คลังข้อมูลผลการจัดเก็บภาษีเพื่อใช้ในการตัดสินใจ

Revenue Data Warehouse for Decision Support

โดย

นายธวัชชัย เพชรวิชิต

รหัส 41067173

H001696

วัน เดือน ปี 25 ส.ค. 2549
เลขทะเบียน 01696
เลขเรียกหนังสือ วพ. ส 395 ค 2543
"ห้องสมุดคณะเทคโนโลยีสารสนเทศ สจล."

อาจารย์ที่ปรึกษา

รศ.ดร.วิเชียร เปรมชัยสวัสดิ์

รายงานนี้เป็นส่วนหนึ่งของวิชาโครงการศึกษากรณีพิเศษ
หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ
ภาคเรียนที่ 1 ปีการศึกษา 2543
คณะเทคโนโลยีสารสนเทศ
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ชื่อหัวข้อ	คลังข้อมูลผลการจัดเก็บภาษีเพื่อใช้ในการตัดสินใจ
นักศึกษา	นายธวัชชัย เพชรวิชิต
อาจารย์ที่ปรึกษา	รศ.ดร.วิเชียร เปรมชัยสวัสดิ์
ระดับการศึกษา	วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ
แขนงวิชา	วิทยาการสารสนเทศ
ปีการศึกษา	2543

บทคัดย่อ

ปัจจุบันระบบฐานข้อมูลมีความจำเป็นต่อหน่วยงานของรัฐบาลมากขึ้นทำให้ต้องมีการพัฒนาคลังข้อมูลผลการจัดเก็บของกรมสรรพากรให้มีความเหมาะสม รวดเร็วในการนำข้อมูลมาใช้ในการตัดสินใจของผู้บริหาร และนำไปใช้เป็นข้อมูลของกรมสรรพากร ในการจัดทำประมาณการการจัดเก็บภาษีของประเทศในปีถัดไป โครงการนี้จะทำการศึกษาทฤษฎีเกี่ยวกับคลังข้อมูลด้าน Star Schema โดยนำมาใช้เป็นระบบสนับสนุนการตัดสินใจ และเครื่องมือที่ใช้ในการวิเคราะห์ข้อมูล ประโยชน์ที่จะได้รับคือการนำระบบคลังข้อมูลผลการจัดเก็บภาษีมาใช้ในงานปัจจุบันขององค์กร และช่วยผู้บริหารในการวางแผนการจัดเก็บภาษี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Title	Revenue Data Warehouse for Decision Support
Student	Mr. Thawatchai Petvichit
Advisor	Assoc.Prof.Dr. Wichian Premchaisawadi
Level of Study	Master of Science in Information Technology
Major	Information Science
Academic Year	2000

ABSTRACT

Data Warehouse Development of Revenue Department to appropriate and Tax access speed up to Decision Support system by executive and take MIS to make estimate tax in nextyear. Inside study in Data Warehouse Theory , Star Schema which support in Decision Support tools analysis. This special project is design Data Warehouse by Star Schema. The Pilot Decision Support Suit Tools is tool for Bureau executive to analysis in summarized complete operation of the Bureau Province and Ampler in Revenue Department. This enable the operation unit to report top management in the preferred method fast to Decision Support System

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

โครงการศึกษากรณีพิเศษนี้สำเร็จได้ด้วยความกรุณาของอาจารย์ที่ปรึกษา รศ.ดร.วิเชียร เปรมชัยสวัสดิ์ และกรรมการสอบ ดร.ประจวบ วานิชชัชวาล, รศ.นุชรี เปรมชัยสวัสดิ์ ขอกราบขอบพระคุณอาจารย์ทุกท่านเป็นอย่างสูง และขอขอบคุณพี่ๆ จากภาค 5 ทุกคนที่ให้ข้อมูล เพื่อน ๆ ทุกคนที่ได้ให้คำแนะนำ ข้อคิดเห็น ที่มีประโยชน์ และทุกท่านที่มีส่วนเกี่ยวข้องในการทำให้โครงการศึกษากรณีพิเศษนี้ สำเร็จได้ด้วยดี

ขอขอบคุณทุกท่าน

ธวัชชัย เพชรวิชิต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

หน้า

บทคัดย่อภาษาไทย..... I

บทคัดย่อภาษาอังกฤษ..... II

กิตติกรรมประกาศ..... III

สารบัญ..... IV

สารบัญตาราง..... VII

สารบัญภาพ..... VIII

บทที่

1. บทนำ..... 1

1.1 ความเป็นมา..... 1

1.2 หน้าที่ของกรมสรรพากร..... 1

1.3 วัตถุประสงค์..... 1

1.4 ขอบเขตของการศึกษา..... 2

1.5 ประโยชน์ที่คาดว่าจะได้รับ..... 2

2. ลักษณะองค์กร..... 3

2.1 ความเป็นมา..... 3

2.2 การจัดรูปองค์กร..... 3

2.3 หน่วยงานที่มีหน้าที่ในการจัดทำรายงานการเก็บภาษีอากรประจำเดือน (บช.8) 5

2.4 โครงสร้างของหน่วยงานที่จัดทำรายการเก็บภาษีอากร (บ.ช.8)..... 6

3. ทฤษฎีที่เกี่ยวข้อง..... 8

3.1 ความหมายของสารสนเทศ..... 8

3.2 ความหมายของสารสนเทศ..... 9

3.3 ลักษณะรายงานสารสนเทศที่สมบูรณ์ 9

3.4 รูปแบบของสารสนเทศ..... 9

3.5 ผู้บริหารระดับสูง..... 11

3.6 ฐานข้อมูลแบบ Star Schema..... 14

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

บทที่	หน้า
3.7 การสร้างคลังข้อมูล	26
4. ระบบคลังข้อมูล (Data Warehousing System)	30
4.1 ความหมายของ Data Warehous	30
4.2 ส่วนประกอบของ Data Warehous	30
4.3 องค์ประกอบในการสร้าง Data Warehouse	33
4.4 คุณลักษณะของข้อมูลใน Data Warehouse	35
4.5 การแปลงข้อมูลเข้าสู่ Data Warehouse	36
4.6 สิ่งที่ต้องคำนึงถึงในการออกแบบ Data Warehoue	38
4.7 Software ที่ใช้ในการจัดการข้อมูล	39
5. ระบบสนับสนุนการตัดสินใจ (Decision Support System)	41
5.1 ความหมายของระบบสนับสนุนการตัดสินใจ	41
5.2 ลักษณะของระบบ	41
5.3 ส่วนประกอบของระบบ	42
5.4 ขั้นตอนของการจัดทำระบบสนับสนุนการตัดสินใจ	42
5.5 คุณลักษณะข้อมูลของระบบสนับสนุนการตัดสินใจ	45
6. เครื่องมือที่ใช้ในการวิเคราะห์ข้อมูล (Tools for Analyse Data)	46
6.1 ข้อกำหนดความต้องการของระบบ	46
6.2 ส่วนประกอบของ Pilot Decision Support Suite	46
6.3 คุณสมบัติของ Pilot Decision Support Suite	48
6.4 การทำงานของ Pilot Decision Support Suite	49
7. การพัฒนาข้อมูลของการจัดเก็บภาษีของภาคเพื่อใช้ในการตัดสินใจ	52
7.1 ขั้นตอนการดำเนินงาน	52
7.2 การวิเคราะห์และออกแบบฐานข้อมูลสำหรับระบบคลังข้อมูล	54
7.3 การออกแบบ Model	54
7.4 ข้อมูลที่นำมาใช้	55
7.5 การเชื่อมโยงฐานข้อมูล	56

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

บทที่	หน้า
7.6 ขั้นตอนการสร้าง Model	58
7.7 Measure Setting	59
7.8 การวิเคราะห์ข้อมูล	60
8. บทสรุป	64
8.1 ข้อเสนอแนะ	64
บรรณานุกรม	66
ประวัติผู้เขียน	67

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่ **หน้า**

3.1 Information Attribute Differences Between Management Levels..... 14

3.2 ลักษณะที่เป็นไปได้ของการขาย..... 17

7.1 แสดงรายการภาษีที่อำเภอ จัดเก็บได้..... 53

7.2 แสดงภาษีที่จังหวัดรวบรวมจากอำเภอ..... 53

7.2 แสดงภาษีที่จังหวัดรวบรวมจากอำเภอ..... 53

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญภาพ

ภาพที่	หน้า
2.1 แสดงโครงสร้างหน่วยงานที่ทำหน้าที่ในการจัดทำรายงานการเก็บภาษีอากร.....	4
2.2 สถาปัตยกรรมแบบ Federated และแบบ Tiered.....	7
3.1 แสดงกรรมวิธีการสร้างสารสนเทศ.....	8
3.2 แสดงรูปแบบของระบบสารสนเทศ.....	10
3.3 แสดงส่วนประกอบของระบบ EIS.....	12
3.4 แสดงให้เห็น Fact และ Dimension.....	15
3.5 ภาพวาดของ Star Schema สำหรับยอดขาย	16
3.6 คุณสมบัติลำดับชั้นที่ผู้ใช้สามารถทำการค้นหาแบบ drill-down และ roll-up ได้..	19
3.7 การกระจายแบบลำดับชั้นในการวิเคราะห์หลายมิติ	20
3.8 Star Schema สำหรับยอดขาย.....	22
3.9 ORDER STAR SCHEMA.....	22
3.10 NORMALIZED DIMENSION TABLE.....	24
3.11 Multiple Fact Tables.....	25
3.12 แนวทางในการสร้างคลังข้อมูล.....	29
4.1 ส่วนประกอบของ Data Warehouse.....	31
4.2 สถาปัตยกรรมแบบ Federated และแบบ Tiered.....	34
6.1 ชุดของโปรแกรม Pilot Decision Support Suite.....	46
6.2.1 Pilot Analysis Server.....	47
6.2.2 Pilot Desktop.....	48
6.2.3 Pilot Designer.....	49
6.4.1 การ Link Configuration.....	50
6.4.2 Model Builder.....	51
7.1 แสดงความสัมพันธ์การจัดเก็บภาษีของภาค.....	52
7.2 การออกแบบคลังข้อมูลผลการจัดเก็บภาษีของภาคแบบ Star Schema.....	55
7.3 รูปแบบโปรแกรม บช.8 ที่ใช้จัดเก็บภาษีของแต่ละจังหวัด.....	56

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญภาพ (ต่อ)

ภาพที่ หน้า

7.4 เข้า Control Panel เลือก ODBC..... 56

7.5 การระบุ ODBC ที่ใช้..... 57

7.6 การ Link Configuration..... 57

7.7 แสดงผลการเชื่อมระหว่างโปรแกรมกับ ODBC 58

7.8 การสร้าง Model โดยใช้ Model Builder (1)..... 58

7.9 การสร้าง Model โดยใช้ Model builder (2)..... 59

7.10 การสร้าง Model โดยใช้ Model builder (3)..... 59

7.11 การระบุสิ่งที่ต้องการการคำนวณ..... 60

7.12 เลือกชื่อ Model ที่ต้องการวิเคราะห์ข้อมูลผลการจัดเก็บ..... 60

7.13 แสดงผลการจัดเก็บภาษีระดับรายเดือนและประเภทภาษี..... 61

7.14 แสดงผลการจัดเก็บแบบแท่ง..... 61

7.15 แสดงผลจัดเก็บตามประเภทภาษี..... 62

7.16 กราฟเส้นแสดงผลจัดเก็บของจังหวัด..... 62

7.17 แสดงผลการจัดเก็บภาษีแบบกราฟรูปวงกลม..... 63

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1
บทนำ

1.1 ความเป็นมา

ในปัจจุบันเทคโนโลยีที่ก้าวหน้าทางด้านสารสนเทศได้ถูกนำมาใช้ในธุรกิจต่างๆทั้งภาครัฐและเอกชนทำให้มีเครื่องมือที่สำคัญในการพัฒนาองค์กรให้บรรลุวัตถุประสงค์ง่ายขึ้นมีความได้เปรียบทางธุรกิจเพิ่มศักยภาพการขยายตัวในระบบเศรษฐกิจโดยรวมของประเทศ จึงอาจกล่าวได้ว่าเป็นเทคนิคที่มีความสำคัญอย่างยิ่งในโลกปัจจุบัน

1.2 หน้าที่ของกรมสรรพากร

กรมสรรพากรมีหน้าที่จัดเก็บภาษีและปฏิบัติตามคำสั่งของกระทรวงการคลัง หน่วยงานที่ทำหน้าที่โดยตรงคือสรรพากรอำเภอที่มีอยู่ในอำเภอต่างๆ ทั่วประเทศ ส่วนสำนักงานสรรพากรจังหวัด สำนักงานสรรพากรภาค และกองต่างๆในกรมสรรพากรมีหน้าที่ด้านการกำหนดนโยบาย วางแผน และบริหาร กรมสรรพากรได้พัฒนาปรับปรุงวิธีการจัดเก็บและการให้บริการแก่ผู้เสียภาษีให้เกิดความสะดวก รวดเร็ว ถูกต้อง เป็นธรรม และเหมาะสมกับสภาพเศรษฐกิจปัจจุบัน เนื่องจากข้อมูลผลการจัดเก็บภาษีและแบบแสดงรายการจำนวนเงินที่ชำระจะถูกบันทึกไว้ในเครื่องออกใบเสร็จรับเงิน (POS) ซึ่งมีอยู่ทุกอำเภอทั่วประเทศ และข้อมูลเหล่านี้ถูกเก็บไว้ในแผ่นดิสเก็ตส่งให้จังหวัดและภาคสัปดาห์ละ 1-2 ครั้งจังหวัดหรือภาคจะรวบรวมข้อมูลนำเสนอต่อผู้บริหารในกรมเดือนละครั้ง ดังนั้นหากมีวิธีการ Decision Support System ที่ดีก็จะทำให้ผู้บริหารในภาค และกรมสามารถทราบข้อมูลจากอำเภอได้รวดเร็วเป็นการช่วยผู้บริหารในการกำหนดนโยบาย วางแผน และบริหารการจัดเก็บภาษีให้เหมาะสมกับสภาพเศรษฐกิจ ดังนั้นการจัดทำคลังข้อมูลเพื่อใช้ในการตัดสินใจสำหรับผู้บริหารจึงเป็นส่วนที่สำคัญส่วนหนึ่งในการจัดเก็บภาษีเพื่อให้เกิดประสิทธิภาพและเป็นธรรมแก่ผู้เสียภาษี และผลตามมาด้วยการพัฒนาประเทศให้เจริญโดยสมบูรณ์ยิ่งขึ้นต่อไป

1.3 วัตถุประสงค์

1.3.1 วิเคราะห์ระบบงานปัจจุบัน

1.3.2 สร้างระบบการรายงานผลการจัดเก็บภาษีใหม่โดยใช้คลังข้อมูล Data Warehouse

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.3.3 เป็นโครงการนำร่อง และนำไปใช้ได้จริงในอนาคต

1.3.4 ช่วยในการบริหารการจัดเก็บภาษีให้มีประสิทธิภาพมากยิ่งขึ้น

1.3.5 เพื่อความรวดเร็วในการตัดสินใจของผู้บริหาร

1.4 ขอบเขตของการศึกษา

1.4.1 ศึกษาระบบการงานผลการจัดเก็บภาษีภายในสำนักงานสรรพากรของกรมสรรพากร

1.4.2 ศึกษาความเป็นไปได้ในการจัดทำคลังข้อมูลผลการจัดเก็บภาษี

1.4.3 ศึกษาการจัดทำคลังข้อมูล (Data Warehouse)

1.4.4 ศึกษาการจัดทำคลังข้อมูลแบบ Star Schema

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1.5.1 ง่ายต่อการติดตามและควบคุมการจัดเก็บภาษีให้เป็นไปตามแผน

1.5.2 เพิ่มประสิทธิภาพในการวางแผน กำหนดนโยบาย และการบริหาร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2
ลักษณะองค์กร

2.1 ความเป็นมา

กรมสรรพากรเป็นหน่วยงานราชการสังกัดกระทรวงการคลัง มีหน้าที่ในการจัดเก็บภาษีอากรประเภทต่างๆ และยังเป็นองค์กรสำคัญของรัฐบาลในการกระจายรายได้ ส่งเสริมความเจริญเติบโตของธุรกิจ รักษาเสถียรภาพทางเศรษฐกิจ ควบคุมการบริโภคของประชาชน ตลอดจนการสนองนโยบายของรัฐบาล

ปัจจุบันกรมสรรพากรมีหน้าที่ดังนี้

1. จัดเก็บภาษีตามประมวลรัษฎากร และกฎหมายอื่น ได้แก่
    - 1.1 ภาษีเงินได้บุคคลธรรมดา
    - 1.2 ภาษีเงินได้นิติบุคคล
    - 1.3 ภาษีมูลค่าเพิ่ม
    - 1.4 ภาษีธุรกิจเฉพาะ
    - 1.5 ภาษีเงินได้ปิโตเลียม
2. ปฏิบัติงานตามมติ กฎระเบียบ ข้อบังคับ เกี่ยวกับการบริหารการจัดเก็บภาษีอากรประเภทต่างที่อยู่ในความรับผิดชอบ
3. ศึกษาค้นคว้าเพื่อเสนอความเห็นต่อกระทรวงการคลังในเรื่องการปรับโครงสร้างภาษีที่อยู่ในความรับผิดชอบ

2.2 การจัดรูปแบบองค์กร

กรมสรรพากรจัดโครงสร้างองค์กร โดยมีรูปแบบเพื่อให้มีการกระจายอำนาจออกไปยังสำนักงานส่วนภูมิภาค มีการแบ่งส่วนราชการ ดังนี้

1. ราชการบริหารส่วนกลาง
    - 1.1 สำนักงานเลขานุการกรม
    - 1.2 กองการเจ้าหน้าที่
    - 1.3 กองคลัง

ตรวจสอบภาษี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.5 กองฝึกอบรม

1.6 ศูนย์เอกสารกลาง

1.7 สำนักกฎหมาย

1.8 สำนักเทคโนโลยีสารสนเทศ

1.9 สำนักแผนภาษี

1.10 สำนักมาตราฐานกรรมวิธีภาษี

1.11 สำนักงานสรรพากรภาค

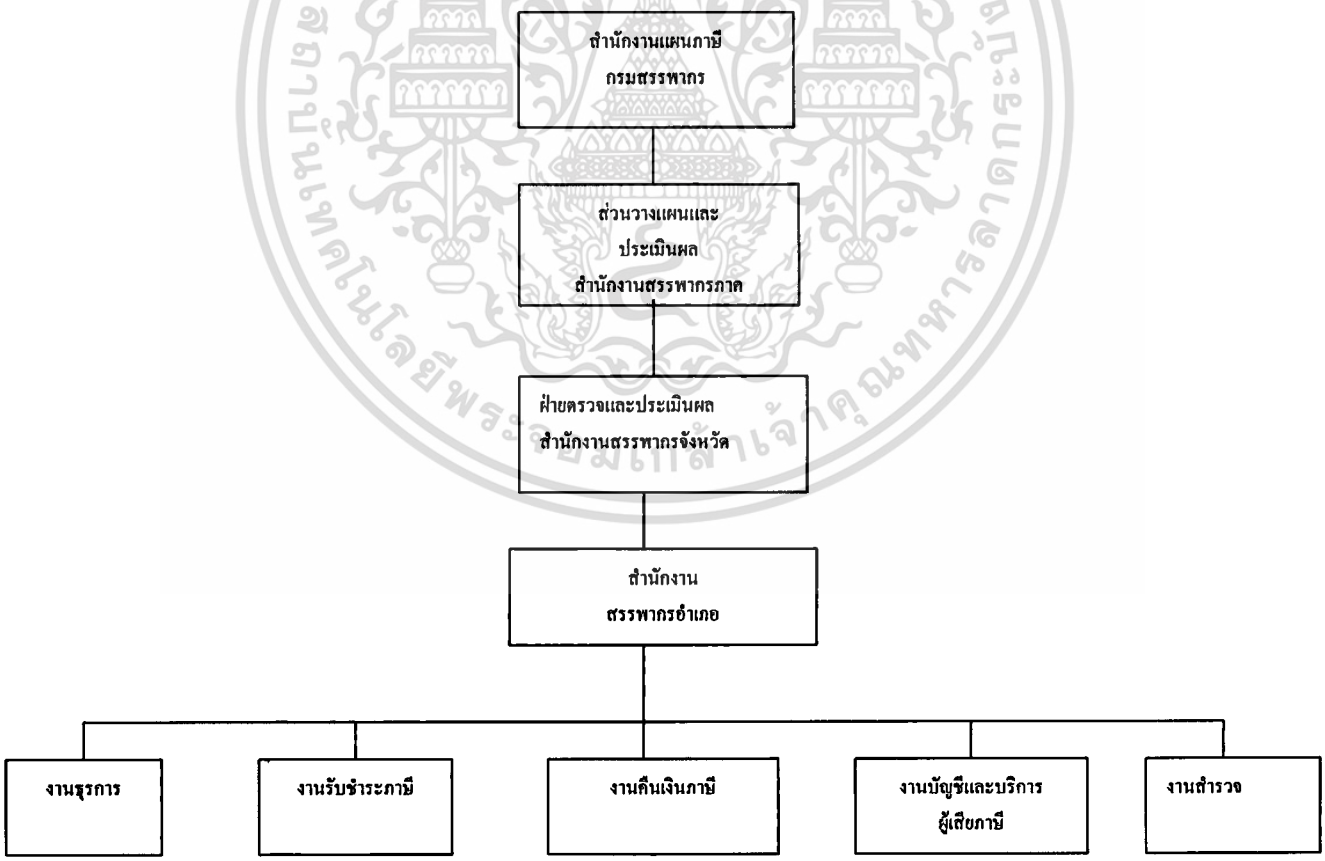
1.12 สำนักงานสรรพากรพื้นที่ 1- 16

2. ราชการบริหารส่วนภูมิภาค

2.1 สำนักงานสรรพากรจังหวัด

2.2 สำนักงานสรรพากรอำเภอ

มีการแบ่งส่วนราชการของกรมสรรพากรตามภาพที่ 2.1



ภาพที่ 2.1 แสดงโครงสร้างหน่วยงานที่ทำหน้าที่ในการจัดทำรายงานการเก็บภาษีอากร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3 หน่วยงานที่มีหน้าที่ในการจัดทำรายงานการเก็บภาษีอากรประจำเดือน (บ.ช.8)

บ.ช.8 หมายถึงรายงานการเก็บภาษีอากรประจำเดือน ซึ่งจะมีรายละเอียดดังนี้

- แต่ละอำเภอจัดเก็บภาษีได้เดือนละเท่าไร
- แต่ละอำเภอจัดเก็บภาษีประเภทใด ได้ในเดือนนั้น และจัดเก็บได้กี่รายเช่น ภาษีเงินได้นิติบุคคล ภาษีมูลค่าเพิ่ม ฯลฯ
- แต่ละอำเภอได้รับเงินนำส่งจากหน่วยราชการอื่น เช่น กรมศุลกากร กรมสรรพสามิต ฯลฯ เดือนละเท่าไร ดูรายละเอียดในภาคผนวก ฯลฯ

ที่มาของ บ.ช.8

บ.ช.8 เป็นรายงานการเก็บภาษีของสำนักงานสรรพากรอำเภอ ซึ่งแต่ละอำเภอจะนำข้อมูลมาจากงบหน้าประเภทภาษีอากรประจำวัน (บ.ช.5) ดูรายละเอียดในภาคผนวกซึ่งมีรายละเอียดเหมือนกับ บ.ช.8 แต่จะเป็นการนำมารวบรวมแล้วนำเสนอเป็นบ.ช.8 บ.ช.5 ปัจจุบันได้มาจากการพิมพ์ จากเครื่องออกใบเสร็จรับเงินชนิด POINT OF SALE (POS) ซึ่งเป็นเครื่องออกใบเสร็จรับเงินทั่วไปแต่มีหน่วยความจำและหน่วยความจุข้อมูล เช่นเดียวกับเครื่องคอมพิวเตอร์ ฉะนั้นจึงมีคุณสมบัติเป็นทั้งเครื่องออกใบเสร็จรับเงินและเครื่องคอมพิวเตอร์รวมกัน

เครื่องออกใบเสร็จรับเงิน POS คือเครื่องออกใบเสร็จรับเงินตามที่กรมสรรพากรกำหนดให้ใช้รับชำระภาษีอากรและเงินรายรับ แล้วให้เลขคุมเอกสาร (DLN) แบบแสดงรายการภาษีอากรทุกประเภทภาษีมีหน่วยบันทึกข้อมูลและติดต่อสื่อสารกับเครื่องคอมพิวเตอร์ที่กรมสรรพากรใช้อยู่ ซึ่งอยู่ในความดูแลรับผิดชอบของสำนักงานสรรพากรอำเภอหรือหัวหน้างานรับภาษี กองคลัง กรมสรรพากร

ส่วนประกอบของเครื่องออกใบเสร็จรับเงิน POS มีดังนี้

1. Central Processing Unit (CPU) คือหน่วยประมวลผลกลาง ภายในจะประกอบไปด้วยอุปกรณ์ต่างๆเช่น อุปกรณ์เก็บข้อมูล (Harddisk), อุปกรณ์ประมวลผล (Micro Processor) อุปกรณ์เชื่อมต่อต่างๆ (I/O Interface card) ทั้งหมดนี้ประกอบเข้าด้วยกันเพื่อร่วมกันทำหน้าที่ในการจัดเก็บและประมวลผลทั้งหมด

2. Keyboard Module คือแป้นพิมพ์จะประกอบไปด้วยปุ่มฟังก์ชันอำนวยความสะดวกต่างๆ

3. Keyboard Module คือแป้นพิมพ์หรืออุปกรณ์เสริมอื่นๆ เช่นถาดใส่เหรียญแต่ในระบบงานของกรมสรรพากรจะติดตั้งแป้นพิมพ์ชุดที่สอง

4. Display Module คือจอภาพแสดงผล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5. Display Module จอภาพแสดงผลแบบ 2 แถวซึ่งใช้ในการแสดงเลขประจำตัวผู้เสียภาษีอากรและจำนวนเงิน

6. Printer Module เครื่องพิมพ์ซึ่งใช้ในการพิมพ์ใบเสร็จรับเงินและสำเนาเทป

7. Cash Drawer เป็นอุปกรณ์เสริมไม่ได้ติดตั้งในระบบงานของสรรพากร

8. Magnetic Strip Reader เป็นอุปกรณ์เสริมไม่ได้ติดตั้งในระบบ

9. External Printer เป็นอุปกรณ์เสริม

จำนวนของเครื่อง POS ปัจจุบันทุกสำนักงานสรรพากรอำเภอทั่วประเทศจะมีเครื่อง POS อย่างน้อยหนึ่งเครื่อง แต่ในเขตกรุงเทพฯ จะมีเฉลี่ย สำนักงานสรรพากรอำเภอละ 3 เครื่อง

หน้าที่ของเครื่อง POS คือ ทำหน้าที่ออกใบเสร็จรับเงิน ณ. จุดรับชำระภาษีของอำเภอต่างๆ ทั่วประเทศ ประโยชน์ของเครื่อง POS คือ

1. ออกใบเสร็จรับเงิน
2. เก็บข้อมูลรายการรับชำระภาษีต่างๆ ทุกรายในแต่ละวัน
3. พิมพ์รายงานงบหน้าประเภทภาษีอากรประจำวัน (บ.ช.5)
4. พิมพ์รายงานการเก็บภาษีอากรประจำเดือน (บ.ช.8)
5. พิมพ์ใบนำส่งแบบแสดงรายการ

2.4 โครงสร้างของหน่วยงานที่จัดทำรายการเก็บภาษีอากร (บ.ช.8)

บ.ช.8 เป็นรายงานการเก็บภาษีอากร ใช้เป็นแบบรายงานผลการจัดเก็บภาษีอากรประจำเดือนของสำนักงานสรรพากรอำเภอรายงานสรรพากรจังหวัดและสำนักงานสรรพากรจังหวัดรายงานสำนักงานสรรพากรภาค รวมทั้งสำนักแผนภาษี เพื่อให้ทราบผลการจัดเก็บภาษีอากรและรายได้อื่นเป็นรายเดือน สำนักงานสรรพากรเขต (อำเภอในเขตกรุงเทพมหานคร) และสำนักงานสรรพากรอำเภอมีหน้าที่ในการจัดเก็บภาษีอากรตามประมวลรัษฎากรและกฎหมายอื่นๆ และจัดทำรายงานการเก็บภาษีอากร(บ.ช.8) เพื่อนำส่งตามลำดับดังภาพที่ 2.2

1. สำนักงานภาค 1-3

1.1 สำนักงานสรรพากรเขต (อำเภอในเขตกรุงเทพมหานคร) รายงานการเก็บภาษีอากรให้สำนักงานสรรพากรพื้นที่

1.2 สำนักงานสรรพากรพื้นที่ รวมผลจัดเก็บที่สำนักงานสรรพากรเขตรายงานการเก็บภาษีอากรมาเป็นผลการจัดเก็บของสำนักงานสรรพากรพื้นที่ให้สำนักงานสรรพากรภาค

1.3 สำนักงานสรรพากรภาค รวมผลจัดเก็บที่สำนักงานสรรพากรพื้นที่รายงานการเก็บภาษีอากรมาเป็นผลการจัดเก็บของทั้งภาคแล้วรายงานให้สำนักแผนภาษี ทราบ

2. สำนักงานสรรพากรภาค 4 – 9

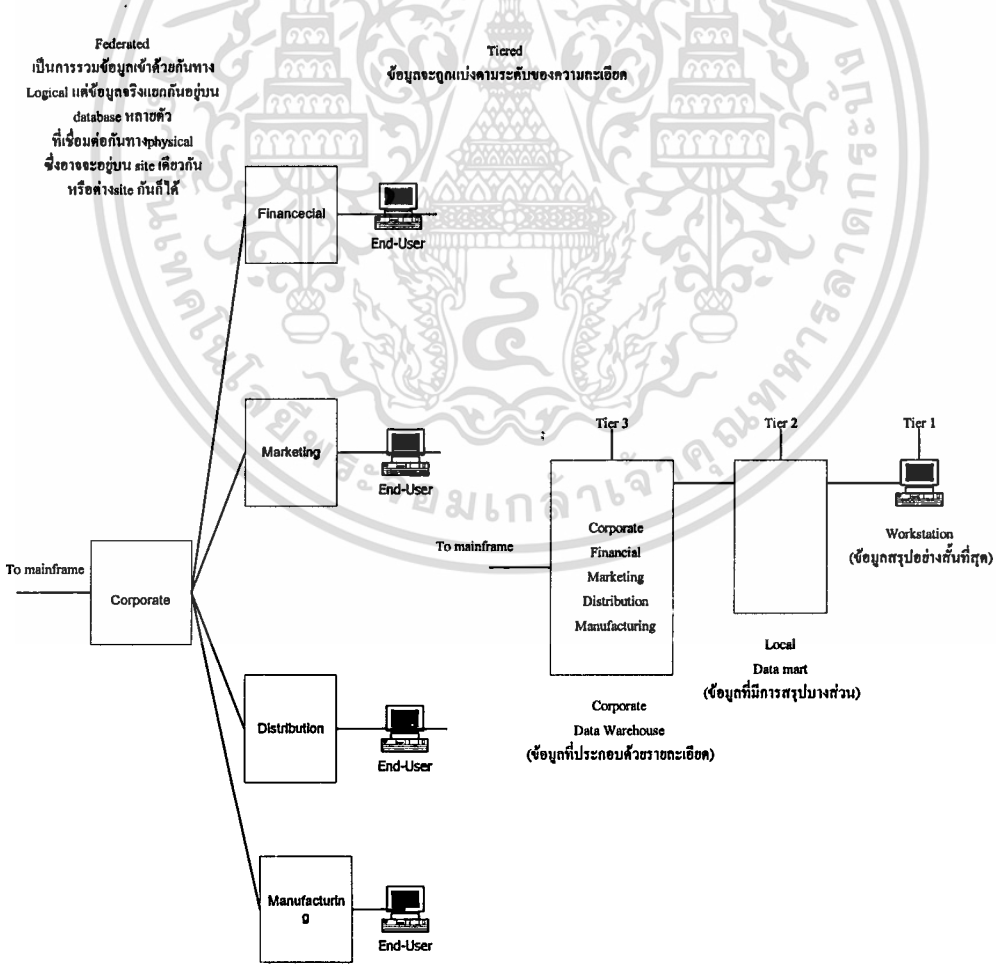
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.1 สำนักงานสรรพากรอำเภอรายงานการเก็บภาษีอากรให้สำนักงานสรรพากรจังหวัด

2.2 สำนักงานสรรพากรจังหวัด รวมผลการจัดเก็บที่สำนักงานสรรพากรอำเภอรายงานการเก็บภาษีอากรเป็นผลการจัดเก็บของทั้งจังหวัดให้สำนักงานสรรพากรภาค

2.3 สำนักงานสรรพากรภาค รวมผลจัดเก็บที่สำนักงานสรรพากรอำเภอรายงานการเก็บภาษีอากรเป็นผลการจัดเก็บของทั้งภาคแล้วรายงานสำนักแผนภาษีทราบ

สำนักแผนภาษี กรมสรรพากรจะได้รับรายงาน บ.ช.8 จากทั่วประเทศประมาณวันที่ 10 ของเดือนถัดไปแล้วนำข้อมูลจาก บ.ช.8 มารวบรวมนำเสนอผู้บริหารระดับสูง (อธิบดี) ด้วยสถิติการจัดเก็บภาษีสรรพากรเบื้องต้นทุกประเภททั่วประเทศ ภายในวันที่ 15 ของเดือนถัดไป



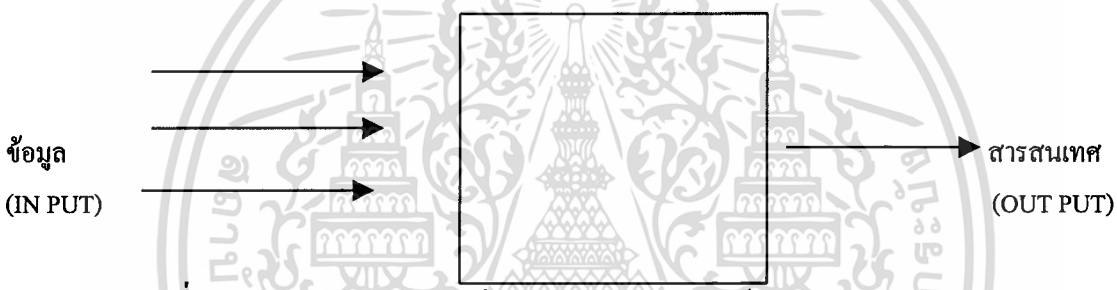
ภาพที่ 2.2 สถาปัตยกรรมแบบ F7ederated และแบบ Tiered

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3
ทฤษฎีที่เกี่ยวข้อง

3.1 ความหมายของสารสนเทศ

สารสนเทศ (Information) คือข้อมูลและข้อเท็จจริงที่ถูกนำมาประมวลขึ้นเพื่อให้สะท้อนหรือแสดงภาพของเหตุการณ์ที่เกิดขึ้น สารสนเทศได้มาจากข้อมูล ข้อเท็จจริงลักษณะต่างๆ ที่สนใจซึ่งสารสนเทศอาจช่วยให้การดำเนินงานมีประสิทธิภาพมากขึ้น ซึ่งมีกรรมวิธีการสร้างสารสนเทศดังแสดงในภาพที่ 3.1



ภาพที่ 3.1 แสดงกรรมวิธีการสร้างสารสนเทศกรรมวิธีข้อมูล(Data Processing)

ขั้นตอนการสร้างสารสนเทศ ประกอบด้วย

1. ข้อมูลนำเข้า Input คือข้อมูลหรือข้อเท็จจริงลักษณะต่างๆ ที่สนใจจำนวนมาก ซึ่งป้อนเข้าสู่กรรมวิธีข้อมูล

2. กรรมวิธีข้อมูล Data Processing คือ กระบวนการประมวลผลข้อมูลโดยการแปลงข้อมูลสารสนเทศ ซึ่งมีขั้นตอนดังนี้

2.1 การเก็บข้อมูล

2.2 การแปลงข้อมูลเข้าสู่ระบบคอมพิวเตอร์

2.3 การบันทึกข้อมูลลงบนสื่อข้อมูล

2.4 การประมวลผลและการคำนวณ

2.5 การแสดงผลลัพธ์

2.6 การอัดสำเนาผลลัพธ์และรายงาน

2.7 การสื่อสาร

2.8 การสำรองข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.9 การบีบอัดข้อมูล

2.10 การเข้ารหัสข้อมูล

3. สารสนเทศที่ได้คือข้อมูลสรุปที่ได้จากการประมวลผลโดยผ่านกรรมวิธีข้อมูล

3.2 ความหมายของสารสนเทศ

ระบบสารสนเทศ (Information System) คือ ระบบที่นำระบบคอมพิวเตอร์มาใช้ โดยมีวัตถุประสงค์ที่จะแปลงข้อมูลไปเป็นสารสนเทศตามความต้องการของผู้บริหาร โดยการจัดทำรายงานสารสนเทศต่อผู้บริหารและผู้ปฏิบัติงาน เพื่อให้ผู้บริหารและผู้ปฏิบัติงานนำสารสนเทศไปใช้ในการควบคุม และตัดสินใจให้การปฏิบัติงานขององค์กรดำเนินงานได้อย่างมีประสิทธิภาพ

รูปแบบการรายงานของสารสนเทศ แบ่งเป็นประเภทได้ดังนี้

1. รายงานสรุปแยกกลุ่มตามความสนใจของผู้บริหาร
2. รายงานยกเว้น เป็นรายงานที่แสดงให้เห็นถึงปัญหาที่เกิดขึ้น
3. รายงานแนวโน้ม เป็นรายงานที่แสดงให้เห็นถึงทิศทางของเหตุการณ์ที่เกิดขึ้น
4. รายงานคาดหมายพยากรณ์ เป็นรายงานที่แสดงให้เห็นถึงเหตุการณ์ที่อาจเกิดขึ้น

3.3 ลักษณะรายงานสารสนเทศที่สมบูรณ์ ควรมีลักษณะดังนี้

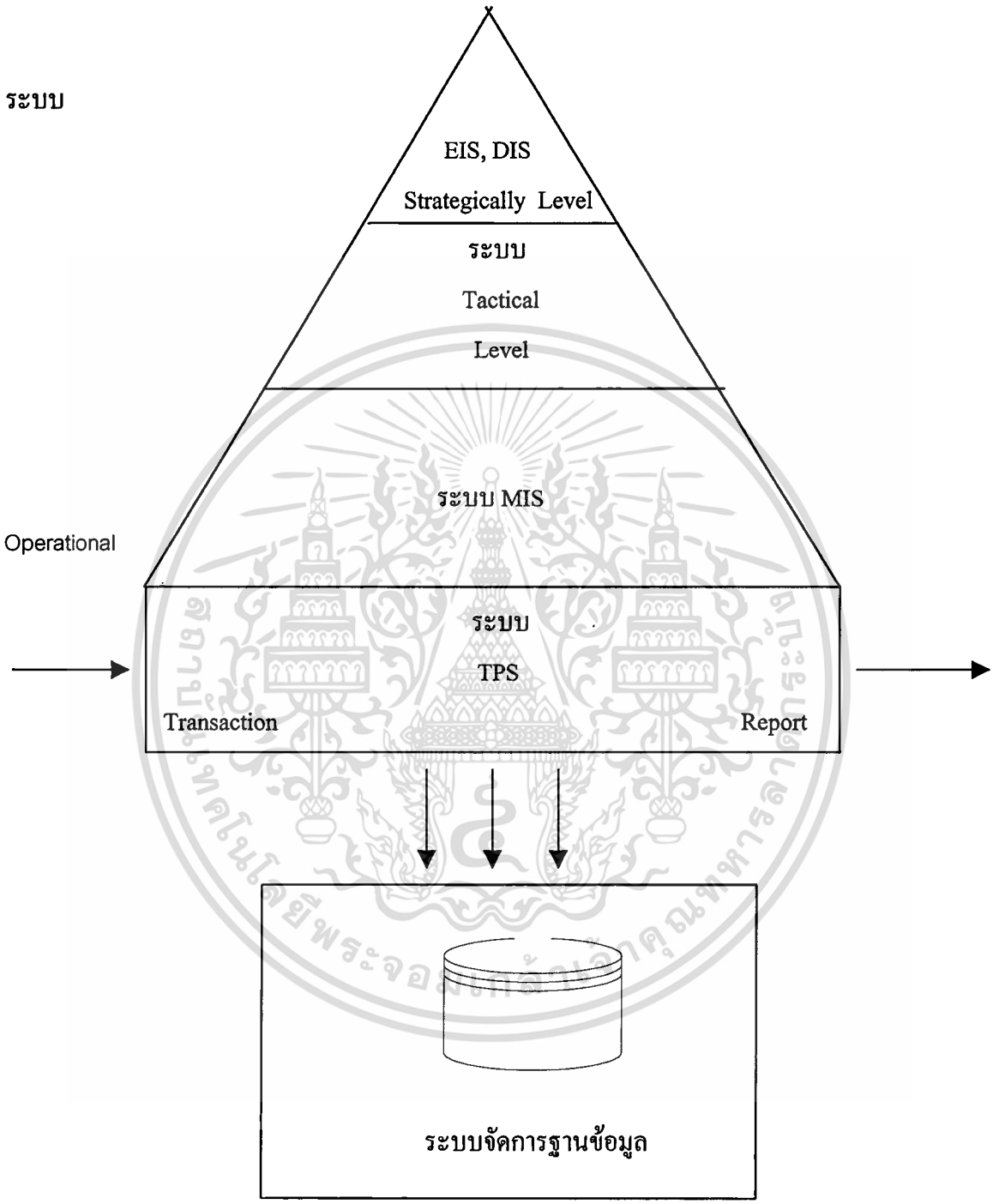
1. ข้อมูลที่นำมาประมวลผลให้เกิดสารสนเทศนั้น ต้องเป็นข้อมูลปัจจุบัน
2. รายงานสารสนเทศ ต้องมีความพร้อมต่อการนำเสนอต่อผู้บริหารตลอดเวลา
3. รายงานสารสนเทศ สามารถแสดงผลได้หลายรูปแบบ

รายงานที่นำเสนอต่อผู้บริหารนั้น ต้องนำเสนอในรูปแบบที่ง่ายต่อการเข้าใจ และสะดวกในการนำสารสนเทศไปใช้งาน อาจเป็นลักษณะของข้อมูลตารางเปรียบเทียบ และกราฟลักษณะต่างๆ รายงานสารสนเทศอาจเป็นลักษณะของ ออนไลน์หรือรายงานก็ได้

3.4 รูปแบบของสารสนเทศ

สารสนเทศมีหลายรูปแบบ แต่ละรูปแบบเป็นการนำเสนอสารสนเทศต่อผู้บริหารเพื่อช่วยในการตัดสินใจของผู้บริหารเป็นส่วนใหญ่ ซึ่งรูปแบบของสารสนเทศประกอบด้วยระบบ TPS ระบบ MIS ระบบ DSS และระบบ EIS ดังรายละเอียดในภาพที่ 3.2 รูปแบบต่างๆ ของระบบสารสนเทศ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 3.2 แสดงรูปแบบของระบบสารสนเทศ

1. ระบบ TPS (Transaction Processing)คือระบบที่รับ Transaction (รายการที่แก้ไขเปลี่ยนแปลงรายการข้อมูลในฐานข้อมูล) มาประมวลผลเพื่อให้เกิด Transaction Documents หรือ Report เป็นระบบพื้นฐานสำคัญของระบบสารสนเทศอื่นๆ และเป็นระบบที่ผลักดันให้การดำเนินธุรกิจก้าวหน้าไปได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คุณลักษณะของระบบ TPS

1.1 ต้องสามารถรับ Transaction ได้อย่างรวดเร็ว

1.2 สามารถดำเนินการกับ Transaction ได้อย่างรวดเร็ว

1.3 สามารถบันทึก Transaction ได้ และมีระบบตรวจสอบความถูกต้องของ Transaction

1.4 เป็นระบบที่ทำงานเป็นประจำและต่อเนื่อง

1.5 มักเป็นระบบที่ประมวลผลในระบบ ONLINE เพราะมีการเพิ่มรายการข้อมูลและดึงข้อมูลจากฐานข้อมูลมาประมวลผล

1. ระบบ MIS (Management Information System) ระบบสารสนเทศสำหรับงานจัดการทั่วไปเป็นระบบที่นำเอาข้อมูล Transaction จากฐานข้อมูลมาประมวลผลให้เกิดสารสนเทศในรูปแบบของรายงานโดยสรุปที่เข้าใจง่าย เพื่อช่วยในการตัดสินใจของผู้บริหารระดับต้น (Operation Level) นอกจากนี้ระบบ MIS จะช่วยให้ผู้บริหารมองเห็นปัญหาขององค์กร และโอกาสในการทำงานที่ทำให้บรรลุวัตถุประสงค์ขององค์กร สารสนเทศที่ได้จากระบบMIS มักใช้ในการวางแผนการปฏิบัติงานในองค์กรระยะสั้น อาจเป็นแผนประจำวัน แผนประจำสัปดาห์ หรือแผนประจำเดือน

2. ระบบ DSS (Decision Support System) คือระบบที่ช่วยในการสนับสนุนการตัดสินใจของผู้บริหาร อยู่ในรูปแบบของแบบจำลองการตัดสินใจในสภาวะต่างๆ เพื่อให้ผู้บริหารสามารถทดสอบได้ โดยการแสดงผลลัพธ์ ข้อดี ข้อเสียของทางเลือกในแบบจำลอง

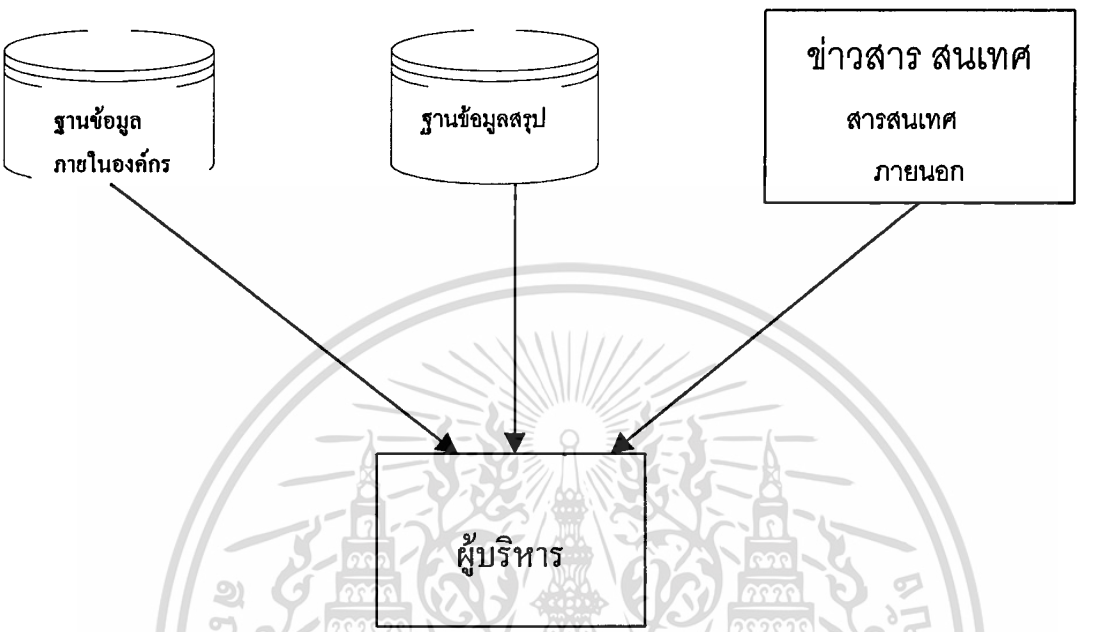
3.5 ผู้บริหารระดับสูง

ความหมายของคำว่า ผู้บริหารระดับสูงในแต่ละองค์กร จะมีความแตกต่างกันทั้งนี้ขึ้นอยู่กับคำจำกัดความของแต่ละองค์กร เราจะพบว่ามีคุณสมบัติเพียงเล็กน้อยที่สามารถช่วยให้เราเข้าใจว่าใครคือผู้บริหารระดับสูง หน้าที่ของเขาคืออะไร และพวกเขามีความแตกต่างอย่างไรกับผู้บริหารระดับผู้จัดการ คุณสมบัติเหล่านี้ประกอบด้วย

1. ผู้บริหารระดับสูงจะมีขอบเขตความรับผิดชอบครอบคลุมทั้งองค์กร ซึ่งจะแตกต่างจากผู้บริหารระดับผู้จัดการที่ขอบเขตความรับผิดชอบอยู่ในขอบเขตที่แคบกว่า โดยจะเป็นแผนกๆหนึ่งในองค์กร

2.บริหารระดับสูง จะมีหน้าที่รับผิดชอบในหลายๆ หน้าที่และในบางองค์กร ธุรกิจจะมอบหมายตำแหน่งผู้บริหารระดับสูงให้กับผู้บริหารที่ดูแลกิจกรรมหลักของธุรกิจ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 3.3 แสดงส่วนประกอบของระบบ EIS

3. บริหารระดับสูง จะเป็นผู้วางแผนกลยุทธ์ขององค์กร ซึ่งจะเป็นการวางแนวทางการเติบโตของธุรกิจที่มีระยะเวลามากกว่า 1 ปีขึ้นไป หรืออาจพูดได้ว่าผู้บริหารระดับสูงจะเป็นผู้มองไปสู่อนาคตมากกว่าผู้บริหารในระดับผู้จัดการที่เน้นการวางแผนในระดับปฏิบัติการ (Operation Plan)

4. บริหารระดับสูง จะเป็นผู้กำหนดนโยบายขององค์กร และเป็นผู้แถลงนโยบายต่อสาธารณะชน รวมทั้งรับผิดชอบต่อความสำเร็จหรือล้มเหลวของการปฏิบัติตามนโยบายนั้น

5. ผู้บริหารระดับสูงเป็นบุคคลที่มีความสำคัญต่อองค์กร การดำเนินการหรือการตัดสินใจของผู้บริหารระดับสูงมีผลกระทบต่อความสำเร็จหรือความล้มเหลวของการปฏิบัติตามนโยบายนั้น ผู้บริหารระดับสูง จะเป็นผู้ที่ติดต่อกับคณะกรรมการบริษัท เพื่อรายงานผลการดำเนินการของบริษัทจากคณะกรรมการของบริษัทก่อนนำออกไปใช้ปฏิบัติต่อไป

3.5.1 ลักษณะงานของผู้บริหารระดับสูง

หน้าที่ของผู้บริหารทุกคนจะทำหน้าที่บริหารกิจการ จัดโครงสร้างขององค์กรจัดสรรบุคลากรในตำแหน่งที่สำคัญ กำหนดแผนงาน ดำเนินงานตามแผน ควบคุมการปฏิบัติงานตามแผน ในขณะที่ผู้ที่อยู่ในตำแหน่งผู้บริหารระดับสูงจะใช้เวลาส่วนใหญ่กับการวางแผนทั้งแผนการเจริญเติบโตของ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิจการ แผนกลยุทธ์เพื่อให้สามารถชนะคู่แข่งขันในตลาดได้ การวางแผนกลยุทธ์จะเป็นการตัดสินใจใช้ทรัพยากรต่างๆ ขององค์กรประสานการทำงานกันเพื่อให้บรรลุผลตามแผนการเติบโตของบริษัท ผู้บริหารระดับสูง จะเป็นผู้สร้างและดูแลรักษาวัฒนธรรมในองค์กรวัฒนธรรมในองค์กรเปรียบเสมือนข้อตกลงร่วมกันของบุคลากรในองค์กร เพื่อกำหนดลักษณะการทำงานร่วมกัน การสร้างจิตสำนึกในความเป็นเจ้าขององค์กรเพื่อให้เกิดความรู้สึกมีส่วนร่วมในความสำเร็จเรือล้มเหลวขององค์กรผู้บริหารระดับสูงมีความสำคัญจนอาจเทียบได้กับเป็นผู้ถือเงินเดิมพันขององค์กร เพื่อแข่งขันกับคู่ต่อสู้ทางธุรกิจอื่นๆ และจะเป็นผู้รักษาความสมดุลย์ของทรัพยากรที่สำคัญของบริษัทอันประกอบด้วยเงินทุน

1. ตลาดผลิตภัณฑ์
2. ความมั่นคงขององค์กร

โดยจะเป็นผู้วางแผนจัดสรรทรัพยากรเหล่านี้ เพื่อให้บรรลุความสำเร็จที่กำหนดในแผนกลยุทธ์ เพราะธรรมชาติของธุรกิจทั่วไปคือ การมีทรัพยากรอย่างจำกัด เมื่อเทียบกับความต้องการใช้งาน ผู้บริหารระดับสูงจะเป็นผู้จัดกระบวนการทำงานขององค์กรโดยการวางแผน การจัดสรรงบประมาณ การควบคุม การสื่อสารกับบุคลากรทั้งในและนอกองค์กร

3.5.2 สารสนเทศผู้บริหาร : อะไรคือสิ่งที่ผู้บริหารต้องการ

ผู้บริหารใช้ข้อมูลสารสนเทศเพื่อวัตถุประสงค์ในหลายๆ ประการ เช่น

- เพื่อติดตามในส่วนที่กำลังสนใจ
- เพื่อค้นหาสาเหตุของปัญหา
- เพื่อกำหนดทางเลือกและปฏิบัติตามทางเลือก (Masom and Mitroff, 1981 ;Simon, 1965) ข้อมูลสารสนเทศที่ผู้บริหารต้องการ โดยแยกตามระดับของผู้บริหารออกเป็น 2 ระดับ ดังแสดงในตารางที่ 3.1

เราสามารถจำแนกระบบสารสนเทศผู้บริหารออกได้เป็น 2 ประเภทใหญ่ๆ คือ

1. การใช้ระบบสารสนเทศในลักษณะ Direct User เช่นการใช้งานจอเทอร์มินอลหรือคอมพิวเตอร์ส่วนบุคคล เพื่อการรับ-ส่ง Electronic-Mail การใช้งานโปรแกรมช่วยการคำนวณเพื่อการสนับสนุนการตัดสินใจ

2. การใช้ระบบสารสนเทศในลักษณะผู้ใช้สารสนเทศ เพื่อการบริหารขององค์กร เช่นการใช้สารสนเทศเพื่อกำหนดแนวทางในการดำเนินงาน เป็นต้นจากลักษณะการใช้สารสนเทศผู้บริหารจะเห็นได้ว่า ความหมายของคำว่า Computer Literacy มีความหมายแคบเกินไป เราจึงต้องใช้ความหมายของ Information Literacy เป็นคุณ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.6 ฐานข้อมูลแบบ Star Schema

Star Schema เป็นเทคนิคของข้อมูลที่รวบรวมแนวทางการตัดสินใจต่างๆ ของข้อมูลใน Database เหตุที่ต้องการพัฒนา Star Schema ก็เพื่อใช้เป็นเทคนิคที่มีความสัมพันธ์กับ E_R และ Normalization ที่ยังไม่เพียงพอต่อการให้บริการจากความต้องการในอนาคต

ตารางที่ 3.1 Information Attribute Differences Between Management Levels

Attribute	Lower – Level Managers	Higher - Level Mangers
Accuracy	High	Low
Timeliness	High	Low
Scope	Narrow	Broad
Time Horizon	Past,present	Future
Relevance	High	Low
Level of detail	High	Low
Level of summarization	Low	High
Orientation	Internal	External
Source	Written	Verbal
Quantifiability	High	Low

สมบัติของผู้บริหารที่ใช้ข้อมูลสารสนเทศไปใช้ได้อย่างมีประสิทธิภาพ

โดยเหตุที่สารสนเทศมีความสำคัญต่อการตัดสินใจของผู้บริหารและสารสนเทศทั้งหลายล้วนได้มาจากข้อมูล อันเป็นข้อเท็จจริงเกี่ยวกับสถานภาพ และเหตุการณ์ต่างๆ ดังนั้นจึงมีความจำเป็นอย่างยิ่งที่ทั้งข้อมูลและสารสนเทศต้องมีคุณสมบัติอย่างน้อยคือ

1. ความเที่ยงตรง (Accuracy) กล่าวคือ สารสนเทศจะต้องไม่ทำให้เกิดความเข้าใจผิด (Mistake) และมีข้อผิดพลาด (Error) สารสนเทศนั้นจะต้องชัดเจน (Clear) และเที่ยงตรง

2. การทันต่อการใช้งาน (Timelines) สารสนเทศที่ผู้บริหารต้องการในช่วงเวลาที่กำหนดต้องได้รับทันทีและสามารถนำไปใช้ประโยชน์ได้

3. ตรงต่อความต้องการ (Relevancy) คือสารสนเทศนั้นจะต้องสามารถตอบคำถามที่ผู้รับเจาะจงเช่น อะไร ทำไม ที่ไหน เมื่อไรและอย่างไร ได้ตรงประเด็นหรือไม่

Star Schema เป็นแบบที่นำไปใช้ได้ง่ายกับการวิเคราะห์ ข้อมูลที่มีหลายมิติ และยังคงมีการ

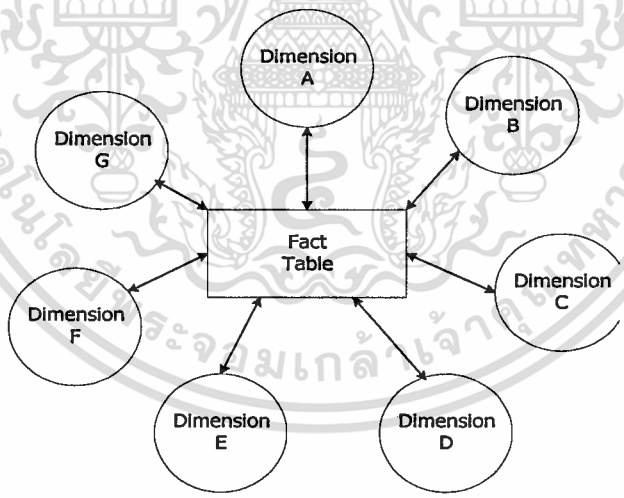
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สร้างฐานข้อมูลที่มีโครงสร้างแบบสัมพันธ์กัน Relational พื้นฐานของ Star Schema มีองค์ประกอบ 4 อย่างคือ

1. Fact ข้อเท็จจริง
2. Dimension มิติ
3. Attribute คุณสมบัติเฉพาะ
4. Attribute hierarchies คุณสมบัติลำดับชั้น

1. Fact ข้อเท็จจริง

Fact เป็นค่าตัวเลขที่ได้จากการวัด แสดงถึงความคาดหวังที่ได้จากกิจกรรมใดๆ ตัวอย่างเช่นตัวเลขที่ได้จากการขายผลิตภัณฑ์ หรือการบริการสินค้า Fact เป็นการวิเคราะห์ข้อมูลทางสถิติเช่น จำนวนสินค้า, ต้นทุน, ราคา และรายรับ Fact เป็นการเก็บ Fact table ซึ่งจะรวมอยู่ใน Star Schema ดังภาพที่ 3.4



ภาพที่ 3.4 แสดงให้เห็น Fact และ Dimension

Fact table ประกอบด้วยการเชื่อมต่อถึงกันของ Dimension Fact อาจได้จากการคำนวณเป็นช่วงเวลา บางครั้งเรียกว่า Metrics ตาราง Fact จะปรับปรุงทุกๆวัน, สัปดาห์, เดือน ฯลฯ พร้อมๆ กับข้อมูลในฐานข้อมูล

2. Dimension มิติ

เป็นการทำให้มีคุณสมบัติที่สามารถเพิ่มเข้าไปใน Fact ลักษณะมิตินี้มีความน่าสนใจมากเพราะข้อมูล DSS ส่วนใหญ่มักจะมีความสัมพันธ์กับข้อมูลอื่นๆ ตัวอย่างเช่นการเปรียบเทียบยอดขายจากเขต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ต่างๆ และจากช่วงเวลาหนึ่งถึงช่วงเวลาหนึ่ง ปัญหาของDSS อาจเป็นการเปรียบเทียบยอดขายของสินค้า x เขตหนึ่งในหนึ่งไตรมาสของปี 1990 จนถึงปี1997 ในตัวอย่างนี้ยอดขายสินค้า แหล่งที่ขาย และมิติเวลา มิติต่างๆ จะถูกเก็บตามปกติใน dimension table โดยมี ชนิดของสินค้า สถานที่ และมิติเวลา

3. Attribute คุณสมบัติเฉพาะ

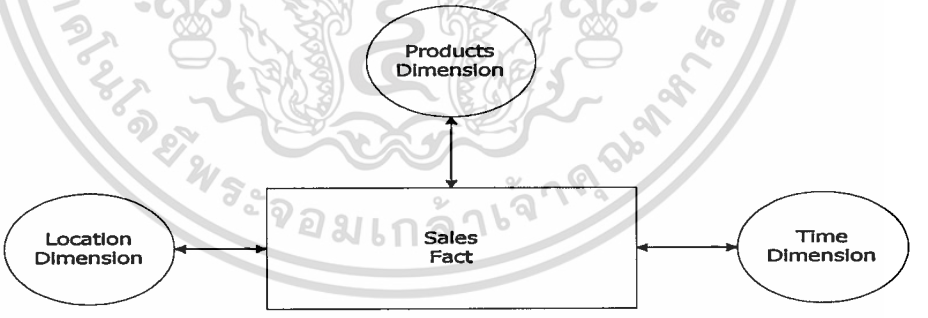
ตารางมิติประกอบด้วยคุณสมบัติต่างๆ ที่มักจะใช้ในการค้นหา, กลั่นกรองหรือแยกแยะ Fact มิติจะมีคำอธิบายลักษณะเฉพาะเกี่ยวกับ Fact รวมถึงคุณสมบัติด้วย ดังนั้นนักออกแบบคลังข้อมูลจึงต้องกำหนดคุณสมบัติทางธุรกิจให้ชัดเจน เพื่อให้นักวิเคราะห์ ค้นหาได้ง่ายขึ้น เช่น กลุ่มข้อมูลข่าวสารหรือรายละเอียดของมิติ ใช้การขายเป็นตัวอย่าง เราสามารถแยกแยะคุณสมบัติแต่ละมิติได้

3.1 มิติของผลิตภัณฑ์ : ID ของผลิตภัณฑ์, รายละเอียดผลิตภัณฑ์, ชนิดของผลิตภัณฑ์, โรงงานที่ผลิต

3.2 มิติของสถานที่ : บริเวณ, รัฐ, เมืองและจำนวนที่จัดเก็บ

3.3 มิติของเวลา : ปี, ไตรมาส, เดือน, สัปดาห์และวัน

ผลิตภัณฑ์, สถานที่ และมิติเวลาเหล่านี้ แสดงให้เห็น Fact ของ Sale facts ดังภาพที่ 3.5การวิเคราะห์ข้อมูลในปัจจุบันมีความสัมพันธ์กับรูปร่างของผลิตภัณฑ์ ในบริเวณที่กำหนด และตามเวลาที่กำหนด



ภาพที่ 3.5 เป็นภาพวาดของ Star Schema สำหรับยอดขาย

Star Schema เข้าถึง Fact และ Dimension ทำให้มีรูปแบบข้อมูลตามต้องการ และไม่ต้องเก็บข้อมูลที่ไม่จำเป็นไว้มากๆ เช่น จำนวนข้อมูล, จำนวนที่สั่งซื้อ, สถานภาพ ฯลฯ เหล่านี้เป็นเพียงพื้นฐานการใช้ข้อมูลโดยทั่วไป ผลที่ได้รับเป็นมิติต่างๆ มากมายจากการศึกษา Fact

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตาราง 3.2 ลักษณะที่เป็นไปได้ของการขาย

ชื่อมิติ	ความหมาย	ลักษณะที่เป็นไปได้
สถานที่	บริเวณที่จัดไว้เป็นสถานที่เช่น เขต, เหนือ, ใต้ ฯลฯ	ขอบเขต, สถานที่จัดเก็บ, รัฐ ฯลฯ
ผลผลิต	สิ่งต่างๆ ที่มีไว้จำหน่าย เช่น ผลิตภัณฑ์ด้านดูแลเส้นผม,แชมพู, ขวด	ชนิดของผลิตภัณฑ์, หมายเลขผลิตภัณฑ์, ยี่ห้อสินค้า, การห่อหุ้ม, การนำเสนอ, สี, ขนาด
เวลา	ช่วงของเวลาในการขาย เช่น ปี 1996 เดือน พฤษภาคม วันที่ 5/29/1996 เวลา 4.46 pm	ปี, ไตรมาส, เดือน, สัปดาห์,วัน ช่วงเวลาใน 1 วัน

แนวคิดด้านการขาย ตัวอย่างการขายแบบหลายมิติ มีรูปแบบนำเสนอรูปลูกบาศก์ 3 มิติ แน่นอนเราไม่สามารถชี้ให้เห็นถึงตัวเลขที่จำกัดของมิติที่สัมพันธ์กับตาราง Fact ไม่มีการจำกัดจำนวนตัวเลขที่เราได้คำนวณ อย่างไรก็ตาม การใช้รูปแบบ 3 มิติจะทำให้เรามองเห็นปัญหาง่ายขึ้น ในตัวอย่างเป็นรูป 3 มิติง่ายๆ แสดงถึงการจัดเก็บทางกายภาพ ใน Data Warehouse เครื่องจัดเก็บข้อมูล ROLAP ใน RDBMS และใช้ข้อมูลพื้นฐานในการวิเคราะห์ และ ให้ผู้ใช้ใช้ GUI ในการวิเคราะห์ ระบบการจัดเก็บข้อมูลแบบ MOLAP ใน MDBMS ใช้เทคนิคการผสมของ array เพื่อหลีกเลี่ยงแบบมิติต่างๆ ของลูกบาศก์

การวิเคราะห์ข้อมูลมิติต่างๆ มีผลอย่างไรกับระบบฐานข้อมูลที่เบาบาง (sparsity)Sparsity เป็นการวัดความหนาแน่นของข้อมูลในลูกบาศก์ Sparsity คำนวณได้จากการแบ่งจำนวนจริงทั้งหมดในลูกบาศก์จากจำนวนเซลว่าง กลับไปมองตัวอย่างการขาย อาจมีบางผลิตภัณฑ์ที่ยังขายไม่ได้ ณ เวลาหนึ่งในสถานที่หนึ่ง ในความเป็นจริงคุณจะพบบ่อยประมาณ 50 % จะเป็นเซลที่ใช้บ่อย ในบางกรณีฐานข้อมูลแบบหลายมิติจะคอยควบคุมผลกระทบเล็กน้อยที่มีแนวโน้มจะทำให้ลดการประมวลผลส่วนเกินและ ลดความต้องการปัจจัยลงได้

อย่างไรก็ตามในด้านเทคนิคของฐานข้อมูล บทบาทที่สำคัญอันหนึ่งของหลายมิติสามารถเน้นบางส่วนของลูกบาศก์ได้ตัวอย่างเช่น ผู้จัดการฝ่ายผลิตภัณฑ์อาจสนใจการตรวจสอบการขายผลิตภัณฑ์ โดยที่ที่จะตรวจสอบคือหน่วยเก็บที่สนใจ การใช้หลายมิติเป็นกลุ่ม สามารถตัดส่วนย่อยของลูกบาศก์ซึ่งเรียกว่า slice และ dice (ส่วนและก้อน) เป็นภาพของ slice และ dice แต่ละส่วนที่ถูกตัดทำให้ได้ส่วนย่อยเป็นส่วนๆ การตัดเป็นส่วนๆทำให้ได้ลูกบาศก์เล็กๆ ทำให้เกิด “dice” เป็นส่วนหนึ่งของการทำ slice และ dice

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การทำให้เป็นส่วนๆ และเป็นก้อน เราสามารถทำการตรวจสอบแต่ละลูกบาศก์ได้โดยใช้คุณสมบัติของมิติ ตัวอย่างเช่น ใช้มิติสถานที่ เราต้องกำหนดคุณสมบัติ STORE_ID เพื่อทำให้ตรงกับหน่วยจัดเก็บ

กำหนดค่าที่ต้องการให้กับ slice และ dice ในสิ่งแวดล้อม เขียนให้เห็นถึงการเพิ่มและ การเพิ่มในแบบ 3 มิติ เข้าไปใน sale fact ดังนั้นการเซ็ตขั้นตอนในการกำหนดเป็นหนทางใหม่ในการค้นหา แยกแยะ และเป็นข่าวสารที่เป็นไปได้ทั้งหมด ตัวอย่างเช่น มิติของสถานที่ ได้เพิ่มรูปภาพ 3 มิติ ของสถานที่ที่ขาย, บริเวณไหน,รัฐใด เมือง และคลังเก็บ คุณสมบัติทั้งหมดที่แต่ละส่วนมี จะช่วยในการตัดสินใจของผู้ใช้ ดังนั้นจะทำให้ผู้ขายศึกษาคุณสมบัติแต่ละมิติได้

จำไว้ว่ามิติของเวลามีความสำคัญ มิติของเวลาจะมีข่ายการทำงาน การขาย สามารถจะวิเคราะห์, ศึกษาความเป็นไปได้, คาดการล่วงหน้าได้ เช่นกัน มิติของเวลามีบทบาทที่สำคัญต่อการขายทั้งหมดที่เราสนใจ (ในสัปดาห์, เดือน, ไตรมาส, ฯลฯ) ให้มิติของเวลาจากข้อมูลเวลาใกล้เคียงกับความจริง ผู้ขายจะเพิ่มเข้ามาโดยอัตโนมัติ จากการจัดการของมิติเวลา เป็นลักษณะหนึ่งที่ได้จากคลังข้อมูล (Data Warehouse)

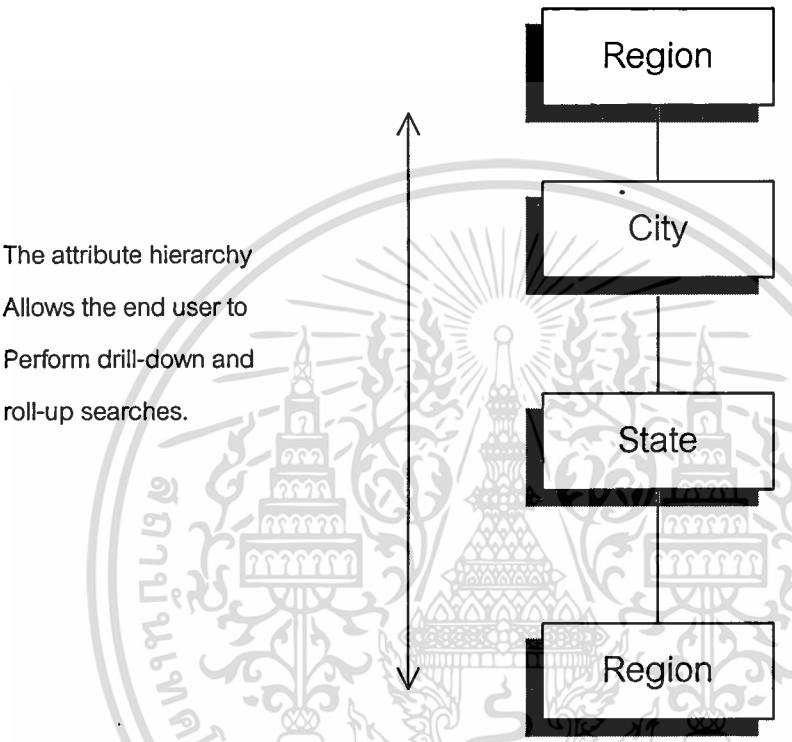
4. Attribute Hierarchy คุณสมบัติแบบลำดับชั้น

เป็นมิติหนึ่งที่สามารถกำหนดคุณสมบัติแบบลำดับได้ Attribute Hierarchy เป็นรูปแบบข้อมูลในระดับบนและล่างเป็นวัตถุประสงค์ที่สำคัญ การรวบรวม และการวิเคราะห์ข้อมูลระดับล่างและบน ตัวอย่างตามภาพที่ 3.6 แสดงถึงสถานที่ที่ สามารถจัดให้อยู่ในแบบลำดับชั้นตาม เขต, รัฐ, เมือง, คลัง

คุณสมบัติลำดับชั้นของคลังข้อมูล

คลังข้อมูลจะมีความสามารถในการค้นระดับล่างและบน ตัวอย่างเช่น การค้นหาข้อมูลดูคำตอบจากคำถาม การขายรายเดือนในปี 1996 เปรียบเทียบกับการขายรายเดือนในปี 1995 การวิเคราะห์ยอดขายที่ลดลงในเดือน มีนาคมในปี 1996 การตัดสินใจในการลดลงของการขายในเดือนมีนาคมต้องเข้าไปดูว่าเกิดจากเขตใดโดยทำการเปรียบเทียบกับปีก่อนหน้านี้ การทำเช่นนี้จะเป็นการเตือน ถึงผลกระทบจากยอดขาย ที่ลดลงในทุกเขตหรือเขตใดโดยเฉพาะ การทำ drill-down สามารถเพิ่มได้เรื่อยๆ จนได้ข้อมูลแยกออกมาเป็นแบบมาตราฐาน ภาพที่ แสดงเป็นคุณสมบัติที่เป็นไปได้ที่จะจัดข้อมูลเป็นลำดับชั้นในคลังข้อมูล และ OLAP เพื่อเป็นแนวทางตรวจสอบการรวบรวมข้อมูลในการ drill-down และ roll up จำเป็นว่า คุณสมบัติทั้งหมดจะเป็นส่วนหนึ่งของคุณสมบัติแบบลำดับชั้น บางคุณสมบัติอาจยังใหม่ และคำอธิบายสั้นๆ แต่ให้ระลึกเสมอว่า ลักษณะที่แตกต่างของมิติ สามารถนำมารวมเข้ากับแบบลำดับชั้นได้ในรูปที่ 3.7 จะเห็นถึงภาพการศึกษาวิเคราะห์การขายและ ใช้มิติของผลิตภัณฑ์, มิติเวลา มิติสถานที่ ในตัวอย่างนี้เป็นมิติของผลิตภัณฑ์เป็นผลิตภัณฑ์ทั้ง

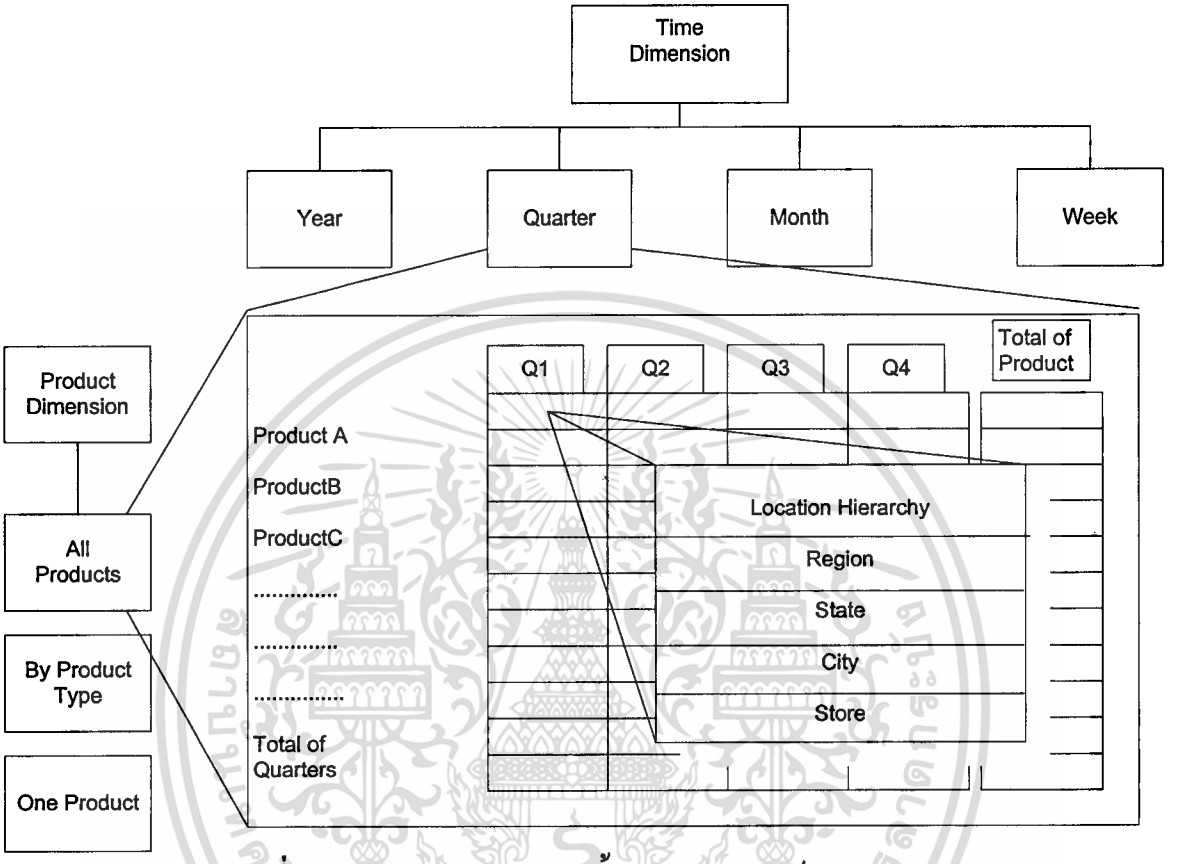
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 3.6 คุณสมบัติลำดับชั้นที่ผู้ใช้สามารถทำการค้นหาแบบ drill-down และ roll-up ได้

หมด หมายถึงแกน y จะเป็นผลิตภัณฑ์ทั้งหมดส่วนแกน x เป็นมิติเวลา ไตรมาส คือข้อมูลที่ได้รวบรวมเป็นไตรมาส ตัวอย่างเช่น ยอดขายทั้งหมดของ ABC ในไตรมาสที่ 1, 2, 3 และ 4 ในที่สุดมิติของที่ต่างๆ ได้ตั้งขึ้นเป็นเซต เพื่อให้แน่ใจว่าแต่ละ cell มียอดขายในเซตและไตรมาสภาพการวิเคราะห์ข้อมูลตัวอย่างดังภาพที่ 3.7 มีแนววิเคราะห์ 3 ทางที่แตกต่างกัน ตามมิติของผลิตภัณฑ์แกน y ในการวิเคราะห์ข้อมูลสามารถขอดูผลิตภัณฑ์, กลุ่มของผลิตภัณฑ์ต่างๆ หรือเพียงผลิตภัณฑ์เดียว ในมิติเวลาแกน x การวิเคราะห์ข้อมูล สามารถขอข้อมูลในเวลาต่างๆ ที่มีจำนวนรวมที่ต่างกันเช่น ปี, ไตรมาส, เดือน หรือ สัปดาห์ แต่ละยอดขายแสดงค่าเริ่มต้น ยอดขายของเซตในแต่ละผลิตภัณฑ์ เมื่อเราใช้ GUI click คลิ๊กในเซลที่สามารถวิเคราะห์ข้อมูลเพื่อ drill-down เพื่อหายอดขายของรัฐภายในเซต click อีกครั้งที่ค่าของ state จะได้ยอดขายแต่ละเมืองในรัฐและไปเรื่อยๆ เช่นเดียวกับตัวอย่างข้างต้น เป็นการกำหนดคัณสมบัติของข้อมูลในแบบลำดับชั้นในคลังข้อมูลที่จะดึงออกมา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 3.7 การกระจายแบบลำดับชั้นในการวิเคราะห์หลายมิติ

แสดง ข้อมูลแบบลำดับชั้นถูกเก็บไว้ใน dictionary ของ DBMS และถูกใช้โดย OLAP ในการเข้าถึงคลังข้อมูล ความมั่นใจในการเข้าถึงข้อมูลแต่ละครั้งเครื่องมือในการใช้ถามควรอยู่ใกล้คลังข้อมูล metadata และมีความสามารถในการวิเคราะห์อย่างเต็มที่

3.6.1 การแสดงให้เห็น Star schema

Fact และ dimension เป็นการนำเสนอแบบปกติโดยใช้ตารางในคลังข้อมูล ตาราง fact สัมพันธ์กับตารางมิติในลักษณะ (M:1) พูดอย่างหนึ่ง Fact หลายแถวสัมพันธ์กับแต่ละแถวของมิติ ใช้ยอดขายเป็นตัวอย่างเราสามารถสรุปได้ว่าแต่ละผลิตภัณฑ์จะมีได้หลายครั้งในตาราง Fact ตาราง Fact และตารางมิติสัมพันธ์กันโดย FK และอยู่ภายใต้ PK และกฎของ FK PK จะอยู่อีกด้านหนึ่ง ตารางมิติจะเก็บส่วนของ PK ไว้หลายที่ในตาราง Fact เพราะว่าตาราง Fact สัมพันธ์กับตารางมิติหลายตาราง PK ของตาราง Fact ประกอบด้วยPK ตามภาพที่ 3.8 เป็นความสัมพันธ์ระหว่างตาราง fact sales และ ผลิตภัณฑ์, สถานที่ และตารางมิติเวลา แสดงให้เห็นถึงความง่ายของ Star schema ที่สามารถขยาย

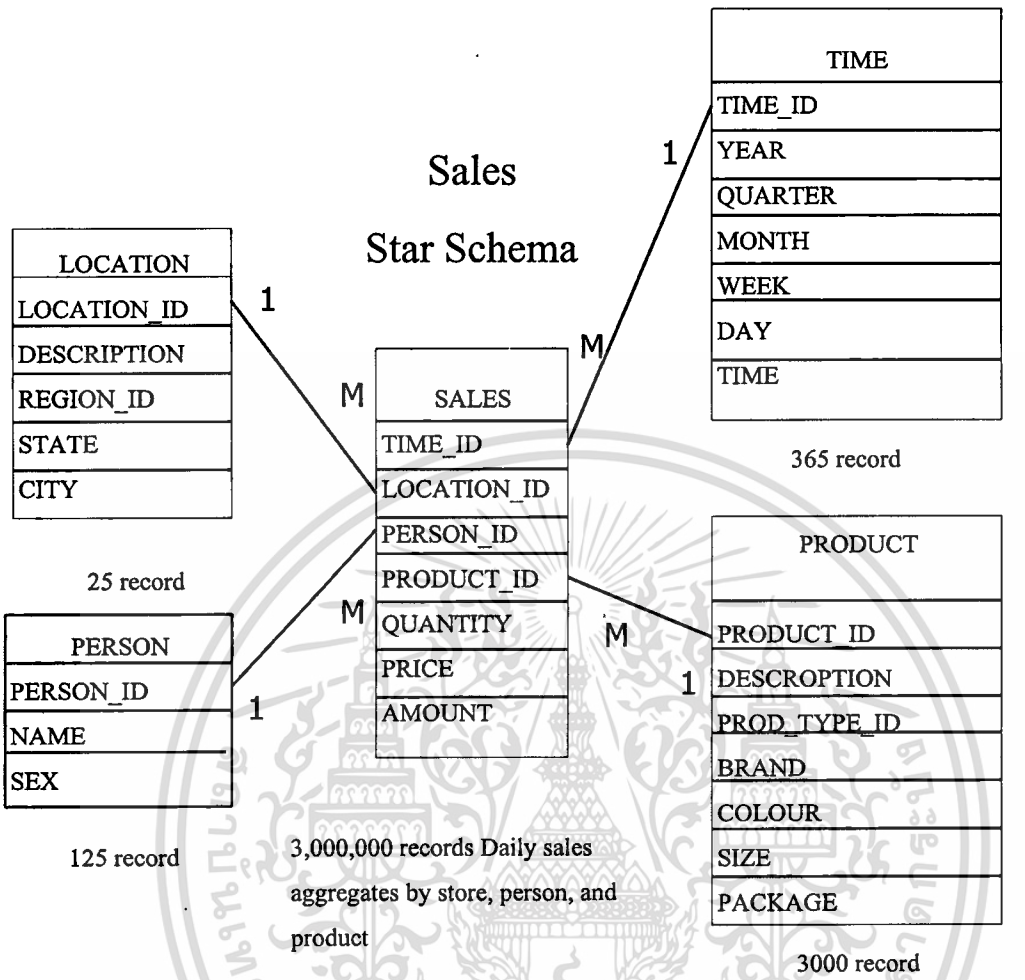
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ออกได้ เราได้เพิ่มมิติ PERSON เข้าไปเฉยๆ ทำให้เกิดการรวม PERSON_ID ในตาราง fact sale และการเพิ่ม PERSON เข้าไปในฐานข้อมูลPK ของตาราง sale fact ประกอบด้วย TIME_ID , LOCATION_ID , PERSON_ID และ PRODUCT_ID แต่ละ record ในตาราง sale table เป็นหนึ่งเดียวจากการรวมค่าแต่ละค่าในตาราง fact FK เป็นค่าปกติ PK ของตาราง fact มักจะเกิดจากการรวมกันของ FK ที่ชี้ไปยังแต่ละตารางเพื่อเชื่อมความสัมพันธ์ระหว่างกัน ในกรณีนี้แต่ละrecord ของ sale จะแสดงถึงผลผลิตที่ขายจากตัวแทนขาย 1 คน ในเวลาหนึ่งและสถานที่หนึ่ง ในตารางนี้ ตารางมิติเวลาที่แสดงเป็นรายวันดังนั้น ตาราง fact เกิดจากยอดขายสินค้ารายวันจากผู้ขายเพราะว่าตาราง fact เป็นค่าที่ใช้จริงในการใช้สนับสนุนการตัดสินใจ ค่านี้จะใช้ซ้ำหลายๆครั้งในตาราง fact ดังนั้น ตาราง fact จึงมักจะเป็นตารางที่มีขนาดใหญ่ใน Star schema เพราะตารางมิติไม่มีข่าวสารที่ซ้ำๆกัน (ผู้ขายไม่ซ้ำกัน, ผลิตภัณฑ์ต่างชนิดกัน) ตารางมิติมักจะเล็กกว่าตาราง fact ในแผนภูมิของ Star schema มี record ของมิติสัมพันธ์กันเป็นจำนวนมากกับ fact record ตัวอย่างเช่น "Widget" มีเพียงครั้งเดียวในมิติผลิตภัณฑ์ แต่มีเหมือนกันในตาราง fact นี้เป็นคุณสมบัติของการเรียกข้อมูลกลับมา เพราะว่าการวิเคราะห์มิติเวลาส่วนใหญ่จะดูตั้งแต่ fact จนถึงคุณสมบัติของมิติดังนั้นในการใช้ DSS จากคลังข้อมูลขั้นแรกคือเข้าไปหาในตารางที่เล็กกว่าก่อนเข้าไปค้นในตารางใหญ่ คลังข้อมูลมักจะมีตาราง fact มาก แต่ละตารางถูกออกแบบเพื่อใช้ตัดสินใจในการตอบคำถามในการตัดสินใจ ตัวอย่างเช่น สมมติว่าสนใจความเป็นลำดับ ในขณะที่เริ่มแรกสนใจยอดขาย ตามภาพนี้เราจะแสดงให้เห็นถึงลำดับของตาราง fact และตาราง fact ของ sales ในคลังข้อมูลเดียวกัน ถ้าลำดับมีแนวโน้มเป็นคีย์แรกที่สนใจ ลำดับของตาราง fact จะเป็นศูนย์กลางของ Star schema ที่มีตัวแทนขาย, ผลิตภัณฑ์ และมิติเวลาในกรณีนี้ สนใจยอดขายจาก ผู้ขายรายใหม่แทนโดย ตาราง vender ใหม่ในฐานข้อมูล มิติของผลผลิตใช้ตารางเดียวกับที่ใช้ใน Star schema เริ่มต้น อย่างไรก็ตามมีความสนใจในลำดับมากกว่ายอดขาย จึงต้องให้ความสนใจกับมิติเวลา ถ้าในส่วนของลำดับ ใช้ช่วงเวลาเดียวกันกับช่วงการขาย สามารถแสดงเวลาโดยใช้ตารางเวลาเดียวกัน ถ้าช่วงเวลาที่ใช้แตกต่างกัน เราต้องสร้างตารางเวลาขึ้นมาใหม่ บางครั้งชื่อ ODER_TIME ใช้แทนช่วงเวลาแบบลำดับ ภาพ 3.9 ลำดับของ Star schema มีผลผลิตรวมกัน ตัวแทนขายและมิติเวลา ตาราง fact ต่างๆก็สามารถสร้างได้โดยเหตุผลเดียวกัน ส่วนข้างล่างต่อไปนี้อธิบายการทำงานด้วยเทคนิคที่สูงขึ้นโดยใช้ Star schema

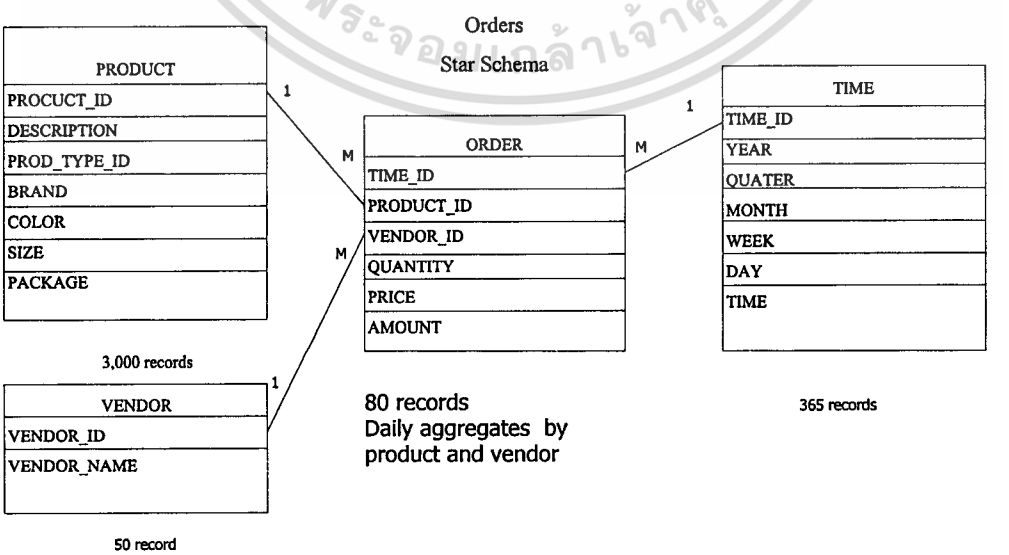
3.6.2 เทคนิคการตรวจสอบ

ทำให้การสร้างฐานข้อมูลรวดเร็วและถูกต้อง ในการวิเคราะห์คลังข้อมูล การออกแบบคลังข้อมูลขั้นต้น ดังนั้นการที่จะทำให้ใช้งานได้เร็วขึ้นโดยใช้โค้ด SQL แทนมิติธุรกิจ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 3.8 Star Schema สำหรับยอดขาย



ภาพที่ 3.9 ORDER STAR SCHEMA

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

มี4 เทคนิคที่ใช้บ่อยในการออกแบบคลังข้อมูลให้สมบูรณ์

1. ตารางมิติ Normalization
2. ตาราง fact ต่างๆ ที่แทนผลรวมระดับที่ต่างกัน
3. การ Denormalization ของตาราง fact
4. แบ่งตาราง และการทำสำเนา

1. ตารางมิติ Normalization

การ Normalization ทำได้ง่าย และใช้หามิติได้ง่ายตัวอย่างเช่น ถ้าตารางในมิติ location มีการเคลื่อนย้ายที่ขึ้นกับเขต, รัฐ และเมืองสามารถแก้ไขความสัมพันธ์นี้โดยใช้ 3 NF ดังแสดงในรูปที่ 3.10 โดยการ Normalization มิติของตาราง เราทำให้ข้อมูลง่ายในการกรองความสัมพันธ์ของมิติ ในตัวอย่างนี้ เขต, รัฐ, เมือง และ location มี record จำนวนน้อยที่มีให้เปรียบเทียบในตาราง fact มีเพียงตาราง fact เท่านั้นที่สัมพันธ์โดยตรงกับ ตาราง fact sales

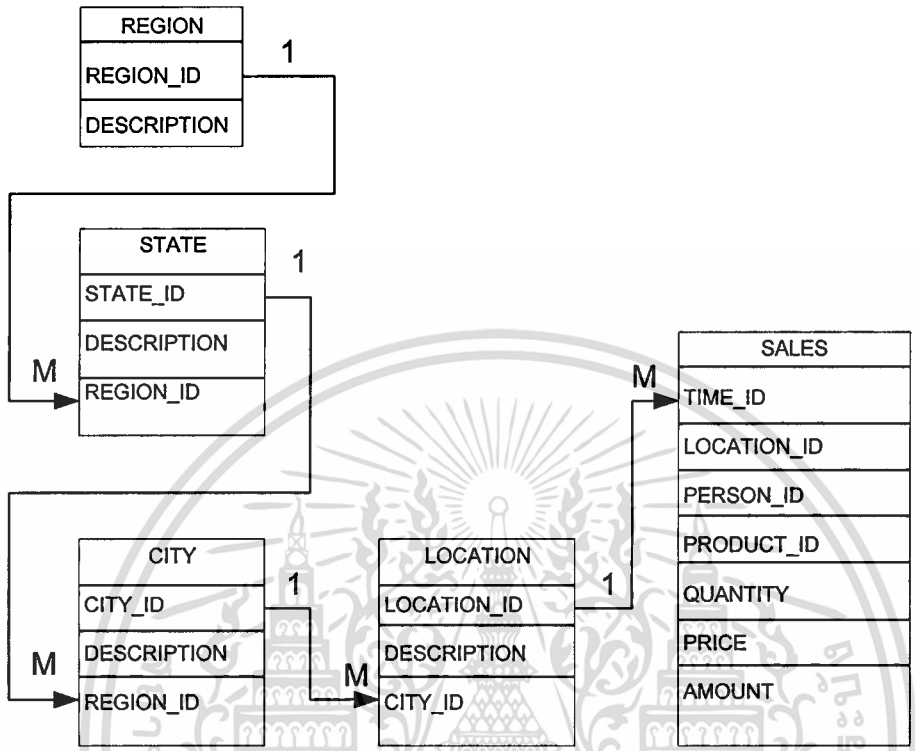
2. ตาราง fact ต่างๆ ที่แทนผลรวมระดับที่ต่างกัน

เราสามารถเพิ่มจำนวนคำถามให้มากขึ้นได้คำนวณล่วงหน้าแล้วในขั้นตอนการ load ข้อมูลแทนที่จะทำขณะ run อยู่ ดังนั้นจึงทำให้วิเคราะห์ข้อมูลได้รวดเร็ว และผู้ใช้สามารถใช้เครื่องมือในการสรุปตาราง fact ได้ดีกว่าการดูค่าในระดับอย่างละเอียด เทคนิคนี้แสดงในภาพที่ 3.11 ใช้การขายเป็นจุดเริ่มต้น เราได้รวมเขต, รัฐ และเมืองรวมเข้าไปไว้ในตาราง fact นักออกแบบคลังข้อมูลต้องแบ่งแยกระดับของการคำนวณล่วงหน้า และเก็บในฐานข้อมูล การรวมตาราง fact ต่างๆ ได้ทำการปรับปรุงขณะที่ load แบบ batch ในแต่ละครั้ง และเป็นเพราะว่าเป้าหมายเพื่อลดเวลาในการเข้าถึงและการประมวลผล นักออกแบบคลังข้อมูลจะต้องเลือกตาราง fact ที่จะใช้และเวลาที่ต้องการในการคำนวณขณะที่ทำการ run

3. การ Denormalization ตาราง fact

Denormalization ตาราง fact เป็นการทดสอบประสิทธิภาพของการเข้าใช้ข้อมูลและประหยัดพื้นที่ที่ใช้เก็บข้อมูล เป็นวัตถุประสงค์สุดท้าย อย่างไรก็ตาม กลายมาเป็นสิ่งที่เกิดขึ้นหลังสุด ต้นทุนการจัดเก็บข้อมูลเพิ่มขึ้นทุกวัน DBMS ก็มีข้อจำกัดของขนาดตาราง ขนาดของ record และจำนวนสูง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



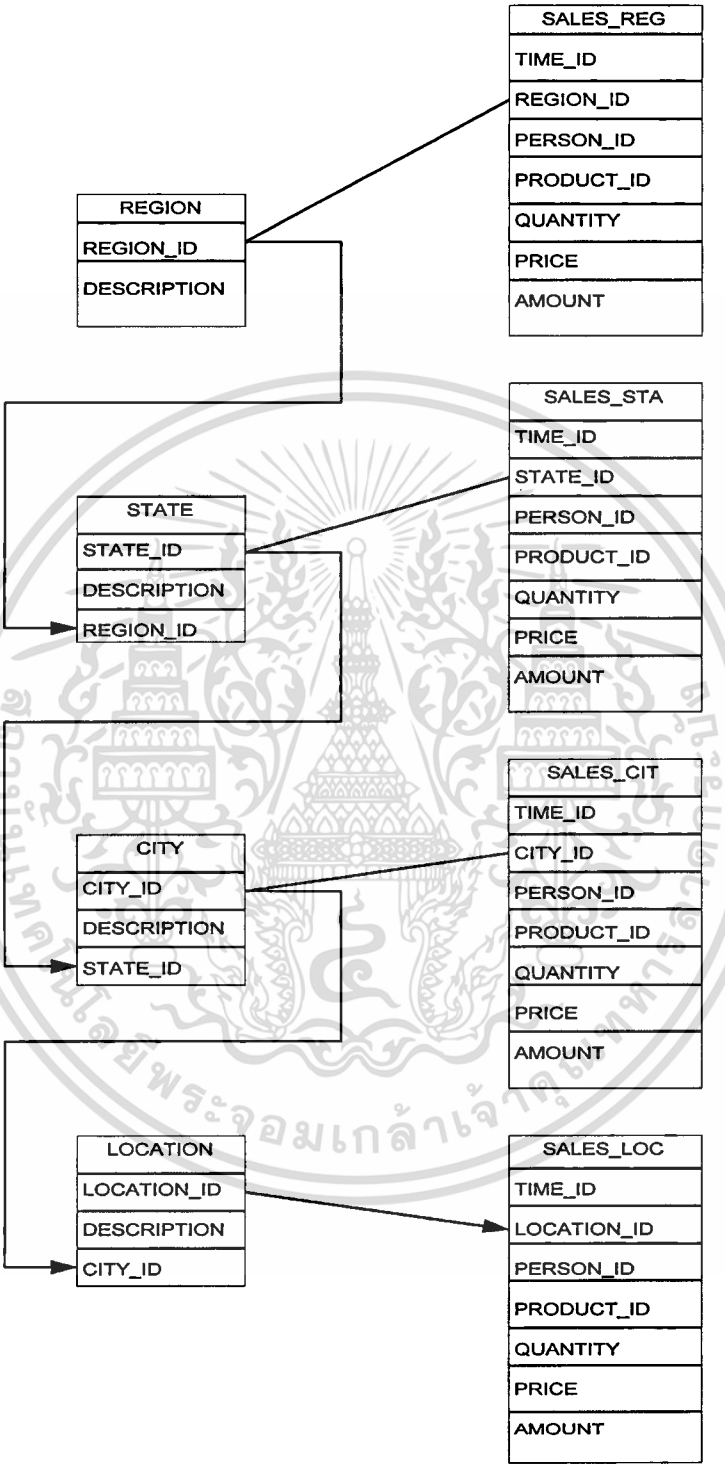
ภาพที่ 3.10 NORMALIZED DIMENSION TABLE

สุดของ record ในหนึ่งตาราง ยังไม่มีเนื้อที่จัดเก็บข้อมูลที่เพียงพอ Denormalization ทดสอบประสิทธิภาพ โดยการจัดเก็บข้อมูล record เดียวที่มักจะคลุมหลายๆ record ตัวอย่างเช่น การคำนวณยอดขายของผลิตภัณฑ์ทั้งหมดของเขตทั้งหมด เราต้องเข้าไปในยอดขายรวมของเขตและสรุป record ทั้งหมดในตาราง ถ้าเรามีสินค้าทั้งหมด 300,000 ชนิด เราต้องทำการสรุป record อย่างน้อย 300,000 แถว ถึงแม้ว่าจะไม่มีการวัด DBMS การเปลี่ยนแปลงเทียบกับเวลาที่ผ่านมาเป็น สิบสิบปีก่อนที่จะเริ่มมีระบบขายเดิมๆ ในกรณีนี้ มีประโยชน์ที่จะใช้ตารางทั้งหมดที่ได้ทำ denormalize ตัวอย่างเช่น ตาราง YEAR_TOTALS ต้องมี field ต่อไปนี้ YEAR_ID, MONTH_1, MONTH_2...MONTHn เช่นเดียวกันกับปีอื่นๆ ดังนั้นตารางธรรมดาๆนี้ เป็นหลักการง่ายๆ สำหรับการเปรียบเทียบปีต่อปี ของเดือนที่สูงสุดในระดับไตรมาส หรือใน 1 ปี อีกอย่างหนึ่งการออกแบบมาตราฐาน เช่นความถี่ของการใช้งาน และประสิทธิภาพที่ต้องการการประเมิน ปริมาณข้อมูลที่ DBMS จะสามารถทำการเชื่อมโยงเป็น Denormalize ได้

4. การแบ่งตารางและการทำสำเนา

เนื่องจากเราได้กล่าวถึงรายละเอียดของ 2 เทคนิคที่กล่าวมาแล้ว เราจะดูเทคนิคนี้ใน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 3.11 Multiple Fact Tables

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ส่วนที่เชื่อมกันเป็นพิเศษกับคลังข้อมูล การแบ่งตารางเป็นส่วนๆ และการทำสำเนาเป็น

ส่วนหนึ่งที่สำคัญเมื่อนำ DSS มาใช้กันอย่างกว้างขวาง การแยกตารางออกเป็นแถว หรือ คอลัมน์ย่อยๆ และนำไปวางไว้ใกล้กับเครื่อง client computer เพื่อทำให้เข้าถึงข้อมูล ได้เร็ว ขึ้น

Replication เป็นการ copy ตารางและนำไปไว้ในที่แตกต่างกัน เพื่อปรับปรุงเวลาในการเข้าถึงข้อมูล ไม่ว่าจะทำให้ตารางมีประสิทธิภาพอย่างไร เวลาเป็นมิติที่สำคัญในธุรกิจที่ใช้การวิเคราะห์ข้อมูล ดังนั้นจึงเป็นเรื่องปกติที่จะมี fact หนึ่งตารางต่อการรวบรวมข้อกำหนดมิติเวลา ตัวอย่างเช่นจากการขายของเรา จะมียอดรวมการขายของตาราง fact 5 ตาราง, ต่อวัน, สัปดาห์ เดือน, ไตรมาส, ปี ตาราง fact นี้จะต้องมีความหมายหรือความชัดเจน Periodicity มักใช้กันมากในปีปัจจุบัน, ปีในอนาคต หรือทุกปี มีข้อมูลข่าวสารเกี่ยวกับเวลาของข้อมูลที่จัดเก็บในตารางปีปัจจุบันหมายถึงตารางที่เก็บยอดขายรวมรายวันเฉพาะในปีนั้น, ปีก่อนๆ จะแสดงถึงปีก่อนหน้าเท่านั้น ทุกปีก็หมายถึงตารางที่จัดเก็บทั้งปีก่อนและปีปัจจุบัน ช่วงสุดท้ายของแต่ละปี ยอดขายรายวันในปีปัจจุบันจะเปลี่ยนไปจากตารางอื่นที่มียอดขายในปีที่แล้ว ตารางนี้ยังคงบันทึกการขายตั้งแต่เริ่มแรกยกเว้นปีปัจจุบัน ข้อมูลในปีปัจจุบันและตารางยอดขายปีที่แล้วที่สมบูรณ์ของบริษัท เราสามารถทำสำเนาตารางยอดขายของสถานที่ต่างๆ เพื่อหลีกเลี่ยงการใช้ข้อมูลยอดขายที่อยู่ห่างไกล ซึ่งทำให้ใช้เวลานานในการตอบสนอง ขนาดตารางที่พอจะเป็นไปได้ แต่คำถามที่ลงตัวที่สุดจะเป็นที่บรรจุค่าของ denormalization

3.7 การสร้างคลังข้อมูล

ระบบสารสนเทศขององค์กรส่วนใหญ่จะมีกฎข้อบังคับจำนวนมาก ซึ่งบางข้อขึ้นกับเงินทุน บางข้อขึ้นกับการจัดการระบบสารสนเทศ และความต้องการข้อมูลและความละเอียดของข้อมูลเพิ่มขึ้น การเพิ่มกฎเกณฑ์ กำหนดโดยระเบียบของวัฒนธรรม และคุณต้องเข้าใจว่าทำไมจึงไม่มีหลายสูตรที่ใช้อธิบายการพัฒนาคลังข้อมูลได้สมบูรณ์ดังนั้นการจัดการในการออกแบบคลังข้อมูลคลังเดียวจะต้องมีการจัดการเป็นพิเศษเราจะระบุองค์ประกอบที่ปรากฏในคลังข้อมูลด้วยวิธีทั่วไป

3.7.1 คลังข้อมูลที่มีลักษณะในการตัดสินใจ

บางครั้งต้องระลึกไว้ว่าฐานข้อมูลไม่ได้เป็นฐานข้อมูลคงที่เป็นการเปลี่ยนแปลงเพื่อการตัดสินใจส่วนใหญ่จะขึ้นกับการกำหนดไว้ล่วงหน้า เนื่องจากเป็นพื้นฐานของ DSS การออกแบบและการปฏิบัติของคลังข้อมูลหมายถึงการออกแบบและการปฏิบัติในการพัฒนาฐานข้อมูลให้สมบูรณ์เพื่อเป็นพื้นฐานในการตัดสินใจของบริษัทโดยทั่วไปแม้ว่าเป็นการง่ายที่จะกำหนดให้คลังข้อมูลเป็นฐานข้อมูลเพื่อเป็นศูนย์กลางเก็บข้อมูล คุณต้องระลึกว่าโครงสร้างพื้นฐานของการตัดสินใจได้รวมไปถึง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

hard ware, software, ตัวบุคคล และขั้นตอนต่างๆ เช่นเดียวกับข้อมูล ข้อโต้แย้งของคลังข้อมูลมีเพียงแต่คำวิจารณ์จากความสำเร็จและ ข้อแย้งของบุคคลทั้งต้องการองค์ประกอบของความคิด

โดยสรุป คลังข้อมูลเป็นองค์ประกอบที่สำคัญของแบบ DSS แต่ไม่ใช่เพียงแค่นั้น การออกแบบและการปฏิบัติแต่ต้องตรวจสอบโครงสร้างพื้นฐานทั้งหมดให้ชัดเจน

3.7.2 ความพยายามของบริษัทเพื่อความเกี่ยวข้องกันทุกระดับในบริษัทเพื่อความชัดเจน องค์กรที่สมบูรณ์ควรได้รับผลระโยชน์จากคลังข้อมูลพื้นฐานที่ใช้ในการตัดสินใจ การออกแบบคลังข้อมูลหมายถึงการให้โอกาสการช่วยพัฒนา และรวบรวมรูปแบบข้อมูลขององค์กรจากข้อมูลจริงในเชิงธุรกิจที่ผู้ใช้ต้องการ คลังข้อมูลตรงข้ามกับส่วนที่เป็นเส้นและกราฟ เพราะว่าคลังข้อมูลพยายามแสดงถึงรูปแบบข้อมูลขององค์กรเหมือนกับคุณได้พบกับองค์ประกอบขององค์กรบางครั้งก็มีเป้าหมายที่ขัดแย้งกัน และแน่นอนว่าจะง่ายในการค้นหาข้อมูลที่แตกต่างกันได้ และนำไปใช้เพื่อความสำเร็จ ความต่อต้านจากผู้ใช้ และอำนาจในการต่อรองในทุกระดับ การสร้างคลังข้อมูลให้สมบูรณ์ไม่เพียงแต่รู้ว่าจะสร้าง Star schema อย่างไรเท่านั้น คลังข้อมูลต้องการการจัดการพื้นฐานเพื่อความชัดเจน ความสำเร็จส่วนใหญ่เกิดจากการออกแบบที่ดี ส่วนน้อยที่เกิดจากการพัฒนาพื้นฐานที่เกี่ยวข้องกับข้อขัดแย้ง, สื่อ และการได้รับอนุญาตในระยะสั้นๆ นักออกแบบต้อง

- มีส่วนร่วมกับผู้ใช้ในการ process
- ผู้ใช้ยอมรับในความปลอดภัยตั้งแต่ต้น
- ยอมรับข้อโต้แย้งของผู้ใช้
- แก้ไขในสิ่งผู้ใช้ต้องการ
- สร้างขั้นตอนการแก้ไข จากเหตุข้อขัดแย้ง

3.7.3 ความเหมาะสมของทั้ง 3 เรื่องมี Data, Analysis และUsers

ความเชี่ยวชาญมากๆไม่มีความจำเป็น แน่นอนในการให้เกิดความพอใจทั้ง 3 อย่าง ลักษณะที่ดีของคลังข้อมูลต้องมีคำบรรยายที่ดีคำคติที่ว่า Input, Process, Output จะมีการกล่าวซ้ำในเทคนิคนี้ นักออกแบบคลังข้อมูลต้องสร้างความพอใจให้เกิดขึ้นใน

- การรวมกลุ่มข้อมูลและภาระที่รับได้
- ความสามารถในการวิเคราะห์ข้อมูลและยอมรับประสิทธิภาพในการตอบข้อซักถามได้
- ความต้องการวิเคราะห์ข้อมูลของผู้ใช้

เทคนิคที่เกี่ยวข้องมากที่สุดในการปฏิบัติด้านคลังข้อมูล คือมีการสนับสนุนการตัดสินใจของผู้ใช้ด้วยประสิทธิภาพที่ดีในการวิเคราะห์ข้อมูล ในช่วงเวลานั้นๆ ในรูปแบบที่ถูกต้อง ใช้ข้อมูลที่ถูกต้องในต้นทุนที่เหมาะสม

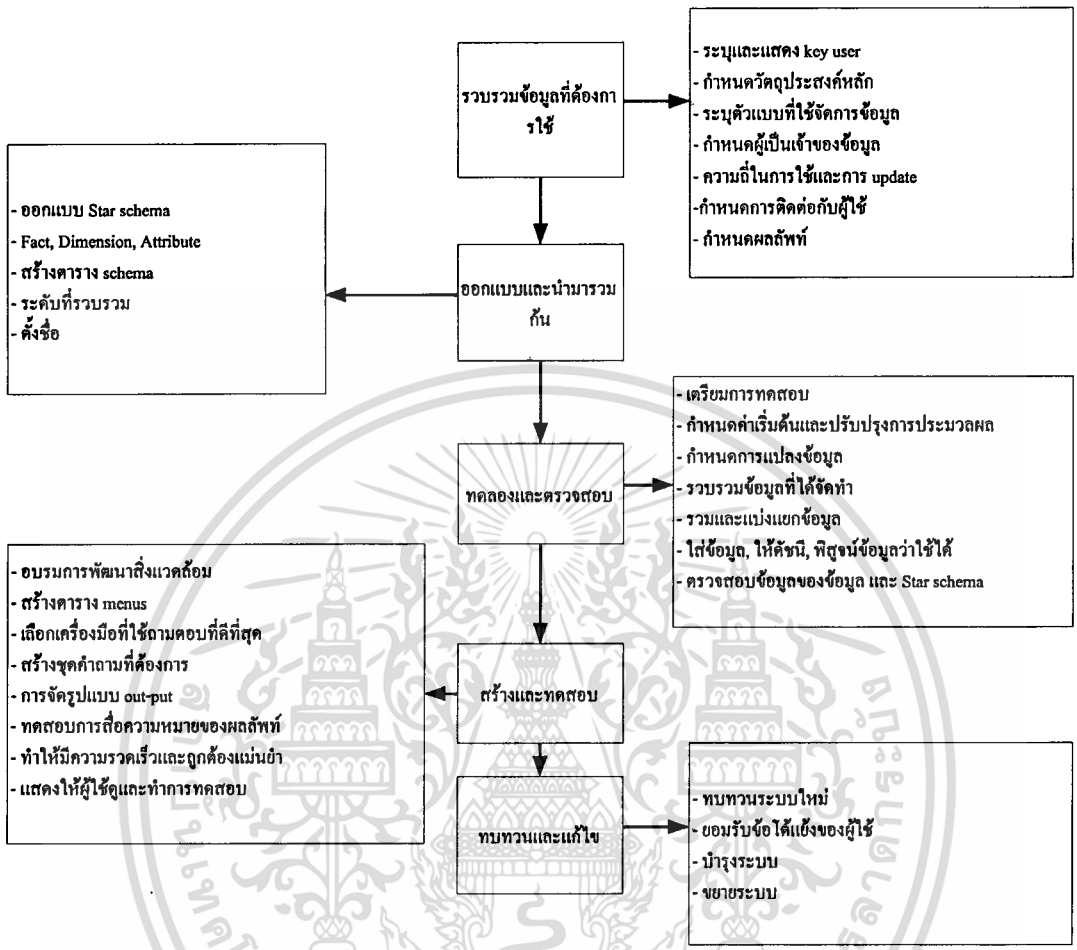
3.7.4 การใช้คลังข้อมูล และการออกแบบขั้นตอนจะต้องปรับให้เข้ากับความ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ต้องการของคลังข้อมูล ถ้าคุณจำได้ว่าคลังข้อมูลเพื่อนำมาในทางการค้าการออกแบบขั้นตอนจะต้องปรับให้เข้ากับความต้องการของคลังข้อมูล ถ้าคุณจำได้ว่าคลังข้อมูลได้มาจากการจัดการกับฐานข้อมูล คุณจะเข้าใจว่าทำไมการออกแบบพื้นฐานอย่างมั่นคงของฐานข้อมูลจึงมีความสำคัญ (มันเป็นการยากที่จะสร้างคลังข้อมูลที่ดี เมื่อการกระทำกับฐานข้อมูลไม่ตรงกับความจริง) ภาพที่วาดให้เห็นในภาพที่ 3.12 แสดงขบวนการง่ายๆ ในการสร้างคลังข้อมูล การพัฒนา, คลังข้อมูลเป็นความต้องการของบริษัทโดยทั่วไปต้องการนำไปใช้เป็นปัจจัย ได้แก่ บุคคล, แหล่งเงินทุน และนักวิชาการ เทคนิคความต้องการ การสนับสนุนการตัดสินของบริษัททั่วไป ต้องการความชำนาญ โครงสร้างด้านเสียงมาผสมเทคโนโลยี และขั้นตอนในการจัดการ สิ่งเหล่านี้ปฏิบัติได้ยาก ตัวอย่างเช่น

- แนวโน้มความต้องการข้อมูลในการตัดสินใจ มีความต้องการ Hardware, Software ล่าสุดที่มีความก้าวหน้าในการคำนวณด้วย cpu หลายๆตัว ความทันสมัยของระบบฐานข้อมูลปัจจุบันนี้ Client/Server เป็นเทคโนโลยีที่เข้ามาช่วยในการสร้างคลังข้อมูลได้ดี
- ต้องจัดการโดยละเอียดเพื่อให้ได้รฐานข้อที่เป็นคลังข้อมูลที่สามาระการควบคุมปริมาณข้อมูลและการคัดเลือกข้อมูล ได้ และข้อมูลมีลักษณะเดียวกัน
- การสร้างและการสนับสนุนโครงสร้างคลังข้อมูล เรามีความต้องการบุคคลที่มีความสามารถในการออกแบบฐานข้อมูลที่ดี, มีลักษณะเดียวกัน และมีการบริหารที่ดี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 3.12 แนวทางในการสร้างคลังข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

ระบบคลังข้อมูล (Data Warehousing System)

Data Warehouse ถือว่าเป็นศูนย์กลางของสถาปัตยกรรมสำหรับระบบสารสนเทศในปี1990 โดย data warehouse จะสนับสนุนการประมวลผลสารสนเทศ มีการรวบรวม จัดเก็บข้อมูลภายในและภายนอกองค์กร ไม่จำกัดข้อมูลในการเก็บบันทึกลงฐานข้อมูลขนาดใหญ่ รวมทั้งจัดรูปแบบของข้อมูลให้สามารถนำมาวิเคราะห์ได้ โดยการอาศัยเทคนิคต่างๆ เช่นการนำความรู้ทางสถิติ หรือคณิตศาสตร์มาประยุกต์ใช้ เพื่อหาแนวโน้มของข้อมูลที่ต้องการ ซึ่งการจัดทำ data warehouse จะทำให้ได้ข้อมูลที่จะนำมาใช้ประกอบการตัดสินใจได้แม่นยำ และถูกต้องยิ่งขึ้น จะส่งผลให้เกิดความถูกต้องในการตัดสินใจ โดยระบบ data warehouse ที่ดีจะเกิดจากการนำข้อมูลจากระดับปฏิบัติการมาผ่านกระบวนการในการแปลงข้อมูลให้เป็นมาตรฐานก่อนจัดเก็บ

Data Warehouse จึงเป็นการนำเสนอแนวทางในการเข้าถึงข้อมูลได้อย่างมีประสิทธิภาพทั้งในแบบท้องถิ่น (local) หรือแบบระยะไกล (remote) รวมทั้งการจัดสภาพแวดล้อมของ data warehouse ยังสามารถสนับสนุนการทำงานของผู้ใช้หลายๆ คน หลายๆ จุดประสงค์ และง่ายต่อการใช้งาน จึงนับได้ว่าเทคโนโลยี data warehouse เป็นสิ่งที่ทำให้องค์กรประสบความสำเร็จในการบริหารงานที่มีประสิทธิภาพ

4.1 ความหมายของ Data Warehouse

Data Warehouse คือ การรวบรวมข้อมูลจากฐานข้อมูลของส่วนปฏิบัติงานต่างๆ นำมาแปลงข้อมูลให้อยู่ในรูปแบบที่เหมาะสมในการจัดเก็บและสะดวกต่อการนำมาใช้งาน แล้วจึงจะนำข้อมูลเหล่านั้นเข้าไปเก็บใน data warehouse ข้อมูลที่อยู่ใน data warehouse เป็นข้อมูลที่องค์กรธุรกิจ นำมาวิเคราะห์ วางแผน และควบคุม ใช้ตัดสินใจในลักษณะของธุรกิจ เพื่อให้การทำ data warehouse เป็นไปอย่างมีประสิทธิภาพ

4.2 ส่วนประกอบของ Data Warehouse

ส่วนประกอบของ Data Warehouse มีดังนี้คือ

1. Source Database
2. Data Extration and Transformation tools

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

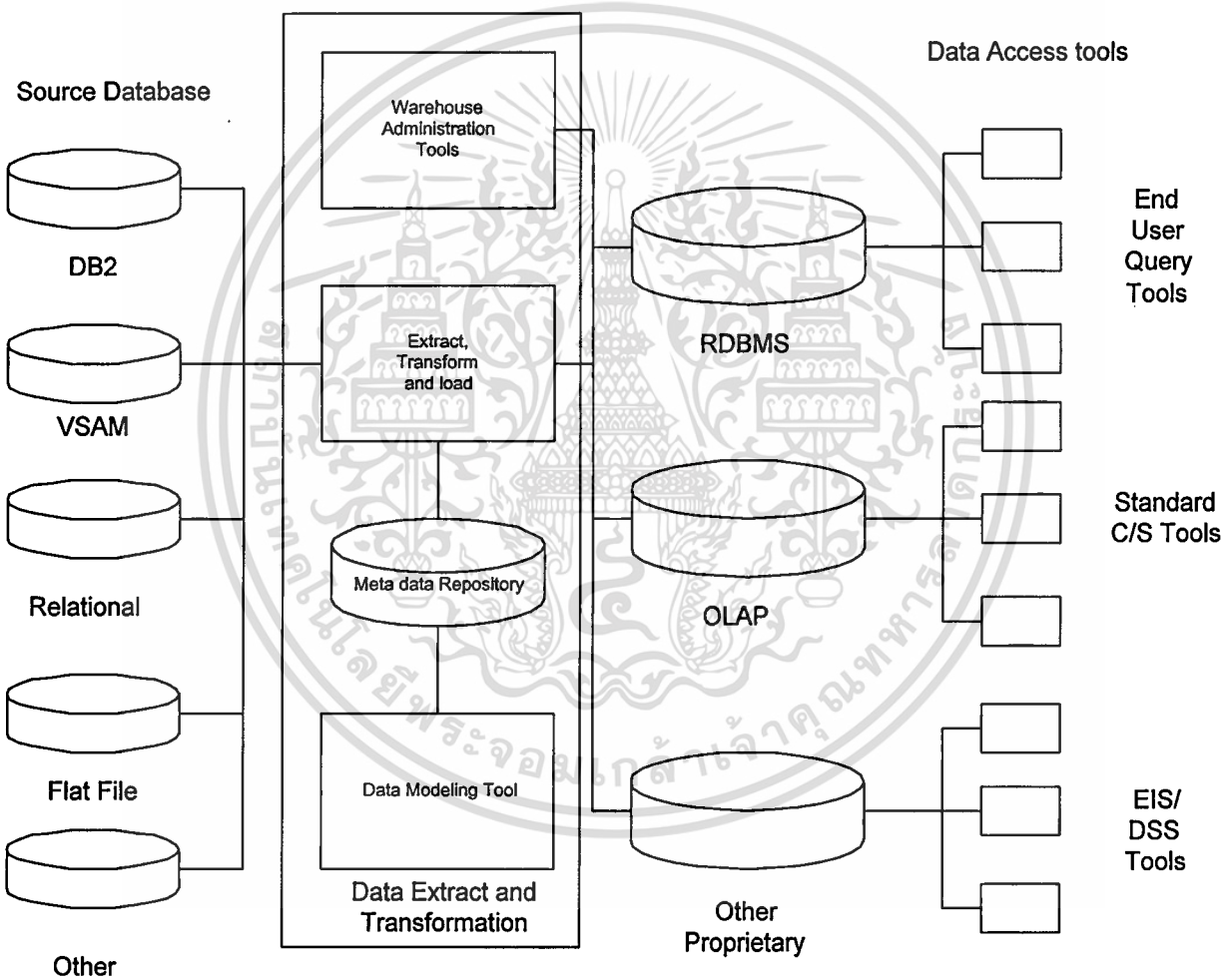
3. Data Modeling tools

4. Metadata Repository

5. Target Database

6. Front-end access tools

ดังภาพที่ 4.1



ภาพที่ 4.1 ส่วนประกอบของ Data Warehouse

คำอธิบายภาพ

1. Source Database

คือแหล่งข้อมูลที่มาจากฐานข้อมูลประเภทต่างๆ ที่อยู่บนเครื่อง Host ได้แก่ข้อ มูลที่มี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ลักษณะเป็นงาน Online transaction processing (OLTP) ข้อมูลที่มาจากส่วนปฏิบัติงานต่างๆ (Operational Data) ข้อมูลการประมวลผลงาน batch และแหล่งข้อมูลภายนอก

2. Data Extract and Transformation Tools

เป็นการนำเครื่องมือที่เหมาะสมเข้ามาช่วยในกระบวนการดึงข้อมูล การแปลงข้อมูล ให้อยู่ในรูปแบบที่เหมาะสมต่อการนำไปใช้งานบน data warehouse โดยที่เครื่องมือที่จะนำมาใช้นั้น จะต้องมีความสามารถที่จะจัดการกับข้อมูลที่มีปริมาณมากได้เป็นอย่างดี และเหมาะสมในการใช้งานกับข้อมูลขององค์กร

2.1 Data Extraction เป็นวิธีการในการดึงข้อมูล โดยทำการแยกข้อมูลบางส่วนออกมาจากฐานข้อมูลขนาดใหญ่ เพื่อประโยชน์ในการวิเคราะห์

2.2 Transformation เป็นการแปลงข้อมูลที่มีอยู่หลายรูปแบบ (format) ให้เป็นรูแบบมาตรฐานเดียวกัน ก่อนที่จะนำเข้าสู่ Data Warehouse

3. Data Modeling Tools

เป็นเครื่องมือที่ใช้ช่วยในการกำหนดแบบจำลองของข้อมูล (Data Model) เพื่อให้การรวบรวมข้อมูลจากแหล่งข้อมูลหลายแห่ง (Source Database) เข้าสู่ฐานข้อมูลปลายทาง (Target Database) เป็นไปอย่างถูกต้อง

4. Metadata Repository

เป็นที่ใช้จัดเก็บรายละเอียดที่สำคัญของข้อมูล เพื่อช่วยในการติดต่อ หรือเชื่อมต่อกับส่วนต่างๆ ของข้อมูลอื่นๆ ที่จะนำเข้าเก็บใน data warehouse เป็นไปได้โดยง่าย ลักษณะที่สำคัญของ Metadata คือ

4.1 เก็บข้อมูลเกี่ยวกับพจนานุกรมข้อมูล (data dictonary)

4.2 เก็บ directory of content ที่ประกอบไปด้วยตำแหน่ง และรายละเอียดเกี่ยวกับ Data Warehouse Server

4.3 กฎเกณฑ์และเงื่อนไขในการที่จะค้นหาข้อมูล

4.4 เก็บ Location ของข้อมูล

4.5 เก็บโครงสร้างของข้อมูล (structure)

5. Target Database

เป็นฐานข้อมูลปลายทางที่ใช้เป็น data warehouse สำหรับจัดเก็บข้อมูลที่ดึงมาจากแหล่งข้อมูล

6. Front-end access Tools

เป็นเครื่องมือที่จะนำมาใช้ในการเข้าถึงข้อมูล (data access) ได้อย่างรวดเร็ว และทำการวิเคราะห์ข้อมูลบน data warehouse ซึ่งได้แก่

6.1 End – User Query Tools

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เป็นเครื่องมือที่ใช้ค้นหาข้อมูล โดยช่วยอำนวยความสะดวกให้ผู้ใช้งาน ในด้านอื่น ด้วย เช่น

- การ interface ได้ง่าย
- การทำรายงาน
- การตอบคำถามแบบ Adhoc
- การคำนวณ
- การนำไปเชื่อมต่อกับ tools อื่นๆ (Export)

6.2 EIS/DSS Tools

เป็นเครื่องมือที่ช่วยสนับสนุนในการตัดสินใจ (Decision Support System – DSS) หรือช่วยในการทำข้อมูลเพื่อช่วยในการบริหารงานของผู้บริหารระดับสูง (Executive Information System – EIS)

4.3 องค์ประกอบในการสร้าง Data Warehouse

ในการสร้าง data warehouse นั้น ต้องมีการเลือกสรร hard ware เพื่อนำมาเป็น server ซึ่งต้องมีการผสมผสาน plateform ต่างๆ ของ server และ set configuration ของระบบให้เหมาะสม รวมทั้งต้องเลือก software ระบบจัดการฐานข้อมูล (DBMS) ที่เหมาะสมมาประกอบกันเป็น data warehouse โดยต้องมีการตัดสินใจอย่างรอบคอบในการเลือกขนาด ประสิทธิภาพ ให้เหมาะสมกับราคา และการใช้งาน องค์ประกอบที่สำคัญในการสร้าง data warehouse ประกอบด้วย 3 ประการ ได้แก่

1. สถาปัตยกรรมรวมของ Data Warehouse
2. สถาปัตยกรรมของ Server
3. สถาปัตยกรรมของ DBMS

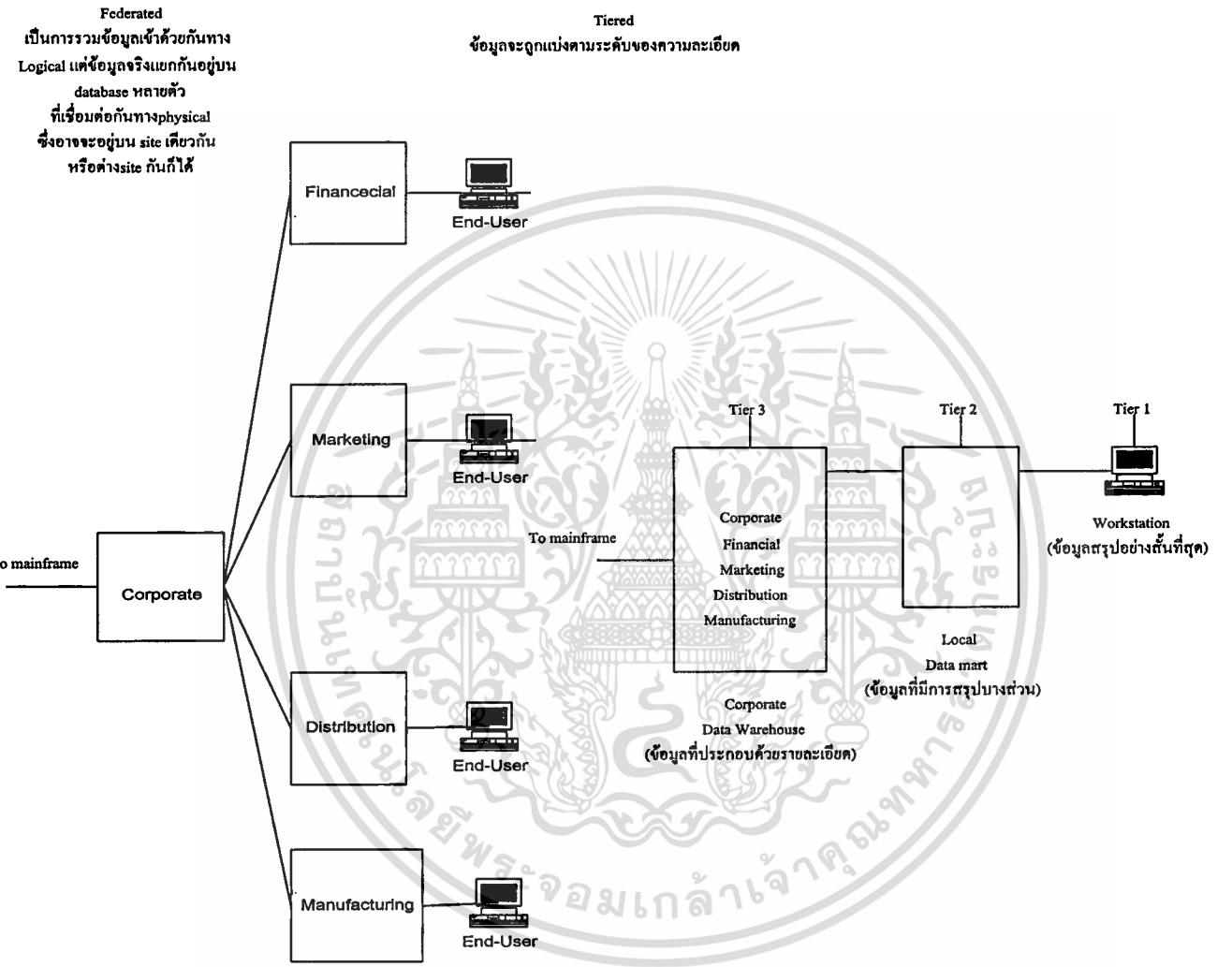
1. สถาปัตยกรรมรวมของ Data Warehouse

การพัฒนาคลังข้อมูล (Data Warehouse) ควรจะเริ่มจากโครงสร้างของฐานข้อมูลสำหรับ warehouse ทั้งด้าน Logical และ physical โดยมีแนวทางในการพิจารณาเลือกใช้ 3 วิธี การจัดเก็บข้อมูล ดังนี้

1.1 การเลือกใช้ platform เป็นแบบรวมศูนย์ (Centralized) ที่ติดตั้งอยู่ในบริเวณเดียวกัน หรือแบบกระจาย (Distributed) ที่แยก warehouse ออกไปตามสำนักงานจังหวัด โดยอาจรวบรวมข้อมูลจากแหล่งภายใน และภายนอกหลายแห่ง มาเก็บไว้ใน data warehouse เดียวกันแนวทางนี้เป็นการรวบรวมข้อมูลเข้าไว้ด้วยกัน ซึ่งจะเป็นการช่วยเพิ่มความสามารถในการประมวลผลข้อมูลได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.2 การกระจายข้อมูลออกไปยัง database หลายตัวตามfunction เป็นแบบ federated แนวทางนี้ข้อมูลบางส่วนอาจถูกจัดเก็บไว้ใน server ของ site แห่งหนึ่งขณะที่ข้อ



ภาพที่ 4.2 สถาปัตยกรรมแบบ Federated และแบบ Tiered

มูลอีกส่วนหนึ่งอาจถูกแยกออกไปอยู่อีกส่วนหนึ่ง

1.3 การเลือกใช้สถาปัตยกรรมแบบลำดับขั้น (Tiered Architecture) แนวทางนี้จะเก็บข้อมูลที่มีข้อมูลกระชับที่สุดบน workstation ของผู้ใช้ในระดับแรก ขณะที่ข้อมูลบน server ในระดับที่สองจะมีรายละเอียดเพิ่มขึ้น และมีรายละเอียดข้อมูลสูงที่สุดในระดับที่สาม ดังภาพที่ 4.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. สถาปัตยกรรมของ Server

สถาปัตยกรรมของ server ที่จะเก็บและส่งผ่านข้อมูล ตั้งแต่ขนาดของ server ที่จะนำมาใช้ การขยายระบบ การดูแลระบบ สิ่งเหล่านี้มีอิทธิพลต่อการเลือกสถาปัตยกรรมของ server ทั้งสิ้น จึงควรพิจารณาให้เหมาะสมในการที่จะนำไปใช้งาน โดยที่สถาปัตยกรรมของ server แบ่งเป็น 3 ประเภท คือ

2.1 Single processor server เป็น server ที่มี processor เพียงตัวเดียว server ประเภทนี้จัดการได้ง่าย แต่มีข้อจำกัดในด้านของการประมวลผล และความสามารถในการขยายระบบ รวมทั้งยังมี single point of failure ทำให้มีความเสี่ยงสูง และมีความน่าเชื่อถือของระบบไม่ดีนัก

2.2 Symmetric multiprocessing server เป็น server ที่มีการเพิ่ม processor ทำให้สามารถเพิ่ม throughput ได้ มากขึ้น โดยที่ยังใช้หน่วยความจำภายใน และอุปกรณ์เก็บข้อมูลเดียวกัน แต่การขยายระบบของ server ประเภทนี้ มีข้อจำกัดที่จำนวนของ processor ที่เครื่องสามารถขยายได้นั้น ขึ้นอยู่กับการออกแบบ backplane และ share bus ของแต่ละรุ่น

2.3 Massively parallel processing server เป็น server ที่ใช้ Link ความเร็วสูงและ bandwidth สูงเชื่อมต่อกับ array processor

3. สถาปัตยกรรมของ DBMS

ระบบจัดการฐานข้อมูล DBMS เป็นสิ่งที่สำคัญ ต้องมีการพิจารณาเลือกใช้ให้เหมาะสม และสามารถสนับสนุนการทำงานของ server ที่ต้องการด้วย ในปัจจุบันมี DBMS ที่นิยมใช้อยู่ 3 ประเภท ได้แก่

3.1 Relational database management system DBMS ประเภทนี้จะมีความสามารถในการเรียกค้นข้อมูลแบบทันที และไม่มีความซ้ำซ้อน

3.2 Multidimensional database เป็นการจัดโครงสร้างฐานข้อมูลแบบหลายมิติโดยใชเทคนิคด้าน storage และ operators เข้ามาช่วย เพื่อให้สนับสนุนลักษณะของโครงสร้างข้อมูล

3.3 Object-relational database DBMS ระบบฐานข้อมูลประเภทนี้ จะช่วยสนับสนุนการจัดการกับข้อมูลชนิดพิเศษ เช่น ภาพ เอกสาร วิดีโอ และข้อมูลที่มีความซับซ้อนรวมทั้งข้อมูลในลักษณะที่เป็น object oriented ด้วย

4.4 คุณลักษณะของข้อมูลใน Data Warehouse

ข้อมูลที่นำเข้ามาเก็บใน data warehouse นั้น จะมาจากแหล่งข้อมูล จากสภาพแวดล้อมต่างกัน จึงต้องมีการกำหนดคุณลักษณะของข้อมูลที่มีการจัดเก็บให้เหมาะสม ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.4.1 Subject oriented ข้อมูลที่สร้างขึ้นใน data warehouse จะประกอบด้วยเรื่องที่น่าสนใจ ซึ่งจะเป็นการเก็บเฉพาะข้อมูลข่าวสารที่จำเป็นต่อกระบวนการตัดสินใจขององค์กร

4.4.2 Integrated การรวบรวมข้อมูลจากแหล่งต่างๆ ซึ่งมีการกำหนดค่าของตัวแปรตัวเดียวกัน ซึ่งเป็นค่าที่ต่างกัน เมื่อนำมาเก็บใน data warehouse แล้วจะต้องกำหนดค่าของตัวแปรเหล่านั้นให้เหมือนกัน เพื่อให้เกิดความสอดคล้องของข้อมูล

4.4.3 Time variancy ข้อมูลที่นำมาจัดเก็บจะเป็นข้อมูลย้อนหลัง ซึ่งมีระยะเวลาประมาณ 5 ถึง 10 ปี เพื่อประโยชน์ในการที่จะนำข้อมูลมาเปรียบเทียบกัน

4.4.4 Nonvolatile ข้อมูลใน data warehouse จะต้องควบคุมเพื่อไม่ให้เกิดการเปลี่ยนแปลงจากผู้ใช้ ได้ง่าย โดยที่ผู้ใช้จะมีความสามารถในการเรียกใช้ และเข้าถึงข้อมูล (access) เท่านั้น

4.5 การแปลงข้อมูลเข้าสู่ Data Warehouse

คุณภาพในการแปลงข้อมูล เป็นสิ่งที่สำคัญมากสำหรับ data warehouse ความซับซ้อนของขบวนการแปลงข้อมูล จะแตกต่างกันไปสำหรับ data warehouse แต่ละระบบ รวมทั้งโครงสร้างของแหล่งข้อมูล ดังนั้นก่อนการแปลงข้อมูลต้องมีการออกแบบ จัดเตรียม physical model ของข้อมูลสำหรับ data warehouse และทำการสร้างขอบเขตของเป้าหมายให้เสร็จสมบูรณ์ก่อนหลังจากนั้นจึงดำเนินการตามกระบวนการ ดังนี้

1. การทำแผนการแปลงข้อมูล
2. การสร้างกฎเกณฑ์ของการแปลงข้อมูล
3. การดึงข้อมูลเข้าสู่ intermediate schema
4. การรวบรวมข้อมูล
5. ความถูกต้องของข้อมูล

1.การทำแผนการแปลงข้อมูล

กระบวนการแรก คือ การพัฒนาแผนการแปลงข้อมูล โดยที่จะต้องเข้าใจในความต้องการพื้นฐานและการไหลของข้อมูล เพื่อให้กระบวนการต่างๆ สามารถเชื่อมโยงเข้าด้วยกันอย่างเหมาะสม ได้แก่

1.1 การย้ายข้อมูลจากระบบเดิมเข้าสู่บริเวณที่ใช้ปรับแต่งข้อมูล (staging area) เพื่อปรับแต่งข้อมูลให้เป็นไปตามความต้องการขององค์กร

1.2 การสร้างเงื่อนไขการปรับปรุงข้อมูล การแปลงและการรวมข้อมูลเข้าด้วยกัน

1.3 การสร้าง primary key และ foreign key ให้เชื่อมโยงกัน

1.4 การย้ายข้อมูลลงสู่ server ของ data warehouse

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.5 การเก็บรายละเอียดของข้อมูลที่มีการปรับปรุง และส่งไปเก็บใน metadata ของ data warehouse

1.6 การเก็บข้อมูล และทำดัชนีอ้างอิงใน warehouse

1.7 การเตรียมการเพื่อจะรองรับข้อมูลที่จะเพิ่มขึ้นใน data warehouse

1.8 การตรวจสอบความถูกต้องของข้อมูลใน data warehouse

2. การสร้างกฎเกณฑ์ของการแปลงข้อมูล

ในกระบวนการสร้างกฎเกณฑ์การแปลงข้อมูลนั้น ต้องมีการวิเคราะห์แหล่งข้อมูลอย่างละเอียดรอบคอบ เพื่อกำหนด attribute ของข้อมูล การ map ข้อมูลจากที่เดิมไปเก็บใน warehouse และ logic ที่ใช้ในการเคลื่อนย้ายข้อมูล รวมทั้งการกำหนดสารสนเทศภายนอก แล้วจึงทำเป็น model ช่วงคราว หลังจากนั้นจึงเป็นการจัดทำโปรแกรม (coding program) ซึ่งต้องประกอบด้วย routine ที่มีการทำหน้าที่ดังนี้

2.1 การแปลงข้อมูลจาก intermediate schema มาเป็นข้อมูลที่พร้อมจะ load

2.2 การรวบรวมข้อมูลใน staging area เข้าด้วยกัน

2.3 การย้ายข้อมูลจาก staging area เข้าสู่ server ของ data warehouse (ถ้าข้อมูลใน staging area ไม่ได้อยู่ใน server เดียวกันกับ warehouse)

2.4 load ข้อมูลลงสู่ data warehouse

3. การดึงข้อมูลเข้าสู่ intermediate schema

Intermediate schema เป็น interface ของข้อมูลที่ดึงมาจากทุกระบบ โดยการดึงข้อมูลจะอ่านข้อมูลจากแหล่งข้อมูล แปลงข้อมูลเข้าสู่ intermediate schema และปรับแต่งข้อมูลที่ดึงมาให้เป็นไปตามความต้องการขององค์กร แล้วจึงทำการ execute routine ในการแปลงข้อมูลเพื่อทำความสะอาดข้อมูล (cleaning data) โดยการทำความสะอาดข้อมูลเป็นการตรวจสอบความถูกต้อง และแก้ไขข้อผิดพลาดใน field ข้อมูล โดยมีองค์ประกอบหลักในการปรับปรุงข้อมูลดังต่อไปนี้

3.1 การตรวจสอบข้อมูล เพื่อดูคุณภาพของข้อมูล (quality of data) รูปแบบของข้อมูล (format of data) และจำนวนของ field ต่างๆ ที่นำมาใช้

3.2 การกระจายข้อมูล เพื่อที่จะค้นหาข้อมูล และจุดหมายของแต่ละองค์ประกอบในแต่ละ field

3.3 การตรวจสอบข้อมูลว่า ตรงกับข้อมูลเดิมที่มีอยู่ และ field ทั้งหมดมีความถูกต้อง หรือได้รับการตรวจแก้ไขโดยอัตโนมัติ

3.4 การตรวจดูว่าข้อมูลใน record ตรงกัน และข้อมูลเป็นชนิดเดียวกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4. การรวบรวมข้อมูล

การรวบรวมข้อมูล เป็นการนำเอาข้อมูลทั้งหมดที่เป็นลักษณะตาราง (dimension table) และข้อมูลจริงลงสู่ data warehouse โดยข้อมูลจะถูกรวบรวมตามการออกแบบใน data warehouse ซึ่งต้องดำเนินการดังนี้

4.1 ทำการ execute ตามลำดับของการจัดเรียง และทำการสร้าง primary key และ foreign key ทำให้ได้ข้อมูลที่เป็นตาราง และข้อมูลจริง

4.2 ย้ายข้อมูลจาก staging area ลงสู่ server ของ data warehouse

5. ความถูกต้องของข้อมูล

การตรวจสอบความถูกต้องของข้อมูล เป็นสิ่งที่ต้องกระทำตลอดกระบวนการแปลงข้อมูล โดยจะต้องมีกระบวนการตรวจแก้ความถูกต้องของข้อมูล กระบวนการสำหรับทำให้ข้อมูลสอดคล้องกับระบบเดิม รวมถึงการตรวจสอบดูสภาพแวดล้อมในปัจจุบัน เพื่อกำหนดวิธีการที่ดีที่สุด เพื่อเคลื่อนย้ายข้อมูลจากแหล่งข้อมูลเดิมไปยัง data warehouse ซึ่งจะทำได้ดังนี้

5.1 การตรวจสอบผลรวมทั้งหมดของจำนวนข้อมูลที่ดึงมาจากข้อมูลเดิม กับข้อมูลที่เพิ่มเข้าไป ใน intermediate schema

5.2 การตรวจแก้ไขข้อมูลในระบบเดิมของแหล่งข้อมูล หรือใน routine ของการแปลงข้อมูล ซึ่งควรจะเก็บข้อมูลจากการตรวจแก้ไว้ใน metadata ของการแปลงข้อมูลด้วย

5.3 การตรวจสอบค่าของข้อมูลให้ถูกต้องในกระบวนการรวบรวมข้อมูล

4.6 สิ่งที่ต้องคำนึงถึงในการออกแบบ Data Warehouse

การออกแบบ data warehouse มีสิ่งที่ต้องพิจารณา 3 ส่วนด้วยกันได้แก่

1. Target database
2. Data Extraction
3. Front-End Access Tools

1. Target database

สิ่งที่ต้องพิจารณาเป็นอันดับแรก คือ ลักษณะข้อมูลที่จะนำมาใช้ว่า จะเป็นแบบจำลอง (model) ชนิดใด เช่น Relational database , Multidimensional database เพื่อการเลือกใช้ DBMS ให้เหมาะสมต่อการใช้งาน

2. Data Extraction Technique

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในการที่จะดึงข้อมูลมาใช้ โดยการแยกข้อมูลบางส่วน จะใช้วิธีการจัดการโดยการใช้ facility ที่มีมากับ DBMS นั้น หรือจะใช้เครื่องมือที่จัดทำมาเพื่อทำการดึงข้อมูลโดยเฉพาะที่ผู้ผลิตได้จัดทำขึ้นมา

2. Front – End Data Access Tools

ประเภทของเครื่องมือที่จะนำมาใช้ในการที่จะเข้าถึงข้อมูล (access) บน data warehouse ควรจะเป็นเครื่องมือที่ใช้ในการค้นหาข้อมูลจาก relation หรือ multidimensional database ได้

4.7 Software ที่ใช้จัดการข้อมูล

เครื่องมือที่ใช้ในการจัดการข้อมูล ไม่ว่าจะเป็นในด้านการดึงข้อมูล การแปลงข้อมูล การคัดลอกข้อมูล และการดูแลฐานข้อมูล จะมีบริษัทต่างๆ ที่พัฒนา software ขึ้นมาช่วยจัดการทางด้านข้อมูล สำหรับ data warehouse ได้แก่

1.Prism Warehouse Manage

1.1 Warehouse Manager ช่วยทำการดึงข้อมูลจากฐานข้อมูลส่วนปฏิบัติงานต่างๆ นำมาทำการแปลงข้อมูล แล้วจึงจะ load ข้อมูลเข้าสู่ data warehouse หลังจากนั้นจะทำการสร้าง และจัดการ metadata ให้โดยอัตโนมัติ ผู้ใช้สามารถที่จะกำหนดการ mapping ระหว่างแหล่งข้อมูลที่ต้องการ กับ data warehouse ได้

1.2 Changed Data Capture เป็นการทำ refresh ข้อมูลโดยทำการ reload ข้อมูลเข้าสู่data warehouse ตามช่วงระยะเวลาที่กำหนด

1.3 Directory Manager ช่วยทำการจัดเก็บข้อมูลต่างๆ ของ metadata ซึ่งได้แก่รายละเอียดของข้อมูลในตารางฐานข้อมูล attribute ต่างๆ และการ mapping

2. IBM's Information Warehouse

2.1 Datapropagator Relation เป็นการถ่ายข้อมูล โดยในการถ่ายข้อมูลจากแหล่งต้นทาง ไปยังปลายทาง จะต้องมีการจัดเก็บข้อมูลที่เปลี่ยนแปลงไว้ในตารางกลางก่อน แล้วจึงคัดลอกข้อมูลส่งไปยังฐานข้อมูลที่ต้องการ

2.2 DataPropagator Nonrelational เป็นการช่วยในการทำให้การถ่ายทอดข้อมูลจากฐานข้อมูลในแบบ nonrelational ไปสู่ฐานข้อมูลในแบบ relational ได้

2.3 Data Refresher เป็นการย้ายข้อมูลที่มีอยู่ทั้งหมดไปสู่ data warehouse โดยทีข้อมูลจะถูก reload และแทนที่ใหม่ทุกครั้ง

2.4 DataJoiner เป็นการใช้เทคโนโลยีด้าน optimizer เพื่อเชื่อมข้อมูลจากระบบฐานข้อมูลที่ต่าง product กันให้เข้ามาเป็นฐานข้อมูลเดียวกัน

2.5 Datahub เป็นเครื่องมือที่ช่วยงานการดูแลฐานข้อมูล ที่มี warehouse

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กระจายอยู่บนระบบที่ต่างกัน ให้ดูแลได้ง่าย และประหยัดเวลา

2.6 DataGuide เป็นเครื่องมือที่ช่วยในการจัดหมวดหมู่สารสนเทศ ทำให้ผู้ใช้สามารถมองเห็นข้อมูลได้ง่าย

2.7 Lotus NotePump เป็นการดึงข้อมูลจากฐานข้อมูลของ Lotusnote เข้าไปสู่ stagingarea ของ DataPropagator เพื่อช่วยในการทำ replication ไปยัง data warehouse ที่ต้องการ

3. Sybase Warehouse Works

3.1 Sybase Replication Server 11 เป็นการทำ replication โดยที่จะทำการอ่าน transaction log ของฐานข้อมูล แล้วจึงจะทำการ replicate ข้อมูลที่มีการเปลี่ยนแปลงไปยัง data warehouse

3.2 Tranzic InfoPump ช่วยในด้านของการจัดการแปลงข้อมูล และจัดทำการปรับปรุงข้อมูล โดยมีการใช้ภาษา script และ trigger

3.3 Sybase IQ ช่วยในการค้นหาข้อมูลสำหรับการ adhoc โดยวิธีการทำ index optimizer ทำให้ได้ผลลัพธ์ที่รวดเร็ว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

ระบบสนับสนุนการตัดสินใจ (Decision Support System)

5.1 ความหมายของระบบสนับสนุนการตัดสินใจ

เป็นระบบที่ช่วยในการสนับสนุนการตัดสินใจของผู้บริหาร โดยจะทำการรวบรวมข้อมูลทั้งหมดที่เกิดขึ้น เพื่อที่จะทำนายผลที่จะเกิดขึ้นในอนาคต เป็นการให้ผู้ตัดสินใจได้รู้ว่าเหตุการณ์ที่เกิดขึ้นจะเป็นอย่างไร โดยในระบบนี้จะมีแบบจำลองการตัดสินใจในสภาวะต่างๆ ที่ช่วยในการตัดสินใจของผู้มีหน้าที่ได้พิจารณาว่า ถ้าหากเลือกที่จะตัดสินใจทางใดทางหนึ่งแล้ว ผลที่เกิดขึ้นจะเป็นอย่างไร ทำให้ผู้บริหารสามารถทดสอบได้ ระบบนี้จึงเป็นระบบที่ต้องจัดการกับข้อมูลที่มีปริมาณมาก โดยจะทำการนำเอาข้อมูลทั้งหมดมาประมวลผล เพื่อหาทางเลือกที่ดีที่สุดที่มีอยู่ ระบบนี้มักจะใช้ในด้านของการตลาดและการขาย

5.2 ลักษณะของระบบ

1. เป็นระบบที่จำลองแบบการตัดสินใจ เพื่อที่จะช่วยทำให้ผู้บริหารสามารถทดสอบการตัดสินใจได้

2. เป็นระบบที่ช่วยสนับสนุนการตัดสินใจ แต่ไม่ใช่ระบบที่ตัดสินใจแทนผู้บริหาร

3. ระบบนี้เหมาะกับการตัดสินใจแบบไม่มีโครงสร้าง และกึ่งโครงสร้าง ซึ่งเป็นสารสนเทศ สำหรับช่วยในการตัดสินใจ

4.เป็นการใช้ข้อมูลจากระบบ Transaction Processing (TPS) และระบบ Management Information System (MIS)

5.ระบบต้องมีความสามารถที่จะค้นหาข้อมูลที่ต้องการได้อย่างรวดเร็ว

6.การค้นหาข้อมูลจะเกี่ยวข้องกับรูปแบบของข้อมูลในหลายๆ รูปแบบ ข้อมูลที่ต้องการค้นหาอาจจะเป็นกราฟ รูปภาพ หรือเป็นภาพเคลื่อนไหว

7.การค้นหาข้อมูลจะมีความซับซ้อนมากกว่าปกติ เนื่องจากข้อมูลที่นำมาใช้นั้น จะมีปริมาณมาก

8.ข้อมูลที่ถูกจัดเก็บไว้ จะไม่มีการเขียนข้อมูลลงไป (update) ส่วนใหญ่จะเป็นการอ่านข้อมูลเพียงอย่างเดียว (read only)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.3 ส่วนประกอบของระบบ

ระบบสนับสนุนการตัดสินใจ มีส่วนประกอบที่สำคัญ 3 ส่วนคือ

1. ฐานข้อมูลจำลอง (Model Base)
2. ฐานข้อมูลจริง (Data Base)
3. ฐานข้อมูลคำสนทนาโต้ตอบกับผู้ใช้ระบบ (Dialogue Base)

5.4 ขั้นตอนของการจัดทำระบบสนับสนุนการตัดสินใจ

ในการที่จะจัดทำระบบสนับสนุนการตัดสินใจ จะมีการแบ่งขั้นตอนของการทำงานออกเป็น 10 ขั้นตอนดังนี้

1. Project planing
2. Gathering data requirement and modeling
3 .Physical database design and development
4. Data sourcing , integrating and mapping
5. Populating the data warehouse
6. Automating the data management process
7. Creating the stator set of reports
8. Data validation and testing
9. Training
10. Rollout

1.Project planing

การวางแผนสำหรับการจัดทำระบบคลังข้อมูล เพื่อช่วยสนับสนุนการตัดสินใจประกอบด้วย

1.1. กำหนดขอบเขตของโครงการที่จะจัดทำให้ชัดเจน

1.2 จัดทำแผนงานของโครงการ

1.3 กำหนดทรัพยากรที่จำเป็นต้องใช้ ได้แก่ ข้อมูลทั้งภายในและภายนอกองค์1

1.4 กำหนดผู้มีส่วนรับผิดชอบในการทำงาน

1.5 กำหนดลักษณะของงานที่ต้องจัดทำทั้งหมด

1.6 กำหนดระยะเวลาต่างๆ ในการทำงาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในขั้นตอน Project planing นี้ จะต้องคำนึงถึงการใช้เทคโนโลยีต่างๆ เช่น ระบบเครือข่ายที่จะนำมาใช้งาน (LAN หรือ WAN) การเลือกใช้ plate from การเลือกใช้ database และการเลือกใช้ front – end access tool

2. Gathering data requirements and modeling

ข้อมูล ในลักษณะต่างๆ จากผู้ใช้ ประกอบด้วย

2.1 การรวบรวมความต้องการในการใช้ข้อมูล โดยการสัมภาษณ์ผู้ใช้

2.2 จัดทำเอกสารที่เกี่ยวข้อง

2.3 กำหนด logical model

2.4 สร้าง dimensional business model

3. Physical database design and development

ขั้นตอนนี้เป็นการออกแบบฐานข้อมูล และการทำ denormalized ข้อมูล รวมถึงการออกแบบฐานข้อมูลในแบบ Star schema ดังนี้

3.1 การออกแบบฐานข้อมูลที่จะใช้

3.2 การทำ denormalized ข้อมูล

3.3 กำหนด key ประเภทต่างๆ ของข้อมูลที่อยู่บนฐานข้อมูล

3.4 สร้าง index key

4. Data sourcing integration and mapping

ขั้นตอนนี้เป็นการเลือกแหล่งข้อมูลที่จะนำมาใช้ การรวมข้อมูล และการดึงข้อมูลได้แก่

4.1 กำหนดแหล่งของข้อมูลที่จะใช้

4.2 กำหนดรูปแบบของข้อมูล

4.3 วิเคราะห์ข้อมูล เพื่อทำการรวบรวมข้อมูลเข้าสู่ data warehouse

4.4 จัดทำการแปลงข้อมูล

4.5 ทำการดึงข้อมูลจากแหล่งต่างๆ

5. Populating the Data Warehouse

เป็นขั้นตอนของการทำ extrace, convert และ load data เข้าสู่ฐานข้อมูลของdata warehouse ประกอบด้วย

5.1 การพัฒนาโปรแกรม หรือใช้เครื่องมือที่จะ extract และ move data

5.2 การพัฒนาและเลือกวิธีการที่จะ load data

5.3 การพัฒนาโปรแกรม หรือเลือกใช้เครื่องมือ เพื่อทำการ convert data

5.4 การพัฒนาและปรับปรุงวิธีการในการ refresh data

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.5 จัดทำการทดสอบกระบวนการต่างๆ ได้แก่ การ extract , การ integration data

6.Automating the data management process

เป็นขั้นตอนในการที่จะทำการ extract, integration และ load data เข้าสู่ data warehouse ให้เป็นไปอย่างอัตโนมัติ ประกอบด้วย

6.1 กระบวนการในการทำ data extract ion อย่างอัตโนมัติตามเวลาที่กำหนด

6.2 กระบวนการในการทำ data conversion อย่างอัตโนมัติตามเวลาที่กำหนด

6.3 กระบวนการในการทำ data load อย่างอัตโนมัติตามเวลาที่กำหนด

6.4 การจัดทำขั้นตอนในการ back up และ recovery

6.5 จัดทำการทดสอบระบบการทำงานที่เป็นไปอย่างอัตโนมัติในทุกขั้นตอน

7. Creating the starter set of reports

ขั้นตอนนี้เป็นการสร้าง และจัดทำรายงานต่างๆ ได้แก่

7.1 จัดสร้างรูปแบบของรายงานเบื้องต้น

7.2 ทดสอบผลที่ได้รับจากการรายงาน

7.3 จัดทำเอกสาร และคู่มือในการใช้งาน

8. Data validation and testing

ขั้นตอนนี้เป็นการตรวจสอบ และทดสอบข้อมูล เพื่อให้ข้อมูลที่ได้มานั้นมีความถูกต้องมากที่สุด ประกอบด้วย

8.1 การตรวจสอบข้อมูล โดยดูจากผลลัพธ์ในรายงานเบื้องต้นที่จัดทำ

8.2 การตรวจสอบข้อมูล โดยใช้วิธีการกำหนดมาตราฐานขึ้นมา

8.3 การทดสอบ โดยการเปลี่ยนแปลงข้อมูล

9. Training

ขั้นตอนนี้เป็นการฝึกอบรมให้ผู้ใช้ได้ทราบวิธีการต่างๆ ในการที่จะใช้งานระบบ เช่น ข้อมูล การใช้ front – end เป็นต้น ซึ่งได้แก่

9.1 ขอบเขตของข้อมูลที่อยู่ใน data warehouse

9.2 วิธีการในการใช้เครื่องมือที่เป็น front – end

9.3 วิธีการที่จะเข้าถึง และดึงข้อมูลไปใช้งาน

9.4 รายงานเบื้องต้นที่ผู้ใช้สามารถกระทำได้

10.Rollout

ขั้นตอนนี้เป็นการติดตั้งระบบ และการบำรุงรักษาระบบ ได้แก่

1.1 จัดทำการติดตั้งระบบให้ผู้ใช้ ได้ใช้งานจริง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.2 จัดทำขั้นตอนในการพัฒนา เพื่อจัดทำรายงานในรูปแบบต่างๆ

1.3 จัดทำวิธีการในการที่จะทำการ back up ข้อมูล และระบบงาน

1.4 จัดทำแนวทางในการที่จะทำให้ข้อมูลที่อยู่ในระบบทั้งหมด มีความถูกต้องและสอดคล้องกัน

5.5 คุณลักษณะข้อมูลของระบบสนับสนุนการตัดสินใจ

ข้อมูลของระบบสนับสนุนการตัดสินใจ จะเป็นข้อมูลที่มีความแตกต่างจากข้อมูลในระบบอื่น เพราะเป็นข้อมูลที่ได้ผ่านการปรับเปลี่ยนให้อยู่ในรูปแบบที่เหมาะสมต่อการนำมาใช้งาน โดยมีลักษณะดังนี้

1. เป็นข้อมูลที่ได้มาจากแหล่งข้อมูลหลายแห่ง
2. เป็นข้อมูลผลสรุป
3. ข้อมูลเป็นแบบ read – only ไม่มีการ update ลงไป
4. ข้อมูลจะเป็นแบบ denormalized
5. โครงสร้างของข้อมูลมีความยืดหยุ่น
6. ข้อมูลเป็นลักษณะของการวิเคราะห์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 6

เครื่องมือที่ใช้ในการวิเคราะห์ข้อมูล (Tools for Analyse Data)

ในการที่จะนำข้อมูลที่จัดเก็บไว้ใน data warehouse มาใช้ประโยชน์นั้น จำเป็นต้องมีเครื่องมือที่มีความสามารถในการที่จะวิเคราะห์ข้อมูลเหล่านั้น ได้อย่างมีประสิทธิภาพและเหมาะสม รวมทั้งมีความสามารถในการจัดการข้อมูลในลักษณะที่เป็น Online Analytical Processing ด้วย

Pilot Decision Support Suite เป็นโปรแกรมที่ได้มีการจัดเตรียมการวิเคราะห์ข้อมูลในรูปแบบต่างๆ เพื่อสนับสนุนความต้องการใช้ข้อมูลของผู้บริหาร และช่วยในการพัฒนาระบบงานของผู้พัฒนาระบบได้เป็นอย่างดี

ภาพที่ 6.1 ชุดของโปรแกรม Pilot Decision Support Suite

6.1 ข้อกำหนดความต้องการของระบบ

Pilot Decision Support Suite สามารถติดตั้งเพื่อให้ผู้ใช้ ใช้งานได้ในหลายรูปแบบ ไม่ว่าจะเป็น single user, workgroup และ client/server โดยสนับสนุนการทำงานบนเครื่องคอมพิวเตอร์ที่มี platform ต่างๆ ได้แก่ Windows 3.1 ,Windows95, Windows NT และ UNIX

6.2 ส่วนประกอบของ Pilot Decision Support Suite

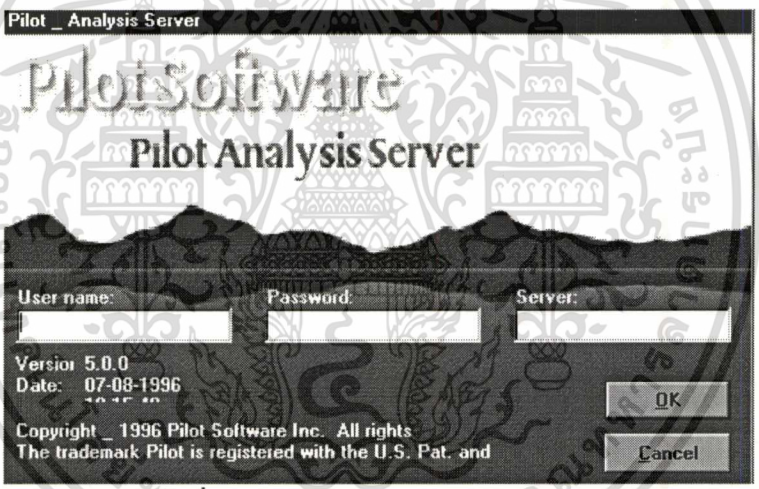
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โปรแกรม Pilot Decision Support Suite เป็นโปรแกรมที่ถูกพัฒนาขึ้นมา เพื่อใช้ในการวิเคราะห์ข้อมูล ซึ่งจะประกอบด้วย โปรแกรมที่สำคัญ 4 ส่วนคือ

1. Pilot Analysis Server
2. Pilot Discovery Server
3. Pilot Desktop
4. Pilot Designer

1. Pilot Analysis Server

เป็นโปรแกรมที่ใช้การสนับสนุนการทำงานของผู้ใช้หลายคน และช่วยในการควบคุมรวมทั้งจัดการฐานข้อมูลที่อยู่บนฐานข้อมูล



ภาพที่ 6.2.1 Pilot Analysis Server

2. Pilot Discovery Server

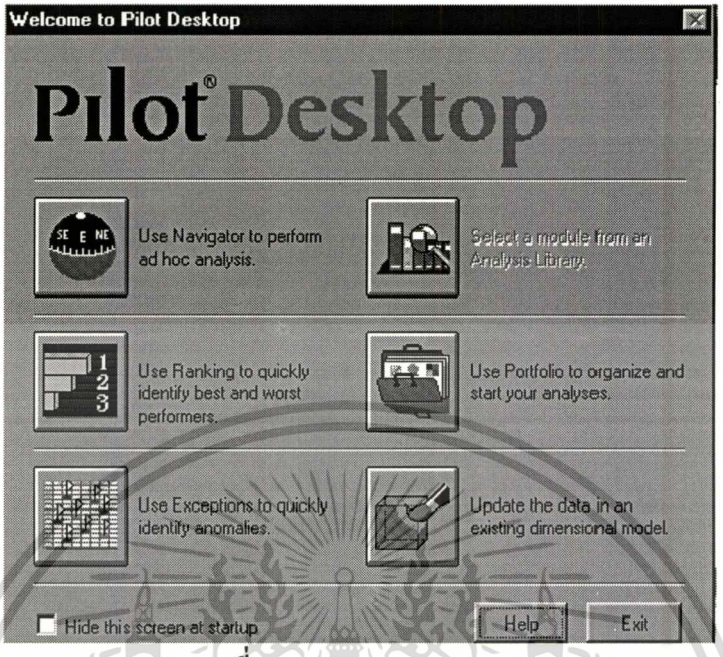
เป็นโปรแกรมที่ใช้ช่วยสนับสนุนการทำงานกับข้อมูลที่อยู่บน data warehouse ที่มีขนาดใหญ่และข้อมูลมีความซับซ้อนมาก

3. Pilot Desktop

เป็นโปรแกรมที่อำนวยความสะดวกให้กับผู้ใช้ เพื่อให้ผู้ใช้สามารถใช้งานได้ง่าย โดยการ

3.1 เลือกการวิเคราะห์ข้อมูลได้หลายรูปแบบ ซึ่งได้แก่ Navigator Adhoc, Ranking Analysis.Exception Analysis และ Portfolio Navigator Adhoc ใช้สำหรับการตอบคำถาม (Adhoc) หรือจัดทำรายงานแบบเร่งด่วนให้ สามารถแสดงผลในรูปแบบของกราฟประเภทต่างๆ ได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 6.2.2 Pilot Desktop

3.2 Ranking Analysis ใช้สำหรับการจัดลำดับของข้อมูลตาม dimension ที่น่าสนใจโดยการเรียงข้อมูลจากต่ำสุดไปสูงสุด หรือในทางกลับกัน จากสูงสุดไปต่ำสุดได้

3.3 Exception Analysis ใช้ในการวิเคราะห์ข้อมูล และนำข้อมูลในแต่ละ Attribu มาเปรียบเทียบเพื่อคำนวณหาแนวโน้มได้

3.4 Portfolio ใช้จัดเก็บผลการวิเคราะห์ข้อมูลในรูปแบบต่างๆ ไว้ เพื่อการเรียกดูหรือออกรายงานได้

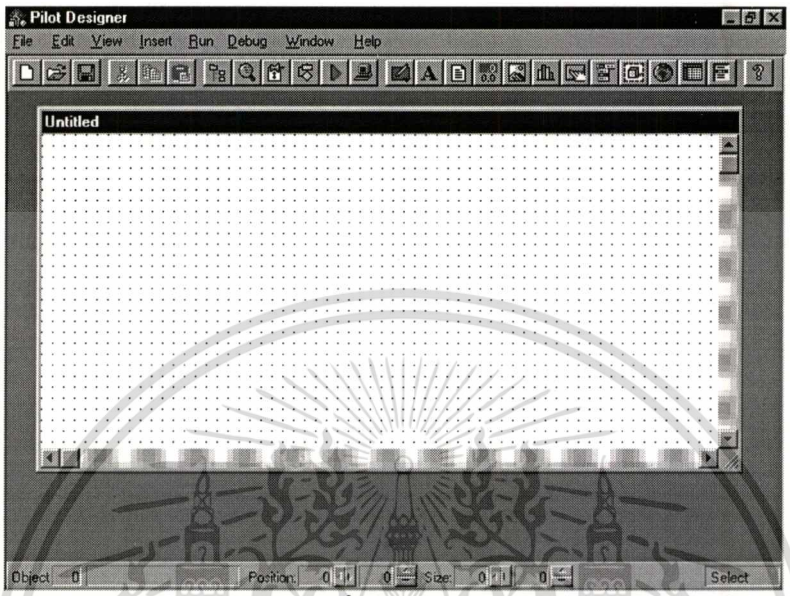
4. Pilot Designer

6.3 โปรแกรมที่ใช้สำหรับการออกแบบ และพัฒนาระบบงาน

ใช้พัฒนาตามรูปแบบที่ผู้ใช้ต้องการ คุณสมบัติของ Pilot Decision Support Suite จะ

1. เป็นระบบ data warehouse ที่เก็บรวบรวมข้อมูลที่ผ่านการประมวลผลแล้วมาใช้
2. สนับสนุนการทำงานระบบ Online Analytical Processing (OLAP)
3. สนับสนุนการค้นหาข้อมูลในแบบ Adhoc query
4. สามารถสร้างฐานข้อมูลแบบหลายมิติได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 6.2.3 Pilot Designer

5. ไม่เหมาะสมสำหรับการจัดการฐานข้อมูลที่เป็นแบบ Transaction processing
6. ผู้ใช้สามารถใช้งานได้ง่าย โดยการใช้ Pilot Desktop หรือพัฒนา Application ได้เองโดยการใช้ Pilot Designer
7. สามารถทำงานกับข้อมูลได้หลายประเภท เช่น
   7.1 Microsoft Excel (แฟ้มข้อมูลที่เป็น .xls)
   7.2 Database file แฟ้มข้อมูลที่เป็น .dbf)
   7.3 ASCII text file (แฟ้มข้อมูลที่เป็น .txt)
   7.4 Relational database เช่น DB2, INFORMIX, INGRES, Oracle, SQL Server

6.4 การทำงานของ Pilot Decision Support Suite

โปรแกรม Pilot Decision Support Suite จะประกอบด้วยขั้นตอนการทำงานที่สำคัญ 4 ส่วนได้แก่

1 Link ID Configuration

2. Model

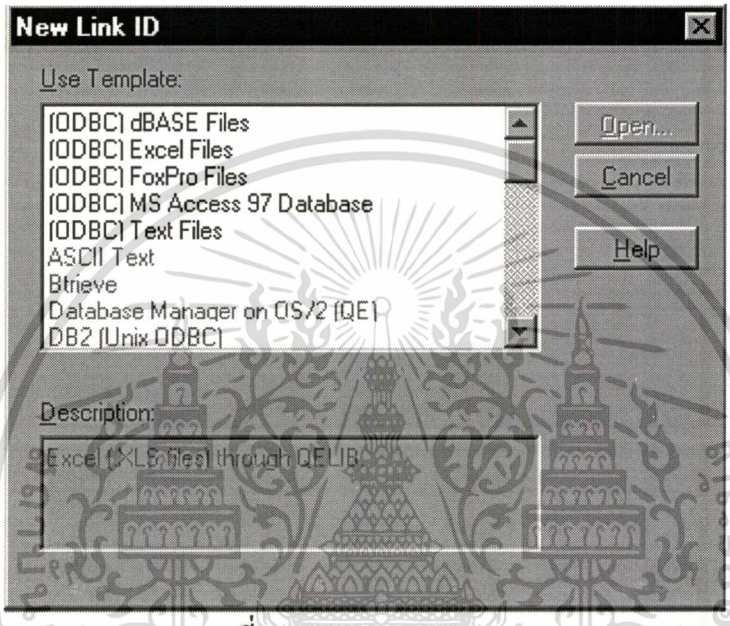
3. Builder

4. การวิเคราะห์ข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. Link ID Configuration

เป็นส่วนที่ทำหน้าที่ในการเชื่อมโยงข้อมูลจากฐานข้อมูลที่ต้องการใช้งานเข้ากับระบบ ซึ่งสามารถเลือกได้ว่า ฐานข้อมูลที่ต้องการใช้งานนั้นมีการจัดเก็บข้อมูลในรูปแบบใด เช่น dBASE, Excel, AXCII, text และ DB2 เป็นต้น



ภาพที่ 6.4.1 การ Link Configuration

2. Model Builder

ทำหน้าที่ในการสร้างรูปแบบ (model)ของข้อมูลที่ต้องการเพื่อการใช้งาน โดยจะเรียกใช้ข้อมูลจากฐานข้อมูลที่ได้เชื่อมโยงไว้แล้ว เข้ามาสู่ model ที่สร้างไว้

1.1 ในการสร้าง model เป็นขั้นตอนในการสร้างและกำหนดชื่อ Model

1.2 Construct Dimensions เป็นขั้นตอนที่จะกำหนด และจัดทำ dimension ของข้อมูล ที่ต้องการจะให้เป็นมุมมองในรูปแบบต่างๆ

1.3 Add Measures เป็นขั้นตอนในการที่จะกำหนด attribute ที่ต้องการให้เป็นค่าที่แสดงผล และกำหนดวิธีการที่ใช้ในการคำนวณ เช่นการหาค่าเฉลี่ย (average) การหาค่าผลรวม (sum) เป็นต้น ในการเลือกใช้ข้อมูลจากฐานข้อมูล สามารถที่จะเลือกใช้ข้อมูลจากตาราง (table) หรือใช้ข้อมูลจากตารางเดียวกันก็ได้ ขึ้นอยู่กับความต้องการใช้ attribute ที่ต้องการใช้ในตารางนั้น และในการสร้าง model จะต้องกำหนดช่วงเวลา (time) ให้กับข้อมูลด้วย

3. Analysis Server

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

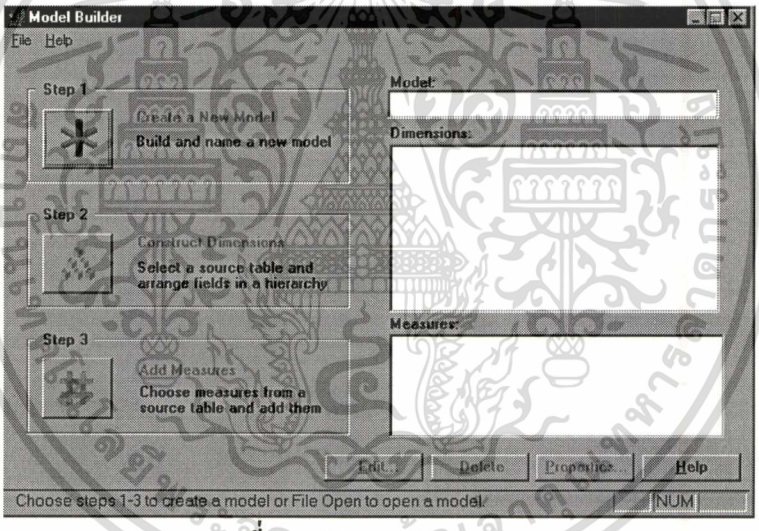
ขั้นตอนนี้เป็นการทำงานของระบบจัดการของโปรแกรม Pilot Decision Support Suite โดยการจัดเก็บข้อมูลที่ได้ผ่านการเชื่อมโยงกับฐานข้อมูล และจัดการกับ model ที่ได้สร้างขึ้นรวมเข้าไว้ด้วยกัน เพื่อใช้ในการวิเคราะห์ต่อไป

4. การวิเคราะห์ข้อมูล

ในการที่จะวิเคราะห์ข้อมูลตาม model ที่ได้สร้างไว้แล้วนั้น ทำได้ 2 วิธีได้แก่

4.1 การใช้ Pilot Desktop ซึ่งจะมี wizard ที่ช่วยให้ผู้ใช้สามารถใช้งานได้ง่ายเพื่อเลือกวิเคราะห์ข้อมูลที่สนใจได้

4.2 การใช้ Pilot Designer เพื่อพัฒนาระบบงานตามความต้องการได้ โดยการใช ภาษา Basic script



ภาพที่ 6.4.2 Model Builder

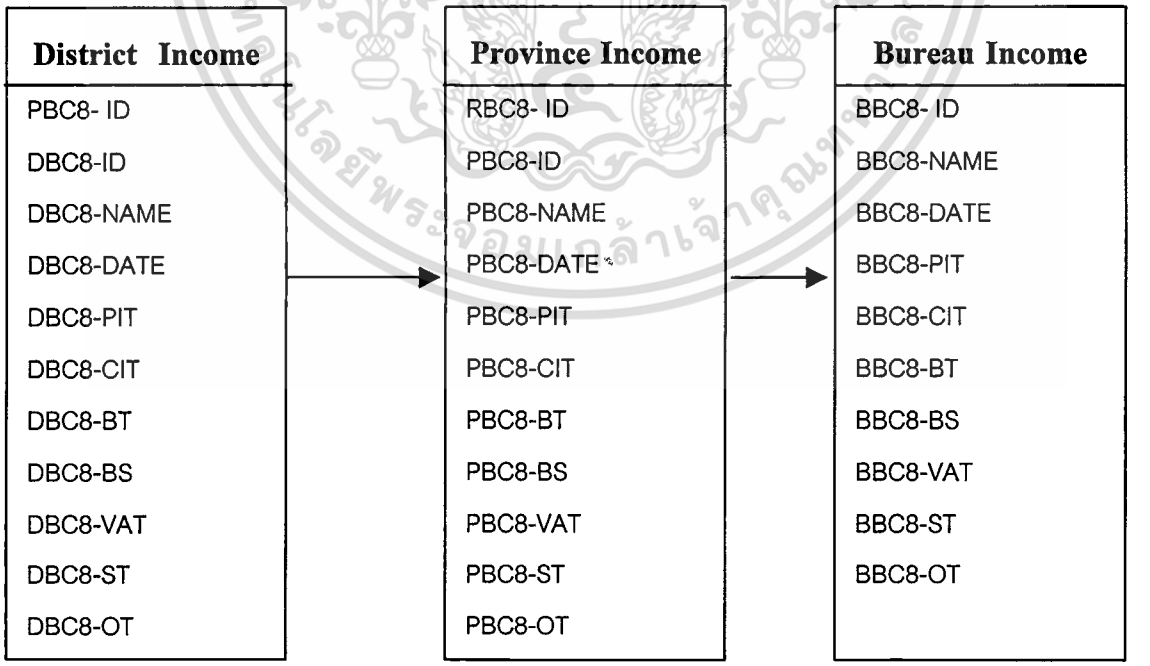
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 7

การพัฒนาระบบคลังข้อมูลผลการจัดเก็บภาษีของภาค

7.1 ขั้นตอนการดำเนินงาน

ศึกษารายละเอียด และโครงสร้างฐานข้อมูลของระบบงานภาษี โครงสร้างฐานข้อมูลของระบบการจัดเก็บภาษีของภาคข้อมูลผลการจัดเก็บภาษีเกือบจะทุกประเภทจะอยู่ที่เครื่อง POS ของทุก ๆ อำเภอ หรือที่รวบรวมเป็น บช.8 ดังนั้นอำเภอจะทำการรวบรวมข้อมูลดังกล่าวแล้วนำส่งให้กับสรรพากรจังหวัด สรรพากรจังหวัด ก็จะรวบรวมแล้วนำส่งให้กับสำนักงานสรรพากรภาค เพื่อประมวลผลและพิจารณาผลการจัดเก็บภาษีโดยส่วนวางแผน เพื่อนำเป็นผลการจัดเก็บภาษีประจำเดือน ไตรมาส ปี และ ยังนำไปใช้วิเคราะห์เปรียบเทียบผลการจัดเก็บกับปีก่อน เปรียบเทียบกับประมาณการ เพื่อเป็นแนวทางในการวางแผนแก้ไขการจัดเก็บภาษี โดยการพิจารณาของผู้บริหารในกรม และภาค แสดงในภาพที่ 7.1



ภาพที่ 7.1 แสดงความสัมพันธ์การจัดเก็บภาษีของภาค

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คำอธิบายภาพ

ตารางที่ 7.1 แสดงรายการภาษีที่อำเภอ จัดเก็บได้

ชื่อ	คำอธิบาย	ชนิดของข้อมูล	ขนาด (byte)
District Income	การจัดเก็บภาษีของอำเภอ	Character	20
PBC8_ ID	รหัสจังหวัดที่อำเภอส่งผลจัดเก็บให้	Dec	5,0
DBC8_ID	รหัสอำเภอที่ส่งผลการจัดเก็บ	Dec	5,0
DBC8_NAME	ชื่ออำเภอที่ส่งผลการจัดเก็บ	Character	20
DBC8_DATE	วันที่รายงานผลการจัดเก็บภาษี	Date	8
DBC8_PIT	ผลการจัดเก็บภาษีเงินได้บุคคลธรรมดา	Float	15
DBC8_CIT	ผลการจัดเก็บภาษีเงินได้นิติบุคคล	Float	15
DBC8_BT	ผลการจัดเก็บภาษีการค้า	Float	15
DBC8_BS	ผลการจัดเก็บภาษีธุรกิจเฉพาะ	Float	15
DBC8_VAT	ผลการจัดเก็บภาษีมูลค่าเพิ่ม	Float	15
DBC8_ST	ผลการจัดเก็บอากรแสตมป์	Float	15
DBC8_OT	ผลการจัดเก็บภาษีอื่นๆ	Float	15

ตารางที่ 7.2 แสดงภาษีที่จังหวัดรวบรวมจากอำเภอ

ชื่อ Attribute	คำอธิบาย	ชนิดของข้อมูล	ขนาด (byte)
Province Income	การรวบรวมผลจัดเก็บภาษีของจังหวัด	Character	20
RBC8_ ID	รหัสภาคที่จังหวัดส่งผลจัดเก็บให้	Dec	5,0
PBC8-NAME	ชื่อจังหวัดที่ส่งผลการจัดเก็บ	Character	20

ตารางที่ 7.3 แสดงภาษีที่ภาครวบรวมจากจังหวัด

ชื่อ Attribute	คำอธิบาย	ชนิดของข้อมูล	ขนาด (byte)
Bureau Income	การรวบรวมผลจัดเก็บภาษีของภาค	Character	20
BBC8_ID	รหัสภาคที่จังหวัดส่งผลจัดเก็บให้	Dec	5,0
BBC8_ID	รหัสภาคที่รวบรวมผลการจัดเก็บ	Dec	5,0
BBC8-NAME	ชื่อภาคที่รวบรวมการจัดเก็บ	Character	20

เอกสารนี้ใช้เพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

7.2 การวิเคราะห์ และออกแบบฐานข้อมูล สำหรับระบบคลังข้อมูล

การออกแบบฐานข้อมูลผลการจัดเก็บภาษีนี้ ได้ใช้เทคนิคการออกแบบแบบสตาร์ (Star schema) ซึ่งเป็นเทคนิคการออกแบบฐานข้อมูลเชิงวิเคราะห์ ที่แสดงความสัมพันธ์ระหว่างตาราง (table) และการรวบรวมกันข้อมูลเพื่อให้ได้ผลการสืบค้นอย่างรวดเร็ว ทำให้ผู้ใช้และผู้วิเคราะห์ระบบเข้าใจการไหลเวียนของข้อมูลได้ง่าย และสะดวกต่อการเปลี่ยนแปลงแก้ไขเพิ่มเติมตลอดช่วงเวลาของการพัฒนาระบบคลังข้อมูล

ในการออกแบบฐานข้อมูลสำหรับคลังข้อมูลผลการจัดเก็บโดยใช้เทคนิคแบบสตาร์นั้น มีตารางอยู่สองชนิดคือ

7.2.1 ตารางข้อเท็จจริง (Fact table) เป็นตารางที่ใช้เก็บข้อมูลที่เป็นข้อเท็จจริงที่เกิดขึ้นจากการประมวลผลผลการจัดเก็บภาษีทุกประเภท

7.2.2 ตารางคุณลักษณะ (Dimension table)เป็นตารางที่เก็บคุณลักษณะต่างๆ ของข้อมูล หรือเป็นคำอธิบายของข้อมูลที่สะท้อนถึงมิติขององค์กร เช่น สรรพากร ชื่อประเภทภาษีและแบบแสดงรายการภาษีที่ชำระ ข้อมูลชื่อหน่วยงานของกรม

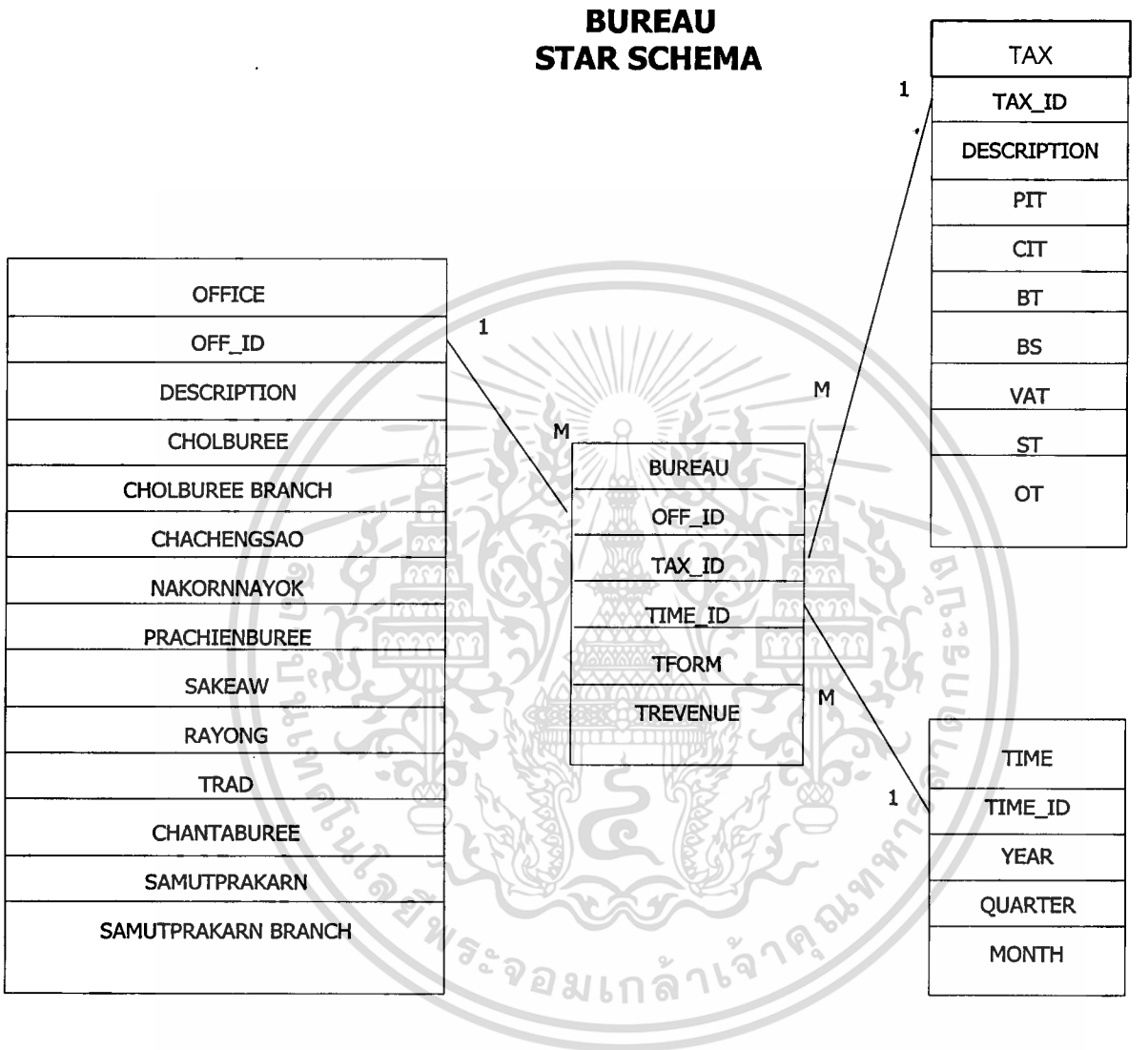
7.3 การออกแบบ Model

ในการออกแบบ Model ภาพที่ 7.2 สำหรับงานภาษีมูลค่าเพิ่มตามลักษณะของข้อมูลที่อยู่ในความสนใจของผู้บริหาร สร้าง dimension ที่ต้องการทราบข้อมูล ได้ออกแบบไว้ 1 Model ภายในมี 1 Fact Table 3 Dimension

7.3.1 Model รวบรวมผลการจัดเก็บภาษีของภาค 5

- BUREAU ตารางรวบรวมผลการจัดเก็บภาษีในภาค 5
- TIME ตารางบอกเวลาผลการจัดเก็บ
- TAX ตารางบอกประเภทภาษีที่จัดเก็บ
- OFF ตารางบอกหน่วยงานส่งผลการจัดเก็บภาษี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

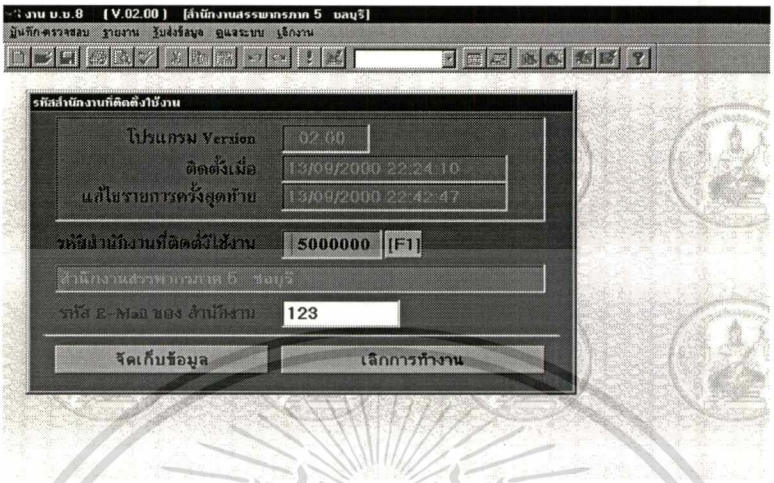


ภาพที่ 7.2 การออกแบบคลังข้อมูลผลการจัดเก็บภาษีของภาคแบบ Star Schema

7.4 ข้อมูลที่นำมาใช้

ได้จากการใช้โปรแกรม บช8 รวบรวมผลจัดเก็บภาษีของแต่ละอำเภอภายในจังหวัด โปรแกรมมีลักษณะดังภาพที่ 7.3

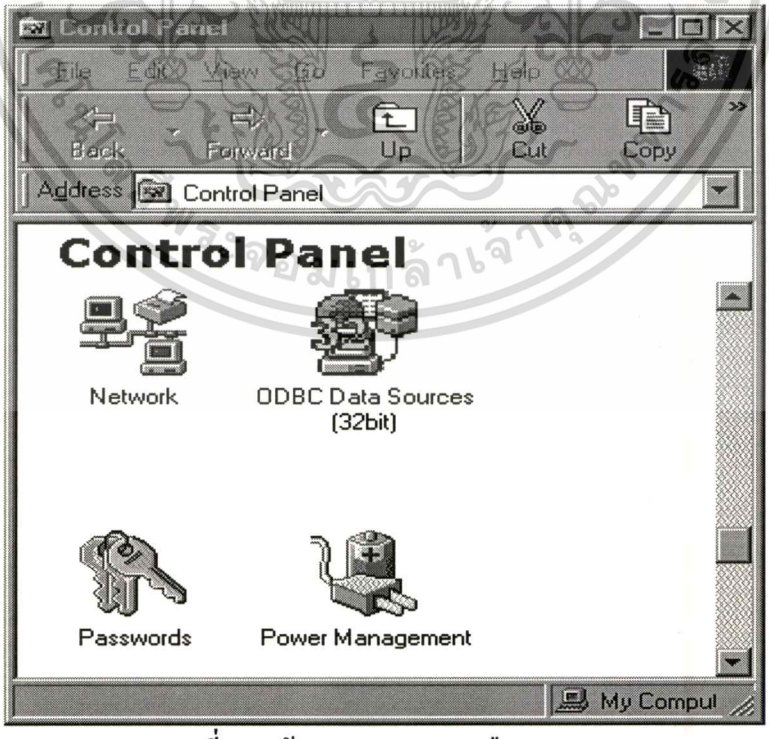
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 7.3 รูปแบบโปรแกรม บช 8 ที่ใช้จัดเก็บภาษีของแต่ละจังหวัด

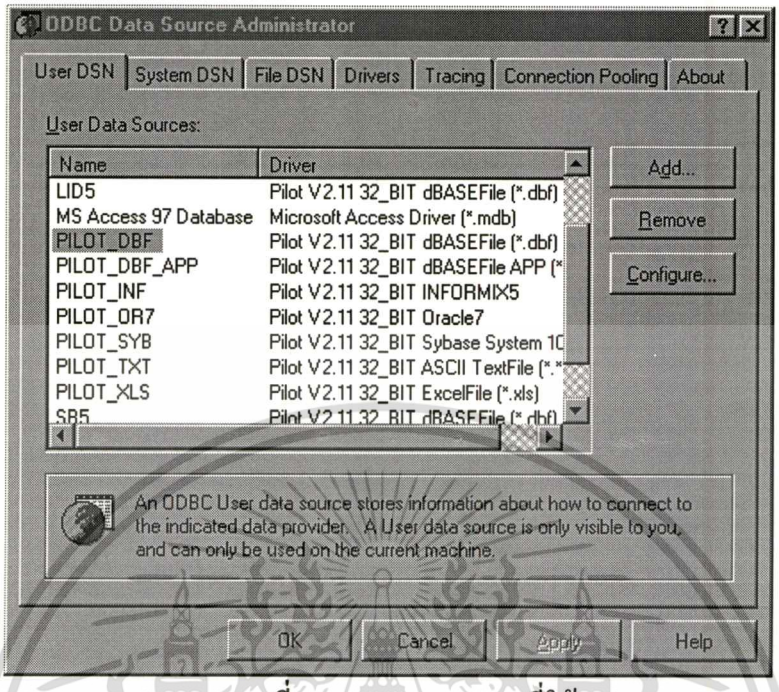
7.5 การเชื่อมโยงฐานข้อมูล

7.5.1 การระบุ ODBC ที่ใช้ ดังภาพที่ 7.4, 7.5



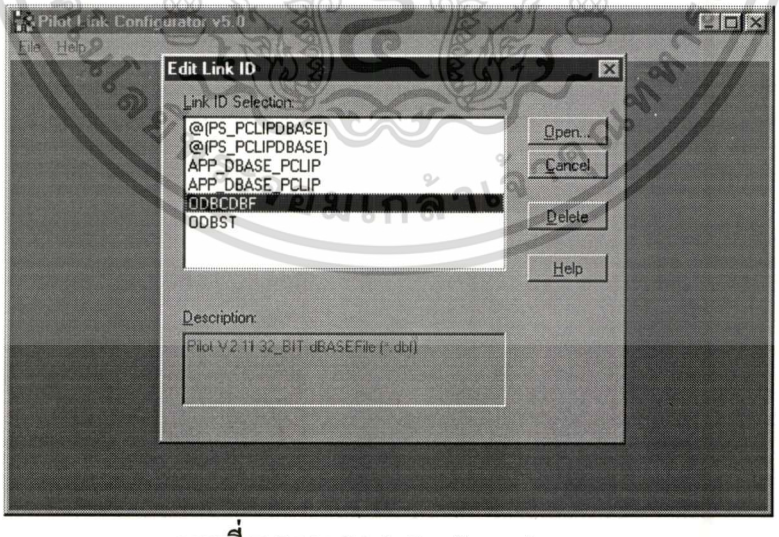
ภาพที่ 7.4 เข้า Control Panel เลือก ODBC

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



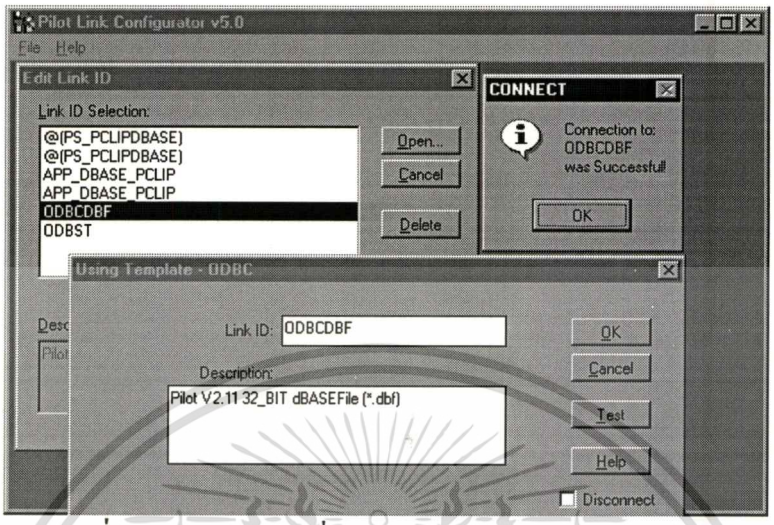
ภาพที่ 7.5 การระบุ ODBC ที่ใช้

7.5.2 การLink กับ ODBCโดยใช้ Pilot Link Configurator ดังภาพที่ 7.6 และ 7.7



ภาพที่ 7.6 การ Link Configuration

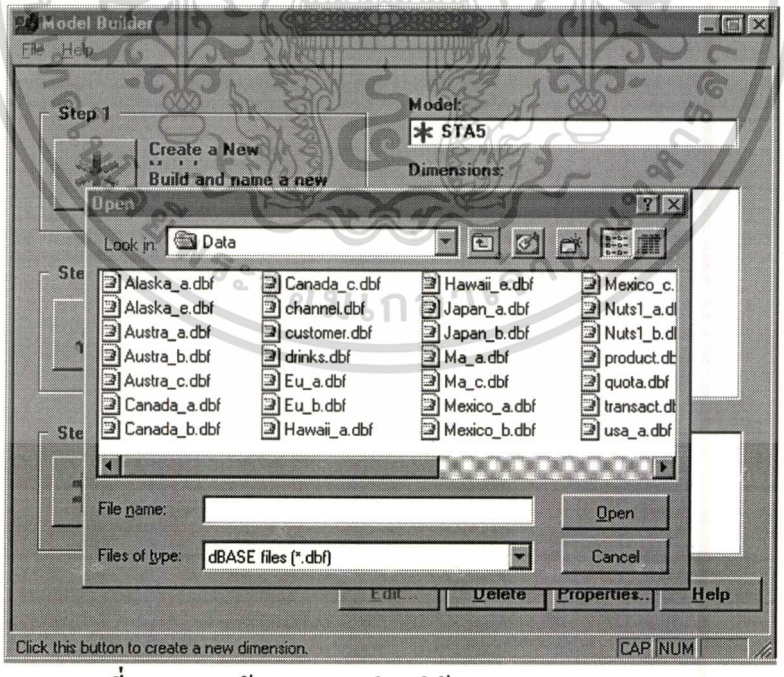
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 7.7 แสดงผลการเชื่อมระหว่างโปรแกรมกับ ODBC

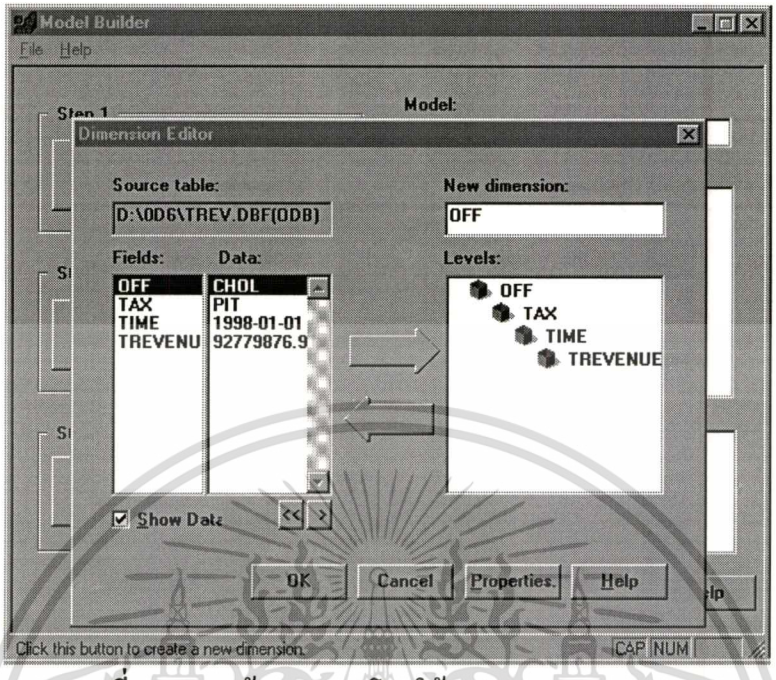
7.6 ขั้นตอนการสร้าง Model

สร้าง Model โดยใช้ Model builder ดังภาพที่ 7.8, 7.9, 7.10

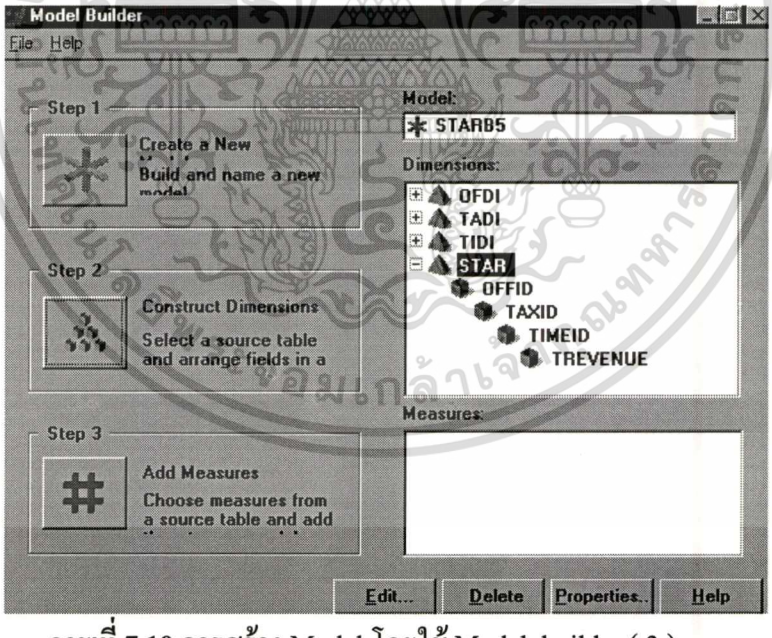


ภาพที่ 7.8 การสร้าง Model โดยใช้ Modele builder (1)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 7.9 การสร้าง Model โดยใช้ Model builder (2)

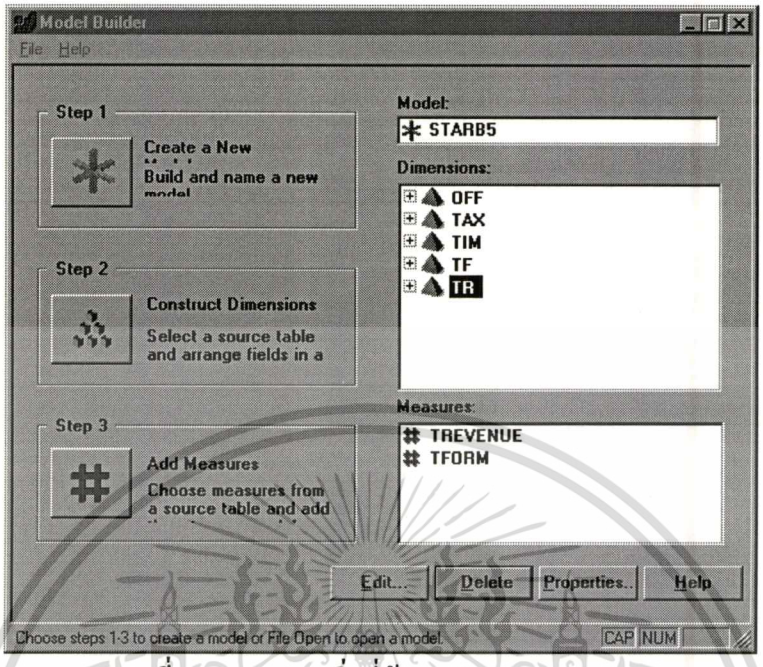


ภาพที่ 7.10 การสร้าง Model โดยใช้ Model builder (3)

7.7 Measure Setting

การระบุสิ่งที่ต้องการการคำนวณ ดังภาพที่ 7.11

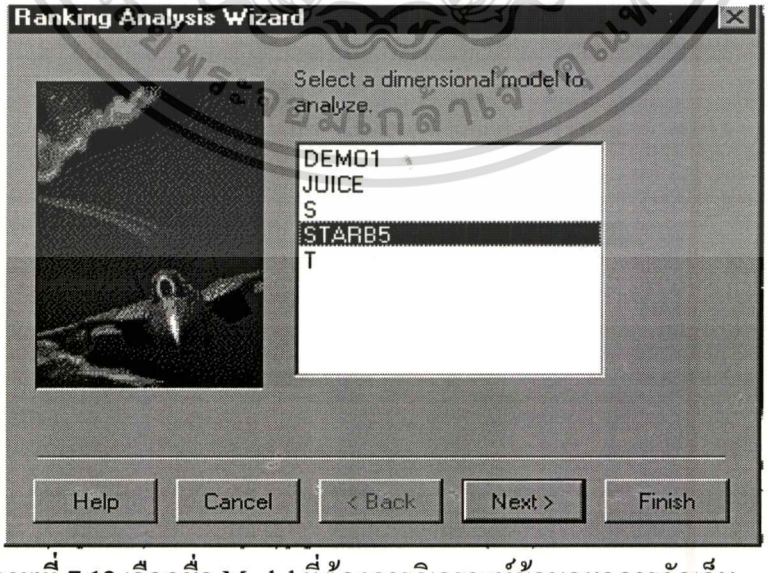
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 7.11 การระบุสิ่งที่ต้องการการคำนวณ

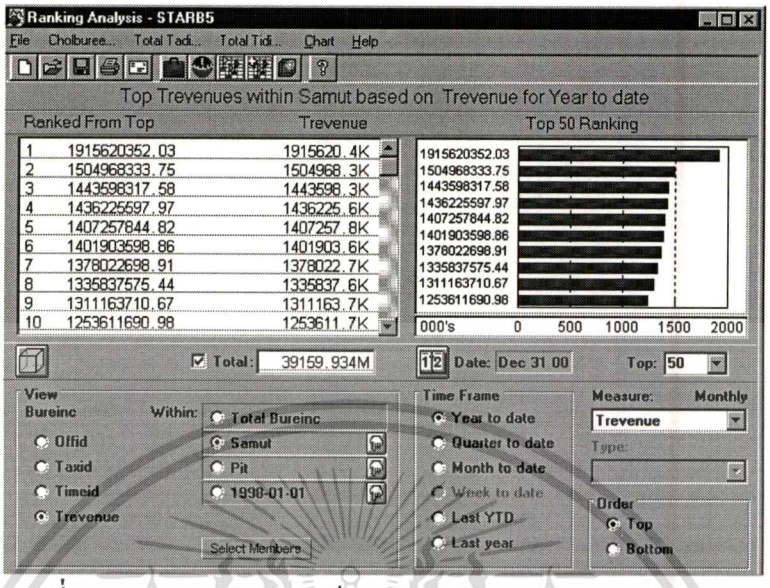
7.8 การวิเคราะห์ข้อมูล

ในการวิเคราะห์ข้อมูลการจัดเก็บภาษีของภาคจะใช้ ทางเลือกการวิเคราะห์โดยใช้ Ranking ในการวิเคราะห์ข้อมูลเพื่อทราบยอดสูงสุดต่ำสุด หรือผลการจัดเก็บในแต่ละจังหวัด หรือประเภทภาษี ดังภาพที่ 7.12, 7.13, 7.14, 7.15, 7.16, 7.17

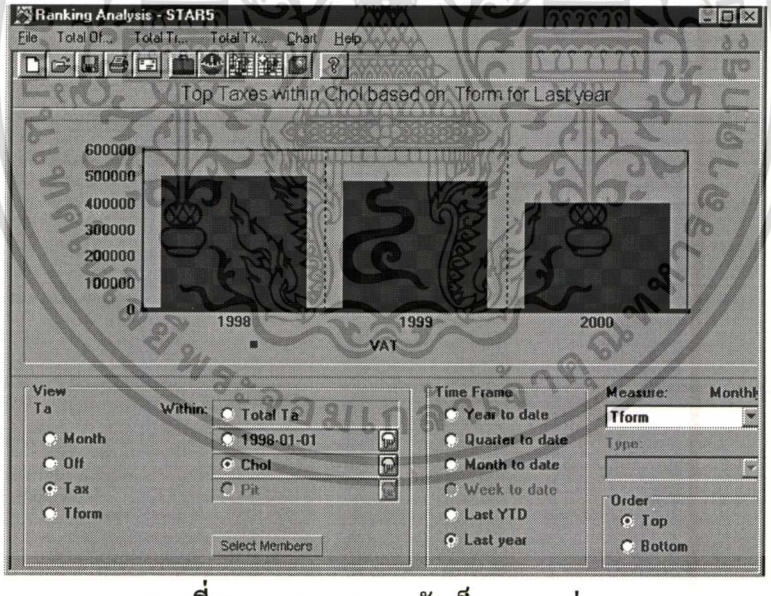


ภาพที่ 7.12 เลือกชื่อ Model ที่ต้องการวิเคราะห์ข้อมูลผลการจัดเก็บ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

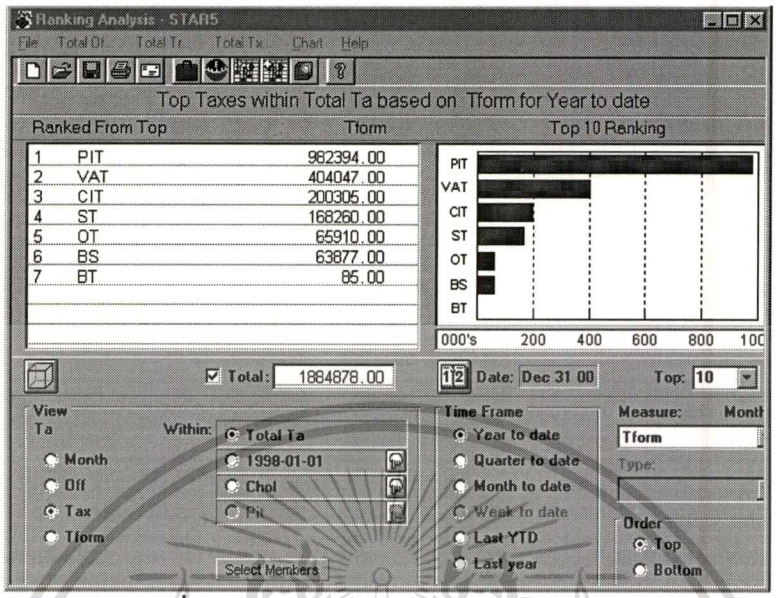


ภาพที่ 7.13 แสดงผลการจัดเก็บภาษีระดับรายเดือนและประเภทภาษี

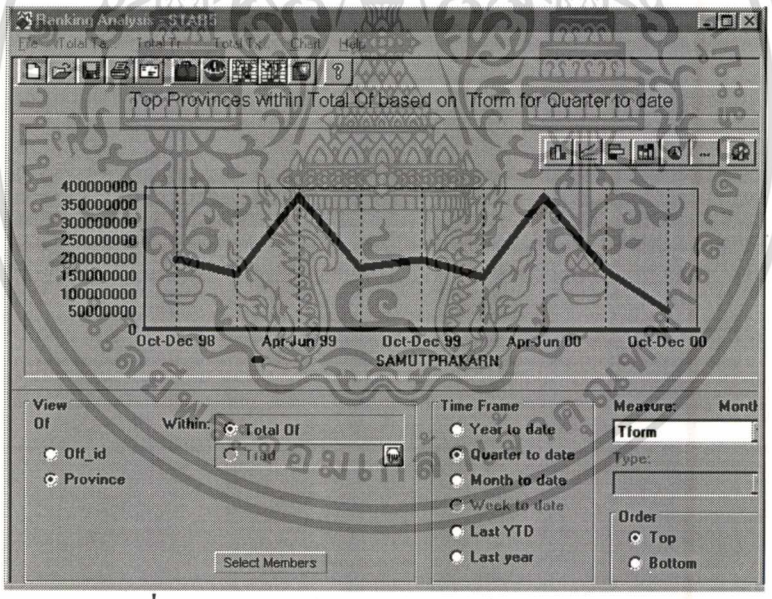


ภาพที่ 7.14 แสดงผลการจัดเก็บแบบแท่ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

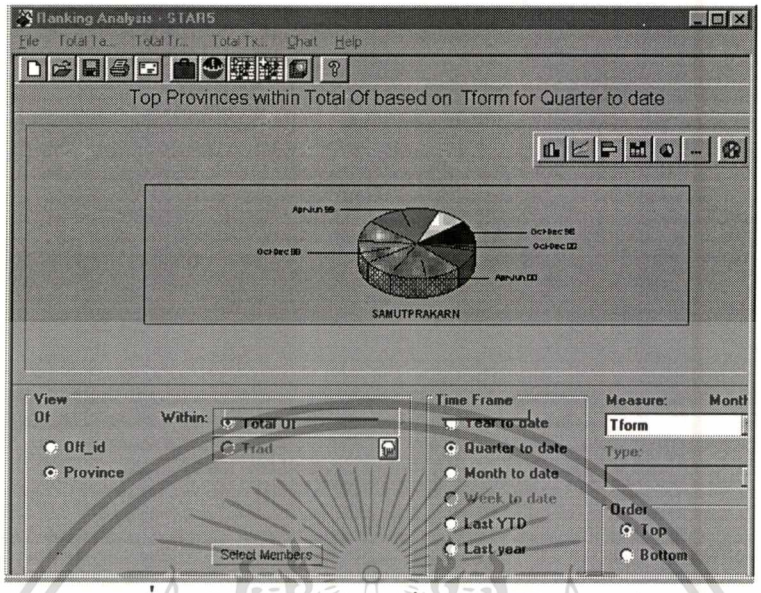


ภาพที่ 7.15 แสดงผลจัดเก็บตามประเภทภาษี



ภาพที่ 7.16 กราฟเส้นแสดงผลจัดเก็บของจังหวัด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 7.17 แสดงผลการจัดเก็บภาษีแบบกราฟรูปวงกลม

การวิเคราะห์ข้อมูลผลการจัดเก็บภาษีโดยการดูข้อมูลในแบบ Ranking Analysis ซึ่งเป็นการเรียกดูข้อมูลในลักษณะของการจัดเรียงลำดับจากน้อยไปมาก หรือจากมากไปน้อยได้ และในการแสดงผลข้อมูล สามารถที่จะดูข้อมูลตาม dimension ที่กำหนดไว้ได้โดยอาศัยมิติของเวลาในการดูข้อมูล เป็นส่วนที่ช่วยจัดการในการเลือกดูข้อมูลได้ ซึ่งเป็น dimension ของข้อมูลแต่ละภาษี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 8
บทสรุป

ในการสร้างระบบคลังข้อมูลผลการจัดเก็บภาษี โดยได้วิเคราะห์ความต้องการการใช้ข้อมูลจากผลการจัดเก็บภาษีได้นำมาทำการออกแบบสร้างฐานข้อมูลบนระบบคลังข้อมูลแบบstar schema เพื่อช่วยให้มีการแสดงผลในรูปแบบต่างๆ ที่ผู้บริหารสามารถนำข้อมูลดังกล่าวไปใช้ในการตัดสินใจ ในการบริหารการจัดเก็บภาษีแต่ละประเภท

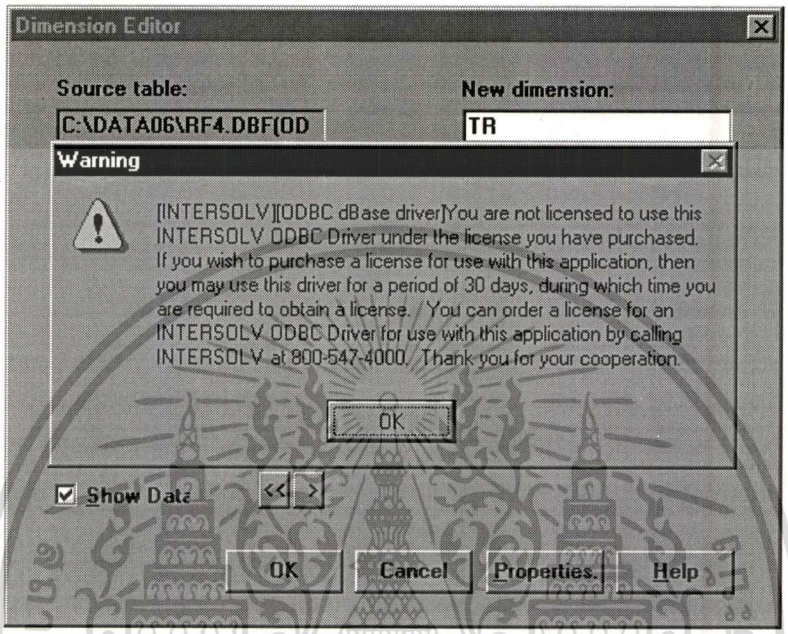
จากการสร้างระบบคลังข้อมูลการจัดเก็บภาษี พบว่าการนำระบบคลังข้อมูลมาใช้งาน มีความเหมาะสมกับระบบงานในปัจจุบันมากแม้ว่าปริมาณข้อมูลจะมีเพิ่มมากขึ้นและผู้บริหารมีความต้องการข้อมูลที่มากขึ้นและมีอยู่ตลอดเวลา ทั้งยังเป็นข้อมูลที่ผู้บริหารให้ความสนใจ และต้องการทราบความเปลี่ยนแปลงของข้อมูลอยู่ตลอดเวลา ประกอบกับข้อมูลที่ถูกจัดเก็บอยู่บนฐานข้อมูลปัจจุบัน จะมีการกระจายข้อมูลไปในตารางที่มีความสัมพันธ์กันหลายตาราง จึงไม่สนับสนุนการจัดทำข้อมูลที่เป็นลักษณะของการวิเคราะห์ เมื่อต้องการใช้ข้อมูลทำให้ต้องทำ การเชื่อมโยงข้อมูลระหว่างตาราง ซึ่งจะใช้เวลาในการจัดทำข้อมูลนานมาก และจากการที่ได้ออกแบบระบบคลังข้อมูล แล้วนำโปรแกรม Pilot Desktop มาใช้วิเคราะห์ข้อมูลจะช่วยทำให้การสืบค้นข้อมูลผลการจัดเก็บได้รวดเร็ว ซึ่งถ้าผู้บริหารมีความต้องการใช้ข้อมูล หรือต้องการให้สร้างระบบงานที่มีการวิเคราะห์ข้อมูลที่มีความซับซ้อน ก็สามารถที่จะพัฒนาระบบงานโดยการใช้ Pilot Designer ซึ่งมีอยู่ใน Pilot Desion Support Suit Tools

8.1 ข้อเสนอแนะ

ในการที่จะนำเทคโนโลยีของระบบคลังข้อมูลมาใช้ในองค์กรได้ดีนั้น จำเป็นจะต้องมีบุคคลากรที่มีความรู้ ความสามารถในการออกแบบระบบฐานข้อมูล และเข้าใจระบบงานเป็นอย่างดีจึงจะสามารถสร้างระบบฐานข้อมูลที่เหมาะสมกับลักษณะขององค์กร เพื่อให้ได้ฐานข้อมูลที่มีโครงสร้างเข้าใจง่าย ช่วยสนับสนุนการวิเคราะห์ข้อมูลปริมาณมากๆได้ ในเวลาที่รวดเร็ว และยังต้องมีการเลือกใช้เครื่องมือ ประเภท software Tools มาใช้รวมถึงสิทธิในการใช้โปรแกรมนั้นหากไม่ได้โปรแกรมมาโดยสมบูรณ์ก็จะทำให้ไม่สามารถวิเคราะห์ข้อมูลให้ถูกต้องได้ปัญหาที่จะเกิดขึ้นในโปรแกรมนี้ดังภาพที่8.1 ในการวิเคราะห์ข้อมูลที่เหมาะสมง่ายต่อการใช้งาน และพัฒนาระบบงานได้ดี จึงจะช่วยให้การวิเคราะห์ข้อมูลเป็นไปอย่างมีประสิทธิภาพ มีความถูกต้อง ตรงกับระเภท

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อมูลที่ใช้และตรงตามความต้องการของผู้บริหารในองค์เพื่อความถูกต้อง และการพัฒนาต่อไปในอนาคต



ภาพที่ 8.1 ปัญหาที่เกิดขึ้นจากโปรแกรม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม

นงลักษณ์ พลอยปลื้ม 2540 "การแปลงข้อมูลเข้าสู่ Data Warehouse"วารสาร BCM..ปีที่ 9ฉบับที่ 103 กันยายน

Barquin, Ramon and Herb Edelstein. 1997. **Planing and Designing the Data Warehouse**.New Jersey:Prentice Hall.

Orgali, Robert, Dan Harkey and Jeri Edwards 1996. **The Essential Client/Server survival guide**. England : John Wiley & Sons.

Poe Vidette, Patricia Klauer and Stephen Brobst 1997. **Building a Data Warehouse for Decision Support**.2 nd ed New Jersey : Prentice Hall.

Peter Rob, Carlos Coronel 1997. **Database System**.Course Technology.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประวัติผู้เขียน

ชื่อ – นามสกุล	นายธวัชชัย เพชรวิชิต
วันเกิด	11 ธันวาคม 2510
ภูมิลำเนา	จังหวัดสุราษฎร์ธานี
การศึกษา	ประถม : โรงเรียนวัดแจ้ง อ.เกาะสมุย จ.สุราษฎร์ธานี มัธยมต้น และมัธยมปลาย : โรงเรียนเกาะสมุย อ.เกาะสมุย จ.สุราษฎร์ธานี ปริญญาตรี : คณะเศรษฐศาสตร์ วิชาเอก เศรษฐศาสร์ระหว่าง-ประเทศ มหาวิทยาลัยรามคำแหง กรุงเทพมหานคร
ปัจจุบัน	รับราชการ ตำแหน่งนักวิชาการภาษี 4 กรมสรรพากร จ.ชลบุรี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้