ห้องสมุดคณะเทคโนโลยีสารสนเทศ สจล.

ระบบคลังข้อมูลเพื่อช่วยการตัดสินใจในงานภาษีมูลค่าเพิ่ม กรมสรรพากร

Data Warehousing System for Decision Support in Value Added Tax of Revenue Department

โดย

นางสาวศศกร ปัทมวิภาต

รหัส 39067325

วัน เดือน ปี 07 ธ.ค. 2550
เลขทะเบียน 02524
เลขเรียกหนังสือ วพ. ศ169ร 2541
"ห้องสมุดคณะเทคโนโลยีสารสนเทศ สจล."

อาจารย์ที่ปรึกษา

อาจารย์อิสระ บุรินทรามาตย์

H002524

รายงานนี้เป็นส่วนหนึ่งของวิชาโครงการศึกษากรณีพิเศษ
หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ
ภาคเรียนที่ 1 ปีการศึกษา 2541
คณะเทคโนโลยีสารสนเทศ
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ชื่อหัวข้อ	ระบบคลังข้อมูลเพื่อช่วยการตัดสินใจในงานภาษีมูลค่าเพิ่มกรมสรรพากร
นักศึกษา	นางสาวศศกร ปัทมวิภาต
อาจารย์ที่ปรึกษา	อาจารย์อิสระ บุรินทรามาตย์
ระดับการศึกษา	วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ
แขนงวิชา	การจัดการเทคโนโลยีสารสนเทศ
พ.ศ.	2541

บทคัดย่อ

กรมสรรพากรเป็นหน่วยงานจัดเก็บภาษีของประเทศ ซึ่งได้มีการนำระบบคอมพิวเตอร์เข้ามาใช้ช่วยในการจัดเก็บและประมวลผลข้อมูล โดยที่ภาษีมูลค่าเพิ่มจัดเป็นภาษีประเภทหนึ่งที่มีปริมาณข้อมูลจำนวนมาก ทำให้การที่จะวิเคราะห์ข้อมูลเป็นไปได้อย่างล่าช้า ซึ่งในปัจจุบันข้อมูลนับว่าเป็นส่วนที่สำคัญมากสำหรับทุกองค์กร การรับรู้ข้อมูลที่ทันต่อสถานการณ์ปัจจุบันก่อนผู้อื่นทำให้เกิดความได้เปรียบในการจัดการวางแผนกลยุทธ์และกำหนดทิศทางของธุรกิจ ดังนั้นองค์กรต่างๆ จึงได้ให้ความสนใจในการจัดทำคลังข้อมูล (Data Warehouse) เพื่อทำการรวบรวมข้อมูลมาใช้ในการบริหารงานให้มีประสิทธิภาพอย่างกว้างขวาง

การพัฒนาระบบคลังข้อมูล (Data Warehouse) ขึ้นมาใช้ในงานภาษีมูลค่าเพิ่ม นับเป็นแนวทางในการที่จะทำให้เกิดข้อมูลที่จะสนองตอบความต้องการของผู้บริหาร ที่จะนำข้อมูลไปใช้ประกอบการตัดสินใจ เพื่อการวางแผนและบริหารการจัดเก็บภาษีอากรได้อย่างมีประสิทธิภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Title	Data Warehousing System for Decision Support in Value Added Tax of Revenue Department
Student	Miss Sasakorn Pattamavipat
Advisor	Mr. Isara Burintramart
Level of Study	Master of Science in Information Technology
Major	Information Technology Management
Year	1998

ABSTRACT

Revenue Department is a tax collection organization of our country where using computer system as a tool for data collection and processing. Because Value Added Tax is a type of tax which has high volume of data, the data analysis can be performed slowly and as nowadays for all organizations, correcting and updating data is the highest importance. Possessions of data before others is the best competitive tool to enable a successful business strategy to be achieved. Therefore, all organizations are interested in having a Data Warehouse. This will help them manage their resources efficiently.

The development of Data Warehousing system for Value Added Tax will be an approach to provide adequate data to serve tax management needs so that it can support tax executive in decision making , planning and governing for tax collection.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

ในการจัดทำโครงการศึกษากรณีพิเศษ เรื่อง "ระบบคลังข้อมูลเพื่อช่วยการตัดสินใจในงานภาษีมูลค่าเพิ่ม กรมสรรพากร" ผู้เขียนต้องขอขอบพระคุณอาจารย์อิสระ บุรินทรามาตย์ อาจารย์ที่ปรึกษาที่ได้กรุณาให้คำปรึกษา คำแนะนำ ช่วยเป็นธุระจัดหาซอฟต์แวร์และคู่มือการใช้งานที่จะใช้ในการพัฒนาระบบงาน และให้แนวทางในการแก้ไขปัญหาต่างๆที่เกิดขึ้น ทำให้โครงการศึกษากรณีพิเศษฉบับนี้สำเร็จได้ด้วยดี

ศศกร ปัทมวิภาต

กันยายน 2541

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

หน้า

บทคัดย่อภาษาไทย..... I
บทคัดย่อภาษาอังกฤษ..... II
กิตติกรรมประกาศ..... III
สารบัญ..... IV
สารบัญตาราง..... VI
สารบัญภาพ..... VII
บทที่
1. บทนำ..... 1
1.1 ความเป็นมา..... 1
1.2 วัตถุประสงค์..... 2
1.3 ขั้นตอนการศึกษา..... 2
1.4 ขอบเขตของการศึกษา..... 3
1.5 ประโยชน์ที่คาดว่าจะได้รับ..... 3
2. ระบบคลังข้อมูล (Data Warehousing System)..... 4
2.1 ความหมายของ Data Warehouse..... 4
2.2 ส่วนประกอบของ Data Warehouse..... 5
2.3 องค์ประกอบในการสร้าง Data Warehouse..... 7
2.4 คุณลักษณะของข้อมูลใน Data Warehouse..... 10
2.5 การแปลงข้อมูลเข้าสู่ Data Warehouse..... 11
2.6 สิ่งที่ต้องคำนึงถึงในการออกแบบ Data Warehouse..... 14
2.7 Software ที่ใช้จัดการข้อมูล..... 15
3. ระบบสนับสนุนการตัดสินใจ (Decision Support System)..... 17
3.1 ความหมายของระบบสนับสนุนการตัดสินใจ..... 17
3.2 ลักษณะของระบบ..... 17
3.3 ส่วนประกอบของระบบ..... 18

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หน้า

3.4 ขั้นตอนของการจัดทำระบบสนับสนุนการตัดสินใจ..... 18
3.5 คุณลักษณะข้อมูลของระบบสนับสนุนการตัดสินใจ..... 22
4. เครื่องมือที่ใช้ในการวิเคราะห์ข้อมูล (Tools for Analyse Data)..... 23
4.1 ข้อกำหนดความต้องการของระบบ..... 23
4.2 ส่วนประกอบของ Pilot Decision Support Suite..... 24
4.3 คุณสมบัติของ Pilot Decision Support Suite..... 26
4.4 การทำงานของ Pilot Decision Support Suite..... 27
5. ภาษีมูลค่าเพิ่ม (Value Added Tax)..... 30
5.1 ภาษีมูลค่าเพิ่ม..... 30
5.2 ผู้มีหน้าที่จดทะเบียนภาษีมูลค่าเพิ่ม..... 30
5.3 ประเภทของแบบแสดงรายการภาษีมูลค่าเพิ่มที่เกี่ยวข้อง..... 31
5.4 การประมวลผลระบบงานภาษีมูลค่าเพิ่ม..... 31
5.5 โครงสร้างฐานข้อมูลของระบบงานภาษีมูลค่าเพิ่ม..... 31
6. การพัฒนาระบบคลังข้อมูลกับงานภาษีมูลค่าเพิ่ม..... 35
6.1 ขั้นตอนการดำเนินงาน..... 35
6.2 ระบบคลังข้อมูลภาษีมูลค่าเพิ่ม..... 53
6.3 รูปแบบการวิเคราะห์ข้อมูล..... 54
7. บทสรุป..... 65
7.1 สรุปผลการศึกษา..... 65
7.2 ประโยชน์ที่ได้รับ..... 66
7.3 ข้อเสนอแนะ..... 67
บรรณานุกรม..... 68
ประวัติผู้เขียน..... 69

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

หน้า

ตารางที่

2.1 Data Warehouse Decision Matrix..... 10

6.1 โครงสร้างข้อมูลของรายการในการยื่นแบบแสดงรายการภาษีมูลค่าเพิ่ม..... 37

6.2 โครงสร้างข้อมูลของการยื่นแบบ ภพ.30..... 38

6.3 โครงสร้างข้อมูลของรายการบัญชีของผู้เสียภาษี..... 40

6.4 การวิเคราะห์แนวโน้มของข้อมูลภาษีมูลค่าเพิ่ม..... 41

6.5 โครงสร้างข้อมูลการขอคืนเงินภาษีมูลค่าเพิ่ม..... 44

6.6 โครงสร้างข้อมูลจำนวนการยื่นแบบแสดงรายการภาษีมูลค่าเพิ่ม..... 45

6.7 โครงสร้างข้อมูลประเภทกิจการ..... 45

6.8 โครงสร้างข้อมูลของสำนักงานสรรพากร..... 46

6.9 โครงสร้างข้อมูลของประเภทแบบแสดงรายการภาษีมูลค่าเพิ่ม..... 46

6.10 โครงสร้างข้อมูลประเภทของการขอคืนเงินภาษีมูลค่าเพิ่ม..... 47

7.1 เปรียบเทียบข้อดี ข้อเสียของโปรแกรม Pilot Desktop และ Pilot Designer..... 66

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญภาพ

หน้า

ภาพที่

2.1 ส่วนประกอบของ Data Warehouse..... 5

2.2 สถาปัตยกรรมแบบ Federated และแบบ Tiered..... 8

2.3 กระบวนการแปลงข้อมูล..... 13

4.1 ชุดของโปรแกรม Pilot Decision Support Suite..... 23

4.2 Pilot Analysis Server..... 24

4.3 Pilot Desktop..... 25

4.4 Pilot Designer..... 26

4.5 การ Link Configuration..... 27

4.6 Model Builder..... 28

6.1 แสดงความสัมพันธ์ของตารางข้อมูลภาษีมูลค่าเพิ่ม..... 36

6.2 ขั้นตอนการนำข้อมูลภาษีมูลค่าเพิ่มจากฐานข้อมูลบนเครื่อง IBM ES/9000 มาจัดทำฐานข้อมูลบนเครื่อง IBM SP/2..... 41

6.3 การออกแบบคลังข้อมูลของการขอคืนเงินภาษีมูลค่าเพิ่ม..... 43

6.4 การออกแบบคลังข้อมูลของการยื่นแบบแสดงรายการภาษีมูลค่าเพิ่ม..... 43

6.5 การรวบรวมข้อมูลภาษีมูลค่าเพิ่มจากแหล่งข้อมูล..... 48

6.6 การจัดทำข้อมูลสรุปของภาษีมูลค่าเพิ่ม..... 49

6.7 รูปแบบ hierarchy ของ Dimension Business..... 50

6.8 รูปแบบ hierarchy ของ Dimension Office..... 50

6.9 รูปแบบ hierarchy ของ Dimension Refund type..... 50

6.10 รูปแบบ hierarchy ของ Dimension Form..... 51

6.11 Model การขอคืนเงินภาษีมูลค่าเพิ่ม..... 52

6.12 Model การยื่นแบบแสดงรายการภาษีมูลค่าเพิ่ม..... 52

6.13 การ Logon เข้าใช้ระบบ..... 53

6.14 แสดง Menu VAT REFUND..... 54

6.15 แสดง Menu VAT RETURN..... 54

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หน้า

ภาพที่

6.16 แสดงข้อมูลการขอคืนเงินภาษีมูลค่าเพิ่มของประเภทกิจการ 55

6.17 แสดงผลการขอคืนเงินภาษีมูลค่าเพิ่ม ในระดับภาค (region) และการเลือกดูข้อมูลตามประเภทของการขอคืนเงินภาษี..... 56

6.18 แสดง 10 จังหวัดที่มีการขอคืนเงินภาษีมูลค่าเพิ่มสูงสุด..... 56

6.19 แสดงผลการขอคืนเงินภาษีมูลค่าเพิ่ม ในระดับ (Amphur)..... 57

6.20 การ drill down ข้อมูล..... 57

6.21 การเลือกข้อมูลภาค เพื่อดูข้อมูลจำแนกรายจังหวัด..... 58

6.22 การเลือกข้อมูลจังหวัด เพื่อดูข้อมูลจำแนกรายอำเภอ..... 58

6.23 การขอคืนเงินภาษีมูลค่าเพิ่มเป็นเงินสด รายภาค..... 59

6.24 แสดงการเปรียบเทียบข้อมูลการขอคืนเงินภาษีมูลค่าเพิ่ม ในช่วงเวลา 2 เดือน..... 59

6.25 แสดงการเปรียบเทียบการขอคืนเงินภาษีมูลค่าเพิ่มของประเภทกิจการ..... 60

6.26 แสดงกราฟวงกลม..... 60

6.27 การกำหนดมุมมองของข้อมูล (View)..... 61

6.28 แสดงข้อมูสในระดับสูงสุด (Total Office)..... 61

6.29 แสดงผลจากการ Drill down ข้อมูลในรูปแบบของ Navigator Adhoc..... 62

6.30 แสดงผลการขอคืนเงินภาษีมูลค่าเพิ่มของแต่ละประเภทกิจการ รายจังหวัด..... 62

6.31 แสดงผลการขอคืนเงินภาษีมูลค่าเพิ่มของแต่ละประเภทกิจการ รายอำเภอ..... 63

6.32 เปรียบเทียบการยื่นแบบ ภพ.01 และ ภพ.09 รายภาค..... 64

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1
บทนำ

1.1 ความเป็นมา

กรมสรรพากรเป็นหน่วยงานสังกัดกระทรวงการคลัง ที่มีหน้าที่ในการจัดเก็บภาษีอากร ซึ่งระบบภาษีอากรที่กรมสรรพากรจัดเก็บประกอบด้วย ภาษีเงินได้บุคคลธรรมดา ภาษีเงินได้นิติบุคคล ภาษีเงินได้หัก ณ ที่จ่าย ภาษีธุรกิจเฉพาะ และภาษีมูลค่าเพิ่ม โดยระบบภาษีมูลค่าเพิ่มได้เริ่มนำมาใช้ตั้งแต่ปี พ.ศ. 2535 ซึ่งได้มีการติดตั้งโปรแกรม และพัฒนาระบบงานบนเครื่องคอมพิวเตอร์ Mainframe IBM ES/9000 รวมทั้งทำการประมวลผลข้อมูลภาษีมูลค่าเพิ่มบนเครื่องดังกล่าวด้วย ข้อมูลภาษีมูลค่าเพิ่มทั้งหมดจะถูกจัดเก็บอยู่บนฐานข้อมูลของเครื่อง ES/9000 โดยข้อมูลที่มีอยู่เป็นข้อมูลที่มีการประมวลผลเป็นประจำวัน หรือที่เรียก Operational Data ซึ่งนับวันจะมีมากขึ้น เนื่องจากยังไม่มีการนำข้อมูลบางส่วนไปเก็บไว้ในสื่อคอมพิวเตอร์ (media) อื่นๆ เมื่อผู้บริหารมีความต้องการที่จะได้ข้อมูลภาษีมูลค่าเพิ่มในลักษณะต่างๆ เพื่อนำไปช่วยประกอบการตัดสินใจในการวางแผน การควบคุม และการบริหารงานจัดเก็บภาษีอากร ทำให้ต้องใช้เวลาในการประมวลผลข้อมูลนานมาก และยังส่งผลกระทบไปถึงระบบการทำงานของเครื่องคอมพิวเตอร์ที่ต้องประมวลผลประจำวันอีกด้วย

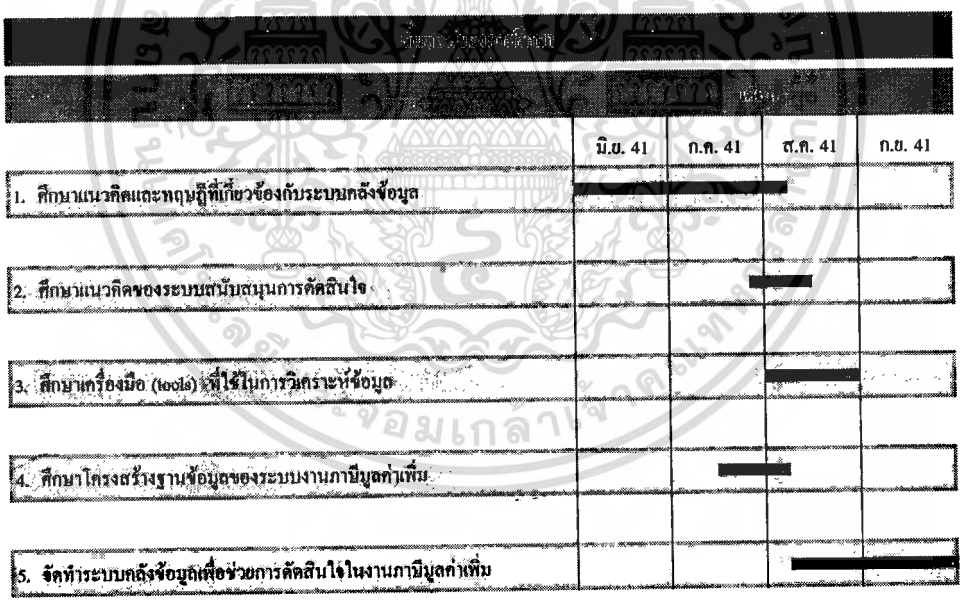
ในยุคปัจจุบันทุกองค์กรได้มีการนำระบบฐานข้อมูลเข้ามาใช้บริหาร และจัดการควบคุมข้อมูล เพื่อนำข้อมูลมาใช้ในการบริหารให้เกิดประสิทธิภาพสูงสุด ประกอบกับความก้าวหน้าในการพัฒนาของเทคโนโลยีทั้ง hardware และ software ที่จะช่วยในการจัดการข้อมูล จึงทำให้เกิดแนวความคิดในการจัดทำระบบคลังข้อมูล (Data Warehouse) ขึ้นมาใช้ ซึ่งตามความต้องการทางด้านสารสนเทศระบบงานภาษีมูลค่าเพิ่มของกรมสรรพากรนั้น จำเป็นที่จะต้องมีฐานข้อมูลเพื่อการรองรับข้อมูลที่ผู้บริหารต้องการได้อย่างเหมาะสม ดังนั้นการพัฒนาระบบคลังข้อมูลขึ้นมาใช้กับระบบงานภาษีมูลค่าเพิ่ม จะเป็นการแนวทางที่จะก่อให้เกิดข้อมูลที่จะช่วยสนับสนุนการตัดสินใจของผู้บริหารได้อย่างมีประสิทธิภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.2 วัตถุประสงค์

1. เพื่อเป็นแนวทางในการสร้างคลังข้อมูลสำหรับกรมสรรพากร
2. เพื่อก่อให้เกิดการนำข้อมูลมาใช้ได้อย่างมีประสิทธิภาพสูงสุด
3. เพื่อเป็นการจัดทำข้อมูลที่จะช่วยในสนับสนุนการตัดสินใจของผู้บริหารระดับสูง ได้อย่างรวดเร็วและทันต่อสถานการณ์
4. เพื่อช่วยให้ผู้บริหารสามารถจะนำข้อมูลมาใช้ช่วยในการวางแผนบริหารงาน การจัดเก็บภาษีอากรของกรมสรรพากรได้อย่างมีประสิทธิภาพ

1.3 ขั้นตอนการศึกษา



1.4 ขอบเขตของการศึกษา

1. ศึกษาแนวคิดและทฤษฎีที่เกี่ยวข้องกับระบบคลังข้อมูล
2. ศึกษาแนวคิดของระบบสนับสนุนการตัดสินใจ
3. ศึกษาเครื่องมือ (tools) ที่ใช้ในการวิเคราะห์ข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4. ศึกษาโครงสร้างฐานข้อมูลของระบบงานภาษีมูลค่าเพิ่ม

5. การจัดทำระบบคลังข้อมูลเพื่อช่วยการตัดสินใจในงานภาษีมูลค่าเพิ่ม กรมสรรพากร โดยมีขอบเขต ดังนี้

5.1 การออกแบบและจัดสร้างฐานข้อมูล เพื่อเป็นคลังข้อมูลของระบบงานภาษีมูลค่าเพิ่ม

5.2 เลือกดูข้อมูลที่สนใจ

5.3 ทำการวิเคราะห์ผลสรุปของข้อมูล

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1. เป็นการนำแนวความคิดของการจัดทำระบบคลังข้อมูลมาประยุกต์เข้าใช้กับงานปัจจุบันขององค์กรที่กระทำอยู่

2. เป็นการนำข้อมูลที่จัดเก็บอยู่บนฐานข้อมูลมาปรับใช้ได้อย่างเหมาะสมกับความต้องการของผู้ใช้

3. เป็นการจัดทำข้อมูล เพื่อช่วยผู้บริหารในการตัดสินใจ และวางแผนการบริหารงานจัดเก็บภาษีอากรได้อย่างมีประสิทธิภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2
ระบบคลังข้อมูล (Data Warehousing System)

Data Warehouse ถือว่าเป็นศูนย์กลางของสถาปัตยกรรมสำหรับระบบสารสนเทศในปี 1990 โดย data warehouse จะสนับสนุนการประมวลผลสารสนเทศ มีการรวบรวม จัดเก็บข้อมูลภายในและภายนอกองค์กร ไม่จำกัดข้อมูลในการเก็บบันทึกลงฐานข้อมูลขนาดใหญ่ รวมทั้งจัดรูปแบบของข้อมูลให้สามารถนำมาวิเคราะห์ได้ โดยเป็นการอาศัยเทคนิคต่างๆ เช่น การนำความรู้ทางสถิติ หรือคณิตศาสตร์มาประยุกต์ใช้ เพื่อหาแนวโน้มของข้อมูลที่ต้องการ ซึ่งการจัดทำ data warehouse จะทำให้ได้ข้อมูลที่จะนำมาใช้ประกอบการตัดสินใจให้แม่นยำ และถูกต้องยิ่งขึ้น จะส่งผลให้เกิดความคุ้มค่าในการตัดสินใจ โดยระบบ data warehouse ที่ดีจะเกิดจากการนำข้อมูลจากระดับปฏิบัติการมาผ่านกระบวนการในการแปลงข้อมูลให้เป็นมาตรฐานก่อนจัดเก็บ

Data Warehouse จึงเป็นการนำเสนอแนวทางในการเข้าถึงข้อมูลได้อย่างมีประสิทธิภาพ ทั้งในแบบท้องถิ่น (local) หรือ แบบระยะไกล (remote) รวมทั้งการจัดสภาพแวดล้อมของ data warehouse ยังสามารถสนับสนุนการทำงานของผู้ใช้หลายๆ คน หลายๆ จุดประสงค์ และง่ายต่อการใช้งาน จึงนับได้ว่าเทคโนโลยี data warehouse เป็นสิ่งที่ทำให้องค์กรประสบความสำเร็จในการบริหารงานได้อย่างมีประสิทธิภาพ

2.1 ความหมายของ Data Warehouse

Data Warehouse คือ การรวบรวมข้อมูลจากฐานข้อมูลของส่วนปฏิบัติงานต่างๆ นำมาแปลงข้อมูลให้อยู่ในรูปแบบที่เหมาะสมในการจัดเก็บและสะดวกต่อการนำมาใช้งาน แล้วจึงจะนำข้อมูลเหล่านั้นเข้าไปเก็บใน data warehouse

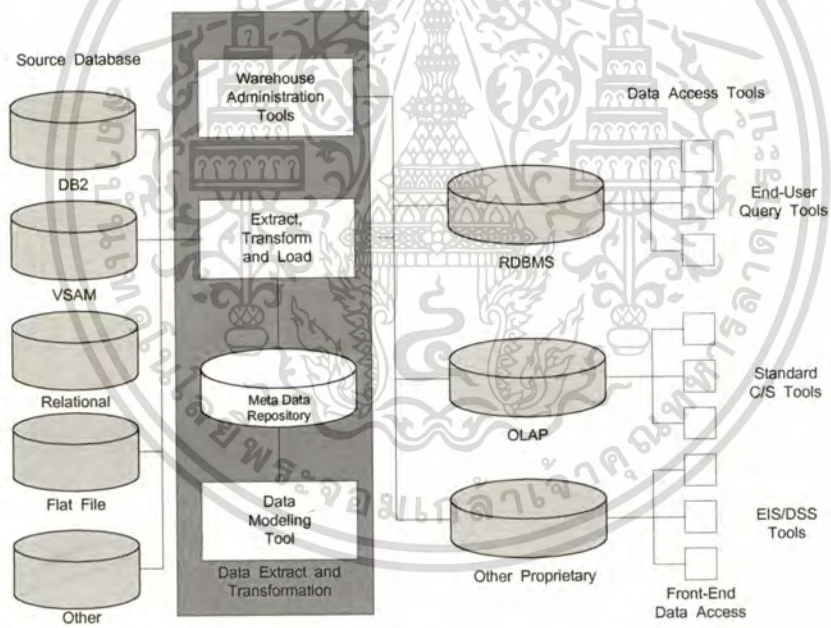
ข้อมูลที่อยู่ใน data warehouse เป็นข้อมูลที่องค์กรธุรกิจ นำมาวิเคราะห์ วางแผน และควบคุม ใช้ตัดสินใจในการดำเนินการทางธุรกิจอย่างถูกต้อง ดังนั้นผู้ที่จะทำการออกแบบ data warehouse ต้องมีความเข้าใจในลักษณะของธุรกิจ เพื่อให้การทำ data warehouse เป็นไปอย่างมีประสิทธิภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2 ส่วนประกอบของ Data Warehouse

ส่วนประกอบของ Data Warehouse มีดังนี้ คือ

1. Source Database
2. Data Extraction and Transformation tools
3. Data Modeling tools
4. Metadata Repository
5. Target Database
6. Front-end access tools



ภาพที่ 2.1 ส่วนประกอบของ Data Warehouse

1. Source Database

แหล่งข้อมูลที่มาจากฐานข้อมูลประเภทต่างๆ ที่อยู่บนเครื่อง Host ได้แก่ ข้อมูลในลักษณะที่เป็นงาน Online transaction processing (OLTP) ข้อมูลที่มาจากส่วนปฏิบัติงานต่างๆ (Operational Data) ข้อมูลการประมวลผลงาน batch และแหล่งข้อมูลภายนอก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. Data Extraction and Transformation Tools

เป็นการนำเครื่องมือที่เหมาะสมเข้ามาช่วยในกระบวนการดึงข้อมูล การแปลงข้อมูล ให้อยู่ในรูปแบบที่เหมาะสมต่อการนำไปใช้บน data warehouse โดยที่เครื่องมือที่จะนำมาใช้นั้น จะต้องมีความสามารถที่จะจัดการกับข้อมูลที่มีปริมาณมากได้เป็นอย่างดี และเหมาะสมกับการใช้งานกับข้อมูลขององค์กร

- Data Extraction เป็นวิธีการในการดึงข้อมูล โดยทำการแยกข้อมูลบางส่วน ออกมาจากฐานข้อมูลขนาดใหญ่ เพื่อประโยชน์ในการวิเคราะห์
- Transformation เป็นการแปลงข้อมูลที่มีอยู่หลายรูปแบบ (format) ให้เป็นรูปแบบมาตรฐานเดียวกัน 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 เป็นไปได้อย่างง่าย ลักษณะที่สำคัญของ Metadata คือ

- เก็บข้อมูลเกี่ยวกับพจนานุกรมข้อมูล (data dictionary)
- เก็บ directory of content ที่ประกอบไปด้วยตำแหน่ง และรายละเอียดเกี่ยวกับ Data Warehouse Server
- กฎเกณฑ์และเงื่อนไขในการที่จะค้นหาข้อมูล
- เก็บ Location ของข้อมูล
- เก็บโครงสร้างของข้อมูล (structure)

5. Target Database

เป็นฐานข้อมูลปลายทางที่ใช้เป็น data warehouse สำหรับจัดเก็บข้อมูลที่ดึงมาจากแหล่งข้อมูลต่าง ๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6. Front-end access Tools

เป็นเครื่องมือที่จะนำมาใช้ช่วยในการเข้าถึงข้อมูล (data access) ได้อย่างรวดเร็ว และทำการวิเคราะห์ข้อมูลบน data warehouse ซึ่งได้แก่

- End-User Query Tools

เป็นเครื่องมือที่ใช้ในการค้นหาข้อมูล โดยช่วยอำนวยความสะดวกให้ผู้ใช้งาน ในด้านอื่น ๆ ด้วย เช่น

- การค้นหาข้อมูลได้อย่างรวดเร็ว
- การ interface ได้ง่าย
- การทำรายงาน
- การตอบคำถาม (Adhoc)
- การคำนวณ
- การนำไปเชื่อมต่อกับ tools อื่น ๆ (Export)

- EIS/DSS Tools

เป็นเครื่องมือที่ช่วยทางด้านการทำข้อมูลในลักษณะที่เป็นการช่วยสนับสนุนในการตัดสินใจ (Decision Support System - DSS) หรือช่วยในการทำข้อมูลเพื่อช่วยในการบริหารงานของผู้บริหารระดับสูง (Executive Information System - EIS)

2.3 องค์ประกอบในการสร้าง Data Warehouse

ในการสร้าง data warehouse นั้น ต้องมีการเลือกสรร hardware เพื่อนำมาเป็น server ซึ่งต้องมีการผสมผสาน platform 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

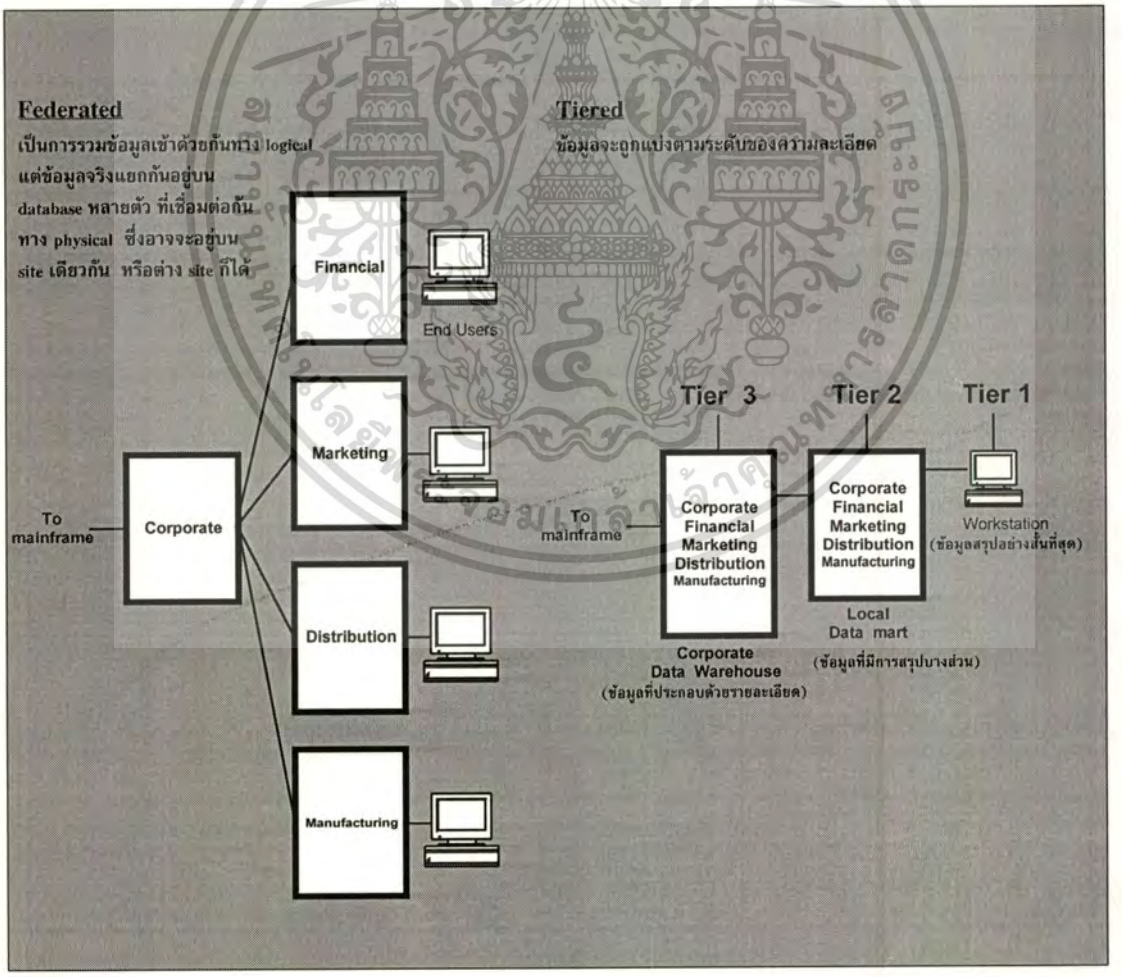
1. สถาปัตยกรรมข้อมูลของ Data Warehouse
2. สถาปัตยกรรมของ Server
3. สถาปัตยกรรมของ DBMS

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. สถาปัตยกรรมข้อมูลของ Data Warehouse

การพัฒนาคลังข้อมูล (Data Warehouse) ควรจะเริ่มจากโครงสร้างของฐานข้อมูลสำหรับ warehouse ทั้งด้าน logical และ physical โดยมี 3 แนวทางในการพิจารณาเลือกใช้วิธีการจัดเก็บข้อมูล ดังนี้

1.1 การเลือกใช้ platform เป็นแบบรวมศูนย์ (Centralized) ที่ติดตั้งอยู่ในบริเวณเดียวกัน หรือแบบกระจาย (Distributed) ที่แยก warehouse ออกไปตามสำนักงาน จังหวัด โดยอาจรวบรวมข้อมูลจากแหล่งภายใน และภายนอกหลายแหล่ง มาเก็บไว้ใน data warehouse เดียวกัน แนวทางนี้เป็นการรวบรวมข้อมูลเข้าไว้ด้วยกัน ซึ่งจะเป็นการช่วยเพิ่มความสามารถในการประมวลผลข้อมูลได้



ภาพที่ 2.2 สถาปัตยกรรม แบบ Federated และแบบ Tiered

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.2 การกระจายข้อมูลออกไปยัง database หลายตัวตาม function เป็นแบบ federated แนวทางนี้ข้อมูลบางส่วนอาจถูกจัดเก็บไว้ใน server ของ site แห่งหนึ่ง ขณะที่ข้อมูลอีกส่วนหนึ่งอาจถูกแยกออกไปอยู่อีกแห่งหนึ่ง

1.3 การเลือกใช้สถาปัตยกรรมแบบลำดับขั้น (Tiered Architecture) แนวทางนี้จะเก็บข้อมูลสรุปที่มีข้อมูลกระชับที่สุดบน workstation ของผู้ใช้ในระดับแรก ขณะที่ข้อมูลบน server ในระดับที่สองจะมีรายละเอียดเพิ่มขึ้น และมีรายละเอียดข้อมูลสูงที่สุดในระดับที่สาม

2. สถาปัตยกรรมของ Server

การพิจารณาเลือก 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 เพียง 1 ตัว server ประเภทนี้จัดการได้ง่าย แต่มีข้อจำกัดในด้านของการประมวลผล และความสามารถในการขยายระบบ รวมทั้งยังมี single point of failure ทำให้มีความเสี่ยงสูง และมีความน่าเชื่อถือของระบบไม่ดีนัก

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

2.3 Massively parallel processing server เป็น server ที่ใช้ link ความเร็วสูง และ bandwidth สูง เชื่อมต่อกับ array processor

3. สถาปัตยกรรมของ DBMS

ระบบจัดการฐานข้อมูล (DBMS) เป็นสิ่งที่สำคัญ ซึ่งต้องมีการพิจารณาเลือกใช้ให้เหมาะสม และสามารถสนับสนุนการทำงานของ server ที่ต้องการด้วย ในปัจจุบันมี DBMS ที่นิยมใช้อยู่ 3 ประเภท ได้แก่

3.1 Relational database management system DBMS ประเภทนี้จะมีความสามารถในการเรียกค้นข้อมูลแบบทันที และทำให้ข้อมูลไม่มีความซ้ำซ้อน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2 Multidimensional database เป็นการจัดโครงสร้างฐานข้อมูลแบบหลายมิติ โดยใช้เทคนิคด้าน storage และ operators เข้ามาช่วย เพื่อให้สนับสนุนลักษณะของโครงสร้างข้อมูล

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

Data Warehouse Decision Matrix					
For these environments			Choose		
Business requirements	**User population**	**Systems support**	**Architecture**	**Server**	**DBMS**
Scope : departmental **Uses** : data analysis	Small; single location	Minimal local; average central	Consolidated; turnkey package	Single-processor or SMP	MDDB
Scope : departmental **Uses** : analysis plus information	Large; analysts at single location; information users dispersed	Minimal local; average central	Tiered; detail at central; summary at local	Clustered SMP for central; SP or SMP for local	RDBMS for central; MDDB for local
Scope : enterprise **Uses** : analysis plus information	Large; geographically dispersed	Strong central	Centralized	Clustered SMP	Object-relational; Web support
Scope : departmental **Uses** : exploratory	Small; few sites	Strong central	Centralized	MPP	RDBMS with parallel support

ตารางที่ 2.1 Data Warehouse Decision Matrix

2.4 คุณลักษณะของข้อมูลใน Data Warehouse

ข้อมูลที่นำเข้ามาเก็บใน data warehouse นั้น จะมาจากแหล่งข้อมูล จากสภาพแวดล้อมต่างกัน จึงต้องมีการกำหนดคุณลักษณะของข้อมูลที่จะจัดเก็บให้เหมาะสม ดังนี้

- Subject-Oriented ข้อมูลที่สร้างขึ้นใน data warehouse จะประกอบด้วยหัวข้อเรื่องที่สนใจ ซึ่งจะเป็นการเก็บเฉพาะข้อมูลข่าวสารที่จำเป็นต่อกระบวนการตัดสินใจขององค์กร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- Integrated การรวบรวมข้อมูลจากแหล่งต่างๆ ซึ่งมีการกำหนดค่าของตัวแปรตัวเดียวกัน ซึ่งเป็นค่าที่ต่างกัน เมื่อนำมาลงใน data warehouse แล้วจะต้องกำหนดค่าของตัวแปรเหล่านั้นให้เหมือนกัน เพื่อให้เกิดความสอดคล้องของข้อมูล
- Time variancy ข้อมูลที่นำมาจัดเก็บจะเป็นข้อมูลย้อนหลัง ซึ่งมีระยะเวลาประมาณ 5 ถึง 10 ปี เพื่อประโยชน์ในการที่จะนำข้อมูลมาใช้เปรียบเทียบกันได้
- Nonvolatile ข้อมูลใน data warehouse จะต้องมีการควบคุม เพื่อไม่ให้ถูกเปลี่ยนแปลงได้ง่ายจากผู้ใช้ โดยที่ผู้ใช้จะมีความสามารถในการเรียกใช้ และเข้าถึงข้อมูล (access) เท่านั้น

2.5 การแปลงข้อมูลเข้าสู่ Data Warehouse

คุณภาพในการแปลงข้อมูล เป็นสิ่งที่สำคัญมากสำหรับ data warehouse ความซับซ้อนของกระบวนการแปลงข้อมูล จะแตกต่างกันไปสำหรับ data warehouse แต่ละระบบ รวมทั้งโครงสร้างของแหล่งข้อมูล ดังนั้นก่อนการแปลงข้อมูลต้องมีการออกแบบ จัดเตรียม physical model ของข้อมูลสำหรับ data warehouse และทำการสร้างขอบเขตของเป้าหมายให้เสร็จสมบูรณ์ก่อน หลังจากนั้นจึงดำเนินการตามกระบวนการ ดังนี้

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

1. การทำแผนการแปลงข้อมูล

กระบวนการทำงานแรก คือ การพัฒนาแผนการแปลงข้อมูล โดยที่จะต้องเข้าใจในความต้องการพื้นฐาน และการไหลของข้อมูล เพื่อให้กระบวนการต่างๆ สามารถเชื่อมโยงเข้าด้วยกันอย่างเหมาะสม ซึ่งได้แก่

- การย้ายข้อมูลจากระบบเดิมเข้าสู่บริเวณที่ใช้ปรับแต่งข้อมูล (staging area) เพื่อปรับแต่งข้อมูลให้เป็นไปตามความต้องการขององค์กร
- การสร้างเงื่อนไข การทำความสะอาด การแปลงและรวมข้อมูลเข้าด้วยกัน
- การสร้าง primary key และ foreign key ให้เชื่อมโยงกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- การย้ายข้อมูลลงสู่ server ของ data warehouse
- การเก็บรายละเอียดของข้อมูลที่มีการปรับปรุง และส่งออกไปไว้ใน metadata ของ data warehouse
- การเก็บข้อมูล และ ทำดัชนีอ้างอิงใน warehouse
- การเตรียมการณ์เพื่อรองรับข้อมูลที่จะมีเพิ่มขึ้นบน data warehouse
- การตรวจสอบความถูกต้องของข้อมูลใน data warehouse

2. การสร้างกฎเกณฑ์ของการแปลงข้อมูล

ในกระบวนการสร้างกฎเกณฑ์ของการแปลงข้อมูลนั้น ต้องมีการวิเคราะห์แหล่งข้อมูลอย่างละเอียดรอบคอบ เพื่อกำหนด attribute ของข้อมูล การ map ข้อมูลจากแหล่งเดิมไปลง warehouse และ logic ที่ใช้ในการเคลื่อนย้ายข้อมูล รวมทั้งการกำหนดสารสนเทศภายนอก แล้วจึงทำเป็น model ชั่วคราว หลังจากนั้นจึงเป็นการจัดทำโปรแกรม (coding program) ซึ่งต้องประกอบด้วย routine ที่มีการทำหน้าที่ ดังนี้

- การดึงข้อมูลจากระบบเดิมเข้าสู่ intermediate schema ที่อยู่ใน staging area
- การแปลงจาก intermediate schema มาเป็นข้อมูลที่พร้อมจะ load
- การรวบรวมข้อมูลใน staging area เข้าด้วยกัน
- การย้ายข้อมูลจาก staging area เข้าสู่ server ของ data warehouse (ถ้าข้อมูลใน staging area ไม่ได้อยู่ใน server เดียวกันกับ warehouse)
- load ข้อมูล ลงสู่ data warehouse

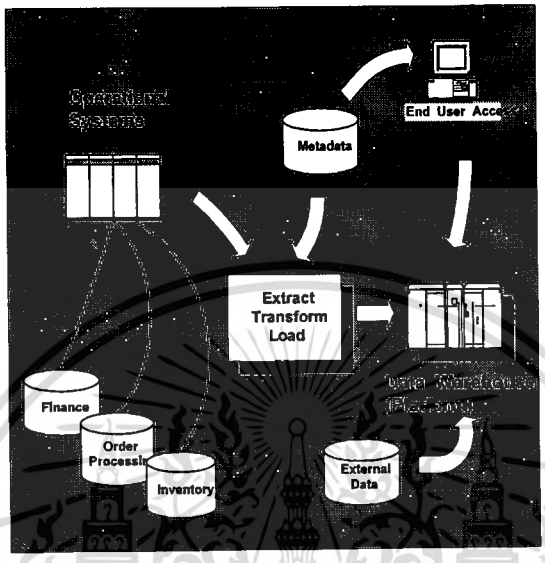
3. การดึงข้อมูลเข้าสู่ intermediate schema

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

- การตรวจสอบข้อมูล เพื่อดูคุณภาพของข้อมูล (quality of data) รูปแบบของข้อมูล (format of data) และจำนวนของ field ต่างๆที่นำมาใช้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- การกระจายข้อมูล เพื่อที่จะค้นหาข้อมูล และจุดหมายของแต่ละองค์ประกอบในแต่ละ field



ภาพที่ 2.3 กระบวนการแปลงข้อมูล

- การตรวจสอบข้อมูลว่า ตรงกับข้อมูลเดิมที่มีอยู่ และ field ทั้งหมดมีความถูกต้อง หรือได้รับการตรวจแก้โดยอัตโนมัติ
- การตรวจดูว่าข้อมูลใน record ตรงกัน และข้อมูลเป็นชนิดเดียวกัน

4. การรวบรวมข้อมูล

การรวบรวมข้อมูลเป็นการนำเอาข้อมูลทั้งหมดที่เป็นลักษณะตาราง (dimension table) และข้อมูลจริงลงสู่ data warehouse โดยข้อมูลจะถูกรวบรวมตามการออกแบบใน data warehouse ซึ่งต้องดำเนินการดังนี้

- ทำการ execute ตามลำดับของการจัดเรียง และทำการสร้าง primary key และ foreign key ทำให้ได้ข้อมูลที่เป็นตาราง และข้อมูลจริง
- ย้ายข้อมูลจาก staging area ลงสู่ server ของ data warehouse

5. ความถูกต้องของข้อมูล

การตรวจสอบความถูกต้องของข้อมูล เป็นสิ่งที่ต้องกระทำตลอดกระบวนการแปลงข้อมูล โดยต้องมีกระบวนการตรวจแก้ความถูกต้องของข้อมูล กระบวนการสำหรับทำให้

เอกสารนี้สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อมูลสอดคล้องกับระบบเดิม รวมถึงการตรวจสอบดูสภาพแวดล้อมในปัจจุบัน เพื่อกำหนดวิธีการที่ดีที่สุด เพื่อเคลื่อนย้ายข้อมูลจากแหล่งข้อมูลเดิมไปยัง data warehouse ซึ่งจะทำได้ ดังนี้

- การตรวจสอบผลรวมทั้งหมดของจำนวนข้อมูลที่ดึงมาจากแหล่งข้อมูล กับข้อมูลที่เพิ่มเข้าไป ใน intermediate schema
- การตรวจแก้ไขข้อมูลในระบบเดิมของแหล่งข้อมูล หรือใน routine ของการแปลงข้อมูล ซึ่งควรจะเก็บข้อมูลในการตรวจแก้ไว้ใน metadata ของการแปลงข้อมูลด้วย
- การตรวจสอบค่าของข้อมูลให้ถูกต้อง ในกระบวนการรวบรวมข้อมูล

2.6 สิ่งที่ต้องคำนึงถึงในการออกแบบ Data Warehouse

ในการออกแบบ data warehouse มีสิ่งที่ต้องพิจารณา 3 ส่วนด้วยกัน ได้แก่

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

1. Target Database

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

2. Data Extraction Technique

ในการที่จะทำการดึงข้อมูลมาใช้ โดยการแยกข้อมูลบางส่วน จะใช้วิธีการจัดการโดยการใช้ facility ที่มีมากับ DBMS นั้น หรือจะใช้เครื่องมือที่จัดทำมาเพื่อทำการดึงข้อมูลโดยเฉพาะ ที่ผู้ผลิตได้จัดทำขึ้นมา

3. Front-End Data Access Tools

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

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.7 Software ที่ใช้จัดการข้อมูล

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

1. Prism Warehouse Manager
2. IBM's Information Warehouse
3. Sybase Warehouse Works

1. Prism Warehouse Manager

- Warehouse Manager ช่วยทำการดึงข้อมูลจากฐานข้อมูลส่วนปฏิบัติงานต่างๆ นำมาทำการแปลงข้อมูล แล้วจึงจะ load ข้อมูล เข้าสู่ data warehouse หลังจากนั้นจะทำการสร้าง และจัดการ metadata ให้โดยอัตโนมัติ ผู้ใช้สามารถที่จะกำหนดการ mapping ระหว่างแหล่งข้อมูลที่ต้องการ กับ data warehouse ได้อย่างง่ายๆ
- Changed Data Capture เป็นการทำ refresh ข้อมูล โดยทำการ reload ข้อมูลเข้าสู่ data warehouse ตามช่วงระยะเวลาที่กำหนด
- Directory Manager ช่วยทำการจัดเก็บข้อมูลต่างๆ ของ metadata ซึ่งได้แก่ รายละเอียดของข้อมูลในตารางฐานข้อมูล attribute ต่างๆ และการ mapping

2. IBM's Information Warehouse

- DataPropagator Relational เป็นการถ่ายข้อมูล โดยในการถ่ายข้อมูลจากแหล่งต้นทาง ไปยังปลายทาง จะต้องมีการเก็บข้อมูลที่เปลี่ยนแปลงไว้ในตารางกลางก่อน แล้วจึงคัดลอกข้อมูลส่งไปยังฐานข้อมูลที่ต้องการ
- DataPropagator NonRelational เป็นการช่วยในการทำให้การถ่ายทอดข้อมูลจากฐานข้อมูลในแบบ nonrelational ไปสู่ฐานข้อมูลในแบบ relational ได้
- DataRefresher เป็นการย้ายข้อมูลที่มีอยู่ทั้งหมดไปสู่ data warehouse โดยที่ข้อมูลจะถูก reload และ แทนที่ใหม่ทุกครั้ง
- DataJoiner เป็นการใช้เทคโนโลยีด้าน optimizer เพื่อเชื่อมข้อมูลจากระบบฐานข้อมูลที่ต่าง product กัน ให้เข้ามาเป็นฐานข้อมูลเดียวกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- Datahub เป็นเครื่องมือที่ช่วยงานการดูแลฐานข้อมูล ที่มี warehouse กระจายอยู่บนระบบที่ต่างกัน ให้ดูแลได้ง่าย และประหยัดเวลา
- DataGuide เป็นเครื่องมือที่ช่วยในการจัดหมวดหมู่ข้อมูลสารสนเทศ ทำให้ผู้ใช้สามารถมองเห็นข้อมูลได้ง่าย
- Lotus NotesPump เป็นการดึงข้อมูลจากฐานข้อมูลของ Lotus Notes เข้าไปสู่ staging area ของ DataPropagator เพื่อช่วยในการการทำ replication ไปยัง data warehouse ที่ต้องการ

3. Sybase Warehouse Works

- Sybase Replication Server 11 เป็นการทำ replication โดยที่จะทำการอ่าน transaction log ของฐานข้อมูล แล้วจึงจะทำการ replicate ข้อมูลที่มีการเปลี่ยนแปลงไปยัง data warehouse
- Tranzic InfoPump ช่วยในด้านของการจัดการแปลงข้อมูล และจัดทำการ cleansing ข้อมูล โดยมีการใช้ภาษา script และ trigger
- Sybase IQ ช่วยในการค้นหาข้อมูลสำหรับการ adhoc โดยวิธีการทำ index optimizer ทำให้ได้ผลลัพธ์ที่รวดเร็ว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ห้องสมุดคณะเทคโนโลยีสารสนเทศ สจล.

บทที่ 3
ระบบสนับสนุนการตัดสินใจ (Decision Support System)

3.1 ความหมายของระบบสนับสนุนการตัดสินใจ

เป็นระบบที่ช่วยในการสนับสนุนการตัดสินใจของผู้บริหาร โดยจะทำการรวบรวมข้อมูลทั้งหมดที่เกิดขึ้น เพื่อที่จะทำนายผลที่จะเกิดในอนาคต เป็นการให้ผู้ตัดสินใจได้รู้ว่าเหตุการณ์ที่เกิดขึ้นจะเป็นอย่างไร โดยในระบบนี้จะมีแบบจำลองการตัดสินใจในสภาวะต่างๆ ที่ช่วยในการตัดสินใจให้กับผู้ที่จะตัดสินใจได้พิจารณาว่า ถ้าหากเลือกที่จะตัดสินใจทางใดทางหนึ่งแล้ว ผลที่เกิดขึ้นจะเป็นอย่างไร ทำให้ผู้บริหารสามารถทดสอบได้ ระบบนี้จึงเป็นระบบที่ต้องจัดการกับข้อมูลที่มีปริมาณมาก โดยจะทำการนำเอาข้อมูลทั้งหมดมาประมวลผล เพื่อหาทางเลือกที่ดีที่สุดที่มีอยู่ ระบบนี้มักจะใช้ในด้านของการตลาดและการขาย

3.2 ลักษณะของระบบ

1. เป็นระบบที่จำลองแบบการตัดสินใจ เพื่อที่จะช่วยทำให้ผู้บริหารสามารถทดสอบการตัดสินใจได้
2. เป็นระบบที่ช่วยสนับสนุนการตัดสินใจ แต่ไม่ใช่ระบบที่ตัดสินใจแทนผู้บริหาร
3. ระบบนี้เหมาะสำหรับการตัดสินใจแบบไม่มีโครงสร้าง และกึ่งโครงสร้าง ซึ่งเป็นสารสนเทศ สำหรับช่วยในการตัดสินใจ
4. เป็นการใช้ข้อมูลจากระบบ Transaction Processing (TPS) และระบบ Management Information System (MIS)
5. ระบบต้องมีความสามารถที่จะค้นหาข้อมูลที่ต้องการได้อย่างรวดเร็ว
6. การค้นหาข้อมูลจะเกี่ยวข้องกับรูปแบบของข้อมูลในหลายๆ รูปแบบ ข้อมูลที่ต้องการค้นหาอาจจะเป็นกราฟ รูปภาพ หรือเป็นภาพเคลื่อนไหว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

7. การค้นหาข้อมูลจะมีความซับซ้อนมากกว่าปกติ เนื่องจากข้อมูลที่นำมาใช้นั้น จะมีปริมาณมาก

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

3.3 ส่วนประกอบของระบบ

ระบบสนับสนุนการตัดสินใจ มีส่วนประกอบที่สำคัญ 3 ส่วน คือ

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

3.4 ขั้นตอนของการจัดทำระบบสนับสนุนการตัดสินใจ

ในการที่จะจัดทำระบบสนับสนุนการตัดสินใจ จะมีการแบ่งขั้นตอนของการทำงานออกเป็น 10 ขั้นตอน ดังนี้

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

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. Project planning

การวางแผนสำหรับการจัดทำระบบคลังข้อมูล เพื่อช่วยสนับสนุนการตัดสินใจ ประกอบด้วย

- กำหนดขอบเขตของโครงการที่จะจัดทำให้ชัดเจน
- จัดทำแผนงานของโครงการ
- กำหนดทรัพยากรที่จำเป็นต้องใช้ ได้แก่ ข้อมูลทั้งภายใน และภายนอกองค์กร
- กำหนดผู้มีส่วนรับผิดชอบในการทำงาน
- กำหนดลักษณะของงานที่ต้องจัดทำทั้งหมด
- กำหนดระยะเวลาต่างๆ ในการทำงาน

ซึ่งในขั้นตอนนี้จะต้องคำนึงถึงการใช้เทคโนโลยีต่างๆ เช่น ระบบเครือข่ายที่จะนำมาใช้งาน (LAN หรือ WAN) การเลือกใช้ platform การเลือกใช้ database และการเลือกใช้ front-end access tool

2. Gathering data requirements and modeling

ขั้นตอนนี้เป็นการเก็บรวบรวมความต้องการ ในการที่จะใช้ข้อมูลในลักษณะต่างๆ จากผู้ใช้ ประกอบด้วย

- การรวบรวมความต้องการในการใช้ข้อมูล โดยการสัมภาษณ์ผู้ใช้
- จัดทำเอกสารที่เกี่ยวข้อง
- กำหนด logical model
- สร้าง dimensional business model

3. Physical database design and development

ขั้นตอนนี้เป็นการออกแบบฐานข้อมูล และการทำ denormalized ข้อมูล รวมถึงการออกแบบฐานข้อมูลในแบบ Star schema ดังนี้

- การออกแบบฐานข้อมูลที่จะใช้
- การทำ denormalized ข้อมูล
- กำหนด key ประเภทต่างๆ ของข้อมูล ที่อยู่บนฐานข้อมูล
- สร้าง index key

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4. Data sourcing integration and mapping

ขั้นตอนนี้จะเป็นการเลือกแหล่งข้อมูลที่จะนำข้อมูลมาใช้ การรวมข้อมูล และการดึงข้อมูล ได้แก่

- กำหนดแหล่งของข้อมูลที่จะใช้
- กำหนดรูปแบบของข้อมูล
- วิเคราะห์ข้อมูล เพื่อทำการรวมข้อมูลเข้าสู่ data warehouse
- จัดทำการแปลงข้อมูล
- ทำการดึงข้อมูลจากแหล่งต่างๆ

5. Populating the Data Warehouse

เป็นขั้นตอนของการทำการ extract, convert และ load data เข้าสู่ฐานข้อมูลของ data warehouse ประกอบด้วย

- การพัฒนาโปรแกรม หรือใช้เครื่องมือในการที่จะ extract และ move data
- การพัฒนาและเลือกวิธีการที่จะ load data
- การพัฒนาโปรแกรม หรือเลือกใช้เครื่องมือ เพื่อทำการ convert data
- การพัฒนา และปรับปรุงวิธีการในการ refresh data
- จัดทำการทดสอบกระบวนการต่างๆ ได้แก่ การ extract, การ integration data

6. Automating the data management process

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

- กระบวนการในการทำ data extraction อย่างอัตโนมัติตามเวลาที่กำหนด
- กระบวนการในการทำ data conversion อย่างอัตโนมัติตามเวลาที่กำหนด
- กระบวนการในการทำ data load อย่างอัตโนมัติตามเวลาที่กำหนด
- การจัดทำขั้นตอนในการ backup และ recovery
- จัดทำการทดสอบระบบการทำงานที่เป็นไปอย่างอัตโนมัติในทุกขั้นตอน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

7. Creating the starter set of reports

ขั้นตอนนี้เป็นการสร้าง และจัดทำรายงานต่างๆ ได้แก่

- จัดสร้างรูปแบบของรายงานเบื้องต้น
- ทดสอบผลที่ได้รับจากรายงาน
- จัดทำเอกสาร และคู่มือในการใช้งาน

8. Data validation and testing

ขั้นตอนนี้เป็นการตรวจสอบ และทดสอบข้อมูล เพื่อให้ข้อมูลที่ได้มานั้นมีความถูกต้องมากที่สุด ประกอบด้วย

- การตรวจสอบข้อมูล โดยดูจากผลลัพธ์ในรายงานเบื้องต้นที่จัดทำ
- การตรวจสอบข้อมูล โดยใช้วิธีการกำหนดมาตรฐานขึ้นมา
- การทดสอบ โดยการเปลี่ยนแปลงข้อมูล

9. Training

ขั้นตอนนี้เป็นการฝึกอบรมให้ผู้ใช้ได้ทราบวิธีการต่างๆ ในการที่จะใช้งานระบบ เช่น ข้อมูล การใช้ front-end เป็นต้น ซึ่งได้แก่

- ขอบเขตของข้อมูลที่อยู่ใน data warehouse
- วิธีการในการใช้เครื่องมือที่เป็น front-end
- วิธีการที่จะเข้าถึง และดึงข้อมูลไปใช้งาน
- รายงานเบื้องต้นที่ผู้ใช้สามารถกระทำได้

10. Rollout

ขั้นตอนนี้เป็นการติดตั้งระบบ และการบำรุงรักษาระบบ ได้แก่

- จัดทำการติดตั้งระบบให้ผู้ใช้ ได้ใช้งานจริง
- จัดทำขั้นตอนในการพัฒนา เพื่อจัดทำรายงานในรูปแบบต่างๆ
- จัดทำวิธีการในการที่จะทำการ backup ข้อมูล และระบบงาน
- จัดทำแนวทางในการที่จะทำให้ข้อมูลที่อยู่ในระบบทั้งหมด มีความถูกต้องและสอดคล้องกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.5 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 4

เครื่องมือที่ใช้ในการวิเคราะห์ข้อมูล (Tools for Analyse Data)

ในการที่จะนำข้อมูลที่จัดเก็บไว้บน data warehouse มาใช้ประโยชน์นั้น จำเป็นต้องมีเครื่องมือที่มีความสามารถในการที่จะวิเคราะห์ข้อมูลเหล่านั้นได้อย่างมีประสิทธิภาพและเหมาะสม รวมทั้งมีความสามารถในการจัดการข้อมูลในลักษณะที่เป็น Online Analytical Processing ด้วย

Pilot Decision Support Suite เป็นโปรแกรมที่ได้มีการจัดเตรียมการวิเคราะห์ข้อมูลในรูปแบบต่างๆ เพื่อสนับสนุนความต้องการใช้ข้อมูลของผู้บริหาร และช่วยในการพัฒนาระบบงานของผู้พัฒนาระบบได้เป็นอย่างดี

ภาพที่ 4.1 ชุดของโปรแกรม Pilot Decision Support Suite

4.1 ข้อกำหนดความต้องการของระบบ

Pilot Decision Support Suite สามารถติดตั้งเพื่อให้ผู้ใช้ ใช้งานได้ในหลายรูปแบบ ไม่ว่าจะเป็น single user, workgroup และ client/server โดยสนับสนุนการทำงานบนเครื่องคอมพิวเตอร์ที่มี platform ต่างๆ ได้แก่ Windows 3.1, Windows 95, Windows NT และ UNIX

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

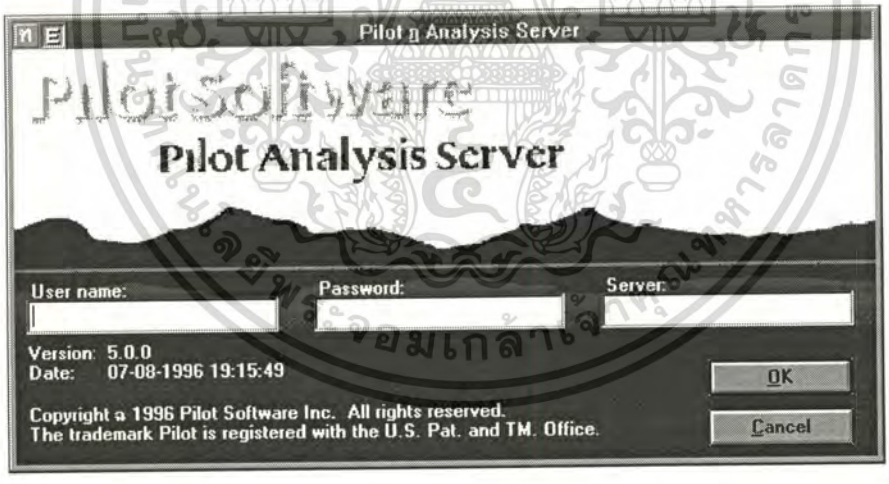
4.2 ส่วนประกอบของ Pilot Decision Support Suite

โปรแกรม Pilot Decision Support Suite เป็นโปรแกรมที่ถูกพัฒนาขึ้นมา เพื่อใช้ในการวิเคราะห์ข้อมูล ซึ่งจะประกอบด้วยโปรแกรมที่สำคัญ 4 ส่วน คือ

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

1. Pilot Analysis Server

เป็นโปรแกรมที่ใช้สนับสนุนการทำงานของผู้ใช้หลายคน และช่วยในการควบคุม รวมทั้งจัดการข้อมูลที่อยู่บนฐานข้อมูล



ภาพที่ 4.2 Pilot Analysis Server

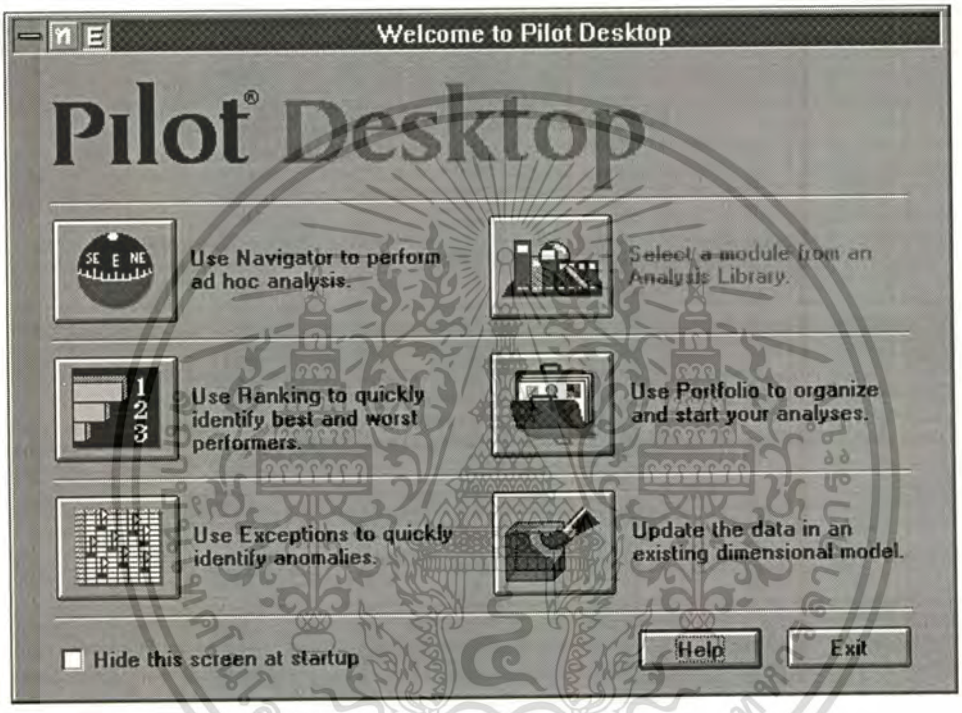
2. Pilot Discovery Server

เป็นโปรแกรมที่ใช้ช่วยสนับสนุนการทำงานกับข้อมูลที่อยู่บน data warehouse ที่มีขนาดใหญ่ และข้อมูลมีความซับซ้อนมาก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. Pilot Desktop

เป็นโปรแกรมที่อำนวยความสะดวกให้กับผู้ใช้ เพื่อให้ผู้ใช้สามารถใช้งานได้ง่าย โดยการเลือกวิเคราะห์ข้อมูลได้ในหลายรูปแบบ ซึ่งได้แก่ Navigator Adhoc, Ranking Analysis, Exception Analysis และ Portfolio



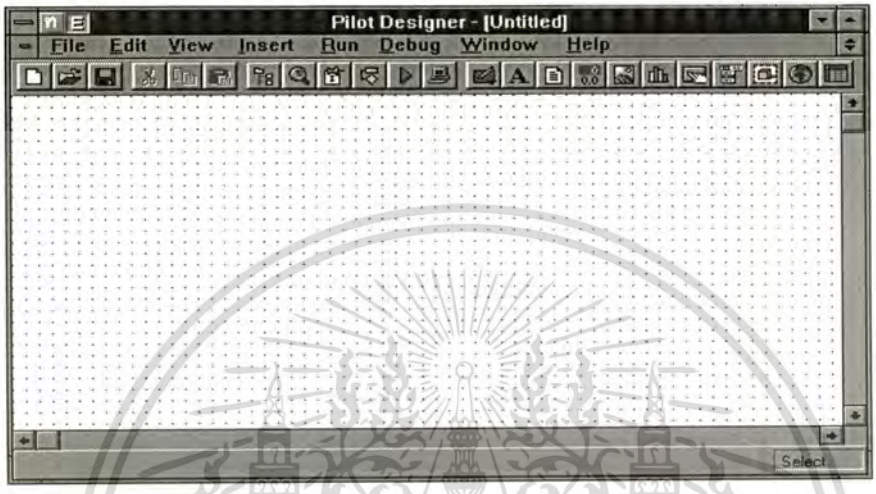
ภาพที่ 4.3 Pilot Desktop

- Navigator Adhoc ใช้สำหรับการตอบคำถาม (Adhoc) หรือจัดทำรายงานแบบเร่งด่วนได้ สามารถแสดงผลในรูปแบบของกราฟประเภทต่างๆได้
- Ranking Analysis ใช้สำหรับการจัดลำดับของข้อมูลตาม dimension ที่สนใจ โดยการเรียงข้อมูลจากต่ำสุดไปสูงสุด หรือในทางกลับกัน จากสูงสุดไปต่ำสุดได้
- Exception Analysis ใช้ในการวิเคราะห์ข้อมูล และนำข้อมูลในแต่ละ attribute มาเปรียบเทียบ เพื่อคำนวณหาแนวโน้มได้
- Portfolio ใช้จัดเก็บผลการวิเคราะห์ข้อมูลในรูปแบบต่างๆ ไว้ เพื่อการเรียกดู หรือออกรายงานได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4. *Pilot Designer*

เป็นโปรแกรมที่ใช้สำหรับการออกแบบ และพัฒนาระบบงานตามรูปแบบที่ผู้ใช้ต้องการ



ภาพที่ 4.4 Pilot Designer

4.3 คุณสมบัติของ Pilot Decision Support Suite

1. เป็นระบบ data warehouse ที่เก็บรวบรวมข้อมูลที่ผ่านการประมวลผลแล้วมาใช้งาน
2. สนับสนุนการทำงานระบบ Online Analytical Processing (OLAP)
3. สนับสนุนการค้นหาข้อมูลในแบบ Adhoc query
4. สามารถสร้างฐานข้อมูลแบบหลายมิติได้ (Multidimensional database)
5. ไม่เหมาะสำหรับการจัดการข้อมูลที่เป็นแบบ Transaction processing
6. ผู้ใช้สามารถใช้งานได้ง่าย โดยการใช้ Pilot desktop หรือพัฒนา application ได้เอง โดยการใช้ Pilot designer
7. สามารถทำงานกับข้อมูลได้หลายประเภท เช่น
   - Microsoft Excel (แฟ้มข้อมูลที่เป็น .xls)
   - Database file (แฟ้มข้อมูลที่เป็น .dbf)
   - ASCII text file (แฟ้มข้อมูลที่เป็น .txt)
   - Relational database เช่น DB2, INFORMIX, INGRES, Oracle, SQL server

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

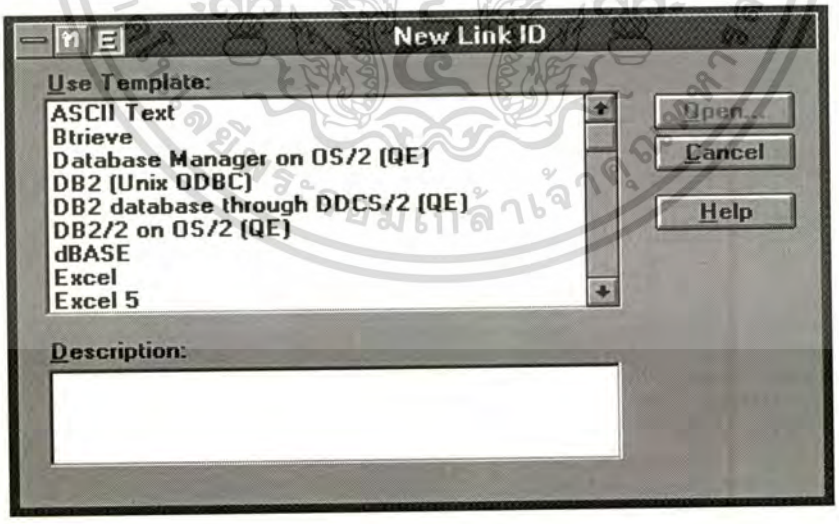
4.4 การทำงานของ Pilot Decision Support Suite

โปรแกรม Pilot Decision Support Suite จะประกอบด้วยขั้นตอนการทำงานที่สำคัญ 4 ส่วน ได้แก่

1. Link ID Configuration
2. Model Builder
3. Analysis Server
4. การวิเคราะห์ข้อมูล

1. Link ID Configuration

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



ภาพที่ 4.5 การ Link Configuration

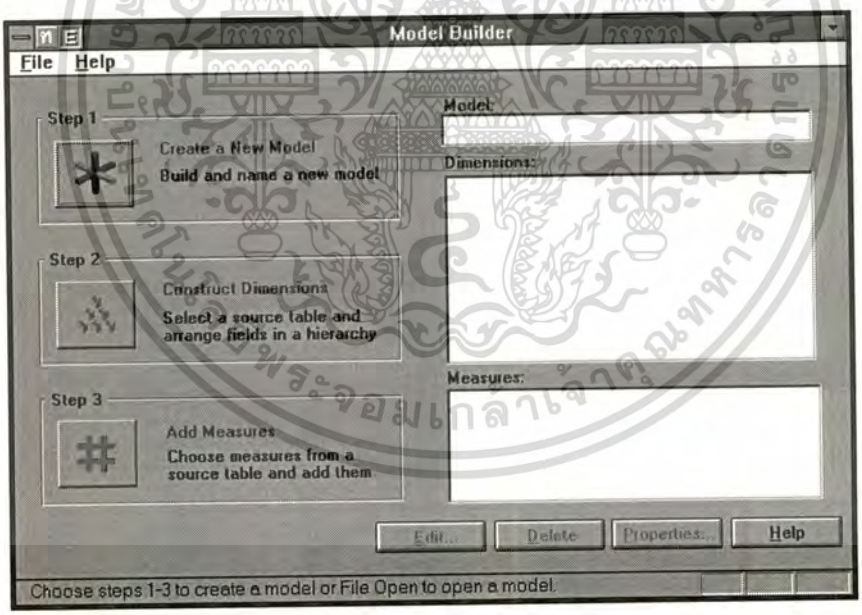
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. *Model Builder*

ทำหน้าที่ในการสร้างรูปแบบ (model) ของข้อมูลที่ต้องการเพื่อการใช้งาน โดยจะเรียกใช้ข้อมูลจากฐานข้อมูลที่ได้เชื่อมโยงไว้แล้ว เข้ามาสู่ model ที่สร้างไว้

ในการสร้าง model จะประกอบด้ วย 3 ขั้นตอน ดังนี้

- Create Model เป็นขั้นตอนในการสร้าง และกำหนดชื่อ model
- Construct Dimensions เป็นขั้นตอนในการที่จะกำหนด และจัดทำ dimension ของข้อมูล ที่ต้องการจะให้เป็นมุมมองข้อมูลในรูปแบบต่างๆ
- Add Measures เป็นขั้นตอนในการที่จะกำหนด attribute ที่ต้องการให้เป็นค่าที่แสดงผล และกำหนดวิธีการที่ใช้ในการคำนวณ เช่น การหาค่าเฉลี่ย (average) การหาค่าผลรวม (sum) เป็นต้น



ภาพที่ 4.6 Model Builder

ในการเลือกใช้ข้อมูลจากฐานข้อมูล สามารถที่จะเลือกใช้ข้อมูลจากตารางได้หลายตาราง (table) หรือใช้ข้อมูลจากตารางเดียวก็ได้ ขึ้นอยู่กับความต้องการใช้ attribute ที่อยู่ในตารางนั้น และในการสร้าง model จะต้องกำหนดช่วงเวลา (time) ให้กับข้อมูลด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. *Analysis Server*

ขั้นตอนนี้เป็นการทำงานของระบบจัดการของโปรแกรม Pilot Decision Support Suite โดยการจัดเก็บข้อมูลที่ได้ผ่านการเชื่อมโยงกับฐานข้อมูล และจัดการ model ที่ได้สร้างขึ้นรวมเข้าไว้ด้วยกัน เพื่อใช้ในการวิเคราะห์ต่อไป

4. *การวิเคราะห์ข้อมูล*

ในการที่จะวิเคราะห์ข้อมูลตาม model ที่ได้สร้างไว้แล้วนั้น ทำได้ 2 วิธี ได้แก่

- การใช้ Pilot desktop ซึ่งจะมี wizard ที่ช่วยให้ผู้ใช้สามารถใช้งานได้ง่ายเพื่อเลือกวิเคราะห์ข้อมูลตามที่สนใจได้
- การใช้ Pilot designer เพื่อพัฒนาระบบงานตามความต้องการได้ โดยการใช้ภาษา Basic script

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5
ภาษีมูลค่าเพิ่ม (Value Added Tax)

5.1 ภาษีมูลค่าเพิ่ม

ตามที่กรมสรรพากรได้นำระบบคอมพิวเตอร์มาใช้พัฒนาระบบภาษีสรรพากรเมื่อประมาณเดือนตุลาคม 2534 ซึ่งขณะนั้นรัฐบาลได้ประกาศให้ใช้ระบบภาษีมูลค่าเพิ่มแทนภาษีการค้า โดยให้มีผลบังคับใช้ตั้งแต่วันที่ 1 มกราคม 2535 เป็นต้นไป กรมสรรพากรจึงได้ทำการพัฒนาระบบภาษีมูลค่าเพิ่มด้วยระบบคอมพิวเตอร์ก่อนระบบอื่น โดยให้ชื่อว่า "ระบบคอมพิวเตอร์เฉพาะกาลสำหรับภาษีมูลค่าเพิ่ม" (Interim Value Added Tax - IVAT) โดยมีขอบเขตในการพัฒนาระบบด้านการจดทะเบียนภาษีมูลค่าเพิ่ม ด้านการประมวลผลแบบแสดงรายการภาษีมูลค่าเพิ่ม และในด้านของการทำรายงานต่างๆ ที่จำเป็น

ระบบภาษีมูลค่าเพิ่ม IVAT ได้พัฒนาเสร็จ และติดตั้งใช้งานเมื่อวันที่ 14 กุมภาพันธ์ 2535 และใช้งานมาจนถึงเดือนพฤษภาคม 2537 กรมสรรพากรจึงได้นำระบบภาษีมูลค่าเพิ่มขั้นจำกัด (Limited Value Added Tax - LVAT) มาใช้ในเดือนมิถุนายน 2537

5.2 ผู้มีหน้าที่จดทะเบียนภาษีมูลค่าเพิ่ม

ผู้ประกอบการที่ขายสินค้าหรือให้บริการในทางธุรกิจหรือวิชาชีพเป็นปกติธุระ ไม่ว่าการขายสินค้าหรือให้บริการดังกล่าวนั้น จะได้รับประโยชน์หรือผลตอบแทนหรือไม่ และไม่ว่าจะประกอบกิจการในรูปของบุคคลธรรมดา คณะบุคคล นิติบุคคล ส่วนราชการหรือองค์การของรัฐบาล ถ้าประกอบกิจการรายใดมีรายรับเกินกว่า 600,000 บาทต่อปี ผู้ประกอบการนั้นมีหน้าที่ต้องจดทะเบียนภาษีมูลค่าเพิ่ม แต่ถ้าหากการประกอบกิจการนั้น เป็นกิจการที่ได้รับการยกเว้นภาษีมูลค่าเพิ่ม ผู้ประกอบการไม่มีหน้าที่ต้องจดทะเบียนภาษีมูลค่าเพิ่ม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.3 ประเภทของแบบแสดงรายการภาษีมูลค่าเพิ่มที่เกี่ยวข้อง

แบบแสดงรายการภาษีที่เกี่ยวข้องกับงานภาษีมูลค่าเพิ่ม จะประกอบด้วยแบบประเภทต่างๆ ที่สำคัญ ได้แก่

1. แบบ ภพ. 01
   - ใช้สำหรับการจดทะเบียนภาษีมูลค่าเพิ่ม
2. แบบ ภพ. 09
   - ใช้ในการเปลี่ยนแปลง หรือการแก้ไขรายการจดทะเบียนภาษีมูลค่าเพิ่ม เช่น การเปลี่ยนแปลงที่อยู่ เปลี่ยนชื่อสถานประกอบการ เป็นต้น
3. แบบ ภพ. 30
   - ใช้ในการยื่นแบบแสดงรายการภาษีมูลค่าเพิ่มที่จดทะเบียน เสียภาษีในอัตรา 10 %
4. แบบ ภพ. 31
   - ใช้ในการยื่นแบบแสดงรายการภาษีมูลค่าเพิ่มที่จดทะเบียน เสียภาษีในอัตรา 1.5 %

5.4 การประมวลผลระบบงานภาษีมูลค่าเพิ่ม

ในการประมวลผลภาษีมูลค่าเพิ่มนั้น จะมีการบันทึกข้อมูลที่สำนักงานสรรพากร แล้วทำการ transfer ข้อมูลเข้ามาประมวลผลบนเครื่องคอมพิวเตอร์ IBM ES/9000 ซึ่งเมื่อประมวลผลแล้ว จะนำรายการข้อมูลไปจัดเก็บบนฐานข้อมูล โดยกระจายลงตามตารางต่างๆ ที่เกี่ยวข้อง

5.5 โครงสร้างฐานข้อมูลของระบบงานภาษีมูลค่าเพิ่ม

โครงสร้างของฐานข้อมูลภาษีมูลค่าเพิ่มที่อยู่บนเครื่อง IBM ES/9000 จะประกอบด้วยฐานข้อมูลประเภทต่างๆ ดังนี้ คือ

1. Taxpayer Database
2. Account Database
3. Audit Database

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4. Taxpayer Form Database
5. MIS Database
6. Reference Database (RDB)
7. System Database
8. Text Database

Database ตามข้อที่ 1 - 4 จะมีการใช้เลขประจำตัวผู้เสียภาษี (TIN) หรือใช้เวลาประทับในการประมวลผลข้อมูล (Registration timestamp - REGTST) เป็น key ที่จะใช้ในการค้นหา และเชื่อมโยงข้อมูลในแต่ละ Table เข้าด้วยกัน

1. Taxpayer Database

เป็น database ที่เก็บข้อมูลรายละเอียดของผู้เสียภาษีที่มีการจดทะเบียนเข้ามาในระบบภาษีมูลค่าเพิ่มทั้งหมด ได้แก่ ข้อมูลการจดทะเบียนของผู้เสียภาษี ข้อมูลสถานประกอบการและสาขาของผู้เสียภาษี ข้อมูลแบบแสดงรายการต่างๆ ทุกประเภทที่ผู้เสียภาษีได้ยื่นแบบเข้ามาในระบบ ซึ่งประกอบด้วยตาราง (table) ที่สำคัญ ดังนี้

- TXPPSN เก็บข้อมูลของบุคคลธรรมดา
- TXPPSNADD เก็บข้อมูลที่อยู่ของบุคคลธรรมดา
- TXPJUR เก็บข้อมูลของนิติบุคคล
- TXPJURNAM เก็บข้อมูลชื่อของนิติบุคคล
- TXPBRAADD เก็บข้อมูลที่อยู่ของสาขา
- TXPFOR เก็บข้อมูลแบบแสดงรายการภาษีมูลค่าเพิ่มทุกประเภทที่ผู้เสียภาษีได้ยื่นเข้ามา
- TXPBUSTYP เก็บข้อมูลประเภทของการประกอบกิจการ
- TXPGOOTYP เก็บข้อมูลประเภทของสินค้าและบริการ
- TXPVAT เก็บข้อมูลอัตราภาษีมูลค่าเพิ่ม
- TXPTAXTYP เก็บข้อมูลรายการประเภทภาษี

2. Account Database

เป็น database ที่จัดเก็บข้อมูลของรายการบัญชีต่างๆ ที่เกิดขึ้นจากการประมวลผลข้อมูลภาษีมูลค่าเพิ่ม เช่น รายการบัญชีจำนวนเงินต่างๆ ที่ผู้เสียภาษีได้ยื่นแบบแสดงรายการภาษีมูลค่าเพิ่มเข้ามา ซึ่งได้แก่ตาราง ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- TACCTRA เก็บข้อมูลรายการบัญชีของผู้เสียภาษี
- TACCCRE เก็บข้อมูลรายการเครดิตภาษีของผู้เสียภาษี
- TACCDUEAMO เก็บข้อมูลการชำระเงินของผู้เสียภาษี

3. Audit Database

เป็น database ที่จัดเก็บข้อมูลที่ใช้ในการตรวจสอบ หรือเพื่อที่จะนำมาใช้ในการคัดเลือกรายที่ต้องการตรวจสอบได้ ประกอบด้วยตาราง ดังนี้

- AUDRST เก็บข้อมูลผลของการตรวจสอบ
- AUDPIK เก็บข้อมูลหลักเกณฑ์ในการคัดเลือกรายการตรวจสอบ

4. Taxpayer Form Database

เป็น database ที่จัดเก็บข้อมูลรายละเอียดของแบบแสดงรายการภาษีมูลค่าเพิ่มต่างๆ ที่มีการประมวลผลบนระบบ เช่น แบบ ภพ.30 แบบ ภพ.31 เป็นต้น ซึ่งประกอบด้วยตาราง ดังนี้

- TRETP30 เก็บข้อมูลแบบ ภพ.30
- TRETP31 เก็บข้อมูลแบบ ภพ.31

5. MIS Database

เป็น database ที่จัดเก็บข้อมูลสารสนเทศเพื่อการบริหาร ซึ่งจะเป็นการจัดเก็บข้อมูลสรุปของการประมวลผลในแต่ละวัน ไม่ว่าจะเป็นข้อมูลผู้เสียภาษี ข้อมูลแบบแสดงรายการ ข้อมูลรายการความผิดพลาด และรายการเตือน (Error and Warning) ประกอบด้วยตาราง ดังนี้

- MISTXP เก็บข้อมูลจำนวนผู้เสียภาษี
- MISFOR เก็บข้อมูลจำนวนแบบที่นำส่ง
- MISERRWAR เก็บข้อมูลรายการความผิดพลาด และรายการเตือน

6. Reference Database (RDB)

เป็น database ที่จัดเก็บข้อมูลรายการรหัสประเภทต่างๆ ที่นำมาใช้อ้างอิงในระบบ เช่น รหัสสำนักงาน รหัสคำนำหน้าชื่อ รหัสประเภทของสินค้าและบริการ รหัสของค่าคงที่ต่างๆ ซึ่งประกอบด้วยตาราง ที่สำคัญ ดังนี้

- RDBOFF เก็บข้อมูลรหัสสำนักงานสรรพากร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- RDBTIT เก็บข้อมูลรหัสคำนำหน้าชื่อ
- RDBGOO เก็บข้อมูลรหัสประเภทสินค้าและบริการ
- RDBBUS เก็บข้อมูลรหัสประเภทกิจการ
- RDBCON เก็บข้อมูลค่าคงที่ของระบบ
- RDBPER เก็บข้อมูลเดือนและปีภาษี
- RDBTRM เก็บข้อมูล terminal number ที่ใช้ในการบันทึกข้อมูล

7. System Database

เป็น database ที่จัดเก็บข้อมูลที่เกี่ยวข้องกับการทำงานของระบบ เช่น จำนวนแบบที่ส่งเข้ามาประมวลผล โปรแกรมที่ใช้ในการประมวลผล เป็นต้น ซึ่งประกอบด้วยตารางดังนี้

- BATFIL เก็บข้อมูลรายการชุดข้อมูล จำนวนแบบที่นำส่งเข้ามาประมวลผล
- SYSPGMTST เก็บข้อมูลโปรแกรมที่ใช้ในการประมวลผล และเวลาประทับการประมวลผลครั้งสุดท้าย

8. Text Database

เป็น database ที่เก็บข้อมูลเกี่ยวกับ error code ประกอบด้วยตาราง ดังนี้

- TXTERR เก็บข้อมูลรายการของ error code ต่างๆ
- TXTMON เก็บข้อมูลรหัสของเดือนภาษี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 6
การพัฒนาระบบคลังข้อมูลกับงานภาษีมูลค่าเพิ่ม
(Development of Data Warehouse and Value Added Tax)

เนื่องจากระบบงานภาษีมูลค่าเพิ่มนั้น เป็นการประมวลผลในลักษณะที่เป็นงาน batch โดยทำการประมวลผลเป็นประจำวัน ข้อมูลที่ผ่านการประมวลผลในระบบทั้งหมด จะถูกจัดเก็บลงบนฐานข้อมูล ในรูปแบบฐานข้อมูลเชิงสัมพันธ์ (Relational database) โดยมีปริมาณข้อมูลจำนวนมาก ซึ่งมีข้อมูลตั้งแต่ปีภาษี 2535 - ปัจจุบัน (ปีภาษี 2541) เมื่อผู้บริหารต้องการทราบข้อมูลสรุปในรูปแบบที่เป็นรายงาน สถิติต่างๆ เพื่อนำข้อมูลไปใช้ช่วยในการตัดสินใจในการบริหารจัดเก็บภาษีอากร โดยเฉพาะข้อมูลที่เกี่ยวข้องกับการขอคืนเงินภาษีมูลค่าเพิ่ม จะทำให้ต้องใช้เวลาในการสืบค้นข้อมูลที่นานมาก และในบางครั้งก็ไม่ทันต่อความต้องการที่จะใช้ข้อมูลของผู้บริหาร

จากปัญหาดังกล่าวจึงได้เกิดแนวความคิดที่จะจัดทำระบบคลังข้อมูล (Data Warehouse) สำหรับงานภาษีมูลค่าเพิ่ม เพื่อสร้างฐานข้อมูล สำหรับใช้ในการสืบค้นข้อมูลได้อย่างรวดเร็ว ช่วยในการตอบคำถาม และช่วยในการตัดสินใจของผู้บริหารได้ ซึ่งจะเป็นการสร้างรูปแบบเบื้องต้น ได้แก่ การขอคืนเงินภาษีมูลค่าเพิ่ม และการยื่นแบบแสดงรายการภาษีมูลค่าเพิ่มประเภทต่างๆ เพื่อเป็นแนวทางเริ่มต้นสำหรับการพัฒนาระบบคลังข้อมูลของกรมสรรพากรต่อไป

6.1 ขั้นตอนการดำเนินงาน

1. ศึกษารายละเอียด และโครงสร้างฐานข้อมูลของระบบงานภาษีมูลค่าเพิ่ม
2. การวิเคราะห์ และออกแบบฐานข้อมูล สำหรับระบบคลังข้อมูล
3. จัดสร้างระบบคลังข้อมูลของภาษีมูลค่าเพิ่ม
4. การออกแบบ Model
5. จัดสร้าง Model
6. จัดทำการวิเคราะห์ข้อมูล

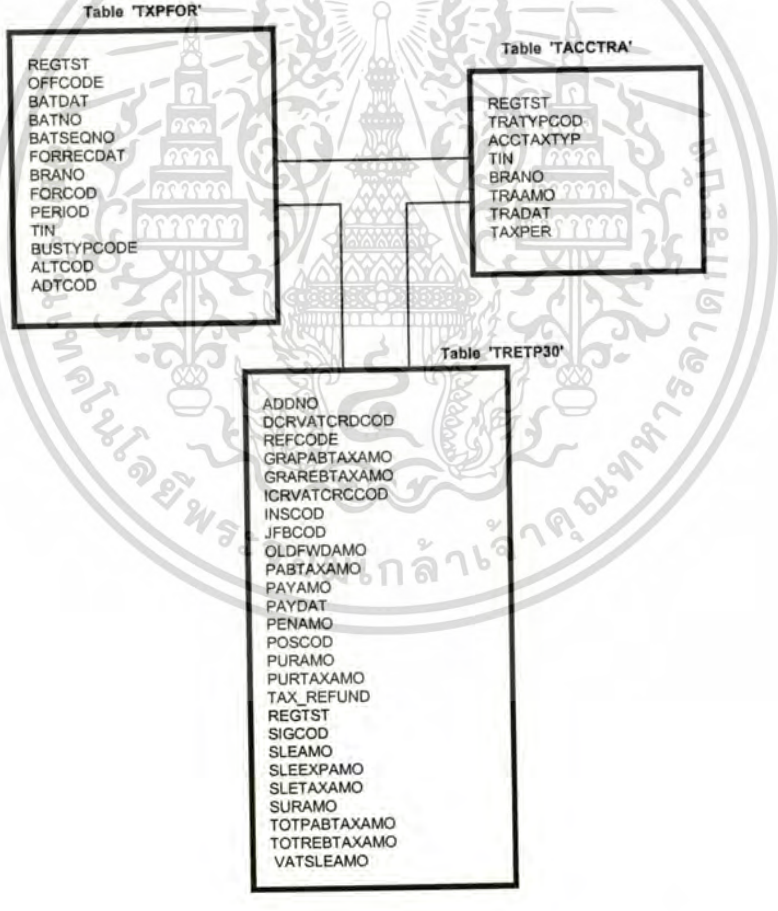
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. ศึกษารายละเอียด และโครงสร้างฐานข้อมูลของระบบงานภาษีมูลค่าเพิ่ม

1.1 โครงสร้างฐานข้อมูลของระบบภาษีมูลค่าเพิ่ม

ข้อมูลภาษีมูลค่าเพิ่มที่ผ่านการประมวลผลแล้ว จะถูกจัดเก็บกระจายไปตามตาราง (table) ต่างๆ จำนวนมาก ขึ้นอยู่กับลักษณะของข้อมูล และความสัมพันธ์ของข้อมูลนั้น ซึ่งข้อมูลที่ผู้บริหารสนใจ ได้ถูกนำมาจัดทำรายงานหรือจัดทำสถิติต่างๆ ส่วนใหญ่จะถูกจัดเก็บอยู่บนตารางที่สำคัญ 3 ตารางดังนี้

1. ตาราง 'TXPFOR'
2. ตาราง 'TRETP30'
3. ตาราง 'TACCTRA'



ภาพที่ 6.1 แสดงความสัมพันธ์ของตารางข้อมูลภาษีมูลค่าเพิ่ม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. ตาราง 'TXPFOR' - ตารางนี้จะจัดเก็บข้อมูลของรายการการยื่นแบบแสดงรายการภาษีมูลค่าเพิ่มของผู้เสียภาษี ทุกประเภทแบบที่ได้มีการยื่นเข้ามาในระบบ ซึ่งประกอบด้วยโครงสร้างของข้อมูล ดังนี้

ลำดับ	ชื่อ Attribute	คำอธิบาย	ชนิดของข้อมูล	ขนาด (bytes)
1.	REGTST	เวลาประทับการประมวลผลข้อมูล	timestamp	
2.	OFFCODE	รหัสที่ตั้งของสำนักงานที่นำส่งข้อมูล	decimal	7,0
3.	BATDAT	วันที่นำส่ง	date	
4.	BATNO	ลำดับที่ของใบนำส่ง	decimal	4,0
5.	BATSEQNO	เลขลำดับที่ของแบบในใบนำส่ง	decimal	3,0
6.	FORRECDAT	วันที่รับแบบ	date	
7.	BRANO	เลขที่สาขา	decimal	4,0
8.	FORMCODE	รหัสแบบแสดงรายการ	character	6
9.	PERIOD	ระยะเวลาภาษี	date	
10.	TIN	เลขประจำตัวผู้เสียภาษี	decimal	10,0
11.	BUSTYPCODE	ประเภทกิจการ	character	1
12.	ALTCOD	รหัสแสดงสถานะของแบบที่ประมวลผล	decimal	1,0
13.	ADTCOD	รหัสแสดงสถานะการยื่นเพิ่มเติม	decimal	1,0

ตารางที่ 6.1 โครงสร้างข้อมูลของรายการในการยื่นแบบแสดงรายการภาษีมูลค่าเพิ่ม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. ตาราง 'TRETP30' - ตารางนี้จะจัดเก็บข้อมูลรายละเอียดของการยื่นแบบแสดงรายการภาษีมูลค่าเพิ่ม (ภพ.30) ของผู้เสียภาษี ซึ่งประกอบด้วยโครงสร้างของข้อมูล ดังนี้

ลำดับ	ชื่อ Attribute	คำอธิบาย	ชนิดของข้อมูล	ขนาด (bytes)
1.	ADDNO	ที่อยู่เลขที่	character	10
2.	DCRVATCRDCOD	รหัสการแก้ไขจำนวนเงินภาษีมูลค่าเพิ่มที่ลดลง	decimal	1,0
3.	REFCODE	รหัสขอคืนภาษี / ขอยกยอดไป	decimal	1,0
4.	GRAPABTAXAMO	จำนวนเงินรวมภาษี	decimal	15,2
5.	GRAREBTAXAMO	จำนวนเงินภาษีที่คืนทั้งสิ้น	decimal	15,2
6.	ICRVATCRCCOD	รหัสยอดขายที่ยื่นเพิ่มเติมในข้อ 1	decimal	1,0
7.	INSCOD	รหัสการผ่อนชำระภาษีเป็นงวด	decimal	1,0
8.	JFBCOD	รหัสการยื่นแบบรวมกันกับสาขา	decimal	1,0
9.	OLDFWDAMO	จำนวนยอดภาษีที่ยกมาจากเดือนก่อน	decimal	15,2
10.	PABTAXAMO	ภาษีที่ต้องชำระ	decimal	15,2
11.	PAYAMO	จำนวนเงินที่ชำระ	decimal	15,2
12.	PAYDAT	วันที่ชำระ	date	
13.	PENAMO	จำนวนเงินเบี้ยปรับ	decimal	15,2
14.	POSCOD	รหัสไปรษณีย์	decimal	5,0

ตารางที่ 6.2 โครงสร้างข้อมูลของการยื่นแบบ ภพ.30

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตาราง 'TRETP30' (ต่อ)

ลำดับ	ชื่อ Attribute	คำอธิบาย	ชนิดของข้อมูล	ขนาด (bytes)
15.	PURAMO	ยอดซื้อ	decimal	15,2
16.	PURTAXAMO	ยอดภาษีซื้อของเดือน	decimal	15,2
17.	TAX_REFUND	ภาษีที่ชำระเกินเดือนนี้	decimal	15,2
18.	REGTST	เวลาประทับการประมวลผลข้อมูล	timestamp	
19.	SIGCOD	รหัสการลงชื่อของผู้เสียภาษี	decimal	1,0
20.	SLEAMO	ยอดขายในเดือนนี้	decimal	15,2
21.	SLEEXEAMO	ยอดขายที่ได้รับยกเว้น	decimal	15,2
22.	SLEEXPAMO	ยอดขายที่เสียภาษีในอัตราร้อยละ 0	decimal	15,2
23.	SLETAXAMO	ภาษีขายเดือนนี้	decimal	15,2
24.	SURAMO	จำนวนเงินเพิ่ม	decimal	15,2
25.	TOTPABTAXAMO	ยอดรวมภาษีที่ต้องชำระ	decimal	15,2
26.	TOTREBTAXAMO	ยอดรวมภาษีที่ชำระไว้เกิน	decimal	15,2
27.	VATSLEAMO	ยอดขายที่ต้องเสียภาษี	decimal	15,2

ตารางที่ 6.2 โครงสร้างข้อมูลของการยื่นแบบ ภพ.30 (ต่อ)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. ตาราง 'TACCTRA' - ตารางนี้จะจัดเก็บข้อมูลของรายการบัญชีการเงินของผู้เสียภาษี ซึ่งประกอบด้วยโครงสร้างของข้อมูล ดังนี้

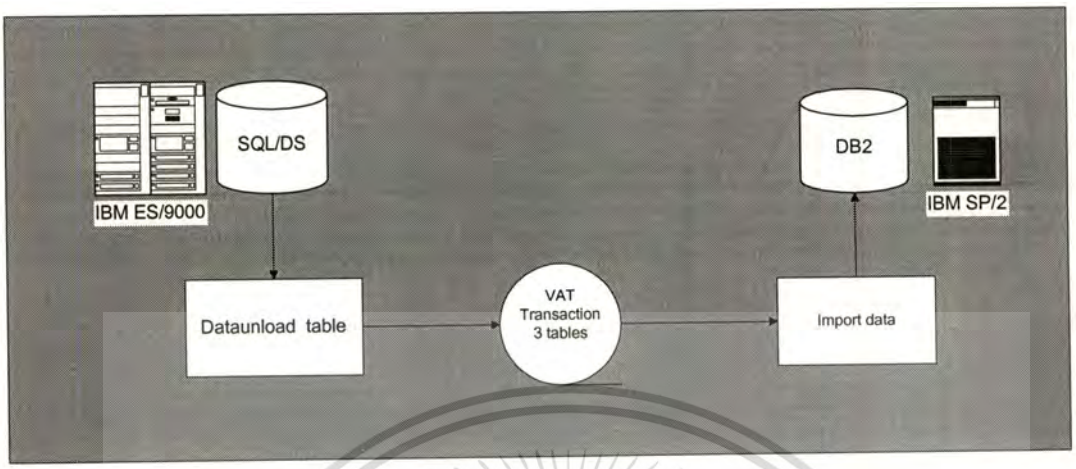
ลำดับ	ชื่อ Attribute	คำอธิบาย	ชนิดของข้อมูล	ขนาด (bytes)
1.	REGTST	เวลาประทับการประมวลผลข้อมูล	timestamp	
2.	TRATYPCOD	รหัสประเภทรายการ	decimal	5,0
3.	ACCTAXTYP	ชนิดของภาษีมูลค่าเพิ่ม	decimal	1,0
4.	TIN	เลขประจำตัวผู้เสียภาษี	decimal	10,0
5.	BRANO	เลขที่สาขา	decimal	4,0
6.	TRAAMO	จำนวนเงินจากรายการ	decimal	15,2
7.	TRADAT	วันที่เกิดรายการ	date	
8.	TAXPER	รอบระยะเวลาภาษี	decimal	6,0

ตารางที่ 6.3 โครงสร้างข้อมูลของรายการบัญชีของผู้เสียภาษี

1.2 การจัดทำฐานข้อมูลภาษีมูลค่าเพิ่ม

เนื่องจากข้อมูลภาษีมูลค่าเพิ่มที่มีการประมวลผลในแต่ละวัน จะถูกจัดเก็บลงบนฐานข้อมูลของเครื่องคอมพิวเตอร์ IBM ES/9000 โดยใช้ DBMS ที่เป็น SQL/DS ซึ่งเมื่อมีปริมาณข้อมูลเพิ่มมากขึ้น ทำให้การจัดการของระบบฐานข้อมูลไม่มีประสิทธิภาพดีเท่าที่ควร รวมทั้งในการสืบค้นข้อมูลก็ต้องใช้เวลานาน และในบางครั้งก็ไม่สามารถกระทำได้ทันต่อความต้องการที่จะใช้ข้อมูลของผู้บริหาร จึงได้มีการนำข้อมูลที่สำคัญจากฐานข้อมูลภาษีมูลค่าเพิ่มมาทำการจัดทำฐานข้อมูลภาษีมูลค่าเพิ่มบนเครื่องคอมพิวเตอร์ IBM SP/2 ซึ่งเป็นเครื่องคอมพิวเตอร์ที่มีขนาดความจุมาก (150 Gb.) โดยใช้ operating system เป็น UNIX ที่เรียก AIX มีระบบจัดการฐานข้อมูลคือ DB2 ซึ่งรองรับปริมาณข้อมูลจำนวนมากได้เป็นอย่างดี โดยทำการปรับปรุงฐานข้อมูลภาษีมูลค่าเพิ่มบนเครื่อง IBM SP/2 เดือนละ 1 ครั้ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 6.2 ขั้นตอนการนำข้อมูลภาษีมูลค่าเพิ่มจากฐานข้อมูลบนเครื่อง IBM ES/9000 มาจัดทำฐานข้อมูลบนเครื่อง IBM SP/2

1.3 การวิเคราะห์แนวโน้มของข้อมูลภาษีมูลค่าเพิ่ม

ในการจัดเก็บข้อมูลภาษีมูลค่าเพิ่มนั้น นับว่าเป็นประเภทภาษีที่มีปริมาณข้อมูลมากที่สุด ซึ่งการนำระบบคลังข้อมูล (Data Warehouse) มาใช้ จึงนับว่าเป็นทางเลือกที่เหมาะสม เพราะระบบคลังข้อมูล จะสนับสนุนการใช้ข้อมูลปริมาณมากได้เป็นอย่างดี และจากตารางที่ 6.4 ได้แสดงให้เห็นถึงการวิเคราะห์แนวโน้มของข้อมูลภาษีมูลค่าเพิ่มในอนาคต ที่จะมีปริมาณเพิ่มมากขึ้นทุกปี

ตาราง (Table)	จำนวนข้อมูล (records)	คิดเป็นเนื้อที่ในการจัดเก็บข้อมูล (Gb.)	อัตราการเพิ่มขึ้นของข้อมูลโดยเฉลี่ย (records/เดือน)	คิดเป็นอัตราการเพิ่มขึ้นของข้อมูลในแต่ละปี (records)
TXPFOR	24,206,887	2.58	431,600	5,179,200
TRETP30	17,396,797	2.89	327,250	3,927,000
TACCTRA	59,720,882	4.96	1,065,270	12,783,240

ตารางที่ 6.4 การวิเคราะห์แนวโน้มของข้อมูลภาษีมูลค่าเพิ่ม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. การวิเคราะห์ และออกแบบฐานข้อมูล สำหรับคลังข้อมูล

2.1 การวิเคราะห์คลังข้อมูล

ในการศึกษาการออกแบบคลังข้อมูล (Data Warehouse) สำหรับผู้บริหาร ได้คำนึงถึงสารสนเทศที่ผู้บริหารระดับสูงสนใจ และได้นำมาใช้ช่วยในการตัดสินใจในการบริหารงานจัดเก็บภาษีอากรของกรมสรรพากร ซึ่งโดยส่วนใหญ่จะเป็นข้อมูลสรุปผลการจัดเก็บภาษีอากรประภทต่างๆ สำหรับงานภาษีมูลค่าเพิ่มนั้น ข้อมูลที่ผู้บริหารสนใจ ได้แก่

- ข้อมูลของการขอคืนเงินภาษีมูลค่าเพิ่ม ทั้งที่เป็นการขอคืนเป็นเงินสดและการขอเครดิตภาษี
- ข้อมูลจำนวนผู้ที่ยื่นแบบแสดงรายการภาษีมูลค่าเพิ่ม โดยเฉพาะแบบ ภพ.01 ภพ.09 ภพ.30 และ ภพ.31

2.2 การออกแบบฐานข้อมูล

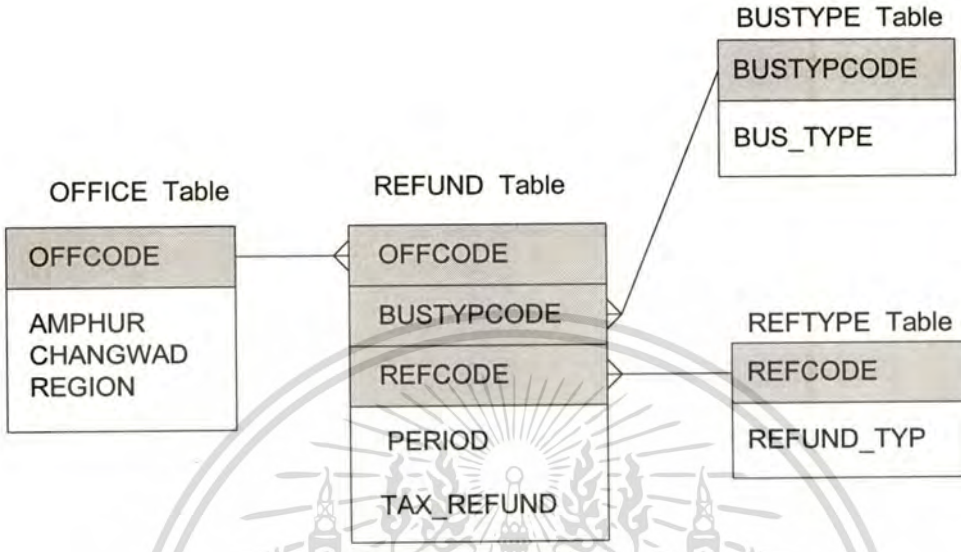
การออกแบบฐานข้อมูลบนคลังข้อมูล (Data Warehouse) ของระบบงานภาษีมูลค่าเพิ่ม ได้ใช้เทคนิคการออกแบบแบบสตาร์ (Star schema) ซึ่งเป็นเทคนิคการออกแบบฐานข้อมูลเชิงวิเคราะห์ ที่แสดงความสัมพันธ์ระหว่างตาราง (table) และการรวมกันของข้อมูล เพื่อให้ได้ผลของการสืบค้นอย่างรวดเร็ว ทำให้ผู้ใช้และนักวิเคราะห์ระบบเข้าใจการไหลเวียนของข้อมูลได้ง่าย และยังสะดวกต่อการเปลี่ยนแปลงแก้ไข หรือเพิ่มเติมตลอดช่วงเวลาของการพัฒนาระบบคลังข้อมูล

ในการออกแบบฐานข้อมูลสำหรับคลังข้อมูลของระบบงานภาษีมูลค่าเพิ่ม โดยใช้แบบสตาร์นั้น ประกอบด้วยตาราง 2 ประเภท คือ

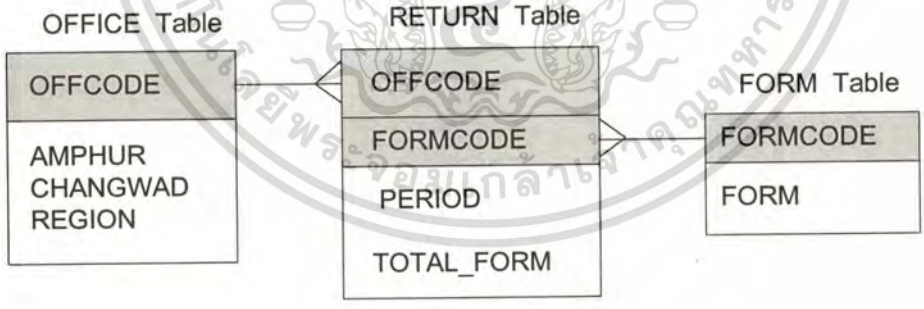
1. ตารางข้อเท็จจริง (Fact table) เป็นตารางที่ใช้ในการจัดเก็บข้อมูลที่เป็นข้อเท็จจริงที่เกิดขึ้นจากการประมวลผลภาษีมูลค่าเพิ่ม เช่น ข้อมูลจำนวนเงินที่มีการขอคืนเงินภาษีมูลค่าเพิ่ม จำนวนแบบแสดงรายการภาษีมูลค่าเพิ่มประเภทต่างๆ ที่ผู้เสียภาษีได้มีการยื่นแบบเข้ามาในระบบ

2. ตารางคุณลักษณะ (Dimension table) เป็นตารางที่เก็บคุณลักษณะต่างๆ ของข้อมูล หรือเป็นคำอธิบายข้อมูลที่สะท้อนถึงมิติขององค์กร เช่น ข้อมูลชื่อหน่วยงานของกรมสรรพากร ชื่อประเภทของแบบแสดงรายการภาษีมูลค่าเพิ่ม ชื่อประเภทของการขอคืนเงินภาษี ชื่อประเภทของการประกอบกิจการ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 6.3 การออกแบบคลังข้อมูลการขอคืนเงินภาษีมูลค่าเพิ่ม



ภาพที่ 6.4 การออกแบบคลังข้อมูลการยื่นแบบแสดงรายการภาษีมูลค่าเพิ่ม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3 โครงสร้างข้อมูลของตารางบนคลังข้อมูล

ตารางข้อมูลบนคลังข้อมูล (Data Warehouse) สำหรับงานภาษีมูลค่าเพิ่ม ตามที่ได้ออกแบบโดยการใช้เทคนิคแบบสตาร์นั้น ประกอบด้วยตาราง 2 ประเภท คือ

1. ตารางข้อเท็จจริง (Fact table) เป็นตารางที่จัดเก็บข้อมูลที่ได้จากการประมวลผลภาษีมูลค่าเพิ่ม ซึ่งมีตารางหลัก 2 ตาราง ได้แก่

● Table 'REFUND' - เป็นตารางที่เก็บข้อมูลสรุปของการขอคืนเงินภาษีมูลค่าเพิ่ม ทั้งการขอคืนเงินที่เป็นเงินสด และการขอเครดิตภาษี ในประเภทกิจการต่างๆ ทั้ง 5 ประเภท แยกตามสำนักงานสรรพากร มีโครงสร้างข้อมูล ดังนี้

ลำดับ	ชื่อ Attribute	คำอธิบาย	ชนิดของข้อมูล	ขนาด (bytes)
1.	OFFCODE	รหัสสำนักงานสรรพากร	character	7
2.	BUSTYPCODE	รหัสแบบแสดงรายการภาษีมูลค่าเพิ่ม	character	1
3.	REFCODE	รหัสการขอคืน	character	1
4.	PERIOD	ระยะเวลาภาษี	date	
5.	TAX_REFUND	จำนวนเงินที่ขอคืนของภาษีมูลค่าเพิ่ม	decimal	17,2

ตารางที่ 6.5 โครงสร้างข้อมูลการขอคืนเงินภาษีมูลค่าเพิ่ม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

● Table 'RETURN' - เป็นตารางที่เก็บข้อมูลสรุปของการยื่นแบบแสดงรายการภาษีมูลค่าเพิ่ม เช่น แบบภพ.01 แบบภ.พ.09 แบบภ.พ.30 และแบบภ.พ.31 ของสำนักงานสรรพากรทั่วประเทศ มีโครงสร้างข้อมูล ดังนี้

ลำดับ	ชื่อ Attribute	คำอธิบาย	ชนิดของข้อมูล	ขนาด (bytes)
1.	OFFCODE	รหัสสำนักงานสรรพากร	character	7
2.	FORMCODE	รหัสแบบแสดงรายการภาษีมูลค่าเพิ่ม	character	2
3.	PERIOD	ระยะเวลาภาษี	date	
4.	TOTAL_FORM	จำนวนแบบทั้งหมด	decimal	7,0

ตารางที่ 6.6 โครงสร้างข้อมูลจำนวนการยื่นแบบแสดงรายการภาษีมูลค่าเพิ่ม

2. ตารางคุณลักษณะ (Dimension table) ประกอบด้วยตารางที่แสดงคุณลักษณะของข้อมูล 4 ตาราง ได้แก่

● Table 'BUSTYPE' - เป็นตารางที่เก็บข้อมูลประเภทของกิจการต่างๆ ของภาษีมูลค่าเพิ่ม มีโครงสร้างข้อมูล ดังนี้

ลำดับ	ชื่อ Attribute	คำอธิบาย	ชนิดของข้อมูล	ขนาด (bytes)
1.	BUSTYPCODE	รหัสประเภทกิจการ	character	1
2.	BUS_TYPE	ประเภทแบบ	character	10

ตารางที่ 6.7 โครงสร้างข้อมูลประเภทกิจการ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

● Table 'OFFICE' - เป็นตารางที่เก็บข้อมูลสำนักงานสรรพากร โดยมีโครงสร้างข้อมูล ดังนี้

ลำดับ	ชื่อ Attribute	คำอธิบาย	ชนิดของข้อมูล	ขนาด (bytes)
1.	OFFCODE	รหัสสำนักงานสรรพากร	character	7
2.	AMPHUR	ชื่ออำเภอ	character	30
3.	CHANGWAD	ชื่อจังหวัด	character	30
4.	REGION	ชื่อภาค	character	7

ตารางที่ 6.8 โครงสร้างข้อมูลของสำนักงานสรรพากร

● Table 'FORM' - เป็นตารางที่เก็บข้อมูลประเภทแบบแสดงรายการภาษีมูลค่าเพิ่ม มีโครงสร้างข้อมูล ดังนี้

ลำดับ	ชื่อ Attribute	คำอธิบาย	ชนิดของข้อมูล	ขนาด (bytes)
1.	FORMCODE	รหัสประเภทแบบแสดงรายการภาษีมูลค่าเพิ่ม	character	2
2.	FORM	ประเภทแบบ	character	5

ตารางที่ 6.9 โครงสร้างข้อมูลของประเภทแบบแสดงรายการภาษีมูลค่าเพิ่ม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

● Table 'REFTYPE' - เป็นตารางที่เก็บข้อมูลประเภทของการขอคืนเงินภาษีมูลค่าเพิ่ม มีโครงสร้างข้อมูล ดังนี้

ลำดับ	ชื่อ Attribute	คำอธิบาย	ชนิดของข้อมูล	ขนาด (bytes)
1.	REFCODE	รหัสประเภทการขอคืนเงินภาษี	character	1
2.	REFUND_TYPE	ประเภทการขอคืนเงินภาษี	character	10

ตารางที่ 6.10 โครงสร้างข้อมูลประเภทของการขอคืนเงินภาษีมูลค่าเพิ่ม

3. จัดสร้างระบบคลังข้อมูลของภาษีมูลค่าเพิ่ม

การจัดสร้างระบบคลังข้อมูล สำหรับระบบงานภาษีมูลค่าเพิ่ม จะประกอบด้วย 3 ขั้นตอน ได้แก่

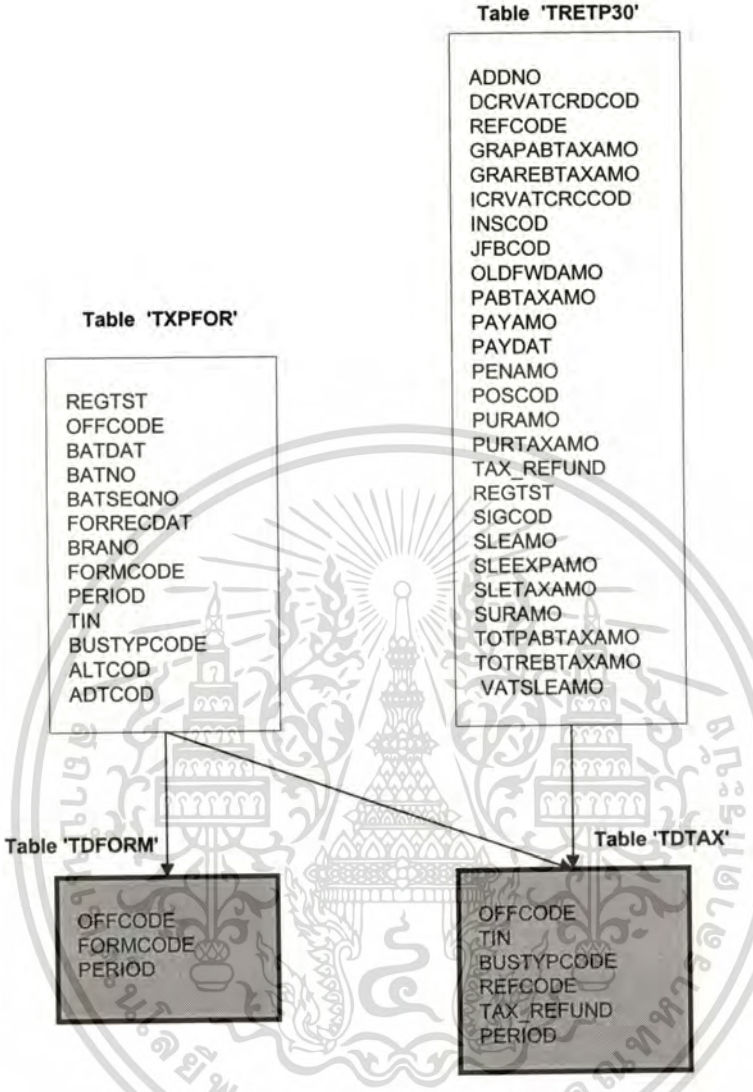
3.1 การรวบรวมข้อมูลจากแหล่งข้อมูลที่เกี่ยวข้อง

เป็นการรวบรวมข้อมูลภาษีมูลค่าเพิ่ม ซึ่งอยู่ในความสนใจของผู้บริหาร ทุก attribute ที่เกี่ยวข้อง และทำการดึงข้อมูลเข้ามาไว้ในคลังข้อมูล โดยการเชื่อมโยงความสัมพันธ์ของข้อมูลระหว่างตาราง 2 ตารางเข้าด้วยกัน (join table) ทำให้ได้ตารางข้อมูลใหม่ 2 ตาราง ดังภาพที่ 6.5 โดยในการดึงข้อมูลได้ทำการคัดเลือกเฉพาะข้อมูลที่จำเป็น ดังนี้

● ข้อมูลการขอคืนเงินภาษีมูลค่าเพิ่ม เป็นการรวบรวมข้อมูลที่ผู้เสียภาษีในแต่ละประเภทกิจการ ได้มีการขอคืนเงินทั้งที่เป็นการขอคืนเงินสด และการเครดิตภาษีตามข้อมูลของแต่ละเดือนภาษี

● ข้อมูลการยื่นแบบแสดงรายการภาษีมูลค่าเพิ่ม โดยเป็นการรวบรวมข้อมูล เฉพาะแบบแสดงรายการภาษีมูลค่าเพิ่มที่อยู่ในความสนใจของผู้บริหาร เพื่อให้ทราบถึงสถานะการยื่นแบบประเภทต่างๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



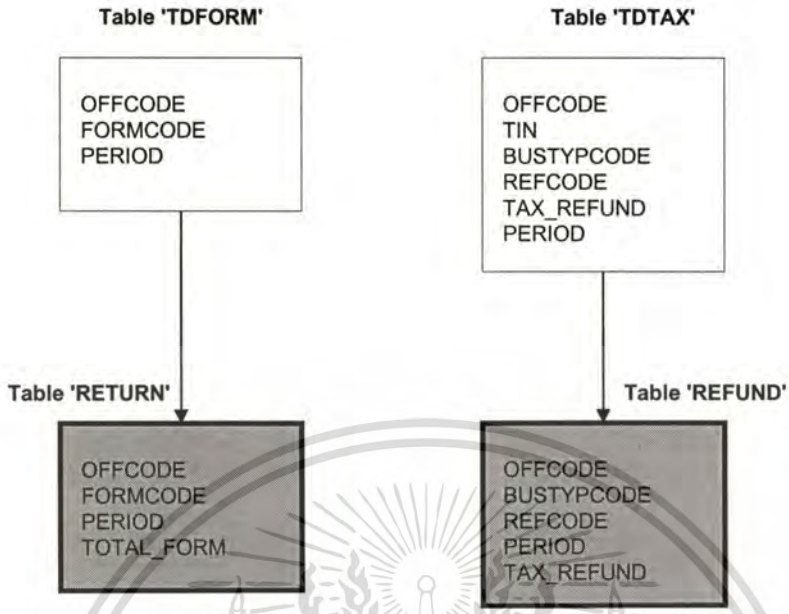
ภาพที่ 6.5 การรวบรวมข้อมูลภาษีมูลค่าเพิ่มจากแหล่งข้อมูล

3.2 การจัดทำข้อมูลสรุป

เป็นการนำข้อมูลที่ได้รวบรวมมาจัดทำเป็นข้อมูลสรุปตามที่ต้องการ โดยข้อมูลที่ได้จะถูกจัดเก็บเป็นตารางผลสรุป เพื่อใช้ในการวิเคราะห์ต่อไป ได้แก่

1. ตาราง 'RETURN'
2. ตาราง 'REFUND'

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 6.6 การจัดทำข้อมูลสรุปของภาษีมูลค่าเพิ่ม

4. การออกแบบ Model

ในการออกแบบ Model สำหรับงานภาษีมูลค่าเพิ่มตามลักษณะของข้อมูลที่อยู่ในความสนใจของผู้บริหาร ต้องมีการสร้าง dimension ตามความต้องการทราบข้อมูลที่เกี่ยวข้อง โดยได้มีการออกแบบ แบ่งเป็น 2 model คือ

4.1 Model การขอคืนเงินภาษีมูลค่าเพิ่ม

4.2 Model การยื่นแบบแสดงรายการภาษีมูลค่าเพิ่ม

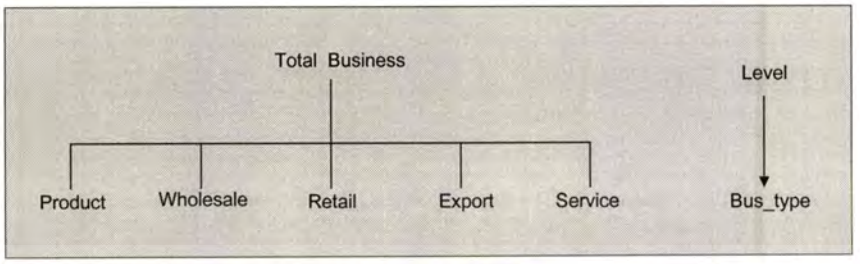
4.1 Model การขอคืนเงินภาษีมูลค่าเพิ่ม

Model สำหรับการขอคืนเงินภาษีมูลค่าเพิ่มนั้น ได้ออกแบบโดยคำนึงถึงส่วนประกอบที่สำคัญ 2 ส่วน คือ Dimension และ Measure

1. Dimenesion ที่ใช้สำหรับ model นี้ จะประกอบด้วย 3 dimension ได้แก่

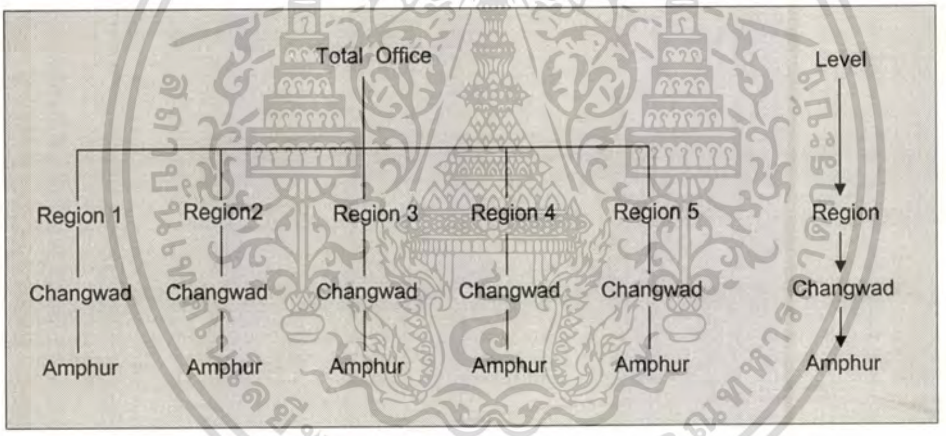
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- BUSINESS เป็นการกำหนด level ของประเภทกิจการ



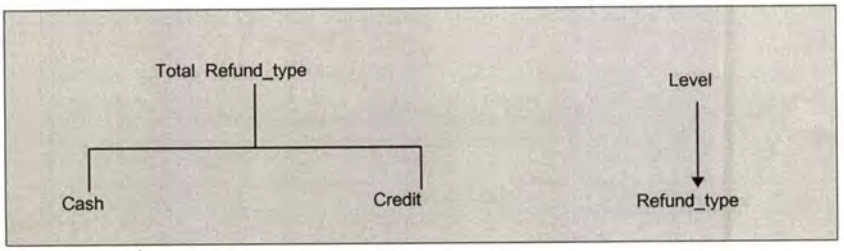
ภาพที่ 6.7 รูปแบบ hierarchy ของ Dimension Business

- OFFICE เป็นการกำหนด level ของสำนักงานสรรพากรทั้งหมด



ภาพที่ 6.8 รูปแบบ hierarchy ของ Dimension Office

- REFUND_TYPE เป็นการกำหนด level ของประเภทของการขอคืนเงินภาษีมูลค่าเพิ่ม



ภาพที่ 6.9 รูปแบบ hierarchy ของ Dimension Refund type

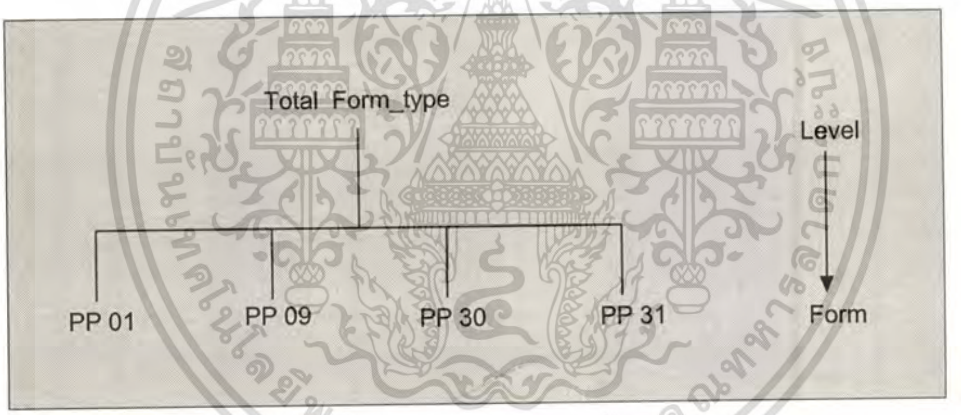
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. Measure สำหรับ model นี้ ข้อมูลที่ต้องการทราบ คือ จำนวนเงินที่ผู้เสียภาษีได้ขอคืนเงินภาษีมูลค่าเพิ่ม และใช้วิธีการคำนวณแบบการหาค่าผลรวม (sum)

4.2 Model การยื่นแบบแสดงรายการภาษีมูลค่าเพิ่ม

1. Dimension สำหรับ model นี้ จะประกอบด้วย 2 dimension ได้แก่

- OFFICE เป็นการกำหนด level ของสำนักงานสรรพากร ซึ่งจะมีรูปแบบ hierarchy ของข้อมูล ดังแสดงในภาพที่ 6.8
- FORM_TYPE เป็นการกำหนด level ของประเภทแบบแสดงรายการภาษีมูลค่าเพิ่ม



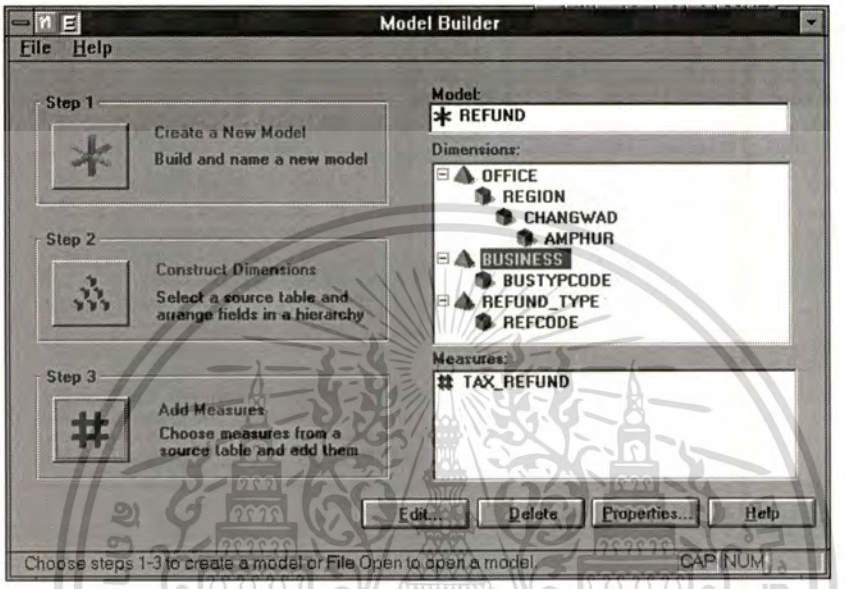
ภาพที่ 6.10 แสดงรูปแบบ hierarchy ของ Dimension Form

2. Measure สำหรับ model นี้ ข้อมูลที่ต้องการทราบ คือ จำนวนแบบแสดงรายการภาษีมูลค่าเพิ่ม (4 ประเภทแบบ) ที่ได้มีการยื่นเข้ามาในระบบ และใช้วิธีการคำนวณแบบการหาค่าผลรวม (sum)

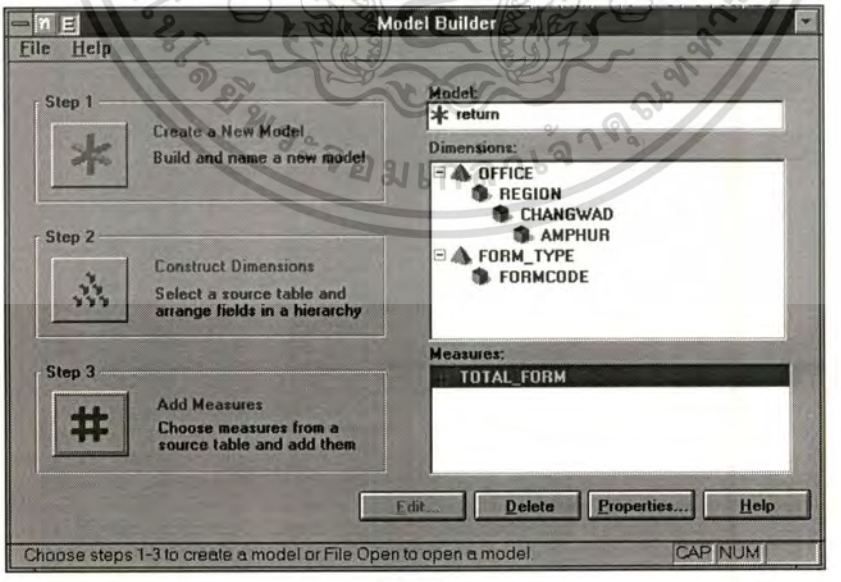
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5. จัดสร้าง Model

ในการจัดสร้าง Model ของงานภาษีมูลค่าเพิ่ม ทั้ง 2 model ได้ใช้ Model Builder ของ Pilot Decision Support มาทำการสร้าง ดังแสดงในภาพที่ 6.11 และภาพที่ 6.12



ภาพที่ 6.11 Model การขอคืนเงินภาษีมูลค่าเพิ่ม



ภาพที่ 6.12 Model การยื่นแบบแสดงรายการภาษีมูลค่าเพิ่ม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6. จัดทำการวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลจะใช้ Pilot Desktop เป็นเครื่องมือในการวิเคราะห์ข้อมูลในรูปแบบต่างๆ โดยทำการวิเคราะห์ข้อมูล ใน 2 รูปแบบ คือ

1. Ranking Analysis
2. Navigator Adhoc

6.2 ระบบคลังข้อมูลภาษีมูลค่าเพิ่ม

การพัฒนาระบบคลังข้อมูลภาษีมูลค่าเพิ่ม ได้มีการนำโปรแกรม Pilot Designer และ Pilot Desktop มาช่วยในการพัฒนาระบบงาน เพื่อให้สามารถวิเคราะห์ข้อมูลตามที่ผู้บริหารต้องการได้

ในการพัฒนาระบบงาน จำเป็นต้องคำนึงถึงการรักษาความปลอดภัยของข้อมูลด้วย จึงต้องมีการ logon เข้าใช้ระบบ เพื่อให้ระบบทำการตรวจสอบสิทธิ์ในการเข้าใช้ข้อมูลก่อน ถ้าเป็นผู้ที่ได้รับอนุญาตให้เข้าใช้ระบบ จึงจะสามารถเข้าใช้ระบบ เพื่อทำการวิเคราะห์ข้อมูลได้

ภาพที่ 6.13 การ Logon เข้าใช้ระบบ

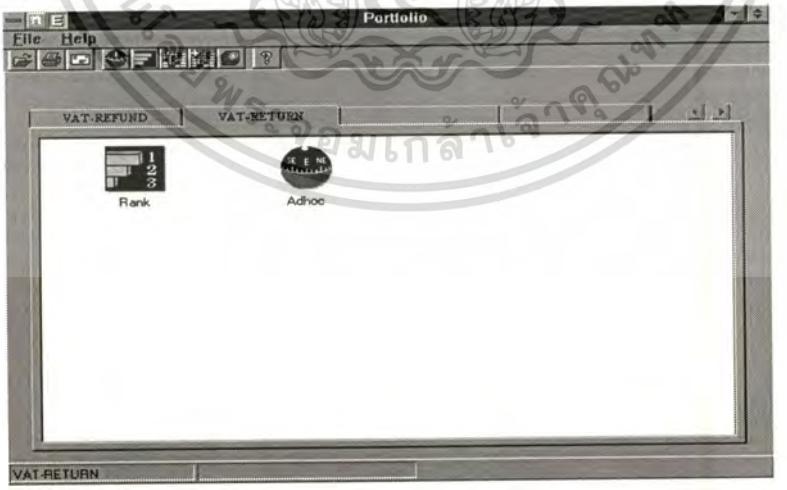
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6.3 รูปแบบการวิเคราะห์ข้อมูล

ในการวิเคราะห์ข้อมูลภาษีมูลค่าเพิ่ม จะสามารถเลือกดูข้อมูลได้ 2 ลักษณะ คือ

1. VAT REFUND เป็นการดูข้อมูลการขอคืนเงินภาษีมูลค่าเพิ่ม
2. VAT RETURN เป็นการดูข้อมูลการยื่นแบบแสดงรายการภาษีมูลค่าเพิ่ม

ภาพที่ 6.14 แสดง Menu VAT REFUND



ภาพที่ 6.15 แสดง Menu VAT RETURN

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

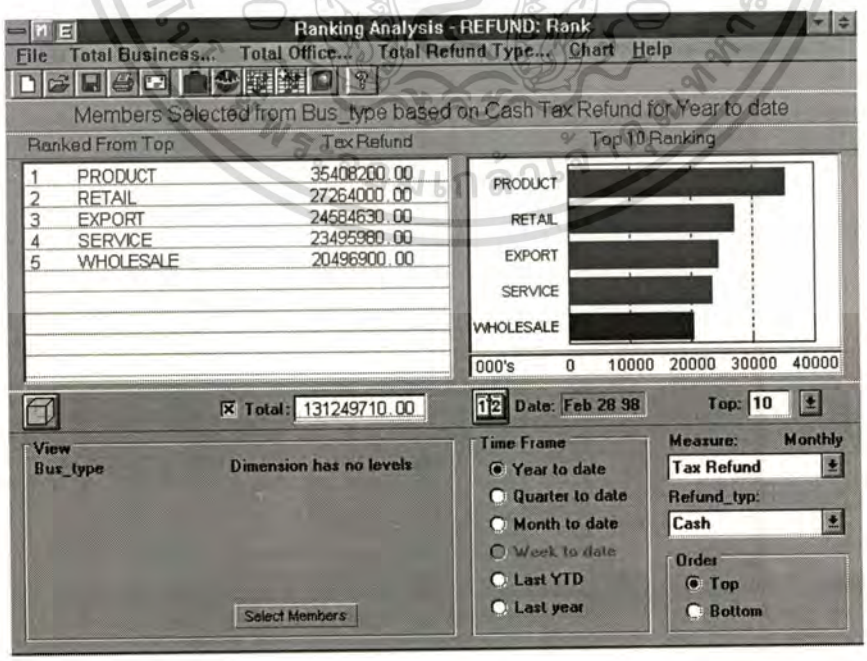
6.3.1 การวิเคราะห์ข้อมูลการขอคืนเงินภาษีมูลค่าเพิ่ม

6.3.1.1 การวิเคราะห์ข้อมูล เพื่อดูข้อมูลในแบบ Ranking Analysis ซึ่งเป็นการเรียกดูข้อมูลในลักษณะของการจัดเรียงลำดับจากมากไปน้อย หรือจากน้อยไปมากได้ และในการแสดงผลข้อมูล สามารถที่จะดูข้อมูลตาม dimendion ที่กำหนดไว้ได้ โดยอาศัยมิติของเวลาที่ต้องการดูข้อมูล เป็นส่วนที่ช่วยจัดการในการเลือกข้อมูลด้วย ซึ่งมีการทำงานได้ ดังนี้

1. การเลือกมุมมองข้อมูล และหลักเกณฑ์การคัดเลือก เป็นการเลือกมุมมองที่ต้องการจะดูข้อมูล และกำหนดหลักเกณฑ์ในการคัดเลือกได้

2. การเลือกวิเคราะห์ข้อมูล ตาม dimension สามารถแสดงผลข้อมูลได้ ดังนี้

- Dimension : BUSINESS จะเป็นการเลือกข้อมูลประเภทของการประกอบกิจการ ซึ่งแบ่งเป็น 5 ประเภท ได้แก่
  - ผลิต (Product)
  - ขายปลีก (Retail)
  - ขายส่ง (Wholesale)
  - ส่งออก (Export)
  - บริการ (Service)

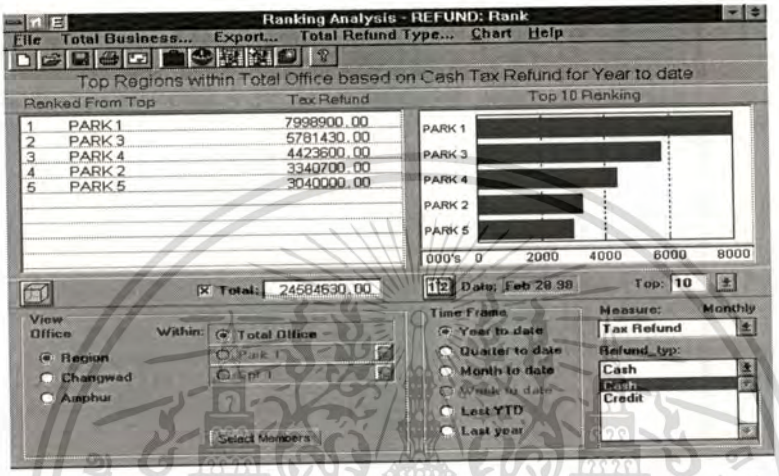


ภาพที่ 6.16 แสดงข้อมูลการขอคืนเงินภาษีมูลค่าเพิ่มของประเภทกิจการ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

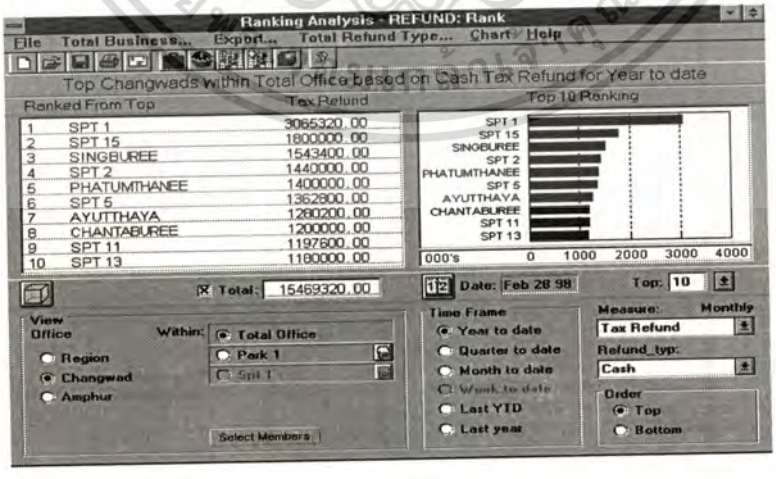
● Dimension : REFUND_TYPE จะเป็นการเลือกข้อมูลประเภทของการขอคืนเงินภาษีมูลค่าเพิ่ม ได้ทั้ง 2 ประเภท คือ

- เงินสด (Cash)
- การขอเครดิตภาษี (Credit)



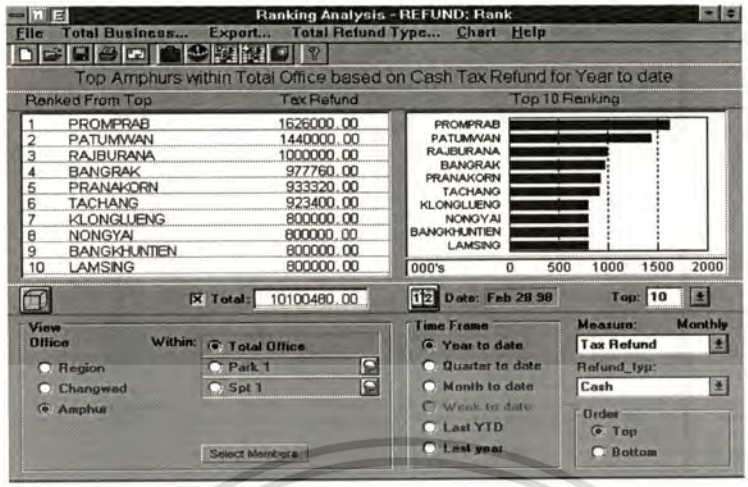
ภาพที่ 6.17 แสดงผลการขอคืนเงินภาษีมูลค่าเพิ่ม ในระดับภาค (region) และการเลือกดูข้อมูลตามประเภทของการขอคืนเงินภาษี

● Dimension : OFFICE เป็นการเลือกดูข้อมูลตามระดับสำนักงาน



ภาพที่ 6.18 แสดง 10 จังหวัดที่มีการขอคืนเงินภาษีมูลค่าเพิ่มสูงสุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

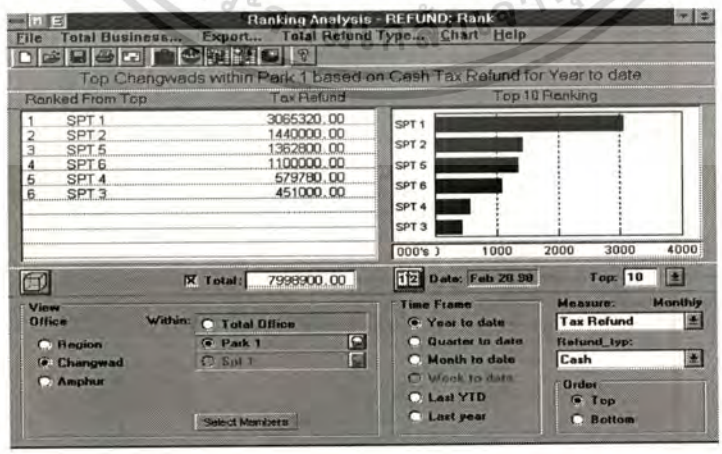


ภาพที่ 6.19 แสดงผลการขอคืนเงินภาษีมูลค่าเพิ่ม ในระดับอำเภอ (Amphur)

3. การเรียกดูข้อมูล ตามลำดับชั้นของ dimension : OFFICE ซึ่งจะทำให้การวิเคราะห์รายละเอียดของแต่ละระดับ กระทำได้ง่าย และรวดเร็ว โดยการเลือกดูข้อมูลเฉพาะกลุ่ม (group) ที่สนใจ ได้แก่

- การ Drill up เป็นการเรียกดูข้อมูลจากกลุ่มใหญ่ไปหากลุ่มย่อย
- การ Drill down เป็นการเรียกดูข้อมูลจากกลุ่มย่อยขึ้นมาหากลุ่มใหญ่

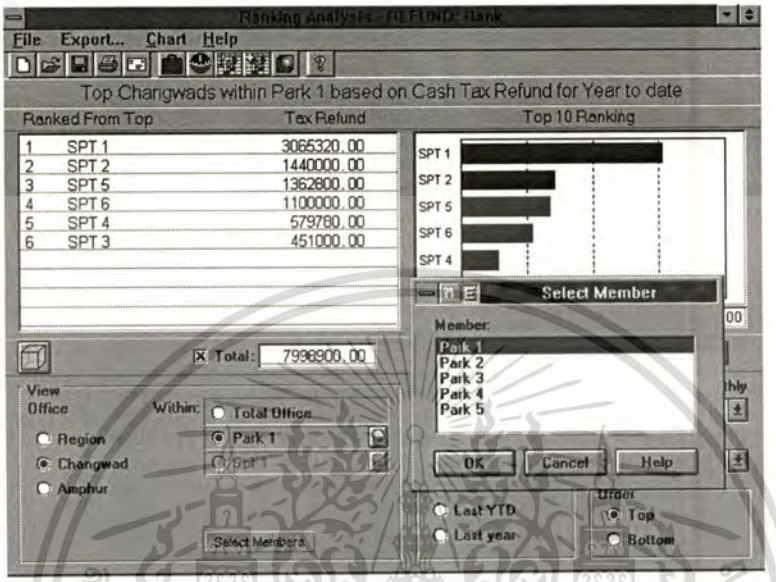
จากภาพที่ 6.20 จะเป็นการ drill down เพื่อดูข้อมูลการขอคืนเงินภาษีมูลค่าเพิ่มในทุกจังหวัดที่อยู่ในภาค 1



ภาพที่ 6.20 การ drill down ข้อมูล

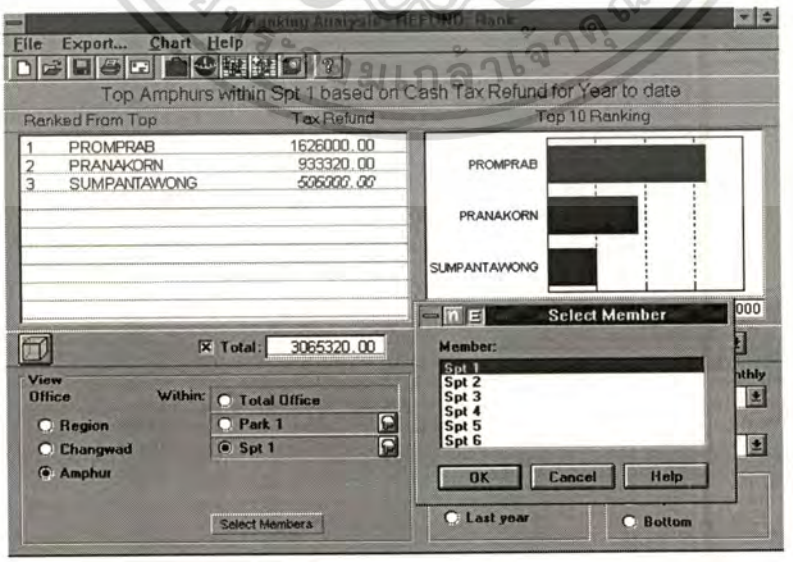
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การเรียกดูข้อมูลของจังหวัดที่อยู่ในแต่ละภาค สามารถเลือกมุมมอง (view) ได้โดยการเลือกข้อมูลภาคที่ต้องการ ดังภาพที่ 6.21



ภาพที่ 6.21 การเลือกข้อมูลภาค เพื่อดูข้อมูลจำแนกรายจังหวัด

ในทำนองเดียวกัน การที่จะเลือกดูข้อมูลทุกอำเภอที่อยู่ภายใต้จังหวัด และภายในภาคที่ต้องการ ก็จะเลือกดูข้อมูลได้ ตามภาพที่ 6.22

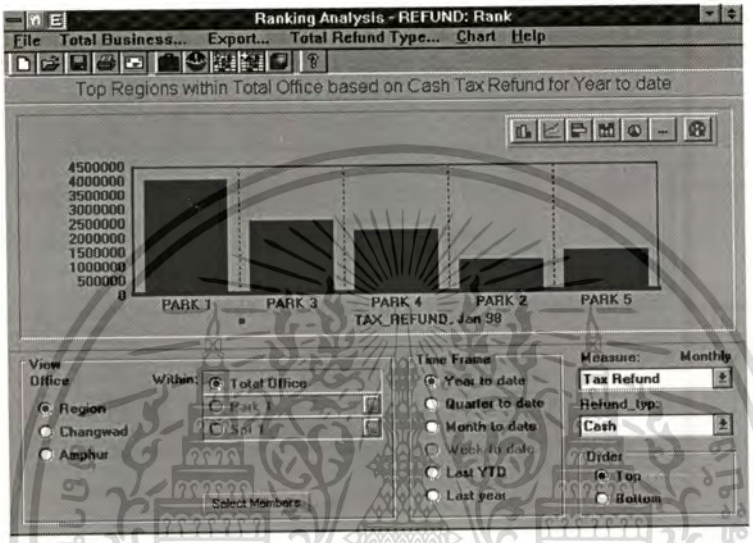


ภาพที่ 6.22 การเลือกข้อมูลจังหวัด เพื่อดูข้อมูลจำแนกรายอำเภอ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

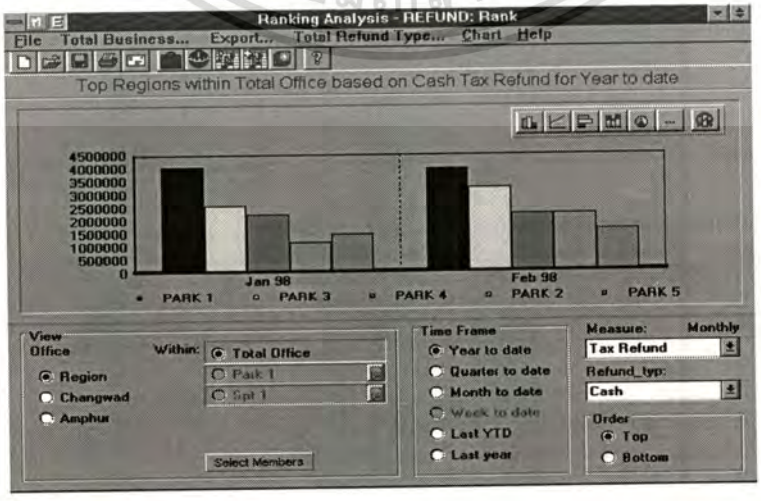
4. การแสดงผลข้อมูลในรูปแบบของกราฟประเภทต่างๆ เพื่อแสดงให้เห็นการเปรียบเทียบในลักษณะที่ต้องการได้ เช่น กราฟแท่ง กราฟวงกลม เป็นต้น

- การแสดงผลของการขอคืนเงินภาษีมูลค่าเพิ่มเป็นเงินสดของภาคต่างๆ ในรูปแบบของกราฟแท่ง ซึ่งจะแสดงให้เห็นการเปรียบเทียบของจำนวนเงินที่มีการขอคืนเงินภาษี



ภาพที่ 6.23 การขอคืนเงินภาษีมูลค่าเพิ่มเป็นเงินสด รายภาค

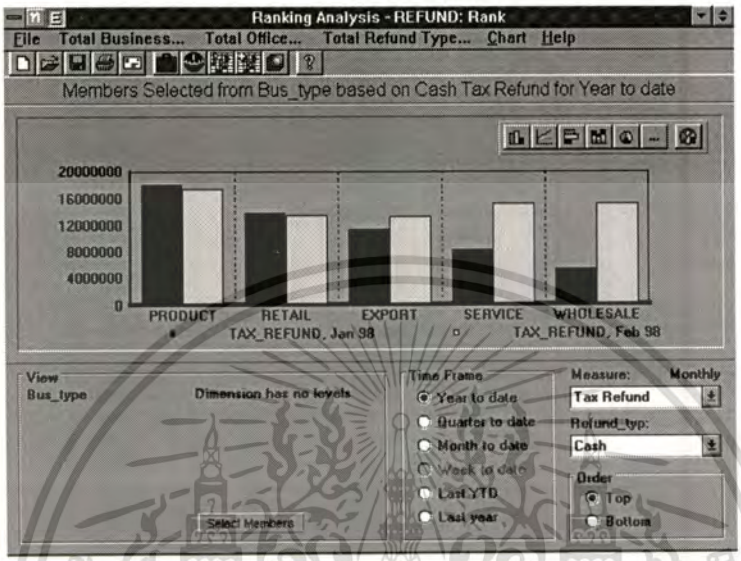
- การแสดงผล เพื่อเปรียบเทียบข้อมูลการขอคืนเงินภาษีมูลค่าเพิ่มของภาคต่างๆ ในช่วงเวลาที่ต่างกัน หรือเลือกกำหนดระยะเวลาตามที่ต้องการได้



ภาพที่ 6.24 แสดงการเปรียบเทียบข้อมูลการขอคืนเงินภาษีมูลค่าเพิ่ม ในช่วงเวลา 2 เดือน

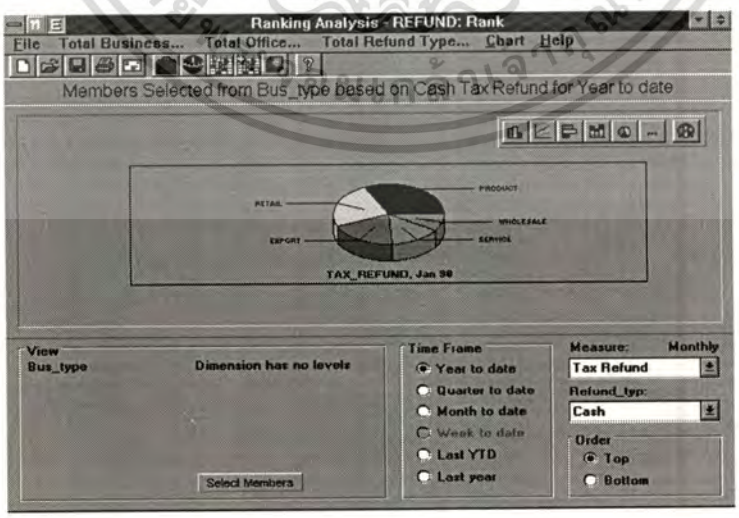
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- การแสดงผลเพื่อเปรียบเทียบการขอคืนเงินภาษีมูลค่าเพิ่มของประเภทกิจการต่างๆ ในช่วงระยะเวลาที่ต่างกัน



ภาพที่ 6.25 แสดงการเปรียบเทียบการขอคืนเงินภาษีมูลค่าเพิ่มของประเภทกิจการ

- การแสดงผลในรูปแบบของกราฟวงกลม ซึ่งเป็นข้อมูลที่จะแสดงให้เห็นสัดส่วนของการขอคืนเงินภาษีมูลค่าเพิ่มของประเภทกิจการต่างๆ ได้ชัดเจน

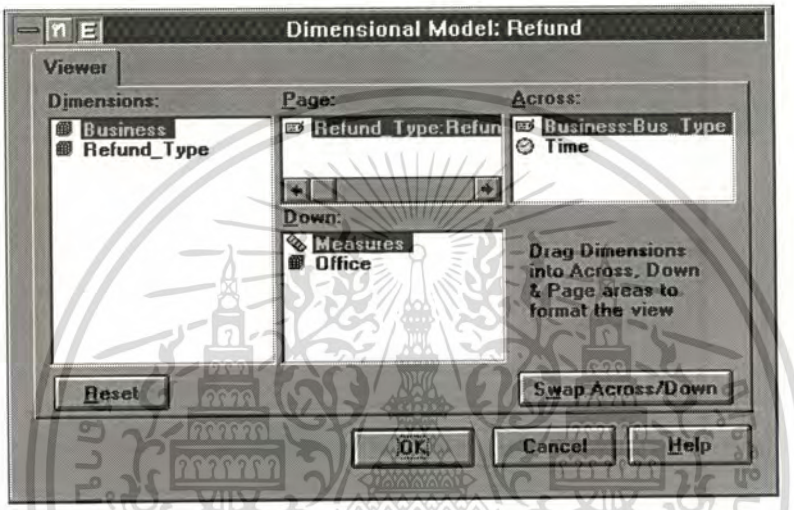


ภาพที่ 6.26 แสดงกราฟวงกลม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

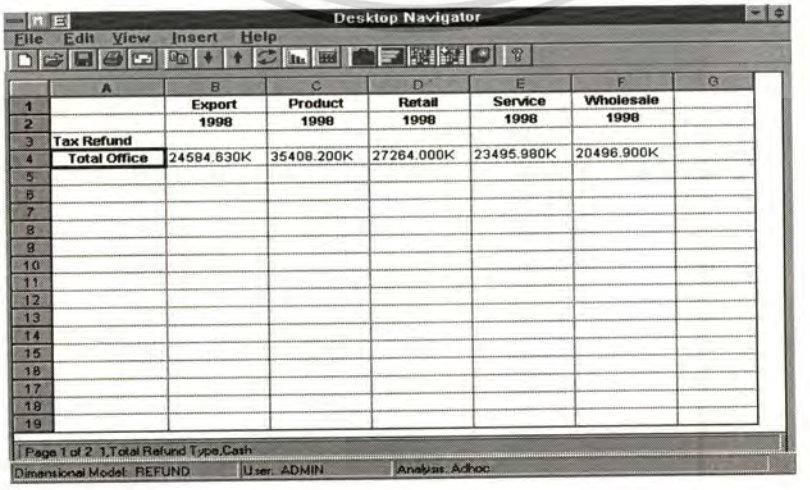
6.3.1.2 การวิเคราะห์ข้อมูล เพื่อการดูข้อมูลในแบบ Navigator Adhoc โดยมีคุณสมบัติในการเรียกดูข้อมูลได้หลายลักษณะ ดังนี้

- การกำหนดมุมมอง ซึ่งจะใช้การหมุน (Rotate) เพื่อสลับมุมมองของการเรียกดูข้อมูลตามที่ต้องการได้ โดยการย้าย dimension ต่างๆ ทำให้เกิดความยืดหยุ่นในการสืบค้นข้อมูลได้มาก ดังภาพที่ 6.27



ภาพที่ 6.27 การกำหนดมุมมองข้อมูล (View)

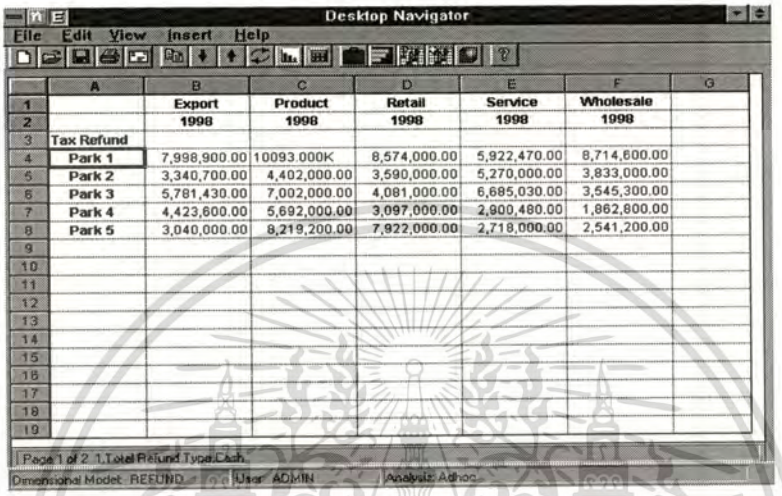
- การ Drill up และการ Drill down ข้อมูล เพื่อดูข้อมูลในระดับกลุ่มใหญ่ กลุ่มย่อยตามลำดับชั้นได้



ภาพที่ 6.28 แสดงข้อมูลในระดับสูงสุด (Total Office)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

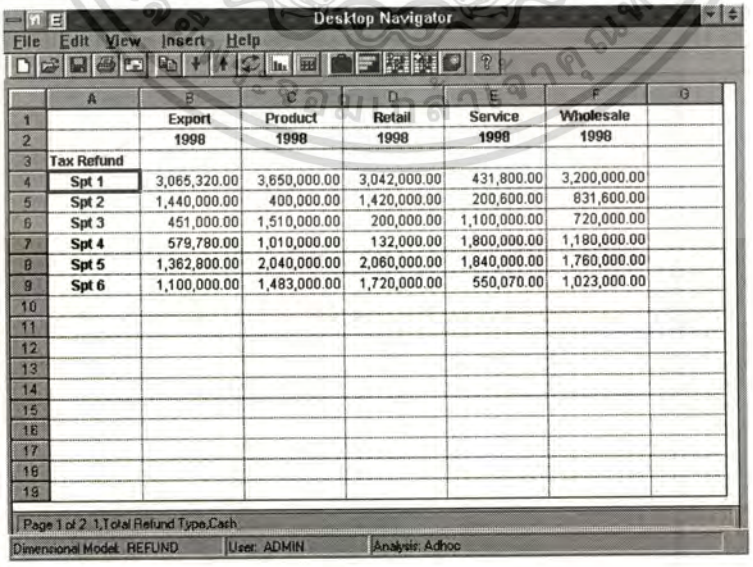
การ drill down ข้อมูล สามารถกระทำได้อย่างง่ายๆ เพื่อ drill down ตาม hierarchy ของข้อมูลได้ โดยการใช้ mouse double click ที่ช่อง Total Office ก็จะปรากฏข้อมูลของทุกภาคขึ้นมา ตามภาพที่ 6.29



	A	B	C	D	E	F	G
1		Export	Product	Retail	Service	Wholesale	
2		1998	1998	1998	1998	1998	
3	Tax Refund						
4	Park 1	7,998,900.00	10093.000K	8,574,000.00	5,922,470.00	8,714,600.00	
5	Park 2	3,340,700.00	4,402,000.00	3,590,000.00	5,270,000.00	3,833,000.00	
6	Park 3	5,781,430.00	7,002,000.00	4,081,000.00	6,685,030.00	3,545,300.00	
7	Park 4	4,423,600.00	5,692,000.00	3,097,000.00	2,900,480.00	1,862,800.00	
8	Park 5	3,040,000.00	8,219,200.00	7,922,000.00	2,718,000.00	2,541,200.00	
9							
10							
11							
12							
13							
14							
15							
16							
17							
18							
19							

ภาพที่ 6.29 แสดงผลจากการ Drill down ข้อมูลในรูปแบบของ Navigator Adhoc

ในการเรียกดูข้อมูลของระดับจังหวัดที่อยู่ภายใต้ภาคที่ต้องการ ก็จะแสดงผลได้ดังภาพที่ 6.30

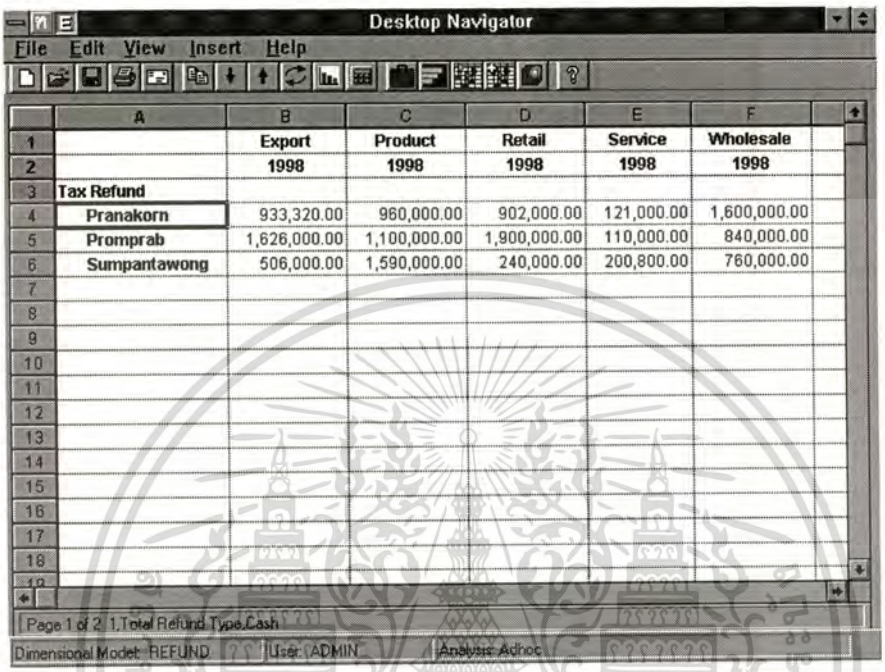


	A	B	C	D	E	F	G
1		Export	Product	Retail	Service	Wholesale	
2		1998	1998	1998	1998	1998	
3	Tax Refund						
4	Spt 1	3,065,320.00	3,650,000.00	3,042,000.00	431,800.00	3,200,000.00	
5	Spt 2	1,440,000.00	400,000.00	1,420,000.00	200,600.00	831,600.00	
6	Spt 3	451,000.00	1,510,000.00	200,000.00	1,100,000.00	720,000.00	
7	Spt 4	579,780.00	1,010,000.00	132,000.00	1,800,000.00	1,180,000.00	
8	Spt 5	1,362,800.00	2,040,000.00	2,060,000.00	1,840,000.00	1,760,000.00	
9	Spt 6	1,100,000.00	1,483,000.00	1,720,000.00	550,070.00	1,023,000.00	
10							
11							
12							
13							
14							
15							
16							
17							
18							
19							

ภาพที่ 6.30 แสดงผลการขอคืนเงินภาษีมูลค่าเพิ่มของแต่ละประเภทกิจการ รายจังหวัด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การเรียกดูข้อมูลกลุ่มย่อยลงไปอีก ซึ่งได้แก่ระดับอำเภอ ใช้วิธีการ drill down เพื่อการเลือกดูข้อมูลของอำเภอที่อยู่ภายใน สพท.1 (ระดับจังหวัด) ดังภาพที่ 6.31



	A	B	C	D	E	F
1		Export	Product	Retail	Service	Wholesale
2		1998	1998	1998	1998	1998
3	Tax Refund					
4	Pranakorn	933,320.00	960,000.00	902,000.00	121,000.00	1,600,000.00
5	Promprab	1,626,000.00	1,100,000.00	1,900,000.00	110,000.00	840,000.00
6	Sumpantawong	506,000.00	1,590,000.00	240,000.00	200,800.00	760,000.00

ภาพที่ 6.31 แสดงผลการขอคืนเงินภาษีมูลค่าเพิ่มของแต่ละประเภทกิจการ รายอำเภอ

6.3.2 การวิเคราะห์ข้อมูลการยื่นแบบแสดงรายการภาษีมูลค่าเพิ่ม

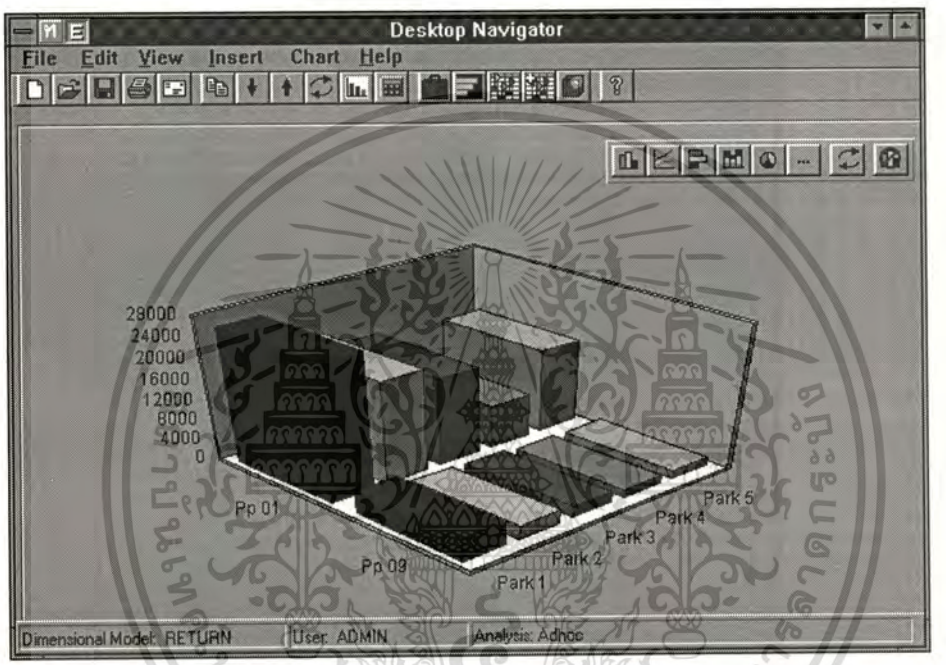
ในการวิเคราะห์ข้อมูล ก็จะแบ่งเป็น 2 ลักษณะ ซึ่งได้แก่ Ranking Analysis และ Navigator Adhoc เหมือนกับการวิเคราะห์ข้อมูลของการขอคืนเงินภาษีมูลค่าเพิ่ม โดยการทำงาน ดังนี้

1. การเรียกดูข้อมูล ตาม dimension ได้แก่

- OFFICE ซึ่งเป็นการเรียกดูข้อมูลของสำนักงานสรรพากร
- FORM_TYPE เป็นการเรียกดูข้อมูลของประเภทของแบบแสดงรายการภาษีมูลค่าเพิ่ม ซึ่งได้แก่
  - แบบ ภพ. 01
  - แบบ ภพ. 09
  - แบบ ภพ. 30
  - แบบ ภพ. 31

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. การแสดงผลในรูปแบบของกราฟประเภทต่างๆ เช่น การแสดงผลข้อมูลในรูปแบบของกราฟแท่ง 3 มิติ เพื่อการเปรียบเทียบให้เห็นถึงผลสรุปของการยื่นแบบแสดงรายการภาษีมูลค่าเพิ่ม 2 ประเภท ได้แก่ แบบ ภพ.01 และแบบ ภพ.09 ของภาคต่างๆ ในปีภาษี 2541 ดังภาพที่ 6.32



ภาพที่ 6.32 เปรียบเทียบการยื่นแบบ ภพ.01 และ ภพ.09 รายภาค

3. การเลือกมุมมองของข้อมูล ในการที่จะเลือกดูมุมมองต่างๆของข้อมูลนั้น จะสามารถเลือกดูได้ตามมิติของเวลาที่ต้องการ เช่น รายปี รายไตรมาส และรายเดือนภาษี โดยมีมุมมอง ดังนี้

- การเลือกดูข้อมูล แบบ ภพ.01 ของสำนักงานสรรพากร ได้ทั้ง 3 ระดับ
- การเลือกดูข้อมูล แบบ ภพ.09 ของสำนักงานสรรพากร ได้ทั้ง 3 ระดับ
- การเลือกดูข้อมูล แบบ ภพ.30 ของสำนักงานสรรพากร ได้ทั้ง 3 ระดับ
- การเลือกดูข้อมูล แบบ ภพ.31 ของสำนักงานสรรพากร ได้ทั้ง 3 ระดับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 7
บทสรุป

7.1 สรุปผลการศึกษา

จากการที่ได้จัดสร้างระบบคลังข้อมูลของงานภาษีมูลค่าเพิ่ม โดยได้ทำการวิเคราะห์ความต้องการใช้ข้อมูลภาษีมูลค่าเพิ่มที่ผู้บริหารสนใจ แล้วนำมาทำการออกแบบ จัดสร้างฐานข้อมูลบนระบบคลังข้อมูล ซึ่งได้เลือกการจัดทำข้อมูลสรุปทางด้านการขอคืนเงินภาษีมูลค่าเพิ่ม และข้อมูลสรุปผลการยื่นแบบแสดงรายการภาษีมูลค่าเพิ่ม และได้มีการนำโปรแกรม Pilot Decision Support Suite ในส่วนที่เป็น Pilot Desktop มาใช้ช่วยในการวิเคราะห์ข้อมูล เพื่อให้ได้การแสดงผลข้อมูลภาษีมูลค่าเพิ่มในรูปแบบต่างๆ ที่ผู้บริหารสามารถจะนำข้อมูลดังกล่าวไปใช้ เพื่อช่วยในการตัดสินใจ และวางแผนการบริหารจัดเก็บภาษีอากรนั้น

พบว่า การนำระบบคลังข้อมูลมาใช้งาน มีความเหมาะสมกับงานภาษีมูลค่าเพิ่มเป็นอย่างมาก เพราะปริมาณข้อมูลของภาษีมูลค่าเพิ่มนับวันจะเพิ่มมากขึ้นเรื่อยๆ และยังเป็นข้อมูลที่ผู้บริหารให้ความสนใจ และต้องการทราบความเปลี่ยนแปลงของข้อมูลอยู่ตลอดเวลา ประกอบกับข้อมูลที่ถูกจัดเก็บอยู่บนฐานข้อมูลปัจจุบัน จะมีการกระจายข้อมูลไปในตารางที่มีความสัมพันธ์กันหลายตาราง จึงไม่สนับสนุนการจัดทำข้อมูลที่เป็นลักษณะของการวิเคราะห์ เมื่อต้องการใช้ข้อมูลทำให้ต้องทำการเชื่อมโยงข้อมูลระหว่างตาราง ซึ่งจะใช้เวลาในการจัดทำข้อมูลนานมาก และจากการที่ได้ออกแบบระบบคลังข้อมูล แล้วนำโปรแกรม Pilot Desktop มาใช้วิเคราะห์ข้อมูลเบื้องต้นนั้น ช่วยทำให้การสืบค้นข้อมูล เป็นสิ่งที่สะดวกต่อการใช้งาน โดยที่ผู้ใช้ไม่จำเป็นต้องมีความรู้ หรือความชำนาญในการใช้ภาษา SQL (Structured Query Language) และไม่ต้องทำการ query ให้เสียเวลา ทำให้ทราบข้อมูลได้อย่างรวดเร็ว ซึ่งถ้าผู้บริหารมีความต้องการใช้ข้อมูล หรือต้องการให้สร้างระบบงานที่มีการวิเคราะห์ข้อมูลที่มีความซับซ้อน ก็สามารถที่จะพัฒนาระบบงานโดยการใช้ Pilot Designer ได้ ซึ่งการใช้โปรแกรม Pilot Desktop และ Pilot Designer จะมีข้อดี ข้อเสียที่แตกต่างกัน ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โปรแกรม	ข้อดี	ข้อเสีย
Pilot Desktop	1. ผู้ใช้สามารถใช้งานได้ง่าย 2. ไม่ต้องใช้เวลาในการพัฒนาระบบงานนานมาก เพราะผู้ใช้สามารถเลือกใช้ Wizard ที่มีให้ได้	1. ผู้ใช้ต้องมีความเข้าใจในการกำหนด Model และการเลือกใช้ Dimension 2. ไม่สามารถจัดทำรูปแบบที่นอกเหนือจากการที่ Wizard มีให้ได้ 3. การวิเคราะห์ข้อมูลบางอย่าง อาจได้ไม่ตรงกับความต้องการใช้งาน
Pilot Designer	1. สามารถพัฒนารูปแบบที่ต้องการได้ 2. ถ้ามีการจัดทำ Menu ที่ดี จะทำให้ผู้ใช้สามารถใช้งานได้ง่าย จึงไม่ต้องใช้เวลาในการเรียนรู้มาก	1. ผู้ใช้ไม่สามารถพัฒนางานได้เองทันที จำเป็นต้องมีผู้พัฒนาระบบงาน จัดสร้างระบบงานให้ก่อน

ตารางที่ 7.1 เปรียบเทียบข้อดี ข้อเสียของโปรแกรม Pilot Desktop และ Pilot Designer

7.2 ประโยชน์ที่ได้รับ

ระบบคลังข้อมูลภาษีมูลค่าเพิ่ม นับได้ว่าเป็นระบบที่สามารถแก้ไขปัญหาที่เกิดขึ้นในองค์กรได้เป็นอย่างดี โดยการนำโปรแกรม Pilot Desktop มาเป็นเครื่องมือในการวิเคราะห์ข้อมูล จะสามารถตอบคำถาม และแสดงผลข้อมูลเบื้องต้นที่ผู้บริหารต้องการได้ในเวลาที่รวดเร็ว รวมทั้งสามารถเลือกการแสดงผลข้อมูลได้หลายรูปแบบ ไม่ว่าจะเป็นรูปแบบของกราฟแท่ง กราฟวงกลม กราฟเส้น หรือกราฟ 3 มิติ โดยเป็นการเลือกมุมมองข้อมูลตาม dimension ที่ต้องการ ซึ่งถูกกำหนดไว้ได้ และยังมีแนวทางในการจัดการข้อมูลได้อย่างเป็นระบบ ซึ่งนับได้ว่าเป็นพื้นฐานในการที่จะพัฒนาการจัดทำระบบคลังข้อมูลของภาษีประเภทอื่นๆ สำหรับกรมสรรพากรต่อไปในอนาคต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

7.3 ข้อเสนอแนะ

ในการที่จะนำเทคโนโลยีของระบบคลังข้อมูลมาใช้ในองค์กรนั้น จำเป็นจะต้องมีบุคลากร ที่มีความรู้ ความสามารถในการออกแบบฐานข้อมูล และเข้าใจระบบงานเป็นอย่างดี จึงจะสามารถสร้างฐานข้อมูลที่เหมาะสมกับลักษณะงานขององค์กร โดยจะต้องเป็นฐานข้อมูลซึ่งมีโครงสร้างข้อมูลที่ผู้ใช้เข้าใจได้ง่าย ช่วยสนับสนุนการวิเคราะห์ข้อมูลปริมาณมากๆได้ ในเวลาที่รวดเร็ว และยังต้องมีการเลือกใช้เครื่องมือ (Software Tools) ในการวิเคราะห์ข้อมูลที่เหมาะสม ง่ายต่อการใช้ และพัฒนาระบบงาน เพื่อช่วยให้การวิเคราะห์ข้อมูล เป็นไปได้อย่างมีประสิทธิภาพ มีความถูกต้อง ตรงกับความต้องการที่จะใช้ข้อมูลขององค์กร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม

นงลักษณ์ พลอยปลื้ม "การแปลงข้อมูลเข้าสู่ Data Warehouse" วารสาร BCM. ปีที่ 9 ฉบับที่ 103 (กันยายน 2540) : 126-129.

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

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

Pilot Software. Introduction Pilot Designer. USA. : n.p., 1996.

Pilot Software. Introduction Pilot Desktop. USA. : n.p., 1996.

Pilot Software. Introduction to Pilot DSS. USA. : n.p., 1996.

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

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประวัติผู้เขียน

ชื่อ สกุล	นางสาวศศกร ปัทมวิภาต
สถานที่เกิด	กรุงเทพฯ
การศึกษา	ศิลปศาสตรบัณฑิต (สถิติ) มหาวิทยาลัยธรรมศาสตร์
ปัจจุบัน	รับราชการตำแหน่ง นักวิชาการคอมพิวเตอร์ 6
	สำนักเทคโนโลยีสารสนเทศ กรมสรรพากร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้