

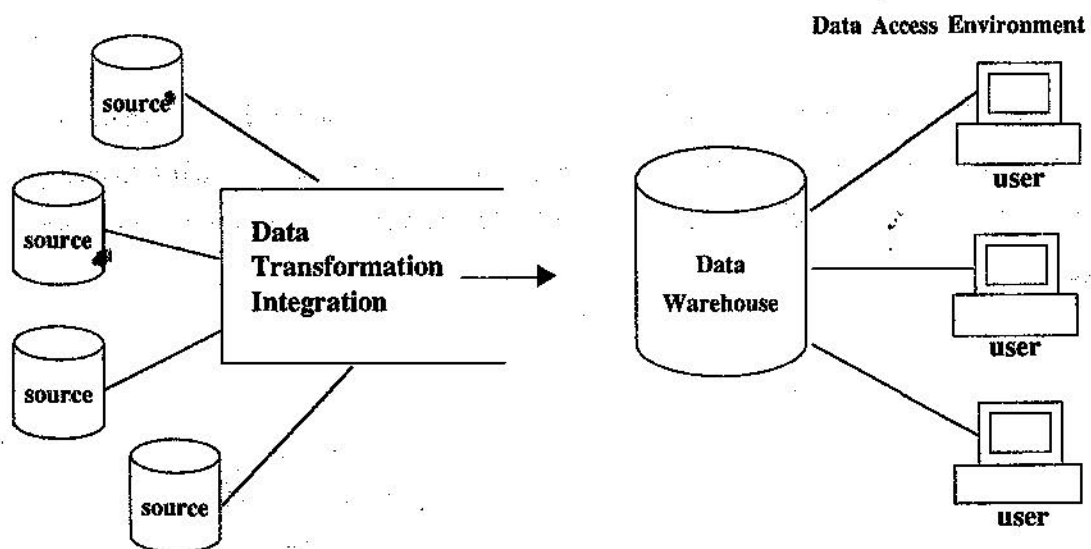
การสร้าง Data Warehouse สำหรับการตัดสินใจ

สมชาย เล็กเจริญ**

อชิป ชวากัณหะ**

ปัจจุบันนี้ธุรกิจมีการแข่งขันกันสูงขึ้นดังนั้นผู้บริหารธุรกิจต้องสามารถคาดการณ์หรือรับรู้แนวโน้มการเปลี่ยนแปลงธุรกิจได้อย่างรวดเร็วและถูกต้อง เพื่อเตรียมตัวรับสถานการณ์ต่างๆ ที่อาจจะเกิดขึ้น และสามารถแก้ไขได้ทันเวลาที่ ฉะนั้นผู้บริหารต้องตื่นตัวในการรับรู้ข้อมูลทั้งภายในและภายนอกองค์กร ปัจจุบันปริมาณและความซับซ้อนของข้อมูลมีมากขึ้น ดังนั้นองค์กรต้องมีแหล่งรวบรวมข้อมูลที่สามารถจัดการและวิเคราะห์ข้อมูลทั้งในอดีตปัจจุบันและอนาคตได้อย่างมีประสิทธิภาพ จึงสามารถทำกำไรหรือผลประโยชน์ทางธุรกิจได้อย่างมหาศาล ด้วยเหตุนี้จึงมีการพัฒนา Data Warehouse ขึ้นมาเป็นส่วนสำคัญของกลยุทธ์ขององค์กร

Data Warehouse เป็นกระบวนการรวบรวมข้อมูลจากหลายแหล่งของฐานข้อมูลหรือ ข้อมูลการทำงานประจำวัน (Operational data) จากนั้นก็แปลงข้อมูลเพื่อให้อยู่ในรูปแบบที่เหมาะสมในการจัดเก็บ และง่ายต่อการนำมาใช้ ข้อมูลที่ผ่านการเปลี่ยนแปลงนั้นถูกจัดเก็บในฐานข้อมูลของ Data Warehouse ดังรูปที่ 1

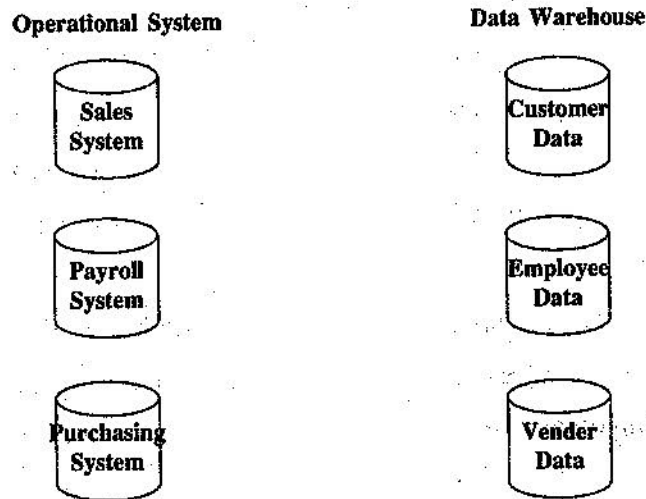


รูปที่ 1

ฉะนั้นพอสรุปได้ว่า Data Warehouse นั้นเป็นเพียงฐานข้อมูลที่ช่วยสนับสนุนการตัดสินใจของผู้บริหาร โดยฐานข้อมูลดังกล่าวจะต้องมีคุณสมบัติดังนี้

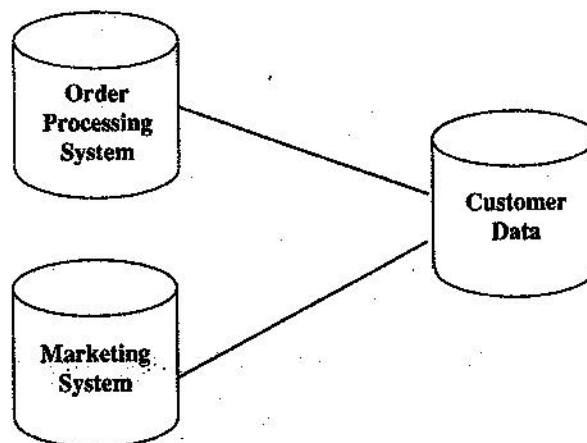
** อาจารย์คณะนิเทศศาสตร์ สาขาการจัดการสารสนเทศ

1. Subject Oriented คือข้อมูลของ Data Warehouse ได้มาจากฐานข้อมูลของบริษัทซึ่งจะเลือกเอาเฉพาะใช้ในการตัดสินใจของผู้บริหารซึ่งจำเป็นต้องไปสัมภาษณ์ผู้บริหารก่อนว่าข้อมูลใดเป็นปัจจัยที่สำคัญในการตัดสินใจเพื่อให้ธุรกิจประสบความสำเร็จ (Critical Success Factor) เช่น ข้อมูลลูกค้า ข้อมูลพนักงาน ข้อมูลของผู้ส่งสินค้า เป็นต้น ดังรูปที่ 2



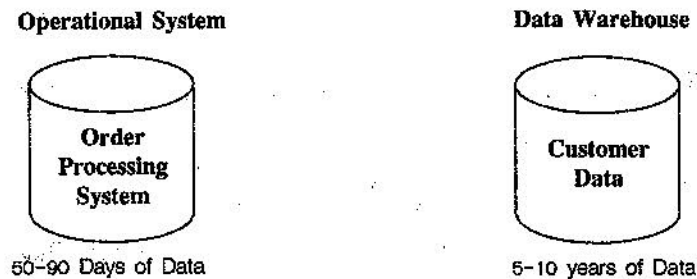
รูปที่ 2

2. Integrated ข้อมูลใน Data Warehouse นั้นมาจากแหล่งข้อมูลปฏิบัติการประจำวัน เช่น ระบบประมวลผลการสั่งซื้อพิจารณาในแง่ของลูกค้าและฐานข้อมูลภายนอก เช่น ระบบตลาด (Marketing System) มีการพิจารณาเกี่ยวกับลูกค้าเหมือนกัน ดังนั้นจะดึงข้อมูลลูกค้าไปไว้ใน Data Warehouse ดังรูปที่ 3



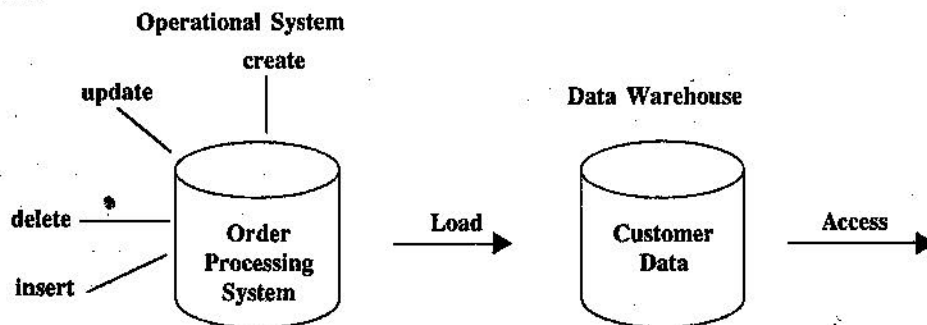
รูปที่ 3

3. Time-Variant ใน Data Warehouse บรรจุข้อมูลในอดีตตั้งแต่ 5 ถึง 10 ปี และข้อมูลเกือบเป็นปัจจุบันเพื่อใช้ในการวิเคราะห์หาแนวโน้มและทำนายผลลัพธ์ในอนาคตได้ ดังรูปที่ 4



รูปที่ 4

4. Non-Volatile ข้อมูลใน Data Warehouse ต้องเปลี่ยนแปลงได้ยาก ซึ่งจุดประสงค์ของ Data Warehouse นั้นต้องการให้ผู้บริหารเข้าใช้ข้อมูลในฐานะข้อมูลได้เพียงอย่างเดียว (Read Only database) ดังรูปที่ 5



รูปที่ 5

5. Accessible ใน Data Warehouse ถูกออกแบบมาเพื่อให้ผู้บริหารเข้าถึงข้อมูลได้ง่ายเพราะคำนึงถึงว่าผู้บริหารนั้นมีเวลาน้อยในการศึกษาระบบและมีความคุ้นเคยกับระบบคอมพิวเตอร์

6. Transformed แหล่งต่างๆ ของข้อมูลปฏิบัติการประจำวันและฐานข้อมูลภายนอก ซึ่งจะมีลักษณะการจัดเก็บที่แตกต่างกัน ฉะนั้นจำเป็นต้องเปลี่ยนแปลงรูปแบบการจัดเก็บก่อนบันทึกลงใน Data Warehouse เพื่อให้ง่ายต่อการเข้าถึงและวิเคราะห์ข้อมูล

7. Management-Oriented จุดประสงค์หลักของ Data Warehouse คือ ช่วยการวิเคราะห์และสนับสนุนในการตัดสินใจ ฉะนั้น Data Warehouse ต้องมีความยืดหยุ่นและจัดเตรียมให้ผู้บริหารสามารถมองและวิเคราะห์ข้อมูลได้หลายรูปแบบเพื่อใช้สนับสนุนการตัดสินใจ ผลลัพธ์ที่ได้อาจจะเป็นในรูปแบบของบทสรุป

การออกแบบฐานข้อมูล Data Warehouse (Poe, Videtee 1995:124-133)

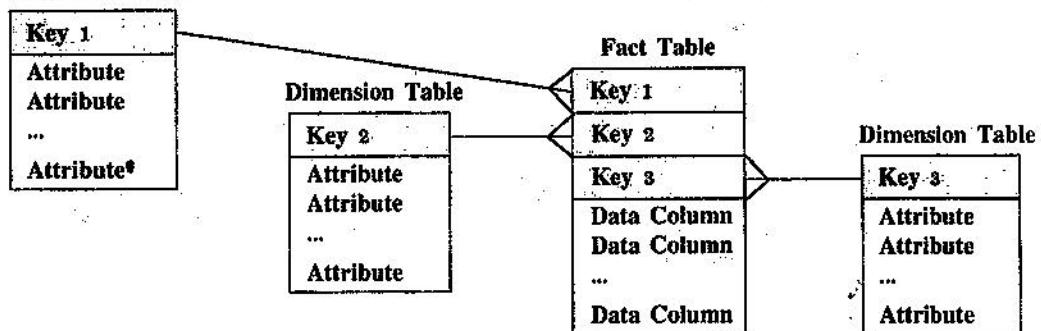
ฐานข้อมูลสนับสนุนการตัดสินใจต้องถูกออกแบบเพื่อการวิเคราะห์ข้อมูลและมีความง่ายต่อการใช้ ตลอดจนมีความรวดเร็วในการดึงข้อมูลออกจากฐานข้อมูล โดยใช้การออกแบบฐานข้อมูลเรียกว่า Star Schema ซึ่งมีประโยชน์ดังนี้

1. เป็นวิธีการออกแบบฐานข้อมูลที่ใช้เวลาตอบสนองรวดเร็ว เช่น การดึงข้อมูลได้รวดเร็ว
2. สามารถปรับปรุงฐานโครงสร้างฐานข้อมูลได้ง่าย
3. มีเครื่องมือ (tools) ในการดึงข้อมูลให้เลือกจำนวนมาก

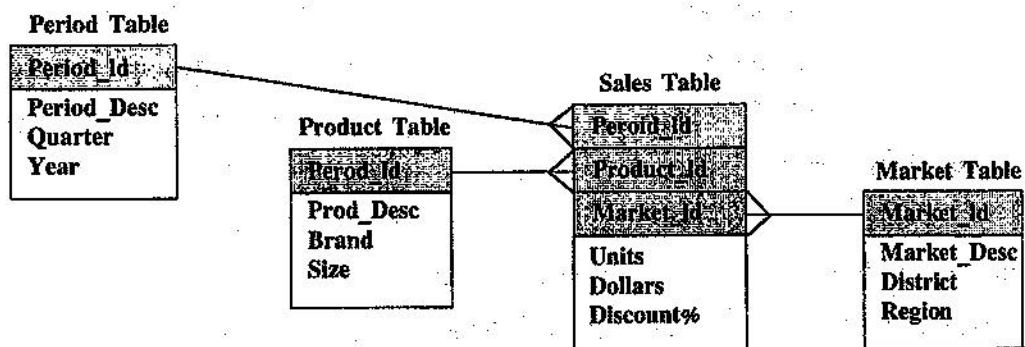
Star Schema ประกอบด้วยตาราง (table) 2 ชนิดคือ fact table เป็นตารางหลักสามารถประกอบด้วยหลาย column และอาจประกอบด้วยจำนวนล้าน Row ก็ได้ และ Dimension table เป็นตารางรองมีข้อมูลน้อยกว่า

ตัวอย่างให้พิจารณาดังรูปที่ 6 Fact Table คือ Sale Table ซึ่งประกอบด้วย primary key ของ Period table, Product table และ Market table ส่วน Attribute จะเป็นเกี่ยวกับรายละเอียดของการขายสำหรับ dimension table คือ Period table, Product table และ Market table

Dimension Table



ตารางความสัมพันธ์ของ fact และ dimension table ใน simple star schema



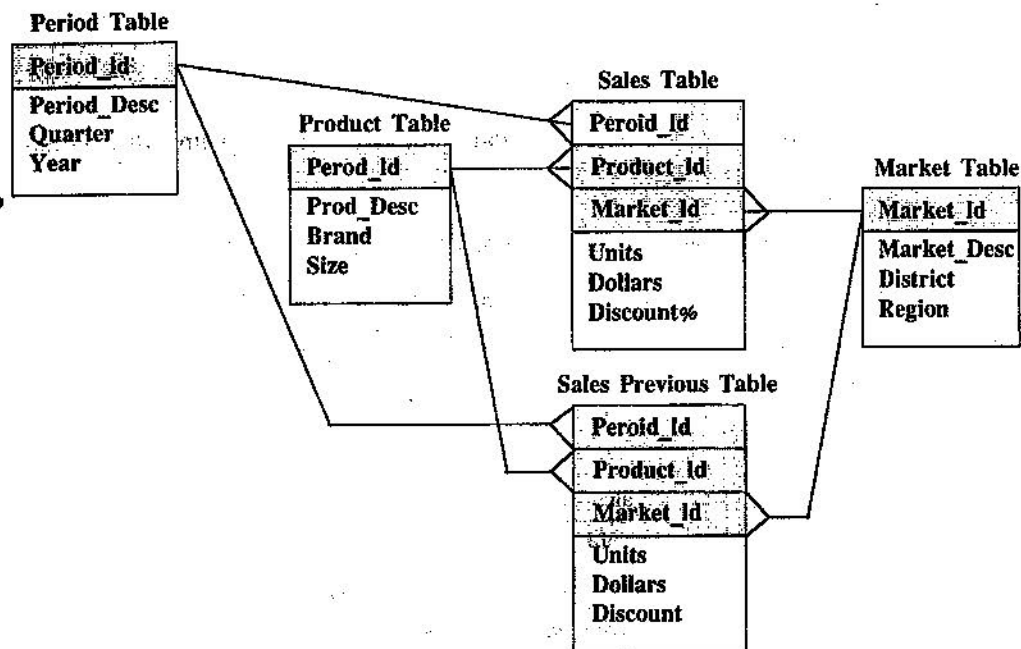
รูปที่ 6

ฐานข้อมูลการขายใน simple star schema

สังเกตพบว่าเป็นความสัมพันธ์แบบ many to one ระหว่าง foreign key ใน fact table และ primary key ของ dimension table เช่น รหัสสินค้าแต่ละ Row ใน product table มีค่าไม่ซ้ำกัน แต่รหัสสินค้าเดียวกันปรากฏใน sale table ได้หลายครั้งโดยเฉพาะสินค้าที่เกิดขึ้นต่างเวลา และตลาด

ในบางครั้งในการออกแบบฐานข้อมูลแบบ star schema สามารถมี fact table หลายๆ ตารางได้ ขึ้นอยู่กับความแตกต่างของเวลา หรือไม่สัมพันธ์กับความเป็นจริงจึงต้องเพิ่ม fact table หรือเพื่อเพิ่มประสิทธิภาพในการค้นคืนข้อมูลวิเคราะห์และออกรายงาน เช่น ต้องการรวบรวมข้อมูลการขายเป็นรายวัน รายปี ฉะนั้น fact table ต้องแยกออกเป็นหลาย table เพื่อเหมาะสมกับงาน

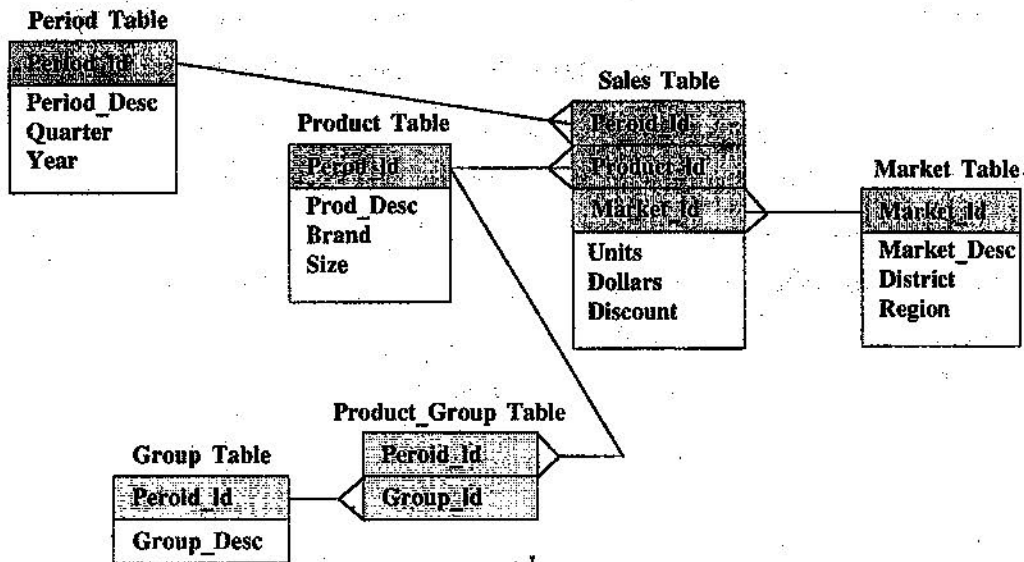
ตัวอย่างในรูปที่ 7 ฐานข้อมูลระบบการขายมีการเพิ่ม fact table เพื่อรวบรวมข้อมูลในอดีตของการขายปีที่ผ่านมา



รูปที่ 7

ฐานข้อมูลการขายโดยเพิ่ม sale previous table ในส่วนของ fact table

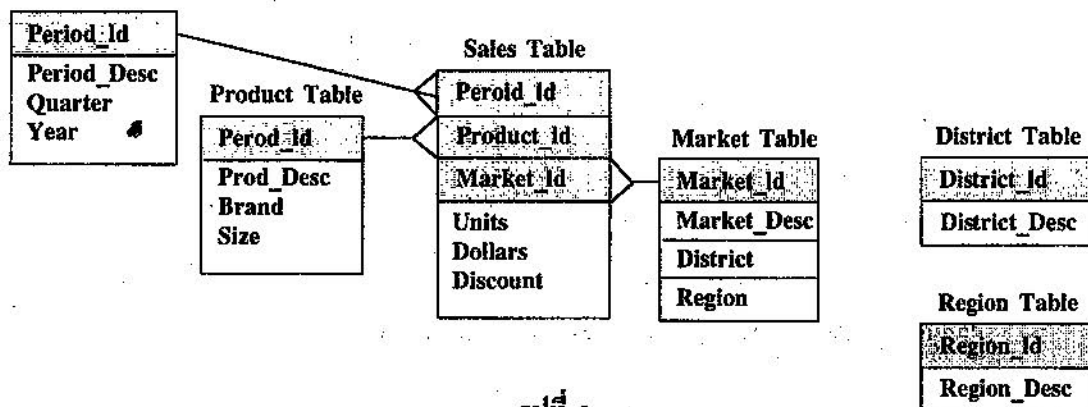
อีกกรณีหนึ่งทำให้มี fact table หลายตารางคือ ความสัมพันธ์แบบ many to many เช่น สินค้าชนิดหนึ่งขึ้นอยู่กับกลุ่มของสินค้าหนึ่งกลุ่มหรือหลายกลุ่ม และกลุ่มแต่ละกลุ่มมีสินค้าหลายชนิดเมื่อพิจารณาพบว่าสินค้ากับกลุ่มเป็นตารางสัมพันธ์กันแบบ many to many ฉะนั้นจึงต้องสร้าง fact table ขึ้นมาซึ่งมี Attribute ดังนี้คือ product-id และ Group-id เป็นตัวเชื่อมระหว่าง 2 table เพื่อกระจาย เป็นแบบ one to many ดังรูปที่ 8



รูปที่ 8

การใช้ fact table ในการแก้ปัญหาลความสัมพันธ์แบบ many to many ระหว่าง group และ products

ในส่วนของ Dimension table สามารถกระจายออกไปได้อีกเรียกว่า Secondary dimension table หรือ outboard table ดังรูปที่ 9 จะพบว่ามี 2 outboard table ประกอบด้วย District table และ Region table ซึ่ง primary key ทั้ง 2 outboard table มาเป็น foreign key ของ Market table



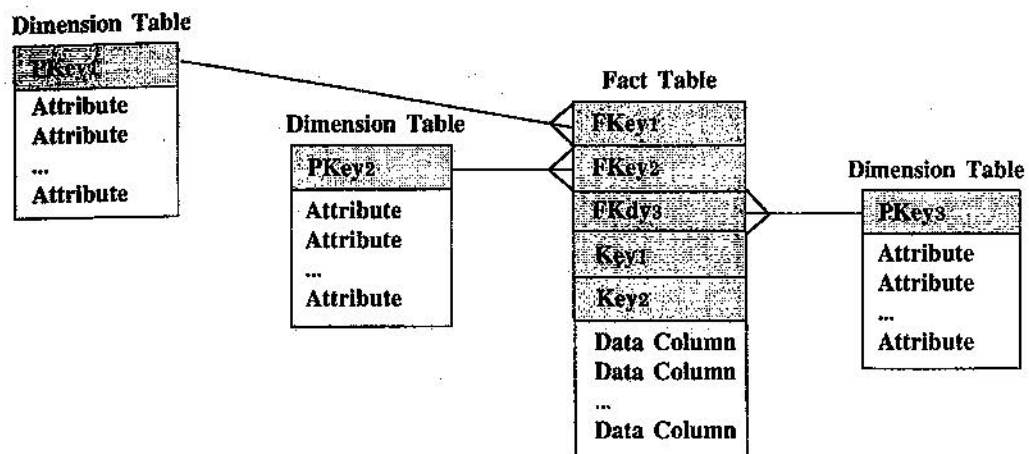
รูปที่ 9

ตัวอย่าง secondary dimension หรือ outboard table

ในบางครั้งไม่สามารถใช้ star schema ซึ่ง primary key ในตาราง fact table มาจาก primary key ของ dimension table ซึ่ง primary key ของ table ไม่สามารถไปกำหนดหรือชี้แต่ละ row ของ fact table จึงต้องใช้ multi-star schema

multi-star schema เป็น fact table ซึ่งประกอบด้วยกลุ่มของ foreign key ซึ่งอ้างอิงไปถึง dimension table และ primary key ร่วมกันสามารถไปกำหนดหรือชี้แต่ละ Row ได้

ดังรูปที่ 10 เป็นการออกแบบโดยสร้าง multi-star schema ในตาราง fact table ประกอบด้วย foreign key คือ Fkey 1, Fkey 2, Fkey 3 ซึ่งแต่ละตัวเป็น primary key ของ dimension table และ primary คือ key 1 กับ key 2



รูปที่ 10

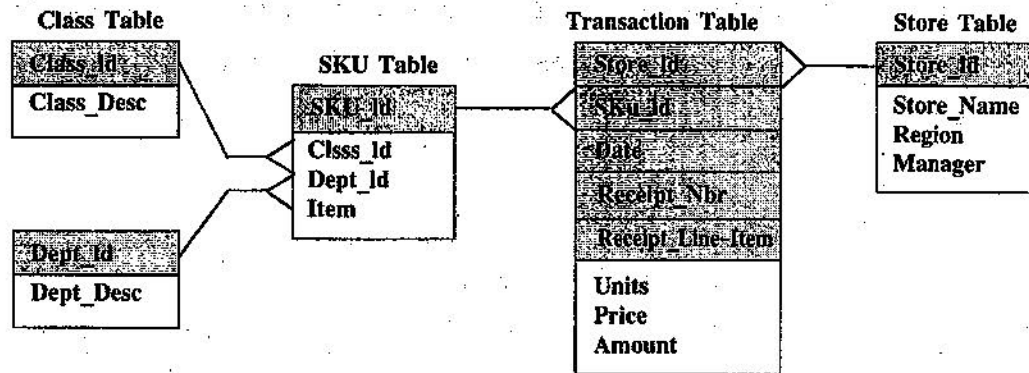
ตัวอย่างการออกแบบ multi-star schema

ในรูปที่ 11 เป็นตัวอย่างการออกแบบฐานข้อมูลขายปลีกซึ่งเป็นแบบ multi-star schema เมื่อพิจารณาพบว่ามีสอง dimension table คือ store table (ตารางร้านค้า) มี Primary key คือ store-id และ SKU table (ตารางคลังสินค้า) มี primary key คือ SKU-id ทั้งสอง primary key ก็เป็น primary key ของ Transaction table (หรืออาจเรียกว่าเป็น foreign key ที่อ้างอิงไปถึง store table และ SKU table แต่ไม่สามารถไปกำหนด row ได้

ตารางที่ 1

store-id	sku-id	units	price	amount
10	15	2	100	200
10	15	5	100	500
11	14	3	200	600

ฉะนั้นจึงจำเป็นต้องเพิ่ม primary key ของ fact table เป็น Receipt-Nbr (เลขที่ใบเสร็จ) Data (วัน เดือน ปี) Receipt-Line-Item (ชนิดของสินค้า) จึงสามารถไปกำหนด Row ได้ ส่วนของ dimension table ก็สามารถกระจายเป็น secondary dimension table ได้อีก เช่น class table และ Dept table



รูปที่ 11

การออกแบบฐานข้อมูลการขายปลีกแบบ multi-star schema และ
สอง secondary dimension table

สรุปได้ว่า การออกแบบฐานข้อมูล Data Warehouse นั้นประกอบด้วย table สองชนิดคือ fact table และ dimension table

ในรูป fact table แบ่งได้สองชนิดคือ simple fact table มี primary key ซึ่งเป็น foreign key ที่อ้างไปถึง dimension table และอีกชนิดก็คือ multi fact table นั้นเหมือนกับ simple fact table ทุกประการแต่ต้องเพิ่ม primary key ของตัวมันเองเพื่อไปกำหนดว่าอยู่ใน Rowใด ในส่วนของ dimension table เก็บข้อมูลทางธุรกิจ (ตัวอย่างเช่น Sales Previous Table รูปที่ 7)

ปัจจุบันเป็นยุคของการแข่งขันทางด้านข้อมูล Data Warehouse เป็นศูนย์รวมการจัดเก็บข้อมูลมากมายมหาศาล แต่มีความสามารถในการลดเวลาการสืบค้นข้อมูล ช่วยในการวิเคราะห์และสนับสนุนการตัดสินใจของผู้บริหารและใช้งานง่ายสามารถแสดงถึงข้อมูลรายละเอียดต่างๆ ในลักษณะเจาะลึก (drill-down)

แต่สิ่งสำคัญที่ไม่สามารถมองข้ามไปได้ คือ การออกแบบฐานข้อมูลของ Data Warehouse ซึ่งผู้ออกแบบฐานข้อมูลจะต้องมีความรู้ทางธุรกิจไม่มากนักน้อย ต้องรู้ว่าข้อมูลอะไรบ้างที่ผู้บริหารต้องการในการตัดสินใจ อาจใช้เทคนิคการสัมภาษณ์ผู้บริหาร (Critical success factor) และต้องมีความรู้การออกแบบฐานข้อมูลของ Data Warehouse ว่ามีลักษณะอย่างไร ตลอดจนต้องศึกษาว่าเครื่องมือใด (tools) มีประสิทธิภาพที่เหมาะสมกับงานธุรกิจของบริษัท

บรรณานุกรม

1. Geraham, Stephen. The foundations of wisdom : a study of the financial data warehousing. Ontario : International Data Corporation (Canada), 1996.
2. Mattison, Rob. Data Warehouse : strategies and techniques. Toronto : McGraw-Hill, c1996.
3. Poe, Vodette. Building a data warehouse for decision support. Englewood Cliffs, N.J. : Prentice Hall PTR., 1996.