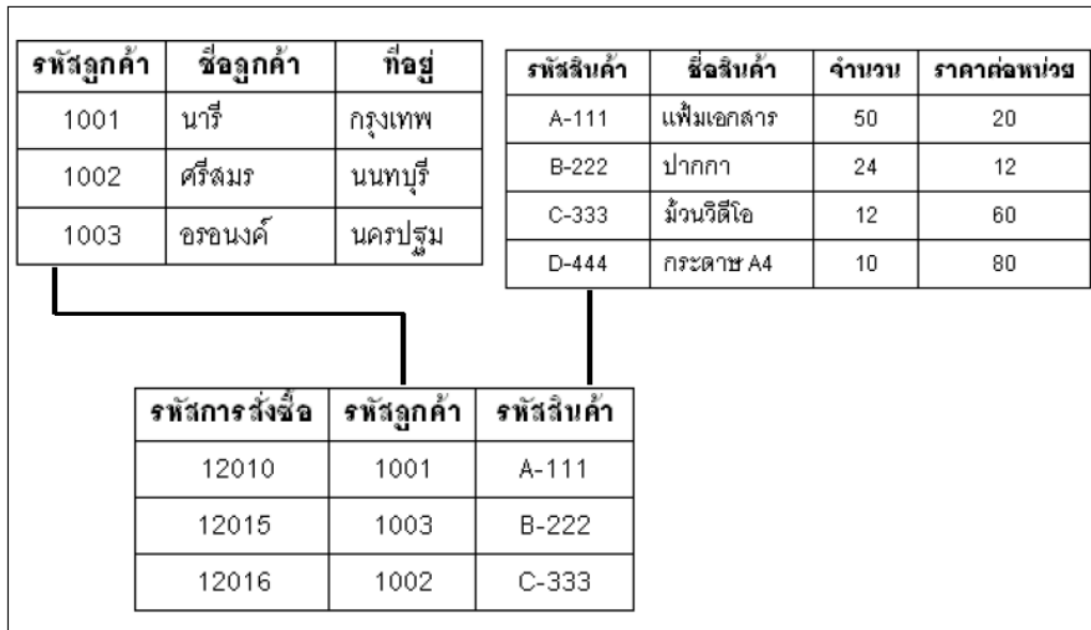


ภาพที่ 2.3 ตัวอย่างข้อมูลในฐานะข้อมูลแบบเครือข่าย  
ที่มา (มณีโชติ สมานไทย, 2546, หน้า 34)

ถึงแม้ว่าการเก็บข้อมูลแบบเครือข่ายจะช่วยแก้ปัญหาความซ้ำซ้อนของข้อมูลให้หมดไปได้ก็ตาม แต่ความสัมพันธ์ของข้อมูลที่โยงกันไปโยงกันมาก็ทำให้ยากต่อการใช้งานและผู้ใช้ยังคงต้องเข้าใจโครงสร้างข้อมูลเพื่อให้สามารถนำข้อมูลมาใช้ได้เหมือนเดิม ดังนั้นแบบจำลองฐานข้อมูลแบบเครือข่ายจึงเหมาะสำหรับโปรแกรมเมอร์ที่คุ้นเคยเป็นอย่างดีกับโครงสร้างข้อมูลแบบต่างๆ ทั้งแบบง่ายๆ และแบบที่ซับซ้อนอย่างที่ใช้ในแบบเครือข่าย แต่ไม่เหมาะสมสำหรับผู้ใช้งานทั่วไป ซึ่งต้องการแบบจำลองข้อมูลที่สามารถทำความเข้าใจและใช้งานได้ง่าย

### 2.3 ฐานข้อมูลเชิงสัมพันธ์ (The Relational Database Model)

เป็นแบบจำลองที่ใช้กันอยู่ในปัจจุบัน โครงสร้างข้อมูลที่ใช้แสดงความสัมพันธ์ของข้อมูลนั้นเป็นตาราง (Table) เป็นตารางข้อมูลที่ประกอบด้วยชุดของคอลัมน์ (Column) และชุดของแถว (Row) ซึ่งเก็บข้อมูลที่มีลักษณะเหมือนกันไว้ ตารางข้อมูลที่เก็บข้อมูลอยู่นั้น ซึ่งไม่จำเป็นต้องรู้ว่าตารางเก็บข้อมูลอย่างไรและเก็บไว้ที่ไหนก็สามารถนำข้อมูลมาใช้ได้ โดยตารางจะมีชื่อเรียกเพื่อให้อ้างถึงเวลาต้องการข้อมูลในตารางนั้น และเมื่อเราต้องการข้อมูลในตารางก็จะใช้วิธีเปรียบเทียบค่าของข้อมูลแทน และสามารถนำมาสร้างความสัมพันธ์ร่วมกันในรูปแบบแนวความคิด ฐานข้อมูลเชิงสัมพันธ์จะผ่านทางระบบจัดการฐานข้อมูลเชิงสัมพันธ์ (Relational Database Management System: RDBMS) และด้วยความยืดหยุ่นในการทำความเข้าใจและง่ายในการนำเข้าสู่ข้อมูล ทำให้แบบจำลองฐานข้อมูลเชิงสัมพันธ์มีผู้ใช้งานกันอย่างแพร่หลายในปัจจุบัน



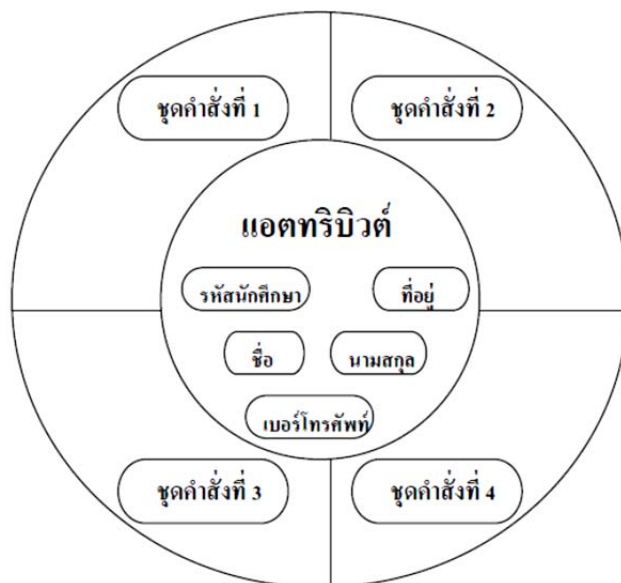
ภาพที่ 2.4 ตัวอย่างข้อมูลในฐานข้อมูลเชิงสัมพันธ์

ทีมา (สมจิตร อาจอินทร์ และ งามนิจ อาจอินทร์, 2543, หน้า 28)

## 2.4 ฐานข้อมูลเชิงวัตถุ (The Object-Oriented Database Model)

เป็นแบบจำลองที่นำมาใช้กับหน่วยงานขนาดใหญ่ จำเป็นต้องใช้บุคลากรที่มีความรู้ ความชำนาญ และประสบการณ์ในการจัดการกับข้อมูลที่มีความสลับซับซ้อน เนื่องจากข้อมูล ในปัจจุบันไม่ได้มีเพียงแต่ข้อความเท่านั้น แต่ยังมีข้อมูลที่เป็นภาพและเสียงเพิ่มมากขึ้น นอกจากนั้นยังมีฐานข้อมูลด้านมัลติมีเดีย งานดีไซน์ที่ใช้คอมพิวเตอร์ช่วยในการออกแบบ (Computer Aid Design : CAD) และฐานข้อมูลไฮเปอร์เท็กซ์ (Hypertext) ซึ่งเป็นเอกสารที่มีลิงค์เชื่อมถึงกันได้ โดยเว็บเพจก็เป็นเอกสารไฮเปอร์เท็กซ์ชนิดหนึ่ง ทำให้ในบางครั้งแบบจำลองฐานข้อมูลเชิงสัมพันธ์ไม่เหมาะสมที่จะนำมาใช้กับงานดังกล่าว ถึงแม้ว่าแบบจำลองฐานข้อมูลเชิงสัมพันธ์จะสามารถเก็บข้อมูลภาพและเสียงได้ แต่จะต้องเพิ่มขั้นตอนเพื่อแปลงข้อมูลให้สามารถเก็บอยู่ในตารางได้ และก็ต้องแปลงกลับเมื่อนำมาใช้อีกด้วย ดังนั้นจึงมีการคิดแบบจำลองข้อมูลเชิงวัตถุขึ้นมาเพื่อเพิ่มความสามารถของฐานข้อมูลให้กับภาษาคอมพิวเตอร์แบบวัตถุ (object-oriented) โดยมองและเก็บข้อมูลเป็นวัตถุ แต่ความสามารถในการค้นหาข้อมูลไม่ดีเท่ากับแบบจำลองฐานข้อมูลเชิงสัมพันธ์ แต่ก็มีประสิทธิภาพในการประมวลผลสูงกว่าเนื่องจากได้ความสามารถของภาษาคอมพิวเตอร์มาใช้

ฐานข้อมูลแบบลำดับชั้น แบบเครือข่าย และฐานข้อมูลเชิงสัมพันธ์ ล้วนจัดเก็บเฉพาะข้อมูล (ข้อมูลและความสัมพันธ์ระหว่างข้อมูล) ไว้ในฐานข้อมูล ส่วนชุดคำสั่งที่ใช้ในการดำเนินการกับฐานข้อมูลจะจัดเก็บไว้ในซอฟต์แวร์ระบบจัดการฐานข้อมูลแยกต่างหาก แต่มีฐานข้อมูลชนิดหนึ่งที่จัดเก็บทั้งข้อมูลและชุดคำสั่งไว้ด้วยกัน ซึ่งการรวมข้อมูลและคำสั่งในการดำเนินการใดๆ เข้าด้วยกัน จะเรียกสิ่งนั้นว่า “วัตถุ (object)” และฐานข้อมูลที่จะนำมาจัดการกับวัตถุก็จะถูกเรียกว่า “ฐานข้อมูลเชิงวัตถุ (object-oriented database)”



ภาพที่ 2.5 การจัดเก็บข้อมูลเชิงวัตถุ

ที่มา (กิตติ ภัคดีวัฒนะกุล, 2547, หน้า 256)

**ฐานข้อมูลเชิงวัตถุ** เป็นฐานข้อมูลที่สามารถจัดเก็บข้อมูลและวิธีการเสมือนวัตถุ ที่สามารถจะดึงไปใช้งานร่วมกันได้โดยอัตโนมัติ ระบบการจัดการฐานข้อมูลเชิงวัตถุสามารถจัดการข้อมูลในลักษณะต่างๆ เช่น แบบเขียน (drawing) รูปภาพ เสียงและภาพเคลื่อนไหวได้อย่างมีประสิทธิภาพ

ถึงแม้ว่าฐานข้อมูลชนิดนี้จะมีประสิทธิภาพในการจัดเก็บและจัดการ แต่ก็พบว่ายังไม่มีการนำมาใช้งานอย่างแพร่หลายเท่ากับฐานข้อมูลเชิงสัมพันธ์ ที่พบก็เพียงการนำแนวคิดเชิงวัตถุมาผสมผสานการทำงานกับแนวคิดฐานข้อมูลเชิงสัมพันธ์ จึงเรียกฐานข้อมูลชนิดนี้ว่า “object-relation database” ที่ผู้ดูแลฐานข้อมูลสามารถใช้ภาษาระบบจัดการฐานข้อมูลชนิดเดียวกับที่ใช้กับฐานข้อมูลเชิงสัมพันธ์ ดังนั้นจึงสามารถจัดการระบบการจัดการฐานข้อมูลได้เช่นเดียวกับฐานข้อมูลเชิงสัมพันธ์ เช่น โปรแกรมออร์ราเคิล (Oracle), ดีบีทู (DB2) และ ไชเบส (SyBase) เป็นต้น

### 2.5 ฐานข้อมูลเชิงวัตถุ-สัมพันธ์ (The Object-Relational Database Model)

ถูกสร้างขึ้นเพื่อให้ฐานข้อมูลเชิงสัมพันธ์ที่ใช้งานอยู่ สามารถเพิ่มคุณสมบัติของแบบจำลองเชิงวัตถุ เข้าไปได้โดยไม่ต้องเสียค่าใช้จ่ายเพิ่มเติมในด้านการออกแบบข้อมูลใหม่ หรือเปลี่ยนแปลงระบบฐานข้อมูลเดิม โดยสิ่งที่เพิ่มขึ้นมาจากแบบจำลองฐานข้อมูลเชิงสัมพันธ์ คือ สามารถสร้างชนิดข้อมูลที่กำหนดเองได้ โดยชนิดข้อมูลที่สร้างขึ้นนี้สามารถเก็บข้อมูลชนิดอื่นๆ ไว้ภายในได้อีก และยังสามารถสร้างวิธี (method) เพื่อจัดการกับข้อมูลภายในได้อีกด้วย ดังนั้นตารางจึงสามารถเก็บข้อมูลที่มีความซับซ้อนได้ตามชนิดข้อมูลที่เพิ่มขึ้นมา ซึ่งตัวอย่างของระบบการจัดการฐานข้อมูลที่ใช้แบบจำลองฐานข้อมูลเชิงวัตถุ-สัมพันธ์ ที่ได้รับความนิยมทั่วโลกก็คือโปรแกรมออร์ราเคิล (Oracle)

### 2.6 คลังข้อมูล (Data Warehouse)

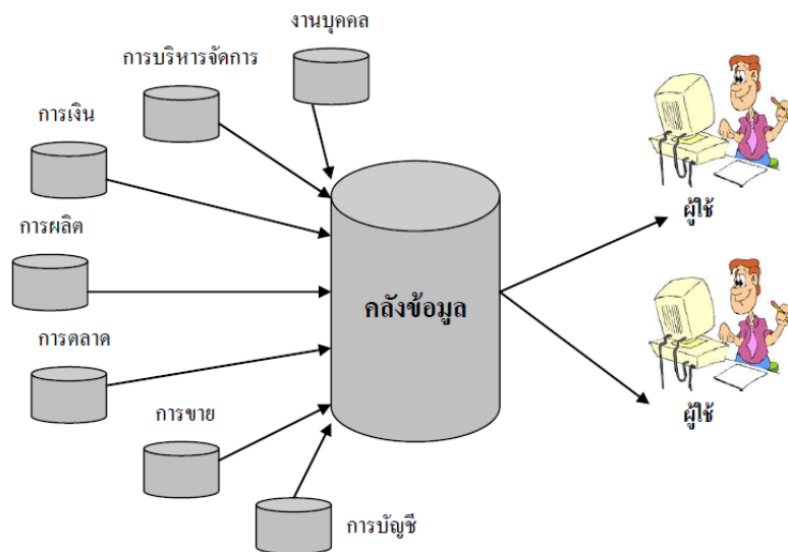
ฐานข้อมูลส่วนใหญ่ในปัจจุบันใช้เพื่อจัดเก็บข้อมูลที่เป็นปัจจุบัน เพื่อนำออกมาใช้งาน ในด้านต่างๆ โดยเฉพาะด้านธุรกิจ เช่น คลังสินค้า การเงิน การผลิต และการขายสินค้า เป็นต้น แต่ ในบางครั้งข้อมูลใน

อดีตก็มีประโยชน์ในการวางแผนการดำเนินงานในอนาคตได้ ซึ่งข้อมูลในอดีตเหล่านั้นอาจจัดเก็บในฐานข้อมูลอื่น จึงจำเป็นจะต้องมีการเชื่อมโยงฐานข้อมูลเพื่อเรียกใช้ข้อมูลที่สัมพันธ์กัน โดยเฉพาะอย่างยิ่งองค์กรขนาดใหญ่ที่มีฐานข้อมูลจำนวนมาก ข้อมูลในฐานข้อมูลที่หนึ่งมีความสัมพันธ์กับข้อมูลในฐานข้อมูลอีกที่หนึ่ง จึงต้องมีการเข้าถึงข้อมูลจากฐานข้อมูลที่แตกต่างกันแต่ข้อมูลสัมพันธ์กัน

ด้วยเหตุผลดังกล่าว ทำให้การพัฒนาฐานข้อมูลในรูปแบบเดิมมีความสามารถมากขึ้น เพื่อรองรับและตอบสนองความต้องการตามที่กล่าวไว้ข้างต้น กลายเป็นฐานข้อมูลที่มีชื่อเรียกว่า “คลังข้อมูล (Data Warehouse)”

**คลังข้อมูล** หมายถึง ฐานข้อมูลที่จัดเก็บข้อมูลที่ได้มาจากการสกัดข้อมูล (extract) จากฐานข้อมูลอื่น ซึ่งอาจจัดโครงสร้างแตกต่างกันหรืออยู่บนระบบปฏิบัติการแตกต่างกันก็ได้ ทั้งนี้ เพื่อประโยชน์ในการวิเคราะห์และตัดสินใจเชิงธุรกิจ

ดังนั้นคลังข้อมูลจึงถือว่าเป็นแหล่งข้อมูลขนาดใหญ่ของทั้งองค์กร ซึ่งแต่ละองค์กรจะต้องมีการแบ่งส่วนการทำงานออกเป็นส่วนย่อยๆ ขึ้นอยู่กับหน้าที่การทำงาน เช่น การแบ่งส่วนออกเป็น การผลิต การเงิน การตลาด การขาย การบัญชี และการบริหารงานบุคคล เป็นต้น ซึ่งในบางกรณีต้องทำสำเนาข้อมูลที่มีความจำเป็นต้องใช้งานของแต่ละส่วนมาจัดเก็บไว้ภายในส่วนการทำงานย่อยๆ เหล่านี้ ซึ่งจะเรียกแหล่งข้อมูลที่จัดเก็บข้อมูลที่ถูกสำเนามาว่า “ดาต้า มาร์ท (data mart)” ดังรูป



ภาพที่ 2.6 ความสัมพันธ์ระหว่างคลังข้อมูล ดาต้ามาร์ท และผู้ใช้

ที่มา (กิตติ ภัคดีวัฒนกุล, 2547, หน้า 258)

สำหรับงานด้านการตัดสินใจแก้ปัญหาานั้น คลังข้อมูลนับว่ามีส่วนช่วยในการจัดหารสารสนเทศในแต่ละด้านได้เป็นอย่างดี นอกจากนี้ คลังข้อมูลยังจะต้องมีความสามารถค้นหาข้อมูลเชิงวิเคราะห์เพื่อให้ได้ข้อมูลที่ีขอบเขตเหมาะสมกับปัญหาได้ชัดเจนขึ้น ซึ่งจะต้องอาศัยเทคนิคอย่างหนึ่งที่เรียกว่า “เหมืองข้อมูล (data mining)”

**เหมืองข้อมูล** บางครั้งเรียกว่า “การค้นพบองค์ความรู้ (knowledge discovery)” เป็นเทคนิคที่ใช้ในการค้นหาข้อมูลเชิงวิเคราะห์ขั้นสูง สำหรับใช้จัดการเก็บข้อมูลที่มีอยู่ในฐานข้อมูลเป็นจำนวนมาก ซึ่ง

ประกอบไปด้วยการค้นหา แยกแยะกลุ่มข้อมูล และคัดเลือกข้อมูลที่มีคุณค่า เป็นต้น เพื่อใช้อธิบายข้อมูลในอดีตและคาดการณ์ข้อมูลในอนาคต

## 2.7 ฐานข้อมูลแบบหลายมิติ (Multidimensional Database)

เป็นฐานข้อมูลที่พัฒนามาจากฐานข้อมูลเชิงสัมพันธ์ ซึ่งแต่ละแถวก็คือ เรคคอร์ดที่ทำให้ทราบคุณลักษณะของข้อมูลใด ๆ การมองข้อมูลที่เห็นเป็นเพียงเรคคอร์ดเท่านั้น เป็นการมองแบบมิติเดียว เช่น สมมติว่ามีตารางข้อมูลยอดขายสินค้าของพนักงานขายแต่ละคนในแต่ละภูมิภาค และใน แต่ละภูมิภาคนั้น แบ่งออกเป็นยอดขายในแต่ละฤดูด้วยและในแต่ละฤดูแบ่งออกเป็นยอดขายในแต่ละสายการผลิต (production line) ดังตารางข้อมูลต่อไปนี้

ตารางที่ 2.1 ข้อมูลยอดขายสินค้าของพนักงานขายแต่ละคนในแต่ละภูมิภาค แต่ละฤดูกาล แต่ละสายการผลิต

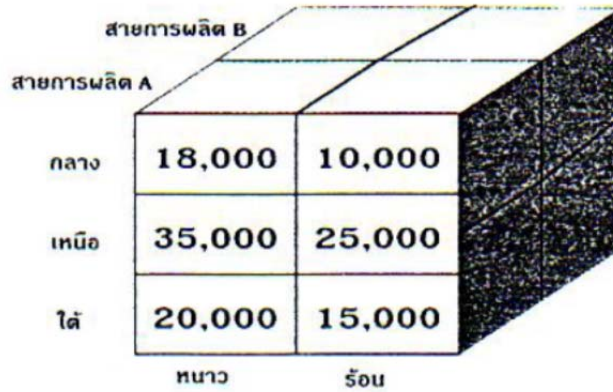
ภูมิภาค	ฤดูกาล	สายการผลิต	ยอดขาย
กลาง	หนาว	A	18,000
กลาง	หนาว	B	15,000
กลาง	ร้อน	A	10,000
กลาง	ร้อน	B	12,500
เหนือ	หนาว	A	35,000
เหนือ	หนาว	B	37,500
เหนือ	ร้อน	A	25,000
เหนือ	ร้อน	B	20,000
ใต้	หนาว	A	20,000
ใต้	หนาว	B	18,000
ใต้	ร้อน	A	15,000
ใต้	ร้อน	B	12,500

ที่มา (กิตติ ภัคดีวัฒนะกุล, 2552, หน้า 16)

จากตาราง จะเห็นว่าในแต่ละเรคคอร์ดนั้นแสดงยอดขาย 1 ค่าในภูมิภาค ฤดู และสายการผลิต ซึ่งคุณลักษณะข้อมูลทั้งหลายนี้ รวมกันอยู่ในเรคคอร์ดเดียว ปัญหาที่เกิดขึ้น ก็คือ การวิเคราะห์ข้อมูลในตารางข้อมูลนั้นทำได้ยาก เช่น ต้องการหายอดขายรวมในฤดูหนาวทุกสายการผลิตและทุกภูมิภาค จากตารางข้อมูลข้างต้น จะทำได้ยากมาก

ดังนั้น จึงต้องมีการจัดเก็บข้อมูลที่มากกว่า 1 มิติ นั่นคือ การจัดเก็บข้อมูลไว้ในฐานข้อมูลแบบหลายมิติ ซึ่งจะเก็บข้อมูลได้มากกว่า 1 มิติ

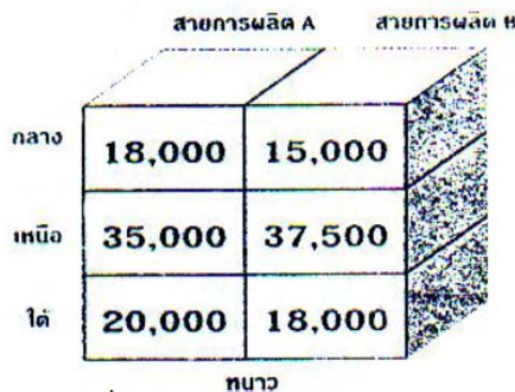




ภาพที่ 2.7 ยอดขายในฐานะข้อมูลแบบหลายมิติ  
 ที่มา (กิตติ ภัคดีวัฒนะกุล, 2552, หน้า 17)

จากภาพที่ 2.7 แสดงข้อมูลจากตารางข้อมูลในฐานะข้อมูลเชิงสัมพันธ์ให้อยู่ในฐานะข้อมูล 3 มิติ มีลักษณะรูปลูกบาศก์ (cube) มิติที่ 1 ของฐานข้อมูลนี้จะแสดงยอดขาย (sale) ในแต่ละภูมิภาค มิติที่ 2 แสดงฤดูกาล (season) และมิติที่ 3 แสดงสายการผลิต (product line)

การแสดงผลข้อมูลหลายมิติ จะทำให้การวิเคราะห์ข้อมูล ยอดขายรวมทุกฤดูกาลและ ทุกภูมิภาคทำได้ง่ายขึ้น ด้วยวิธีการ “เฉือน (slicing)” ลูกบาศก์ออกเป็นส่วนๆ ตามที่ต้องการคำนวณเท่านั้น ดังภาพที่ 2.8



ภาพที่ 2.8 การเฉือน (slicing) ข้อมูล

ที่มา (กิตติ ภัคดีวัฒนะกุล, 2552, หน้า 17)

นอกจากนี้หากผู้ใช้ต้องการวิเคราะห์ข้อมูลยอดขายในฤดูหนาวเพิ่มเติม เช่น หายอดขายรวมในฤดูหนาวเฉพาะสายการผลิต A และเฉพาะในภาคเหนือ ก็สามารถทำได้ด้วยการเฉือนเฉพาะยอดขายในฤดูหนาว ในภาคเหนือ และเฉพาะสายการผลิต A ออกจากลูกบาศก์ที่เฉือนมาแล้วในขั้นตอนที่ผ่านมา และผลที่ได้ก็คือ “ลูกเต๋า (dicing)” ซึ่งในบางครั้งเรียกวิธีการดังกล่าวว่า “slice and dice”



ภาพที่ 2.9 ลูกเต๋าของยอดขายในฤดูหนาว ในภาคเหนือ และเฉพาะสายการผลิต A

ที่มา (กิตติ ภัคดีวัฒนะกุล, 2552, หน้า 17)

วิธีการ “slice and dice” นั้นเป็นวิธีการที่จัดอยู่ในเทคนิคชนิดหนึ่งที่ใช้ในการวิเคราะห์ข้อมูลจากฐานข้อมูล ดาต้ามาร์ท และคลังข้อมูล เทคนิคดังกล่าวเรียกว่า “การประมวลผลเชิงวิเคราะห์แบบออนไลน์ (online analytical processing)” ซึ่งเป็นการวิเคราะห์ข้อมูลจำนวนมากที่จัดเก็บอยู่ในฐานข้อมูลดาต้ามาร์ท และฐานข้อมูลแบบหลายมิติ โดยมีวัตถุประสงค์ก็เพื่อจัดหารสารสนเทศที่ตรงตามความต้องการของผู้ใช้ โดยที่ผู้ใช้ไม่จำเป็นต้องนำไปวิเคราะห์ต่อเลย

## 2.8 ฐานข้อมูลบนเว็บ (Web Database)

ฐานข้อมูลเว็บ คือ การสร้างเว็บเพจที่ผู้ใช้สามารถเลือกดูในสิ่งที่ต้องการได้ ในสมัยก่อนเว็บเพจมีความสามารถเพียงแสดงผลข้อมูลที่ได้มีการจัดเตรียมไว้เรียบร้อยแล้วหรือเรียกว่า “static web” เท่านั้น หากต้องการนำเสนอสิ่งอื่นเพิ่มเติมก็จะต้องทำการสร้างหน้าเว็บเพจขึ้นมาใหม่อยู่เสมอ อีกทั้งทำให้เว็บเพจขาดความน่าสนใจในการใช้งาน เนื่องจากผู้ใช้ทำได้เพียงเป็นฝ่ายรับข้อมูลเท่านั้น แต่เมื่อมีการนำฐานข้อมูลมาประยุกต์กับการสร้างเว็บเพจ ทำให้เว็บเพจมีความสามารถเพิ่มมากขึ้นเป็นอย่างมาก เช่น ผู้ใช้สามารถป้อนข้อความแสดงความคิดเห็นไว้บนเว็บเพจได้ หรือเรียกว่า “เว็บบอร์ด” เป็นต้น

หลักการของเว็บฐานข้อมูล คือ การสร้างฐานข้อมูลไว้ที่เครื่องเซิร์ฟเวอร์ที่ให้บริการเว็บไซท์ที่ต้องการสร้างเป็นเว็บฐานข้อมูล จากนั้นจึงเขียนสคริปต์ (script) ด้วยภาษาในการโปรแกรมมิ่งเว็บฐานข้อมูล เช่น ASP, PHP หรือ CGI เป็นต้น ไว้แทนการสร้างไฟล์ html โดยไฟล์สคริปต์เหล่านี้เมื่อมีผู้เข้ามาเรียกไปแสดงผลยังเครื่องของตนจะถูกประมวลผลโดยเครื่องเซิร์ฟเวอร์ก่อน ซึ่งเซิร์ฟเวอร์จะตรวจสอบคำสั่งที่มีในไฟล์สคริปต์แล้วทำการติดต่อหรืออ่านค่าข้อมูลตามคำสั่งที่มีในไฟล์สคริปต์จากฐานข้อมูล จากนั้นจึงจะนำไปแสดงผลยังเครื่องของผู้ใช้ที่เข้ามาเรียกดูเว็บ

จะเห็นว่าเว็บไซท์ที่สร้างขึ้นโดยใช้เว็บฐานข้อมูลมีมากมายในปัจจุบัน สามารถสังเกตได้จากเว็บเหล่านั้นมักจะมีแบบฟอร์มสำหรับให้ผู้ใช้กรอกข้อมูลหรือเลือกข้อมูลที่ต้องการให้แสดงผล ได้เอง สามารถค้นหาข้อมูลต่าง ๆ ได้อย่างรวดเร็วและง่ายดาย ตัวอย่างเช่น [www.google.co.th](http://www.google.co.th) หรือ [www.yahoo.com](http://www.yahoo.com) เป็นต้น นอกจากนี้ยังมีเว็บไซท์อื่น ๆ อีกมากมายที่ให้บริการในเรื่องของการค้นหาข้อมูลอย่างรวดเร็ว ด้วยกลไกในการค้นหาที่เรียกว่า “search engine” โดยขั้นแรกผู้ใช้ต้องพิมพ์คำค้น (keyword) ที่ต้องการค้นหา จากนั้นเว็บเซิร์ฟเวอร์ของ search engine จะค้นข้อมูลตามที่ต้องการ และแสดงผลการค้นหาออกมา

ในฐานข้อมูลบนเว็บจะประกอบไปด้วยฐานข้อมูลข้อความ และฐานข้อมูลรูปภาพ ซึ่งฐานข้อมูลข้อความ คือ กลุ่มของเอกสารที่ได้รับการเก็บไว้ในฐานข้อมูล โดยที่แต่ละเอกสารและเนื้อความในเอกสารสามารถถูกค้นหาได้โดยอาศัย Catalog หรือบัญชีรายชื่อของสิ่งต่างๆ ที่อยู่ภายในเอกสารนั้นๆ ปัจจุบันเทคนิคนี้ถูกนำไปใช้ในการค้นหาข้อมูลที่ต้องการบนเว็บ ส่วนฐานข้อมูลรูปภาพ คือ ฐานข้อมูลที่ใช้เก็บรายละเอียดต่างๆ เกี่ยวกับรูปภาพซึ่งการใช้วิธีเก็บแบบเดิมจะทำได้ไม่ดีนัก ด้วยความสามารถนี้ทำให้สามารถแสดงภาพของสิ่งต่างๆ ได้ตามต้องการ เช่น การนำข้อมูลรูปภาพดังกล่าวไปใช้งานร่วมกับฐานข้อมูลโดยใช้ภาษาโปรแกรมมิ่งใดๆ ดึงรูปภาพจากฐานข้อมูลรูปภาพขึ้นมาแสดง จากนั้นจึงดึงข้อความออกมาจากฐานข้อมูลข้อความเพื่อนำมาใช้อธิบายรายละเอียดต่างๆ ของรูปภาพดังกล่าว เป็นต้น

## 2.9 ฐานข้อมูลแบบกระจาย (Distributed Database)

เป็นการกระจายการจัดเก็บข้อมูลไว้ในหลายๆ สถานที่ที่เรียกว่า “ไซต์ (site)” ซึ่งแต่ละ ไซต์จะมี เครื่องคอมพิวเตอร์และระบบฐานข้อมูลที่เป็นของตัวเอง เพื่อรองรับการใช้งานต่างๆ ของผู้ใช้ในไซต์นั้นๆ รวมทั้งมีการเชื่อมต่อกับเครื่องคอมพิวเตอร์ของไซต์อื่นด้วยเครือข่ายคอมพิวเตอร์ (Computer Network) เพื่อ ส่งถ่ายข้อมูล (Data Transmission) ไปมาระหว่างฐานข้อมูลของไซต์ต่างๆ ฐานข้อมูลชนิดนี้เหมาะกับธุรกิจ ขนาดใหญ่เช่น ธนาคาร และบริษัทประกันชีวิต เป็นต้น ที่มีสาขาตามที่ตั้งต่างๆ มากมาย

โครงสร้างหลักที่แยกความแตกต่างระหว่างระบบฐานข้อมูลศูนย์กลางออกจากระบบ จัดการ ฐานข้อมูลแบบกระจายคือฮาร์ดแวร์ เพราะระบบจัดการฐานข้อมูลแบบกระจายต้องมีเครื่องคอมพิวเตอร์ หลายเครื่องที่ ไซต์ หรือที่เรียกอีกอย่างหนึ่งว่า โหนด (node) ต่างๆ และแต่ละโหนดจะต้องติดต่อกับ เครือข่ายสื่อสาร เพื่อส่งข้อมูลหรือคำสั่งต่างๆ ระหว่างโหนด โหนดต่างๆ ที่กล่าวถึงนี้อาจจะอยู่ในตึกเดียวกัน หรืออยู่ห่างกันไม่มากนัก และติดต่อกันผ่านระบบเครือข่ายท้องถิ่น (LAN) หรืออาจอยู่ห่างไกลกันคนละ อำเภอซึ่งจะต้องติดต่อผ่านเครือข่ายระยะไกลเช่นสายโทรศัพท์ก็ได้ นอกจากนี้แต่ละโหนดอาจจะมีรูปแบบการ ติดต่อกันที่แตกต่างกันก็ได้ เช่น โหนดที่ 1 ติดต่อกับโหนดที่ 2 และโหนดที่ 2 ติดต่อกับโหนดที่ 3 แต่โหนดที่ 1 ไม่ได้ติดต่อกับโหนดที่ 3 เป็นต้น ในกรณีนี้ถ้าโหนดที่ 1 และโหนดที่ 3 ต้องการติดต่อกันก็ต้อง ติดต่อกันผ่านโหนดที่ 2 ก่อนซึ่งอาจทำให้ประสิทธิภาพการทางานลดลงบ้าง แต่ก็ไม่มีผลกับระบบฐานข้อมูลแต่ อย่างใด

ฐานข้อมูลแบบกระจายทำให้การออกแบบระบบ และการจัดการดูแลฐานข้อมูลมีความยุ่งยากมาก ขึ้น รวมทั้งระบบจัดการฐานข้อมูลที่จะนำมาใช้กับฐานข้อมูลแบบกระจาย ก็จะมี ความซับซ้อนมากกว่าระบบจัดการฐานข้อมูลแบบศูนย์กลาง กล่าวคือต้องมีคุณสมบัติต่างๆ ดังนี้

- เรียกใช้ข้อมูลที่อยู่ห่างไกลออกไปบนโหนดต่างๆ โดยผ่านทางเครือข่ายสื่อสารได้
- เลือกว่าควรเรียกข้อมูลจากที่ใดมาใช้งานบ้าง
- จัดการดูแลการเรียกข้อมูลต่างๆ ได้
- ถ้าระบบใดระบบหนึ่งมีปัญหาไม่ว่าจะด้วยสาเหตุใด เช่น เครื่องมีปัญหา หรือเกิดปัญหาเกี่ยวกับสาร สื่อสาร ระบบจัดการฐานข้อมูลแบบกระจายจะต้องสามารถฟื้นฟูฐานข้อมูลได้เอง

การเลือกใช้งานฐานข้อมูลแบบใดแบบหนึ่ง จะขึ้นอยู่กับปัจจัยหลายประการ ซึ่งต้อง ทำการ พิจารณาให้มีความเหมาะสมกับสภาพขององค์กร โดยปัจจัยที่ทำให้องค์กรใดองค์กรหนึ่ง อาจพิจารณา เลือกใช้ฐานข้อมูลแบบกระจาย คือ

- ข้อมูลบางประเภทจะมีความเหมาะสมกับการใช้ฐานข้อมูลแบบกระจาย เพราะโดยธรรมชาติ ของข้อมูลนั้นอยู่บนที่ต่าง ๆ กันอยู่แล้ว เช่น บริษัทที่มีสาขาอยู่ในจังหวัดต่างๆ หรือธนาคารที่มีหลายๆ สาขา และแต่ละสาขาอยู่คนละแห่งในที่ไกลกันออกไป บริษัทหรือธนาคารดังกล่าวอาจต้องการเก็บข้อมูลของแต่ละ สาขาไว้ที่สาขาเหล่านั้น เนื่องจากงานการเรียกใช้ข้อมูลส่วนมากจะเกิดขึ้นโดยพนักงาน หรือลูกค้าที่สาขา นั้นๆ โดยมีโอกาสที่จะเรียกใช้ข้อมูลข้ามสาขาน้อยมาก แต่ขณะเดียวกันผู้บริหารที่สำนักงานใหญ่อาจ ต้องการเรียกใช้ข้อมูลจากหลายๆ สาขาเพื่อดูข้อมูลสรุปบางอย่างด้วย

- เพื่อเพิ่มความเชื่อถือได้ของข้อมูลและมีความมีประโยชน์ของข้อมูล เพราะการเรียกใช้ข้อมูลจะได้ข้อมูลที่ทันสมัยและถูกต้องอยู่เสมอ ไม่ว่าข้อมูลนั้นจะจัดเก็บอยู่ในที่ใดๆ ก็ตาม และ หากเครื่องคอมพิวเตอร์หรือระบบจัดการฐานข้อมูลในโหนดอื่นจะยังคงใช้งานได้เหมือนเดิม แต่ถ้าใช้ระบบฐานข้อมูลศูนย์กลางแล้ว หากเครื่องที่ศูนย์กลางเกิดปัญหา ก็จะทำให้ไม่สามารถเรียกดูข้อมูล ได้ทั้งระบบ

- ทำให้สามารถใช้ข้อมูลร่วมกัน ในขณะที่การดูแลข้อมูลนั้นแยกจากกันในแต่ละโหนด

- เพิ่มประสิทธิภาพในการทำงาน เนื่องจากการกระจายฐานข้อมูลที่มีขนาดใหญ่ไปในหลาย ๆ แห่ง ทำให้การเรียกใช้ข้อมูลของแต่ละแห่งสามารถทำได้เร็วขึ้น เพราะข้อมูลได้ถูกแบ่งออกเป็นส่วนๆ นอกจากนี้ การประมวลผลของแต่ละโหนดก็จะอิสระจากกัน ทำให้สามารถทำการประมวลผลพร้อมๆ กันได้

## สรุป

กรอบแนวคิดและทฤษฎีเกี่ยวกับฐานข้อมูลต้องเข้าใจถึงสถาปัตยกรรมของฐานข้อมูลใน แต่ละระดับ และรูปแบบของฐานข้อมูลแบบต่างๆ ซึ่งสถาปัตยกรรมของฐานข้อมูลแบ่งเป็น 3 ระดับ คือ 1) ระดับภายในหรือระดับกายภาพ โดยผู้ออกแบบฐานข้อมูล จะพิจารณาว่าควรเก็บข้อมูลไว้ในอุปกรณ์ใด การจัดเก็บข้อมูลจริงๆ ควรทำอย่างไร วิธีการเข้าถึงข้อมูลเพื่อค้นหาหรือปรับปรุงข้อมูล จะใช้วิธีการใด รวมถึงวิธีการบำรุงรักษา และการเพิ่มประสิทธิภาพของฐานข้อมูล 2) ระดับแนวคิดหรือระดับตรรกะ เพื่ออธิบายว่ามีข้อมูลอะไรบ้างที่ถูกจัดเก็บไว้ในฐานข้อมูลและข้อมูลเหล่านี้ มีความสัมพันธ์กันอย่างไร และ 3) ระดับภายนอกหรือระดับวิว เพื่อให้ผู้ใช้สามารถติดต่อกับฐานข้อมูลได้ง่ายขึ้น ส่วนรูปแบบของฐานข้อมูล ซึ่งมีอยู่หลายรูปแบบตามแนวคิดในการพัฒนาการจัดการฐานข้อมูล ซึ่งได้แก่ ฐานข้อมูลแบบลำดับชั้น ฐานข้อมูลแบบเครือข่าย ฐานข้อมูลเชิงสัมพันธ์ ฐานข้อมูลเชิงวัตถุ ฐานข้อมูลเชิงวัตถุ-สัมพันธ์ คลังข้อมูล ฐานข้อมูลแบบหลายมิติ ฐานข้อมูลบนเว็บ และฐานข้อมูลแบบกระจาย นอกจากทราบบแนวคิดและทฤษฎีเกี่ยวกับฐานข้อมูลซึ่งเป็นพื้นฐานในการออกแบบฐานข้อมูลแล้ว จำเป็นต้องศึกษาในเรื่องระบบสารสนเทศที่ใช้ในองค์กรต่างๆ ประกอบด้วย เพื่อจะได้ออกแบบฐานข้อมูลให้เหมาะสม