

## บทที่ 2 แนวคิด ทฤษฎี และเอกสารที่เกี่ยวข้อง

ปัจจุบันองค์กรที่จะประสบความสำเร็จประกอบด้วยปัจจัยหลายอย่าง และปัจจัยหนึ่งคือข้อมูลที่มีอยู่และใช้ประจำวัน (Operational Database) ซึ่งจะเพิ่มขึ้นจนเป็นปัญหาสำหรับองค์กรในการจัดการเก็บข้อมูล ซึ่งข้อมูลเหล่านี้จะถูกเก็บรวบรวมอยู่ในรูปแบบต่างๆ เช่น เทปแม่เหล็ก แผ่นซีดี นอกจากนี้ข้อมูลที่มีจำนวนมากยังไม่เหมาะสมที่จะนำมาใช้ในการวิเคราะห์ทางธุรกิจหรือช่วยด้านบริหารสำหรับการตัดสินใจได้ เนื่องจากมีการใช้เวลาในการประมวลผลที่นานและส่งผลกระทบต่อระบบการทำงานของเครื่องที่ใช้งานประจำวัน เพื่อที่จะทำให้ข้อมูลที่มีอยู่สามารถตอบสนองความต้องการทางธุรกิจได้อย่างรวดเร็วและมีประสิทธิภาพ ดังนั้นจึงเกิดแนวคิดเรื่องคลังข้อมูล (Data Warehouse) ขึ้นเพื่อตอบสนองงานในรูปแบบของคลังเก็บข้อมูลสำหรับการบริหาร สามารถจะนำข้อมูลที่มีอยู่มาใช้อย่างมีประสิทธิภาพจะทำให้องค์กรประสบความสำเร็จเหนือคู่แข่ง โดย กิตติพงษ์ กลมกล่อม(2552) ได้ให้ความหมายของคลังข้อมูลไว้ว่า คลังข้อมูลเป็นระบบข้อมูลขนาดใหญ่ที่เอื้ออำนวยให้ผู้ใช้งานสามารถใช้ข้อมูลได้อย่างมีประสิทธิภาพ ข้อมูลที่จัดเก็บอยู่ในคลังข้อมูลต้องมีปริมาณเพียงพอ และเป็นข้อมูลที่มีคุณภาพเพียงพอต่อการนำไปวิเคราะห์ เพื่อหาคำตอบที่เหมาะสมทางธุรกิจ การจัดเก็บข้อมูลต้องเอื้ออำนวยให้การใช้ข้อมูลเชิงวิเคราะห์เป็นไปได้โดยง่าย สามารถวิเคราะห์ข้อมูลได้อย่างรวดเร็วไม่ยุ่งยากซับซ้อน

### 2.1 ลักษณะของระบบคลังข้อมูล

ลักษณะของระบบคลังข้อมูล ที่แตกต่างจากฐานข้อมูลปฏิบัติการประกอบด้วย 4 ประการดังนี้ (Inmon, 2005: 32)

- 1) การจัดกลุ่มตามเนื้อหาของข้อมูล (Subject-Oriented Data) คลังข้อมูลสามารถนำมาใช้ในการวิเคราะห์เรื่องใดเรื่องหนึ่ง เช่น ในองค์กรแต่ละแห่งมักจะพัฒนาระบบสารสนเทศเพื่อตอบสนองต่อการทำงานขององค์กรที่แตกต่างกัน พัฒนาระบบสารสนเทศการขาย เพื่อตอบสนองต่อการวิเคราะห์ข้อมูลการขายสินค้า พัฒนาระบบสินค้าคงคลัง เพื่อตอบสนองต่อการวิเคราะห์ข้อมูลการสั่งสินค้า เป็นต้น การจัดกลุ่มข้อมูลตามระบบสารสนเทศดังกล่าว เป็นการแบ่งข้อมูลตามกิจกรรมขององค์กรที่แตกต่างกัน ในคลังข้อมูล ข้อมูลจะถูกจัดกลุ่มตามเนื้อหาของข้อมูล โดยการ

พิจารณาว่า ในบรรดาข้อมูลที่มีอยู่ในทุกๆระบบนั้นมีข้อมูลใดบ้างที่บอกเล่าเรื่องราวเดียวกันหรือคล้ายคลึงกัน จะถูกจัดอยู่ในกลุ่มเดียวกันได้

2) การรวมข้อมูลเป็นหนึ่งเดียว (Integrated Data) ข้อมูลที่เข้ามาสู่คลังข้อมูลนั้นมีที่มาจากหลายๆแหล่ง ข้อมูลที่มีความหมายเดียวกันแต่อาจจะมิชนิคข้อมูลที่แตกต่างกันไป เช่น จากระบบปฏิบัติงานประจำวันขององค์กร จากระบบสารสนเทศต่างๆในองค์กร หรือจากแหล่งข้อมูลภายนอก ดังนั้นเมื่อแหล่งข้อมูลมีมากมายหลากหลาย จึงเป็นไปได้ที่จะเกิดความซ้ำซ้อนของข้อมูลขึ้น หน้าที่ของคลังข้อมูล คือ การกำจัดความซ้ำซ้อนของข้อมูล ทำให้ชนิดของข้อมูลที่แตกต่างกันรวมเป็นชนิดเดียวกัน ทั้งนี้เพื่อไม่ให้เกิดความสับสนในการวิเคราะห์ข้อมูล

3) ความสัมพันธ์กับเวลา (Time Referenced Data) ข้อมูลในคลังข้อมูลแตกต่างจากข้อมูลในระบบปฏิบัติการตรงที่ ในข้อมูลปฏิบัติการ จะเก็บเฉพาะข้อมูลในปัจจุบัน ซึ่งมีผลต่อการดำเนินธุรกิจขององค์กร ณ ขณะนั้น แต่ระบบคลังข้อมูลจะมุ่งเน้นไปที่การเก็บข้อมูลเพื่อการวิเคราะห์ข้อมูลตามช่วงเวลา ดังนั้นระบบคลังข้อมูลต้องจัดเก็บข้อมูลทั้งในอดีตและปัจจุบันขององค์กร โดยเมื่อนำมาใช้งานระบบจะสามารถแสดงข้อมูลย้อนหลังได้เช่น 3 เดือน 6 เดือน 12 เดือน

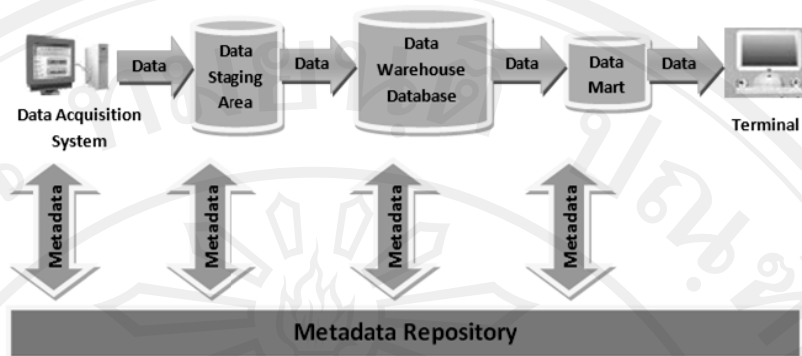
4) ความเสถียรของข้อมูล (Non-Volatile Data) เพื่อรักษาคุณสมบัติที่มีความสัมพันธ์กับเวลา ข้อมูลในคลังข้อมูลจึงเป็นข้อมูลที่คงอยู่ตลอดไปไม่ว่าข้อมูลจะเก่าเพียงใด ก็ยังจะอยู่ในคลังข้อมูลตลอด ไม่เปลี่ยนแปลง เราเรียกคุณสมบัตินี้ว่า ความเสถียรของข้อมูล ทั้งนี้เพื่อให้การวิเคราะห์ข้อมูลแบบช่วงเวลาให้ผลลัพธ์ที่มีประสิทธิภาพ อย่างไรก็ตามถึงแม้ข้อมูลจะไม่ถูกลบออกจากคลังข้อมูล แต่เพื่อการจัดการพื้นที่ในคลังข้อมูลอย่างมีประสิทธิภาพ สามารถถูกเปลี่ยนแปลงรูปแบบไป เพื่อให้ข้อมูลมีขนาดเล็กกลงได้โดยการเปลี่ยนแปลงรูปแบบดังกล่าวไม่กระทบต่อการวิเคราะห์ข้อมูล

## 2.2 สถาปัตยกรรมคลังข้อมูล (Data Warehouse Architecture)

สถาปัตยกรรมคลังข้อมูลเป็นโครงสร้างมาตรฐานที่ใช้อธิบายเพื่อให้เข้าใจแนวคิดและกระบวนการของคลังข้อมูลนั้นๆ ในการศึกษาที่มีรายการดังนี้ องค์ประกอบที่ทำให้หน้าที่ในการเก็บข้อมูล สถาปัตยกรรมแบบต่างๆ (กิตติพงษ์ กลมกล่อม , 2552: 20-26)

### 2.2.1 องค์ประกอบที่ทำให้หน้าที่ในการเก็บข้อมูล

จากความหมายของระบบคลังข้อมูล เป็นระบบข้อมูลขนาดใหญ่ ในทางปฏิบัติสิ่งที่จัดเก็บในคลังข้อมูล ไม่ได้มีข้อมูลเพียงอย่างเดียว หากยังมีข้อมูลสำหรับดำเนินการรวมอยู่ด้วย โดยแสดงในรูปแบบที่ 2.2



รูปที่ 2.1 สถาปัตยกรรมคลังข้อมูลแบบพื้นฐาน

ที่มา: [สไลด์]. 344-662 Data Warehousing and Data Mining. มหาวิทยาลัยสงขลานครินทร์

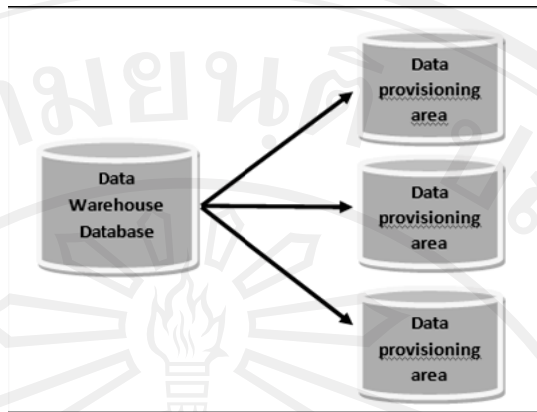
ในรูปที่ 2.1 แสดงองค์ประกอบของคลังข้อมูล ได้แก่

- 1) ฐานข้อมูลคลังข้อมูล เป็นฐานข้อมูลที่ทำหน้าที่ในการจัดเก็บข้อมูลทางธุรกิจขององค์กร จัดเป็นส่วนที่มีความสำคัญที่สุดของคลังข้อมูล
- 2) พื้นที่พักข้อมูล เนื่องจากข้อมูลที่มาจากระบบปฏิบัติงานขององค์กร อาจมีความหลากหลาย ทั้งในแง่ของมาตรฐานและวิธีการจัดเก็บ และอาจมีความซ้ำซ้อนกัน ดังนั้น ข้อมูลจากแหล่งข้อมูลจำเป็นต้องถูกปรับให้เข้ากับมาตรฐานของคลังข้อมูลก่อนที่จะเข้ามาอยู่ในคลังข้อมูล ดังนั้นคลังข้อมูลจึงจำเป็นต้องมีพื้นที่พักข้อมูลซึ่งเป็นฐานข้อมูลที่ทำหน้าที่เป็นที่พักข้อมูล ขณะที่ถูกปรับให้เข้ากับมาตรฐานของคลังข้อมูล
- 3) ดาต้ามาร์ท เนื่องจากฐานข้อมูลคลังข้อมูล มีหน้าที่ในการเก็บข้อมูลต่างๆที่เกี่ยวข้องกับธุรกิจขององค์กรให้ครบถ้วนถูกต้อง ซึ่งคุณลักษณะดังกล่าวอาจทำให้ข้อมูลในฐานข้อมูลคลังข้อมูล มีรูปแบบไม่เหมาะสมต่อการใช้งานหรือนำมาวิเคราะห์ได้ยาก จึงต้องมีดาต้ามาร์ท เพื่อเป็นพื้นที่สำหรับจัดเก็บข้อมูลที่มีโครงสร้างที่เหมาะสมต่อการนำไปใช้งาน

## 2.2.2 สถาปัตยกรรมแบบต่างๆ

ในการแบ่งประเภทของสถาปัตยกรรมของคลังข้อมูล ใช้เกณฑ์การแบ่งจากลักษณะทางกายภาพของฐานข้อมูลคลังข้อมูล

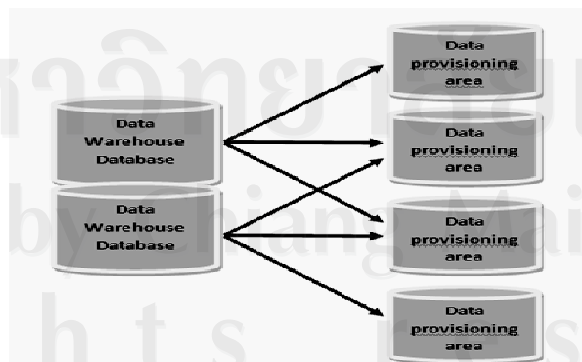
- 1) สถาปัตยกรรมระบบคลังข้อมูลแบบรวม (Centralized Data Warehouse Database Architecture) รูปแบบการเก็บข้อมูลจะถูกรวมอยู่ในฐานข้อมูลเดียวกัน ดังรูปที่ 2.2



รูปที่ 2.2 แสดงรูปแบบสถาปัตยกรรมระบบคลังข้อมูลแบบรวม

ในรูปที่ 2.2 แสดงรูปแบบสถาปัตยกรรมระบบคลังข้อมูลแบบรวม ซึ่งมีข้อดีคือ การรักษาความปลอดภัยและการบำรุงรักษาทำได้ง่าย เนื่องจากเป็นระบบที่มีความซับซ้อนน้อยทำให้การบำรุงรักษาทำได้ง่ายและความเสี่ยงต่อความเสียหายที่อาจเกิดขึ้นน้อยกว่าสถาปัตยกรรมแบบอื่นๆ ข้อดีอีกประการคือสามารถสร้างความเชื่อมโยงเดียวกันของข้อมูลได้ง่ายที่สุด มีความถูกต้องตรงกันของข้อมูล ซึ่งจะเกิดความผิดพลาดเมื่อมีข้อมูลชุดเดียวกัน 2 ที่ ซึ่งข้อมูลชุดนั้นต้องมีค่าเท่ากันทั้งสองที่ ในกรณีสถาปัตยกรรมระบบคลังข้อมูลแบบรวมจะไม่เกิดความผิดพลาดนี้ ข้อเสียของสถาปัตยกรรมนี้คือ มีความเสี่ยงในการสูญเสียข้อมูลมากกว่าสถาปัตยกรรมแบบอื่น ถึงแม้ว่าการรักษาความปลอดภัยจะทำได้ง่ายที่สุด แต่ในกรณีที่เกิดความเสียหายขึ้นหรือระบบป้องกันไม่สามารถป้องกันได้ จะเกิดความเสียหายสูงที่ข้อมูลจะถูกทำลายแล้วส่งผลกระทบต่อการใช้งาน

2) สถาปัตยกรรมระบบคลังข้อมูลแบบกระจาย (Distributed data warehouse database Architecture) หมายถึงสถาปัตยกรรมข้อมูลพื้นฐานข้อมูลมีการกระจายตัวออก มีฐานข้อมูลมากกว่า 1 ตัว ทั้งนี้เพื่อสร้างความคล่องตัวในการใช้งานโดยการกระจายตัวของฐานข้อมูลจะส่งผลให้ข้อมูลตัวใดตัวหนึ่งอาจจะมีอยู่ในฐานข้อมูลคลังข้อมูลเพียงฐานเดียวหรือหลายฐานก็ได้ สามารถอธิบายได้ดังรูปที่ 2.3



รูปที่ 2.3 แสดงรูปแบบสถาปัตยกรรมระบบคลังข้อมูลแบบกระจาย

ในรูปที่ 2.3 แสดงรูปแบบสถาปัตยกรรมระบบคลังข้อมูลแบบกระจาย ซึ่งมีข้อดีคือ เป็นสถาปัตยกรรมข้อมูลที่สร้างง่ายกว่า ดังนั้นจึงเป็นการง่ายที่จะออกแบบและสร้างที่ละส่วนสามารถตอบสนองความต้องการใช้งานข้อมูลจากคลังข้อมูลอย่างเร่งรีบได้ ข้อดีอีกประการคือสามารถกระจายความเสี่ยงในกรณีที่จะเกิดความเสียหายขึ้นกับข้อมูลได้ดี เนื่องจากมีการกระจายตัวของข้อมูล ดังนั้นจึงมีโอกาสที่ข้อมูลชุดหนึ่งในฐานข้อมูลฐานหนึ่งอาจถูกทำลาย แต่ข้อมูลชุดเดียวกันอีกที่หนึ่งยังคงอยู่และสามารถนำมาทดแทนได้ ข้อเสียคือ มีโอกาสที่ข้อมูลขาดความเป็นอันหนึ่งอันเดียวกัน คือมีการยอมให้ข้อมูลชุดเดียวกันมีอยู่ในหลายๆที่ มีโอกาสที่ข้อมูลทั้งสองที่ไม่สอดคล้องกันซึ่งอาจเกิดได้จากหลายสาเหตุ อาทิเช่น เกิดการหยุดของระบบหลังจากการถ่ายโอนข้อมูลชุดเดียวกันจากที่หนึ่งไปยังอีกที่หนึ่งยังไม่เสร็จสมบูรณ์

### 2.3 แนวทางการพัฒนาค้างข้อมูล

การจำแนกประเภทของการพัฒนาจะ พิจารณาจากลักษณะและลำดับเวลา ก่อน-หลัง มี 3 วิธี คือ การพัฒนาค้างข้อมูลด้วยหลักการพัฒนาจากบนลงล่าง การพัฒนาค้างข้อมูลด้วยหลักการพัฒนาจากล่างขึ้นบน และการพัฒนาด้วยหลักการผสม

1) การพัฒนาค้างข้อมูลด้วยหลักการพัฒนาจากบนลงล่าง(Top-Down Data Warehouse Development) เป็นหลักการที่เหมาะสมเพื่อพัฒนาสถาปัตยกรรมแบบรวม จะทำให้ได้คลังข้อมูลรวมของทั้งองค์กรในคราวเดียว โดยวิธีพัฒนาจะเริ่มจากการวิเคราะห์ห้องค์กรธุรกิจและออกแบบแบบจำลองข้อมูลที่เป็นภาพรวมแล้วจึงวิเคราะห์ข้อมูลนำเข้าและผลลัพธ์ เพื่อการออกแบบและสร้างตามาร্তต่อไป

2) การพัฒนาค้างข้อมูลด้วยหลักการพัฒนาจากล่างขึ้นบน(Bottom-Up Data Warehouse Development) ตรงกันข้ามกับแนวทางบนลงล่าง การวิเคราะห์และออกแบบเพื่อสร้างคลังข้อมูลจะเริ่มด้วยวิธีการวิเคราะห์และออกแบบผลลัพธ์และข้อมูลนำเข้าที่ละส่วนแล้วจึงออกแบบแบบจำลองข้อมูลหลังจากนั้นจึงมีการนำข้อมูลที่มีอยู่ในแต่ละส่วนมารวมกัน

3) การพัฒนาค้างข้อมูลด้วยหลักการพัฒนาแบบผสม(Mixed Data Warehouse Development) โดยแยกข้อมูลออกเป็นส่วนๆแล้วพิจารณาเลือกวิธีการวิเคราะห์และออกแบบที่เหมาะสมสำหรับข้อมูลแต่ละส่วน แล้วจึงนำเอาแต่ละส่วนที่ได้พัฒนาขึ้นมาแล้วมารวมกันในภายหลัง

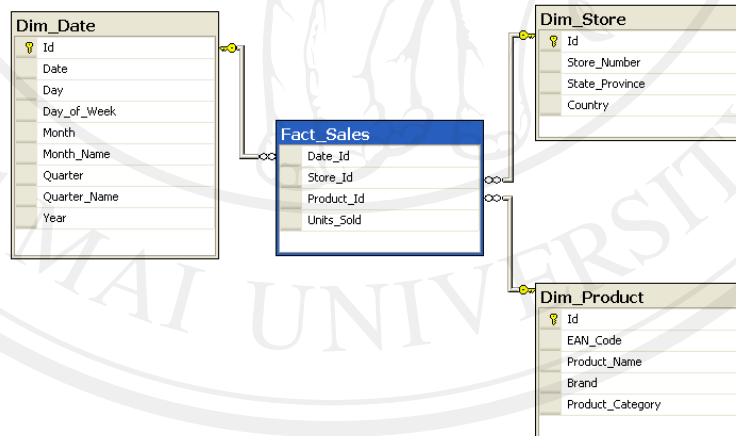
ภายหลัง

## 2.4 แบบจำลองข้อมูลเชิงมิติ (Dimensional Data Model)

จากหลักการและเหตุผลในบทที่ 1 กล่าวว่า การออกแบบแบบจำลองข้อมูลเชิงสัมพันธ์ (Relational Data Model) มีจุดมุ่งหมายที่จะแสดงความสัมพันธ์ที่มีอยู่จริงของข้อมูล มีความซับซ้อนของความสัมพันธ์มาก ทำให้การค้นหาข้อมูลและคำนวณผลลัพธ์ที่ต้องการใช้ในการวิเคราะห์เป็นไปได้ด้วยความยากลำบาก ดังนั้นเพื่อความสะดวกในการใช้ข้อมูลจึงมีแนวคิดแบบจำลองข้อมูลเชิงมิติ ที่เหมาะสำหรับการแสดงผลลัพธ์ที่ต้องการ แบ่งแบบจำลองเชิงมิติออกเป็น 2 กลุ่ม คือ ข้อเท็จจริงและมิติ ข้อเท็จจริงคือค่าที่เกิดขึ้นจริงจากเหตุการณ์หรือกระบวนการรายการ เช่น ข้อเท็จจริงรายการขายหรือรายการผลิต มิติคือข้อมูลด้านคุณภาพ ในมิติหนึ่งอาจรวมถึงคุณสมบัติของสิ่งที่สนใจ เช่น สี ตราสินค้า วันที่ให้บริการ ภูมิภาค เป็นต้น เมื่อแถวของข้อเท็จจริงเชื่อมโยงกับมิติ จะสามารถหาคำตอบได้เช่น สีของสินค้าอะไรที่กำลังเป็นที่นิยม

### 2.4.1 การออกแบบฐานข้อมูลโครงสร้างรูปดาว

หมายถึงแบบจำลองข้อมูลเชิงมิติ ที่มีตารางข้อเท็จจริงขนาดใหญ่เพียงหนึ่งเดียวอยู่ตรงกลาง และมีตารางมิติจำนวนหนึ่งเชื่อมโยงเพื่อกำหนดมุมมองที่จะมีต่อข้อมูลที่ใช้เพื่อการวัดค่าในตารางข้อเท็จจริง โดยจำนวนมุมมองที่มองได้จะเท่ากับจำนวนของตารางมิติที่เชื่อมต่อโดยตรงกับตารางข้อเท็จจริง แสดงโครงสร้างดังรูปที่ 2.4



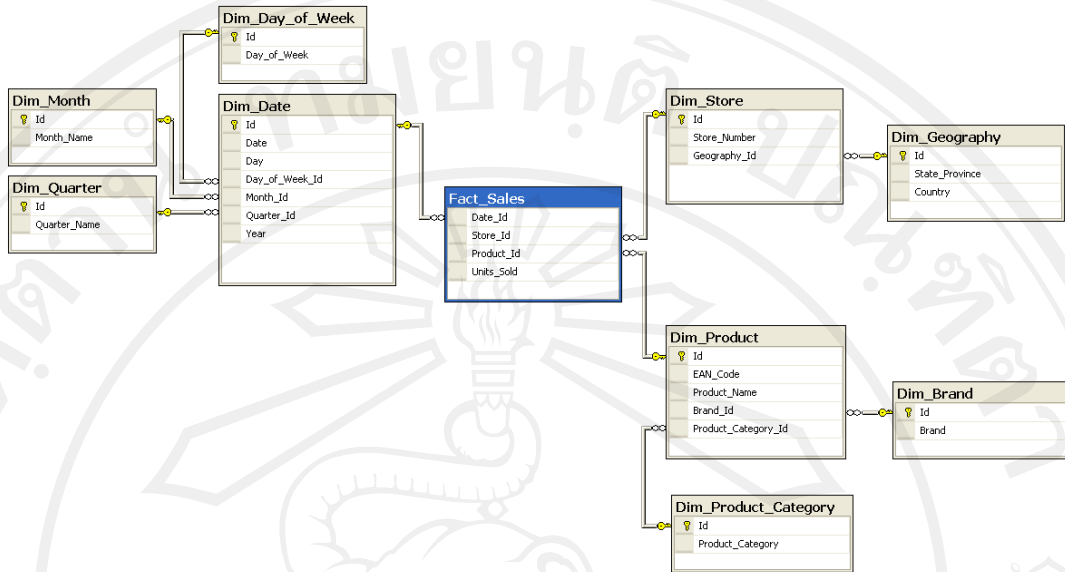
รูปที่ 2.4 การออกแบบฐานข้อมูลโครงสร้างรูปดาว

ที่มา: [ระบบออนไลน์]. Data Warehouse Model. แหล่งที่มา

[http://khomkrit.blogspot.com/2009\\_01\\_01\\_archive.html](http://khomkrit.blogspot.com/2009_01_01_archive.html) (3 พฤษภาคม 2554)

### 2.4.2 การออกแบบฐานข้อมูลโครงสร้างเกล็ดหิมะ

หมายถึงแบบจำลองข้อมูลเชิงมิติ ที่มีตารางข้อเท็จจริงขนาดใหญ่เพียงหนึ่งเดียวอยู่ตรงกลาง และมีตารางมิติจำนวนหนึ่งเชื่อมโยงอยู่เพื่อกำหนดมุมมองที่จะมีต่อข้อมูลที่ใช้เพื่อการวัดค่าในตารางข้อเท็จจริงนั้น แสดงโครงสร้างดังรูปที่ 2.5



รูปที่ 2.5 การออกแบบฐานข้อมูลโครงสร้างเก็ดคิดหิมะ

ที่มา: [ระบบออนไลน์]. **Data Warehouse Model**. แหล่งที่มา

[http://khomkrit.blogspot.com/2009\\_01\\_01\\_archive.html](http://khomkrit.blogspot.com/2009_01_01_archive.html) (3 พฤษภาคม 2554)

## 2.5 กระบวนการเตรียมข้อมูลสำหรับคลังข้อมูล

กระบวนการซึ่งทำหน้าที่แปลงและเคลื่อนย้ายข้อมูล จากที่หนึ่งไปยังอีกที่หนึ่งในคลังข้อมูล หรือเรียกว่า กระบวนการคัดแยกข้อมูล (ETL) โปรแกรมประยุกต์ในปัจจุบันสามารถดำเนินการได้ กับแหล่งที่มาหลายประเภท ชุดของการดำเนินงานของสถาปัตยกรรมคลังข้อมูลที่เกิดในระยะหลัง เป็นที่รู้จักกันในชื่อว่า การคัดแยกข้อมูล (Extraction) การแปลงข้อมูล (Transformation) และการ โอนย้ายข้อมูล (Loading) กระบวนการคัดแยกข้อมูลมีหน้าที่คัดแยกข้อมูลที่มาจากต่างแหล่งที่มา ทำความสะอาดข้อมูลและปรับแต่งข้อมูลให้ถูกต้องสมบูรณ์ เพื่อให้เหมาะสมกับความต้องการทาง ธุรกิจ จากนั้นจึงโอนย้ายข้อมูลเข้าสู่คลังข้อมูล คลังข้อมูลทั่วไปเก็บบันทึกข้อมูลประวัติศาสตร์ที่เป็นผลมาจากการการแปลง การรวมและการรวมกันของรายละเอียดต่างๆของแหล่งที่มา แต่เดิม กระบวนการคัดแยกข้อมูลเกิดปัญหาดังนี้

- ปริมาณข้อมูลขนาดใหญ่ ปริมาณของข้อมูลในการดำเนินงานที่มาขนาดใหญ่ก่อให้เกิด ปัญหาการจัดการข้อมูลในทุกขั้นตอนของกระบวนการ
- คุณภาพของข้อมูล ข้อมูลมักจะไม่ถูกต้องสมบูรณ์เสมอไป ซึ่งจะต้องถูกทำความสะอาด ในกระบวนการ
- วิวัฒนาการการเก็บข้อมูล วิวัฒนาการของแหล่งข้อมูลและคลังข้อมูลสามารถนำไปสู่การ ดำเนินงานบำรุงรักษาทุกวันได้

- ปัญหาเกี่ยวกับประสิทธิภาพ กระบวนการทั้งหมดจะต้องเกิดขึ้นภายในเวลาที่กำหนด ดังนั้นจึงต้องมีการปรับปรุงเวลาปฏิบัติการ (Fon Silvers, 2008: 152-157)

นอกจากนี้ในแต่ละขั้นตอนของกระบวนการคัดแยกข้อมูล มีรายละเอียดดังต่อไปนี้

- 1) การคัดแยกข้อมูล (Extract) คือ การดึงข้อมูลจากแหล่งข้อมูลที่แตกต่างกัน
- 2) การแปลงข้อมูล (Transform) คือ การนำข้อมูลที่ได้อาจจากการ Extract มาจัดรูปแบบให้ถูกต้องสอดคล้องกัน เช่น

- การจับคู่ข้อมูล (Data Mapping) การทำให้ข้อมูลที่มีความหมายเดียวกันแต่อยู่ในรูปแบบที่แตกต่างกันให้อยู่ในรูปแบบเดียวกัน

- การทำความสะอาดข้อมูล (Data Cleansing) การตรวจสอบและแก้ไขข้อมูลให้ถูกต้อง

- 3) การโอนย้ายข้อมูล (Load) คือการนำข้อมูลที่ผ่านการแปลงข้อมูลแล้วเข้าสู่คลังข้อมูล

## 2.6 การประมวลผลในเชิงวิเคราะห์แบบออนไลน์

หมายถึงกระบวนการประมวลผลข้อมูลทางคอมพิวเตอร์ที่ช่วยให้สามารถวิเคราะห์ข้อมูลในมิติต่าง ๆ (Multidimensional Data Analysis) ของข้อมูลได้ง่ายยิ่งขึ้น การดำเนินการประกอบไปด้วย

- 1) การเข้าถึงข้อมูลแบบภาพรวมและแบบเจาะลึก (Roll Up and Drill Down) คือการเปลี่ยนแปลงระดับความละเอียดของการพิจารณาข้อมูล

- แบบภาพรวม หมายถึง การเปลี่ยนแปลงระดับความละเอียดในการพิจารณาข้อมูลจากระดับละเอียดขึ้นไปสู่ระดับที่หยาบมากขึ้น
- แบบเจาะลึก หมายถึง การเพิ่มความละเอียดในการพิจารณาข้อมูลจากระดับที่หยาบไปสู่ระดับที่ละเอียดมากขึ้น

- 2) การเข้าถึงข้อมูลแบบเลือกพิจารณาบางส่วนและแบบการพลิกมิติ (Slice and Dice) คือการแยกข้อมูลออกเป็นส่วนเพื่อพิจารณาเฉพาะส่วนที่ต้องการ โดยเกณฑ์การแยกจะใช้ข้อมูลมิติใด ๆ

- แบบเลือกพิจารณาบางส่วน เป็นการเลือกพิจารณาผลลัพธ์บางส่วนที่เราสนใจ โดยการเลือกเฉพาะค่าที่ถูกกำกับด้วยข้อมูลบางค่าของแต่ละมิติเท่านั้น
- แบบการพลิกมิติ เป็นกระบวนการพลิกแกนหรือมิติข้อมูลให้ตรงความต้องการของผู้ใช้งาน



## 2.7 ระบบธุรกิจอัจฉริยะ (Business Intelligence)

ระบบธุรกิจอัจฉริยะหมายถึงกลุ่มของกิจกรรมที่จะส่งผลให้เข้าใจสถานภาพของธุรกิจ โดยดำเนินการประเภทต่างๆของการวิเคราะห์ข้อมูลของ บริษัท เช่นเดียวกับบนข้อมูลภายนอก จากบุคคลที่สามจะช่วยสร้างกลยุทธ์ ยุทธวิธี และกระบวนการตัดสินใจด้านการดำเนินงาน สามารถสร้างกิจกรรมที่จำเป็นสำหรับปรับปรุงประสิทธิภาพขององค์กร ซึ่งรวมถึง การวิเคราะห์ การทำความเข้าใจ และการจัดการข้อมูลเกี่ยวกับประสิทธิภาพการดำเนินงาน ลูกค้า และกิจกรรมของผู้จัดจำหน่าย การจัดการการเงิน ความเคลื่อนไหวของตลาด คู่แข่ง กฎระเบียบการปฏิบัติงานและการควบคุมคุณภาพ ซึ่งอธิบายคุณลักษณะได้ดังนี้

- การจัดการประสิทธิภาพขององค์กร ซึ่งกล่าวรวมถึงตัวบ่งชี้ประสิทธิภาพการผลิต เช่น การขายประจำวัน การใช้ประโยชน์จากทรัพยากรและต้นทุนหลักในการดำเนินงานสำหรับแต่ละภูมิภาค แต่ละสายการผลิตและแต่ละช่วงเวลา เพื่อที่จะให้องค์กรใช้ยุทธวิธีเพื่อให้ได้ผลการดำเนินงานอยู่บนเป้าหมายที่ต้องการ

- การวิเคราะห์ผลกำไรจากลูกค้า เพื่อที่จะเข้าใจลูกค้าที่ให้ผลกำไรและคุ้มค่าแก่การรักษา และเข้าใจการสูญเสียเงินดังนั้นจึงต้องมีการดำเนินการเกิดขึ้น สิ่งสำคัญคือการจัดสรรต้นทุนให้ดีที่สุดเท่าที่จะเป็นไปได้ในหน่วยการทำงานที่เล็กที่สุดขององค์กร

- การวิเคราะห์ทางสถิติ เช่น โอกาสในการซื้อ หรือ ความสัมพันธ์ของสินค้าที่ลูกค้ามักจะซื้อพร้อมกัน คือกระบวนการวิเคราะห์ข้อมูลการขายในการตัดสินใจซื้อสินค้าพร้อมกัน ความน่าจะเป็นนี้จะแสดงในรูปของค่าทางสถิติ เช่นระดับความเชื่อมั่น ซึ่งการวิเคราะห์ทางสถิติมีความสำคัญต่อกระบวนการขายปลีกและอุตสาหกรรม แต่ก็มีความสำคัญระดับหนึ่งกับธุรกิจบริการด้านการเงิน

- การพยากรณ์ เช่น การพยากรณ์การขาย รายได้และการจัดการต้นทุนสำหรับวางแผนงบประมาณในปีถัดไปโดยคำนึงถึงปัจจัยอื่น ๆ ร่วมด้วย เช่น การเติบโตจากภายใน สถานการณ์ทางเศรษฐกิจ และทิศทางในอนาคตของบริษัท

### 2.7.1 กลุ่มของระบบธุรกิจอัจฉริยะ

จากวิธีการวิเคราะห์เชิงลึกและระดับของความซับซ้อน สามารถแยกกลุ่มของระบบธุรกิจอัจฉริยะ ได้ 3 กลุ่ม ดังนี้

- ระดับรายงาน เช่น ตัวชี้วัดประสิทธิภาพ ตัวเลขยอดขายทั่วโลก จำนวนลูกค้าทั่วโลกและอัตราการใช้ทรัพยากรในสาขาที่แตกต่างกันในหลายประเทศ

- ระดับระบบประมวลผลแบบออนไลน์ เช่น การรวม การเจาะลึกข้อมูล การแบ่งข้อมูลและการเปลี่ยนมุมมอง

- ระดับเหมืองข้อมูล เช่น คุณลักษณะของข้อมูล การจำแนกข้อมูล การวิเคราะห์ความสัมพันธ์ การแยกประเภทข้อมูลและการพยากรณ์ และการจัดกลุ่มข้อมูล

## 2.8 งานวิจัยที่เกี่ยวข้อง

สมนึก วิเศษธรรมรัตน์ (2547) ได้ศึกษาการวิเคราะห์ยอดขายตามเป้าประสงค์ ของบริษัท อาร์ตแอนด์เทคโนโลยี จำกัด ด้วยการแก้ปัญหาธุรกิจอย่างชาญฉลาดของไมโครซอฟท์ ให้สามารถสนับสนุนการวิเคราะห์ยอดขายที่มีประสิทธิภาพ ซึ่งจากการศึกษาพบว่าบริษัทต้องการสารสนเทศสรุปเป็นมิติข้อมูลด้านยอดขายได้ 4 มิติ คือ (1) มิติด้านช่องทางการจัดจำหน่าย (2) มิติด้านลูกค้า (3) มิติด้านช่วงเวลา (4) มิติด้านสินค้า เมื่อนำมาสร้างคลังข้อมูลแล้วได้ทดสอบกับคำถามต่างๆของผู้จัดการฝ่ายขาย ซึ่งสามารถตอบคำถามได้ตามต้องการ ช่วยให้ผู้จัดการฝ่ายขายสามารถเข้าถึงข้อมูลได้ตลอดเวลา ทำการวิเคราะห์ข้อมูลแบบเฉพาะชั่วคราวได้ ส่งผลให้การตัดสินใจเพื่อการบริหารสินค้า การบริหารช่องทางการจัดจำหน่ายมีประสิทธิภาพมากขึ้น

สุประพล เถาวพงศ์ (2553) ได้ศึกษาการพัฒนาคลังข้อมูลด้านนักศึกษาของมหาวิทยาลัยพายัพ การพัฒนาระบบคลังข้อมูลที่มุ่งเน้นการวิเคราะห์ข้อมูลเกี่ยวกับนักศึกษาในด้านต่างๆเพื่อสนับสนุนการตัดสินใจของผู้บริหารนั้นในกระบวนการดำเนินงานต้องเริ่มต้นจากการศึกษาความต้องการของผู้บริหารก่อนว่าต้องการทราบข้อมูลเกี่ยวกับนักศึกษาด้านใดบ้าง จากนั้นทำการวิเคราะห์ระบบงานเดิม ออกแบบคลังข้อมูลแล้วจึงทำการสกัดหรือคัดแยกข้อมูล แปลงข้อมูล และโอนย้ายข้อมูลเข้าสู่คลังข้อมูล โดยการนำเข้าข้อมูลนั้นจะมีกระบวนการคัดกรองข้อมูลให้มีความสอดคล้อง ครบถ้วน และถูกต้องตามความต้องการของผู้บริหาร

กชนทร์ อึ้งสกุล (2552) ได้ศึกษาการพัฒนาาระบบคลังข้อมูลด้านแรงงานของสำนักงานสถิติจังหวัดเชียงใหม่ ว่า การการพัฒนาาระบบคลังข้อมูลดังกล่าวสามารถนำเสนอข้อมูลในรูปแบบต่างๆได้หลากหลาย โครงสร้างของคลังข้อมูลแบบเชิงมิติทำให้ง่ายต่อการสร้างรายงานให้อยู่ในรูปแบบรายงานหลายมิติ สามารถนำรายงานที่ได้ไปใช้ในการวิเคราะห์และกำหนดทิศทางการพัฒนาองค์กรได้อย่างมีประสิทธิภาพ

อุไรพร เจตนาชัย และ ระวีวรรณ เอื้อพันธ์วิริยะกุล (2546) ได้ศึกษางานวิจัยในหัวข้อ การเปรียบเทียบความเร็วในการค้นหาข้อมูลด้วยคำสั่งเอสคิวแอลในคลังข้อมูลที่ออกแบบด้วยโมเดลรูปดาวและโมเดลเกล็ดหิมะ โดยยกฝ่ายการเงินและทะเบียนของมหาวิทยาลัยมหานครเป็นกรณีศึกษา วัดความเร็วในการค้นหาข้อมูลจากคลังข้อมูลที่ออกแบบโดยโมเดลทั้งสอง นำความเร็วที่วัดได้มาวิเคราะห์ความถดถอยและสัมประสิทธิ์สหสัมพันธ์ ทำการสรุปว่าการออกแบบคลังข้อมูลโมเดลแบบดาวมีประสิทธิภาพสูงกว่าแบบโมเดลแบบเกล็ดหิมะ หากมีการเพิ่มลำดับของตารางให้ลึกมากขึ้นก็จะทำให้ความเร็วในการค้นหาข้อมูลลดน้อยลง

Matteo Golfarelli, Stefano Rizzi(1999) การออกแบบระบบคลังข้อมูลมีความแตกต่างจากการออกแบบระบบปฏิบัติการ ในงานวิจัยนี้ได้นำเสนอวิธีปฏิบัติทั่วไปสำหรับการออกแบบระบบคลังข้อมูล โดยที่วิธีปฏิบัตินี้จะอยู่บนพื้นฐานของแบบจำลองเชิงมิติ มีการนำเสนอขั้นตอนทั้งหมด 6 ขั้นตอน ดังนี้

ขั้นตอนที่ 1 วิเคราะห์ระบบสารสนเทศเดิม ศึกษาจากเอกสารในรูปแบบต่างๆ แล้วจึงออกแบบฐานข้อมูลขึ้นมา

ขั้นตอนที่ 2 ระบุความต้องการ นำรูปแบบฐานข้อมูลที่ได้มาวิเคราะห์ภาระงานเบื้องต้น โดยร่วมกันกำหนดกับผู้ใช้งาน

ขั้นตอนที่ 3 ออกแบบแบบจำลองเชิงแนวคิด นำรูปแบบฐานข้อมูลและภาระงานเบื้องต้นที่ได้ มาใช้ในการออกแบบแบบจำลองเชิงมิติ

ขั้นตอนที่ 4 การปรับแต่งภาระงานและการตรวจสอบโครงสร้างแบบจำลอง

ขั้นตอนที่ 5 ออกแบบแบบจำลองแบบเชิงตรรกะ แปลงจากแบบจำลองเชิงมิติให้สอดคล้องกับภาระงาน ในขั้นตอนนี้เป็นส่วนสำคัญคือมีการกำหนดความสัมพันธ์ จับคู่ระหว่างตารางข้อเท็จจริงและตารางมิติ กำหนดลำดับชั้นของมิติแต่ละด้าน ซึ่งอาจจะได้โมเดลหรือแบบจำลองโครงสร้างรูปดาวหรือแบบจำลองโครงสร้างรูปเกล็ดหิมะ

ขั้นตอนที่ 6 ออกแบบแบบจำลองแบบเชิงกายภาพ เป็นขั้นตอนสุดท้ายของคลังข้อมูล ที่กำหนดคุณลักษณะของระบบคลังข้อมูล ชนิดประเภทข้อมูลกำหนดดัชนี