

ห้องสมุดคณะเทคโนโลยีสารสนเทศ สจล.

การประยุกต์ใช้ดาต้าไมนิ่ง
ในการสนับสนุนการวิเคราะห์พฤติกรรมการซื้อสินค้า
The Application of Data Mining to Support Buying Behavior Analysis



วัน เดือน ปี..... 27 ส.ค. 2550
เลขทะเบียน..... 02047
เลขเรียกหนังสือ..... 0๗: ๗/431ก 2546
"ห้องสมุดคณะเทคโนโลยีสารสนเทศ สจล."

รายงานนี้เป็นส่วนหนึ่งของวิชาโครงการพัฒนาระบบงาน
หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ

ภาคเรียนที่ 1 ปีการศึกษา 2546

คณะเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ชื่อหัวข้อ	การประยุกต์ใช้ดาต้าไมน์นิ่งในการสนับสนุนการวิเคราะห์พฤติกรรมการณ์ซื้อสินค้า
นักศึกษา	นายปรัชญา โกศล
อาจารย์ที่ปรึกษา	ดร. ภัทรชัย ลลิตโรจน์วงศ์
ระดับการศึกษา	วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ
แขนงวิชา	วิทยาการสารสนเทศ
ปีการศึกษา	2546

บทคัดย่อ

เพื่อเป็นการศึกษาการพัฒนาดาต้าไมน์นิ่งจากฐานข้อมูลและวิธีการกระบวนการในเชิงประยุกต์ใช้พร้อมทั้งสร้างแอปพลิเคชันเพื่อการจัดการและแสดงผลลัพธ์จากดาต้าไมน์นิ่งที่ได้สร้างมา โดยมีแนวทางในการศึกษาจากการเริ่มรู้จักวิธีการเตรียมข้อมูลที่มีอยู่แล้วจากฐานข้อมูลธุรกิจค้าปลีกให้เหมาะสมในการทำฐานข้อมูลแบบไมน์นิ่ง การออกแบบ โครงสร้างและวิธีการของดาต้าไมน์นิ่งรวมทั้งเครื่องมือและคอมพิวเตอร์ต่างๆที่ช่วยในการสร้างแอปพลิเคชันซึ่งทั้งหมดนี้จะทำให้ได้ระบบที่สนับสนุนการวิเคราะห์พฤติกรรมการณ์ซื้อสินค้าในธุรกิจค้าปลีกเพื่อช่วยในการทำกิจกรรมด้านการตลาดไม่ว่าจะเป็นการออกแบบแคตตาล็อกสินค้า การจัดวางชั้นสินค้า การออกแบบโปรโมชั่น เป็นต้น

Title	The Application of Data Mining to Support Buying Behavior Analysis
Student	Mr. Prachaya Kosol
Advisor	Dr. Pattarachai Lalitrojwong
Level of Study	Master of Science in Information Technology
Major	Information Science
Academic Year	2003

ABSTRACT

For education and development of data mining from database and the method procession of applying, also create the application for management and effect from created data mining. Tendency of education is start from knowledge of method to prepare an origin data from retail business database to appropriate to do the mining kind of database. Framework design and the method of data mining combine with equipment and any component that help to create the application which all will make the system that Support Buying Behavior Analysis of retail business. To help to do the marketing activity such as catalog design, layer of shelf and the promotion design, ect.

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	I
บทคัดย่อภาษาอังกฤษ	II
สารบัญ	III
สารบัญตาราง	V
สารบัญรูป	VI
บทที่	
1. บทนำ	1
1.1 ความสำคัญและที่มาของโครงการ	1
1.2 วัตถุประสงค์ของโครงการ	1
1.3 ขอบเขตของโครงการ	2
1.4 การดำเนินโครงการ	2
1.5 ประโยชน์ที่คาดว่าจะได้รับ	3
2. ทฤษฎีที่เกี่ยวข้อง	4
2.1 การวิเคราะห์พฤติกรรมการณ์ซื้อสินค้า	4
2.2 คาด้าไมน์นิ่ง	10
3. การวิเคราะห์ระบบ	37
3.1 ความต้องการของระบบ	37
3.2 ขั้นตอนการทำคาด้าไมน์นิ่งของระบบ	38
3.3 แผนภาพกระแสข้อมูล	38
4. การออกแบบระบบ	48
4.1 โครงร่างของคิวบีในคาด้าแวร์เฮาส์	48
4.2 การออกแบบคาด้าไมน์นิ่งโมเดล	52
4.3 การออกแบบฐานข้อมูลเชิงสัมพันธ์	54
4.4 การออกแบบส่วนต่อประสานผู้ใช้	64
5. การพัฒนาระบบ	78
5.1 การพัฒนาคาด้าไมน์นิ่ง	78

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
บทที่	
5.2 การพัฒนาระบบฐานข้อมูล	83
5.3 การพัฒนาแอปพลิเคชัน	87
5.4 สรุปผลการพัฒนาระบบ	88
6. สรุปโครงการพัฒนาระบบงาน	89
6.1 สรุปโครงการ	89
6.2 ประโยชน์ที่ได้รับจากโครงการพัฒนาระบบงาน	89
6.3 ปัญหา ข้อจำกัด และข้อเสนอแนะ	89
บรรณานุกรม	91
ประวัติผู้เขียน	92



สารบัญตาราง

ตารางที่	หน้า
2.1 ชั้นวัฏจักรชีวิตครอบครัว	7
4.1 Data Dictionary ของคิวบ์การขาย	49
4.2 โครงสร้างของดาต้าไมน์นึ่งโมเดล Buying Behavior	53
4.3 Data Dictionary ของฐานข้อมูล Membership	56
4.4 Data Dictionary ของฐานข้อมูล ThaiMart	62



สารบัญรูป

รูปที่	หน้า
2.1 แบบจำลองพฤติกรรมผู้ซื้อ	5
2.2 ปัจจัยที่มีอิทธิพลต่อพฤติกรรมผู้บริโภค	5
2.3 วิธีการในการทำค้ำไม้หนึ่ง	12
2.4 โครงสร้างของ MOLAP	14
2.5 โครงสร้างของ ROLAP	14
2.6 โครงสร้างของ HOLAP	15
2.7 สถาปัตยกรรมแบบ Single-Tier	16
2.8 สถาปัตยกรรมแบบ Two-Tier	17
2.9 สถาปัตยกรรมแบบ Three-Tier	17
2.10 Microsoft Analysis Manager	21
2.11 การเลือกชนิดแหล่งข้อมูล	22
2.12 การเลือกตารางข้อมูล	22
2.13 การเลือกเทคนิคของค้ำไม้หนึ่ง	23
2.14 การเลือกคอลัมน์อินพุตและคอลัมน์ทำนายข้อมูล	24
2.15 Relational Mining Model Editor (Schema)	24
2.16 Relational Mining Model Editor (Content)	25
2.17 Dependency Network Browser	26
2.18 The links pointer	26
2.19 การเลือกชนิดแหล่งข้อมูล	28
2.20 การเลือกแหล่งข้อมูลโดยการใช้คิวบ์	28
2.21 Select data-mining technique	29
2.22 การเลือก case	30
2.23 การเลือก “predicted entity”	30
2.24 การเลือกการทำ training ข้อมูล	31
2.25 การสร้างเวอร์ซิวส์ไคเมนชันและคิวบ์	32

สารบัญรูป (ต่อ)

รูปที่	หน้า
2.26 การตั้งชื่อให้กับคาด้าไมน์นึ่งโมเดล	32
2.27 OLAP Mining Model Editor	33
2.28 วิธีการในการที่จะแสดงข้อมูลของเวอร์ชวลคิวบ์	34
2.29 การแสดงข้อมูลของเวอร์ชวลคิวบ์	34
2.30 วิธีการในการที่จะแสดงข้อมูลของไคเมนชั้น	35
2.31 การแสดงข้อมูลของไคเมนชั้นในส่วน MDX	36
2.32 การแสดงขอดีการขายด้วยไคเมนชั้น	36
3.1 DFD ระดับ Context diagram	39
3.2 DFD ระดับ 0	41
3.3 DFD ระดับ 1 ในรายละเอียดของ DATA MINING CREATION	42
3.4 DFD ระดับ 1 ในรายละเอียดของ DATA MINING APPLICATION	44
3.5 DFD ระดับ 1 ในรายละเอียดของ MEMBERSHIP CARD APPLICATION	45
3.6 DFD ระดับ 1 ในรายละเอียดของ PRPDUCT CATALOGUE APPLICATION	47
4.1 โครงสร้างของคิวบ์การขาย	48
4.2 Database Diagram	55
4.3 หน้าจอเมนูหลัก	64
4.4 หน้าจอประมวลผลคาด้าไมน์นึ่งโมเดล	65
4.5 หน้าจอแสดงขั้นตอนการประมวลผลคาด้าไมน์นึ่งโมเดล	66
4.6 หน้าจอข้อมูลของคาด้าไมน์นึ่งโมเดล	67
4.7 หน้าจอการกำหนดเงื่อนไขประเภทบัตรสมาชิก	68
4.8 หน้าจอการเปลี่ยนประเภทบัตรสมาชิก	69
4.9 หน้าจอการส่งเอกสารการเปลี่ยนประเภทบัตรสมาชิก	70
4.10 หน้าจอสถานะการเปลี่ยนประเภทบัตรสมาชิก	71
4.11 หน้าจอใบสมัครบัตรสมาชิก	72
4.12 หน้าจอสร้างแคตตาล็อกสินค้า	73

สารบัญรูป (ต่อ)

รูปที่	หน้า
4.13 หน้าจอจัดส่งแคตตาล็อกสินค้า	75
4.14 หน้าจอดูแคตตาล็อกสินค้า	76
5.1 Buying Behavior Model	78
5.2 Buying Behavior Model เมื่อเลือก Tree color based on แบบบัตรสมาชิกประเภท ธรรมดา	79
5.3 Buying Behavior Model เมื่อเลือก Tree color based on แบบบัตรสมาชิกประเภท ทองแดง	80
5.4 Buying Behavior Model เมื่อเลือก Tree color based on แบบบัตรสมาชิกประเภทเงิน	81
5.5 Buying Behavior Model เมื่อเลือก Tree color based on แบบบัตรสมาชิกประเภททอง	82
5.6 Microsoft SQL Server 2000 รุ่น Standard Edition	84
5.7 การสร้างฐานข้อมูล Membership	84
5.8 การกำหนดคุณสมบัติของไฟล์ข้อมูล	85
5.9 การกำหนดคุณสมบัติของไฟล์ล็อก	85
5.10 การสร้างตารางข้อมูลในฐานข้อมูล Membership	86
5.11 รายการตารางข้อมูลทั้งหมดในฐานข้อมูล Membership	87
5.12 Microsoft Visual Studio .Net	87
5.13 โปรแกรมดูข้อมูลของดาต้าไมน์นิ่งโมเดล	88

บทที่ 1

บทนำ

1.1 ความสำคัญและที่มาของโครงการ

โครงการพัฒนาระบบงานในหัวข้อเรื่อง การประยุกต์ใช้ดาต้าไมน์นิ่งในการสนับสนุนการวิเคราะห์พฤติกรรมการณ์ซื้อสินค้า นั้น มีความสำคัญในแง่ที่ว่า ทำให้ผู้พัฒนาโครงการได้นำเอาความรู้ที่ได้เรียนมาตามหลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ ที่คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง นำมาประยุกต์ปฏิบัติ ฝึกฝน แก้ปัญหา ตลอดจนค้นคว้าหาข้อมูลเพิ่มเติมในบางส่วนที่ยังขาดอยู่ ทำให้เห็นภาพโดยรวมในวิทยาการสารสนเทศที่ว่าด้วยการออกแบบและพัฒนาโครงการระบบงาน ส่วนที่มาของโครงการนี้เกิดจากการที่ผู้พัฒนาโครงการมีความสนใจในเทคโนโลยีของระบบฐานข้อมูล โดยเฉพาะด้านการประยุกต์ใช้ดาต้าไมน์นิ่งและพัฒนาการทางความรู้ด้านการบริหารจัดการการตลาด โดยเฉพาะในด้านการวิเคราะห์พฤติกรรมการณ์ซื้อสินค้าของผู้บริโภค และเมื่อได้นำองค์ความรู้ทั้งสองด้านนี้มารวมกัน ในแง่ที่ว่าดาต้าไมน์นิ่งสามารถสนับสนุนการวิเคราะห์พฤติกรรมการณ์ซื้อสินค้าได้จึงทำให้เกิดโครงการพัฒนาระบบงานนี้ขึ้นมา

1.2 วัตถุประสงค์ของโครงการ

โครงการพัฒนาระบบงานในหัวข้อเรื่อง การประยุกต์ใช้ดาต้าไมน์นิ่งในการสนับสนุนการวิเคราะห์พฤติกรรมการณ์ซื้อสินค้า นั้น มีวัตถุประสงค์หลักดังนี้

- เพื่อให้สามารถประยุกต์ความรู้ที่ได้เรียนมามานำมาพัฒนา ศึกษาค้นคว้าและแก้ปัญหาในโครงการพัฒนาระบบงานได้
- เพื่อให้เกิดความรู้ความเข้าใจในด้านการวิเคราะห์พฤติกรรมการณ์ซื้อสินค้าของผู้บริโภค และด้านการประยุกต์ใช้ดาต้าไมน์นิ่งได้อย่างต้องแท้
- เพื่อให้มีโอกาสได้แลกเปลี่ยนความรู้ความเข้าใจและประสบการณ์กับผู้อื่นที่เกี่ยวข้องในเนื้อหาของโครงการพัฒนาระบบงานที่ทำอยู่ได้
- เพื่อเรียนรู้วิธีการศึกษาค้นคว้า การพัฒนาและการนำเสนอผลงานโครงการพัฒนาระบบงานได้

1.3 ขอบเขตของโครงการ

เพื่อให้โครงการพัฒนาระบบงานในหัวข้อเรื่อง การประยุกต์ใช้ดาต้าไมน์นิ่งในการสนับสนุนการวิเคราะห์พฤติกรรมการณ์ซื้อสินค้า นั้น ได้บรรลุวัตถุประสงค์ที่ได้ตั้งไว้ รวมทั้งสำเร็จเสร็จสิ้นในระยะเวลาที่กำหนดซึ่งเป็นข้อจำกัดที่สำคัญนั้น โครงการพัฒนาระบบงานนี้จึงมีการพัฒนาการทำดาต้าไมน์นิ่งโมเดลจากฐานข้อมูลธุรกิจค้าปลีกที่มีอยู่แล้ว ด้วยการใช้ Microsoft SQL Server Analysis Services และใช้เทคนิควิธีการแบบ Microsoft Decision Trees ซึ่งเป็นวิธีการทำนายข้อมูลแบบความน่าจะเป็น โดยมีเงื่อนไขต่างๆ เป็นตัวกำหนด ซึ่งในส่วนการออกแบบดาต้าไมน์นิ่งโมเดลนั้น จะวิเคราะห์จากความต้องการของระบบ เพราะหนึ่งโมเดลสามารถสนับสนุนความต้องการได้ไม่หลากหลายความต้องการ ต่อจากนั้นจะมีการพัฒนาแอปพลิเคชันขึ้นมาด้วยเครื่องมือพัฒนาและออกแบบแอปพลิเคชัน Microsoft Visual Basic และมีการใช้คอมโพเนนต์ที่สำคัญคือ Decision Support Objects (DSO) ที่มีหน้าที่ในการบริหารจัดการโครงสร้างและข้อมูลของดาต้าไมน์นิ่งโมเดลนั้น โดยแอปพลิเคชันที่พัฒนาขึ้นมา ก็เพื่อเป็นการแสดงข้อมูลต่างๆ จากดาต้าไมน์นิ่งโมเดลที่ได้สร้างไว้แล้ว เพื่อให้ผู้ใช้ได้ทำการวิเคราะห์พฤติกรรมการณ์ซื้อสินค้าและนำไปใช้ในการตัดสินใจในการดำเนินกิจกรรมทางการตลาดต่อไปนั่นเอง

1.4 การดำเนินโครงการ

การดำเนินโครงการพัฒนาระบบงานในหัวข้อเรื่อง การประยุกต์ใช้ดาต้าไมน์นิ่งในการสนับสนุนการวิเคราะห์พฤติกรรมการณ์ซื้อสินค้า นั้น จะเริ่มด้วยการศึกษาทำความเข้าใจทฤษฎีการวิเคราะห์พฤติกรรมการณ์ซื้อสินค้า (Buying Behavior Analysis) ที่ประกอบไปด้วยตั้งแต่

- แบบจำลองพฤติกรรมผู้บริโภค
- ลักษณะเฉพาะที่กระทบต่อพฤติกรรมผู้บริโภค
- ประเภทของพฤติกรรมการณ์ซื้อ
- กระบวนการตัดสินใจซื้อ

จากนั้นจะศึกษาทำความเข้าใจทฤษฎีของดาต้าไมน์นิ่ง ที่ประกอบไปด้วยตั้งแต่

- ดาต้าไมน์นิ่งคืออะไร
- คำนิยามต่างๆที่ใช้ในดาต้าไมน์นิ่ง
- วิธีการในการทำดาต้าไมน์นิ่ง
- ดาต้าไมน์นิ่งในวิธีการของไมโครซอฟท์
- วิธีการของดาต้าไมน์นิ่งและสถิติ
- วิธีการทำดาต้าไมน์นิ่งของไมโครซอฟท์โดยการใช้ Decision Trees

- การสร้าง Decision Trees ด้วย OLAP

จากนั้นจะเป็นการวิเคราะห์ระบบตามความต้องการด้วยการใช้แผนงานกระแสข้อมูล (Data Flow Diagram - (DFD)) ในการอธิบายกระแสการไหลของข้อมูลรวมทั้งการประมวลผลต่างๆ ในระบบ ต่อจากนั้นจะเป็นการออกแบบระบบที่ประกอบไปด้วยตั้งแต่

- โครงสร้างของคิวบ์ (ดาต้าแวร์เฮาส์)
- การออกแบบดาต้าไมนนิ่งโมเดล
- ส่วนต่อประสานผู้ใช้ (User Interface)

จากนั้นเป็นการพัฒนาและการติดตั้งที่เริ่มด้วยการศึกษาคอมพิวเตอร์ Decision Support Objects (DSO) การนำมาใช้ด้วยการเขียนโปรแกรม วิธีการติดตั้งทั้งในส่วนฐานข้อมูลดาต้าไมนนิ่งโปรแกรมที่พัฒนา สุดท้ายของการดำเนินโครงการพัฒนาระบบงานคือการสรุปผล

1.5 ประโยชน์ที่คาดว่าจะได้รับ

สำหรับประโยชน์ที่คาดว่าจะได้รับมีด้วยกันใน 2 ด้านคือ ด้านของผู้พัฒนาโครงการและด้านของโครงการที่ได้พัฒนาขึ้นมา สำหรับในด้านของผู้พัฒนาโครงการมีดังต่อไปนี้

- สามารถประยุกต์ความรู้ที่ได้เรียนมาเพื่อแก้ปัญหาในการพัฒนาและ/หรือการจัดการโครงการพัฒนาระบบงานด้านเทคโนโลยีสารสนเทศที่เกิดขึ้นในสภาพที่เป็นจริงได้
- เพื่อให้เกิดความรู้ความเข้าใจในวิชาที่ศึกษาตามหลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ
- เพื่อให้เกิดทักษะและความรู้ในด้านการวิเคราะห์และออกแบบระบบ รวมทั้งการพัฒนาระบบที่ใช้เทคโนโลยีดาต้าไมนนิ่งในส่วนระบบฐานข้อมูล การโปรแกรมโดยใช้คอมพิวเตอร์ Decision Support Objects (DSO) ในการบริหารจัดการโครงสร้างและข้อมูลของดาต้าไมนนิ่งโมเดลในการเขียนโปรแกรม
- เพื่อให้เกิดทักษะและความรู้ในด้านการวิเคราะห์พฤติกรรมการณ์ซื้อสินค้า

ส่วนในด้านของโครงการที่ได้พัฒนาขึ้นมาที่จะรองรับสนับสนุนการทำงานของระบบมีดังต่อไปนี้

- สนับสนุนกิจกรรมทางการตลาดด้วยการวิเคราะห์พฤติกรรมการณ์ซื้อสินค้าของผู้บริโภคจากฐานข้อมูลที่ได้มาจากการดำเนินงานและปฏิบัติการของธุรกิจค้าปลีกอันนำมาเพื่อขยายและผลกำไรสูงสุด

บทที่ 2

ทฤษฎีที่เกี่ยวข้อง

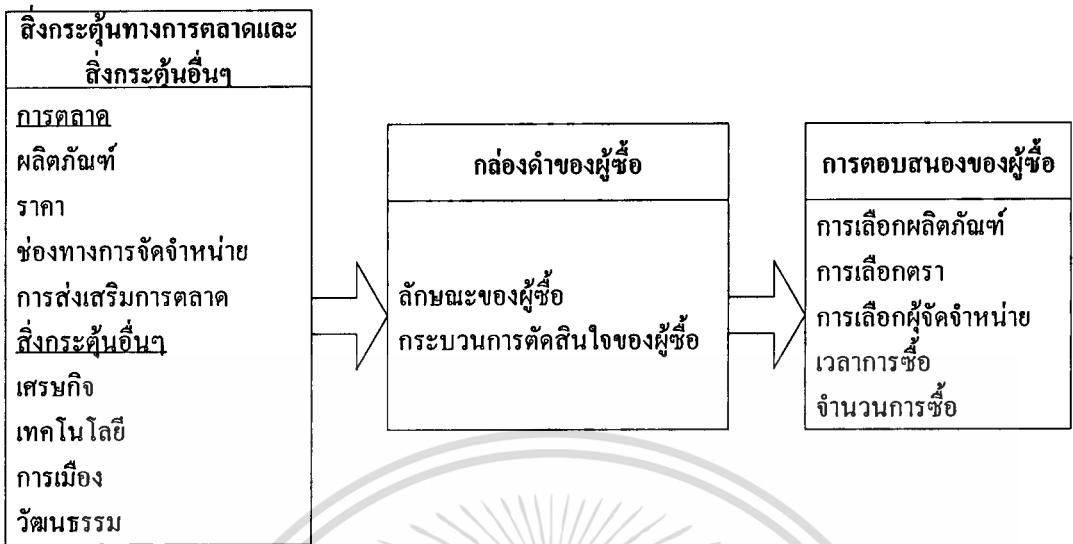
2.1 การวิเคราะห์พฤติกรรมผู้บริโภค

พฤติกรรมของผู้บริโภคเกิดจากปัจจัยต่างๆ ที่มากระทบต่อพฤติกรรมผู้บริโภค การเข้าใจถึงพฤติกรรมผู้บริโภคไม่ใช่เรื่องง่าย และถือเป็นงานที่สำคัญของการจัดการการตลาด พฤติกรรมของผู้บริโภค หมายถึง พฤติกรรมของผู้บริโภคคนสุดท้าย ไม่ว่าจะส่วนบุคคลหรือครัวเรือนที่ทำการซื้อสินค้าและบริการสำหรับการบริโภคส่วนตัว

2.1.1 แบบจำลองพฤติกรรมผู้บริโภค (Kotler, Philip. and Armstrong, Gary. 2001.)

ผู้บริโภคต้องทำการตัดสินใจซื้อมากมายในแต่ละวัน ซึ่งกิจการขนาดใหญ่มักจะทำการวิจัยการตัดสินใจของผู้บริโภคอย่างละเอียด เพื่อตอบคำถามให้ได้ว่าผู้บริโภคซื้ออะไร (what) ซื้อที่ไหน (where) ซื้ออย่างไรและซื้อเท่าใด (how and how much) ซื้อเมื่อใด (when) และซื้อทำไม (why) ซึ่งการจะรู้ว่าผู้บริโภคซื้ออะไรไม่ใช่เรื่องง่ายนัก แบบจำลองสิ่งกระตุ้น-การตอบสนองของพฤติกรรมผู้ซื้อดังรูปที่ 2.1 เพื่อให้กิจการสามารถเข้าใจถึงวิธีการตอบสนองของผู้บริโภคต่อลักษณะผลิตภัณฑ์ ราคา และโฆษณาได้รวมทั้งผู้บริโภคตอบสนองต่อความพยายามทางการตลาดต่างๆ ที่กิจการใช้อย่างไร

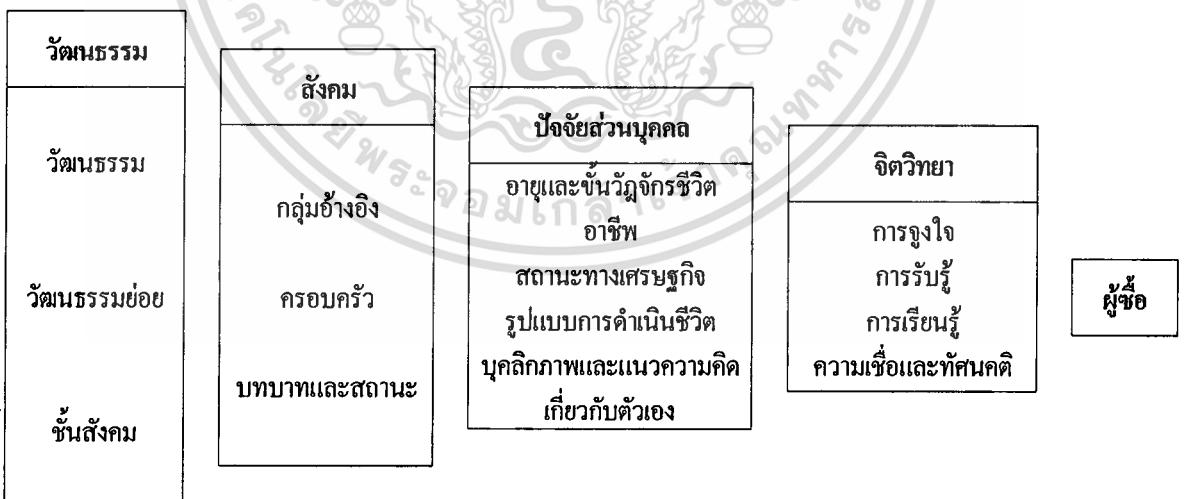
สิ่งกระตุ้นทางการตลาดประกอบด้วยปัจจัย 4 ประการคือ ผลิตภัณฑ์ (Product) ราคา (Price) ช่องทางการจัดจำหน่าย (Place) และการส่งเสริมการตลาด (Promotion) หรือ 4 Ps สิ่งกระตุ้นอื่นๆ จะรวมถึงเหตุการณ์หรือสิ่งกีดขวางในสภาพแวดล้อมของผู้ซื้อ เช่น เศรษฐกิจ เทคโนโลยี การเมืองและวัฒนธรรม ปัจจัยนำเข้าเหล่านี้จะนำเข้าสู่กล่องคำเพื่อจะเปลี่ยนเป็นการตอบสนองของผู้ซื้อ โดยเห็นได้จาก การเลือกผลิตภัณฑ์ ตราผลิตภัณฑ์ ผู้จัดจำหน่าย เวลาในการซื้อ และจำนวนการซื้อ จากรูปที่ 2.1 ในกล่องคำจะทำให้ทราบว่าผู้บริโภคเปลี่ยนแปลงสิ่งกระตุ้นทางการตลาดและสิ่งกระตุ้นอื่นๆ ให้เป็นการตอบสนองได้โดยส่วนแรกคือ ลักษณะของผู้บริโภคที่มีอิทธิพลต่อการรับรู้และการตอบสนองต่อสิ่งกระตุ้นของผู้บริโภค ส่วนที่สองคือ กระบวนการตัดสินใจของผู้บริโภคที่กระทบต่อพฤติกรรมผู้บริโภค



รูปที่ 2.1 แบบจำลองพฤติกรรมผู้ซื้อ

2.1.2 ลักษณะเฉพาะที่กระทบต่อพฤติกรรมผู้บริโภค

ลักษณะทางวัฒนธรรม สังคม ปัจจัยส่วนบุคคลและลักษณะทางจิตวิทยามีอิทธิพลอย่างมากต่อการซื้อของผู้บริโภคดังแสดงในรูปที่ 2.2 ซึ่งโดยส่วนมากแล้วเราไม่สามารถควบคุมปัจจัยเหล่านี้ได้แต่ก็ต้องให้ความสนใจ (Kotler, Philip. and Armstrong, Gary. 2001.)



รูปที่ 2.2 ปัจจัยที่มีอิทธิพลต่อพฤติกรรมผู้บริโภค

- ปัจจัยทางวัฒนธรรม มีอิทธิพลต่อพฤติกรรมผู้บริโภคทั้งในเชิงกว้างและลึก โดยแบ่งได้เป็น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- **วัฒนธรรม (Culture)** เป็นมูลเหตุพื้นฐานที่สุดของพฤติกรรมความต้องการของบุคคล พฤติกรรมของมนุษย์เกิดจากการเรียนรู้ เด็กจะได้เรียนรู้ถึงค่านิยม การรับรู้ความต้องการ และพฤติกรรมพื้นฐานจากครอบครัวและสถาบันทางสังคมที่เด็กเติบโตมา สรุปลแล้ววัฒนธรรมมีอิทธิพลต่อพฤติกรรมการซื้อค่อนข้างมาก
- **วัฒนธรรมย่อย (Subcultures)** ในแต่ละวัฒนธรรมจะประกอบด้วยวัฒนธรรมย่อยหรือกลุ่มบุคคลซึ่งมีระบบค่านิยมร่วมกัน เนื่องจากมีประสบการณ์ชีวิตและอยู่ใต้สถานการณ์อย่างเดียวกัน วัฒนธรรมย่อยรวมถึง สัญชาติ ศาสนา กลุ่มเชื้อชาติและภูมิศาสตร์ วัฒนธรรมย่อยหลายวัฒนธรรมทำให้เกิดส่วนตลาดที่สำคัญเกิดขึ้น
- **ชั้นสังคม (Social classes)** เกือบทุกสังคมจะมีโครงสร้างชั้นสังคม ชั้นสังคมคือการแบ่งสมาชิกสังคมตามลำดับชั้นอย่างถาวร โดยที่สมาชิกในแต่ละชั้นสังคมจะมีค่านิยม ความสนใจและพฤติกรรมที่คล้ายกัน การแบ่งชั้นสังคมจะแบ่งโดยอาศัยปัจจัยหลายตัวร่วมกัน อันได้แก่ รายได้ การศึกษา ฐานะและปัจจัยอื่น บุคคลภายในชั้นสังคมแต่ละชั้นมีแนวโน้มที่จะแสดงพฤติกรรมการซื้อที่คล้ายกัน
- **ปัจจัยทางสังคม (Social factors)** พฤติกรรมผู้บริโภคยังได้รับอิทธิพลจากปัจจัยทางสังคม ได้แก่
 - **กลุ่ม (Group)** หมายถึง บุคคลตั้งแต่สองคนขึ้นไปที่มีปฏิสัมพันธ์กันเพื่อให้บรรลุเป้าหมายของแต่ละคนหรือเป้าหมายร่วม พฤติกรรมของบุคคลได้รับอิทธิพลจากกลุ่มย่อยหลายกลุ่ม กลุ่มที่มีอิทธิพลโดยตรงและมีอิทธิพลต่อคนที่อยู่ในกลุ่มที่เรียกว่า กลุ่มสมาชิก ในทางตรงข้ามกลุ่มอ้างอิงเป็นกลุ่มที่ใช้เปรียบเทียบทั้งทางตรงและทางอ้อม หรือใช้อ้างอิงเพื่อสร้างทัศนคติหรือพฤติกรรมของคนซึ่งจะกระทบต่อการเลือกผลิตภัณฑ์และตราของบุคคล
 - **ครอบครัว (Family)** จะมีอิทธิพลอย่างมากต่อพฤติกรรมการซื้อ ครอบครัวเป็นองค์ประกอบที่สำคัญของการซื้อของผู้บริโภคที่มีความสำคัญมากที่สุดในสังคม มีการวิจัยเกี่ยวกับบทบาทและอิทธิพลของสามี ภรรยาและบุตรต่อการซื้อสินค้าและบริการต่างๆ
 - **บทบาทและสถานะ** แต่ละคนอาจจะอยู่ในหลายกลุ่ม เช่น ครอบครัว ชมรม องค์กรต่างๆ ตำแหน่งของบุคคลในแต่ละกลุ่มสามารถกำหนดได้ในรูปของบทบาทและสถานะ บทบาท (Role) ประกอบด้วยกิจกรรมต่างๆที่บุคคลถูกคาดหวังว่าจะกระทำออกมาโดยสิ่งที่กระทำจะขึ้นอยู่กับบุคคลที่อยู่ห้อมล้อม แต่ละบทบาทจะมีอิทธิพลต่อพฤติกรรมการซื้อบางอย่างและบทบาทจะปฏิบัติตามสถานะซึ่งสะท้อน

ถึงการให้การยอมรับจากสังคม บุคคลมักจะเลือกผลิตภัณฑ์ที่แสดงถึงสถานะในสังคม

- ปัจจัยส่วนบุคคล การตัดสินใจของผู้ซื้อยังได้รับอิทธิพลจากลักษณะเฉพาะส่วนบุคคล ได้แก่
 - อายุและขั้นวัฏจักรชีวิต บุคคลมีการเปลี่ยนแปลงสินค้าหรือบริการที่ซื้อไปตลอดชั่วชีวิต รสนิยมที่มีต่ออาหาร เสื้อผ้า เฟอร์นิเจอร์ และการพักผ่อนมักจะเกี่ยวข้องกับอายุ การซื้อยังสอดคล้องกับขั้นวัฏจักรชีวิตครอบครัว (Family life cycle) ซึ่งเป็นลำดับขั้นของครอบครัวที่สมบูรณ์ที่จะมีขึ้นตลอดเวลา ตามตารางที่ 2.1 แสดงขั้นวัฏจักรชีวิตของครอบครัว (Kotler, Philip. and Armstrong, Gary. 2001.)

ตารางที่ 2.1 ขั้นวัฏจักรชีวิตครอบครัว

วัยหนุ่มสาว	วัยกลางคน	วัยสูงอายุ
คนโสด	คนโสด	คนสูงอายุที่แต่งงาน
คู่สมรสที่ไม่มีบุตร	คู่สมรสที่ไม่มีบุตร	คนสูงอายุที่ไม่ได้แต่งงาน
คู่สมรสที่มีบุตร	คู่สมรสที่มีบุตร	
ผู้เป็นหม้ายที่มีบุตร	คู่สมรสที่บุตรสามารถเลี้ยงตัวเองได้	
	ผู้เป็นหม้ายที่ไม่มีบุตร	
	ผู้เป็นหม้ายที่มีบุตร	
	ผู้เป็นหม้ายที่มีบุตรสามารถเลี้ยงตัวเองได้	

- อาชีพ (Occupation) อาชีพของบุคคลจะส่งผลต่อสินค้าและบริการที่ซื้อ
- สถานะทางเศรษฐกิจ (Economic situation) สถานะทางเศรษฐกิจของบุคคลจะส่งผลต่อการเลือกผลิตภัณฑ์ โดยประกอบด้วยรายได้ เงินออม และอำนาจในการกู้ยืมที่เพียงพอที่จะใช้จ่าย
- รูปแบบการดำเนินชีวิต (Lifestyle) คือ แบบแผนการใช้ชีวิตของบุคคลที่แสดงในรูปของลักษณะจิตนิสัย (Psychographics) รูปแบบการดำเนินชีวิตจะวัดจาก กิจกรรม ความสนใจ และความคิดเห็น ได้แก่ การทำงาน งานอดิเรก การใช้จ่าย กีฬาที่เล่น เหตุการณ์ทางสังคม ความสนใจ ได้แก่ อาหาร แฟชั่น ครอบครัว การพักผ่อนและความคิดเห็นได้แก่ ความคิดเห็นเกี่ยวกับตัวเอง หัวข้อทางสังคม ธุรกิจ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สินค้า รูปแบบการดำเนินชีวิตสามารถช่วยให้เข้าใจถึงการเปลี่ยนแปลงค่านิยมของ ผู้บริโภคและส่งผลต่อพฤติกรรมผู้บริโภคได้

- บุคลิกภาพและแนวความคิดเกี่ยวกับตัวเอง บุคลิกภาพที่แตกต่างของแต่ละบุคคลมี อิทธิพลต่อพฤติกรรมผู้บริโภคของบุคคล บุคลิกภาพ หมายถึง ลักษณะด้านจิตวิทยา ที่มีลักษณะแตกต่างกันในแต่ละบุคคล ซึ่งนำไปสู่การตอบสนองต่อสภาพแวดล้อมของบุคคลนั้นอย่างมั่นคงและถาวร บุคลิกภาพโดยปกติจะอธิบายในรูปของ อุปนิสัย เช่น ความมั่นใจในตนเอง ความมีอำนาจ การชอบเข้าสังคม รักอิสระ ความก้าวร้าว บุคลิกภาพมีประโยชน์อย่างมากต่อการวิเคราะห์พฤติกรรมผู้บริโภค ในการเลือกผลิตภัณฑ์หรือตราผลิตภัณฑ์
- ปัจจัยทางจิตวิทยา การเลือกซื้อของบุคคลได้รับอิทธิพลจากปัจจัยทางจิตวิทยาที่สำคัญ ได้แก่
 - การจูงใจ แต่ละคนมีความต้องการหลายสิ่งหลายอย่างในช่วงเวลาหนึ่ง ความต้องการทางจิตวิทยา (Psychological) เกิดจากความต้องการการยอมรับ การยกย่อง หรือการเป็นเจ้าของทรัพย์สิน ความต้องการส่วนใหญ่อาจไม่มากพอที่จะจูงใจให้ บุคคลกระทำในช่วงเวลานั้น ความต้องการกลายเป็นสิ่งจูงใจเมื่อได้รับการกระตุ้น อย่างเพียงพอจนเกิดเป็นความตึงเครียด สิ่งจูงใจหรือแรงขับเป็นความต้องการที่กดดันจนมากพอที่จะจูงใจให้บุคคลเกิดพฤติกรรมเพื่อตอบสนองความต้องการของตนเอง
 - การรับรู้ เป็นกระบวนการที่บุคคลเลือกรับรู้ จัดการ และตีความหมายข้อมูลต่างๆ เพื่อให้เกิดเป็นความเข้าใจ บุคคลสามารถสร้างการรับรู้ที่แตกต่างกันได้จากสิ่ง กระตุ้นเดียวกันเนื่องจากกระบวนการรับรู้ 3 ขั้นตอน คือ การเลือกให้ความสนใจ การเลือกแปลความหมายบิดเบือน และการเลือกเก็บรักษา การเลือกให้ความสนใจ คือแนวโน้มสำหรับบุคคลที่จะกลั่นกรองข่าวสารข้อมูลออกจากสิ่งที่เห็นหรือได้ยิน การเลือกแปลความหมายบิดเบือน คือแนวโน้มที่บุคคลจะแปลความหมายของ ข่าวสารให้เป็นไปในทางเดียวกับความเชื่อที่มีอยู่แล้ว การเลือกเก็บรักษา คือบุคคล มักจะลืมข่าวสารที่ได้เรียนรู้มา มีแนวโน้มที่จะเก็บรักษาข่าวสารที่สนับสนุนทัศนคติและความเชื่อของตนเอง
 - การเรียนรู้ แสดงถึงการเปลี่ยนแปลงพฤติกรรมของบุคคลที่เกิดจากประสบการณ์ ของแต่ละคน พฤติกรรมมนุษย์โดยมากเกิดจากการเรียนรู้ การเรียนรู้เกิดจาก ปฏิกริยาระหว่างแรงขับ ตัวกระตุ้น ตัวนำ การตอบสนอง และการเสริมแรง

- **ความเชื่อและทัศนคติ** บุคคลเกิดความเชื่อและทัศนคติโดยผ่านการกระทำและการเรียนรู้ซึ่งความเชื่อและการเรียนรู้จะมีอิทธิพลต่อพฤติกรรมการซื้อ ความเชื่อเป็นความคิดที่บุคคลยึดถือเกี่ยวกับสิ่งใดสิ่งหนึ่ง ทัศนคติคือการประเมิน ความรู้สึก และความโน้มเอียงของบุคคลที่มีต่อสิ่งใดหรือความคิดใดอย่างเหนียวแน่น

2.1.3 ประเภทของพฤติกรรมการตัดสินใจซื้อ (Kotler, Philip. and Armstrong, Gary. 2001.)

- **พฤติกรรมการซื้อที่ซับซ้อน (Complex buying behavior)** เมื่อผู้บริโภคได้ทุ่มเทความพยายามสูงในการซื้อสินค้าและรับรู้ว่าแต่ละตราผลิตภัณฑ์มีความแตกต่างกันมาก ราคาแพง มีความเสี่ยงในการซื้อ เป็นผลิตภัณฑ์ที่ซื้อไม่บ่อยนักและเป็นสิ่งที่แสดงถึงตัวผู้ใช้ พฤติกรรมการซื้อของผู้ซื้อจะผ่านขั้นตอนการเรียนรู้โดยเริ่มจากการพัฒนาความเชื่อเกี่ยวกับผลิตภัณฑ์ไปสู่การสร้างทัศนคติและทำการเลือกซื้อ
- **พฤติกรรมการซื้อแบบลดความสงสัย (Dissonance-reducing buying behavior)** เกิดขึ้นเมื่อผู้บริโภคทุ่มเทความพยายามสูงในการซื้อสินค้าราคาแพง ผลิตภัณฑ์ที่ซื้อไม่บ่อย หรือผลิตภัณฑ์ที่มีความเสี่ยงในการซื้อ แต่ผู้บริโภครับรู้ว่าแต่ละตราผลิตภัณฑ์มีความแตกต่างไม่มากนัก โดยผู้ซื้อจะเลือกที่ราคาเหมาะสมหรือสามารถหาซื้อได้สะดวก
- **พฤติกรรมการซื้อที่ติดเป็นนิสัย (Habitual buying behavior)** เกิดขึ้นภายใต้สถานการณ์ที่ผู้ซื้อมีความทุ่มเทพยายามต่ำในการซื้อสินค้าและเห็นว่าแต่ละตราผลิตภัณฑ์มีความแตกต่างไม่มากนัก และเป็นสินค้าที่ซื้อบ่อย มีราคาต่ำ เช่น การซื้อเกลือ ผู้บริโภคสามารถหาซื้อสินค้าได้ง่าย ผู้บริโภคที่ซื้อตราเดิมถือเป็นการซื้อที่ติดเป็นนิสัยมากกว่าซื้อเพราะมีความภักดีสูง พฤติกรรมการซื้อที่ติดเป็นนิสัยนั้นจะไม่เป็นไปตามลำดับการสร้าง ความเชื่อ การสร้างทัศนคติ ไปสู่การเกิดพฤติกรรม
- **พฤติกรรมการซื้อโดยการแสวงหาความหลากหลาย (Variety-seeking buying behavior)** จะเป็นผู้บริโภคที่ทุ่มเทความพยายามต่ำในการซื้อสินค้า แต่รับรู้ว่าแต่ละตราผลิตภัณฑ์มีความแตกต่างกันมาก ในกรณีนี้ผู้บริโภคมักจะเปลี่ยนแปลงการเลือกผลิตภัณฑ์บ่อยแต่เป็นเพราะต้องการหาความหลากหลายมากกว่าความไม่พอใจ

2.1.4 กระบวนการตัดสินใจซื้อ (Kotler, Philip. and Armstrong, Gary. 2001.)

- **การตระหนักถึงความต้องการ** กระบวนการซื้อจะเริ่มต้นที่การตระหนักถึงความต้องการ หมายถึง การที่ผู้บริโภคตระหนักถึงปัญหาหรือความต้องการของตนเอง ในขั้นนี้จะทำวิจัยผู้บริโภคเพื่อค้นหาถึงความต้องการหรือปัญหาที่ทำให้เกิดความต้องการ

- การเสาะหาข้อมูล ผู้บริโภคที่ได้รับการกระตุ้นอาจจะเสาะหาข้อมูลเพิ่มเติมหรือไม่ก็ได้ ถ้าแรงขับมีมากและมีสินค้าที่เป็นพึงพอใจอยู่ใกล้มือผู้บริโภค ผู้บริโภคมักจะซื้อทันที ไม่เช่นนั้นผู้บริโภคจะเก็บความต้องการนั้นไว้ในความทรงจำหรือเสาะหาข้อมูลที่สอดคล้องกับความต้องการนั้น ผู้บริโภคสามารถรับข้อมูลได้จากหลายแหล่ง ทั้งแหล่งบุคคล เช่น ครอบครัว แหล่งพาณิชย์ เช่น การโฆษณา แหล่งสาธารณะ เช่น สื่อมวลชน และจากประสบการณ์ เช่น การใช้สินค้า
- การประเมินทางเลือก ซึ่งหมายถึงวิธีการที่ผู้บริโภคใช้ข้อมูลที่มีนำมาประเมินตราผลิตภัณฑ์ที่อยู่ในกลุ่มตราผลิตภัณฑ์ที่เลือก โดยผู้บริโภคไม่ใช้กระบวนการประเมินง่าย ๆ เพียงกระบวนการเดียวในสถานการณ์การซื้อทุกสถานการณ์
- การตัดสินใจซื้อ โดยปกติผู้บริโภคจะทำการตัดสินใจซื้อตราที่ชอบมากที่สุด อย่างไรก็ตาม ความตั้งใจซื้อและการตัดสินใจซื้อของผู้บริโภคอาจถูกกั้นกลางด้วยทัศนคติของผู้อื่นและปัจจัยทางสถานการณ์ที่คาดไม่ถึง
- พฤติกรรมหลังการซื้อ หลังจากการทำซื้อสินค้าผู้บริโภคอาจรู้สึกพอใจหรือไม่พอใจผลิตภัณฑ์นั้นและทำให้เกิดพฤติกรรมหลังการซื้อที่ต้องให้ความสนใจ สิ่งที่จะนำมาพิจารณาว่าผู้บริโภคจะพอใจหรือไม่พอใจในการซื้อคือการพิจารณาถึงความสัมพันธ์ระหว่างการคาดหวังของผู้บริโภคและการรับรู้ถึงผลการปฏิบัติงานของสินค้า ถ้าสินค้าปฏิบัติงานได้ต่ำกว่าความคาดหวังผู้บริโภคจะรู้สึกผิดหวัง ถ้าสินค้าปฏิบัติงานได้ตามความคาดหวังผู้บริโภคจะรู้สึกพึงพอใจ ถ้าสินค้าปฏิบัติงานได้สูงกว่าความคาดหวังผู้บริโภคจะรู้สึกประทับใจสินค้านั้น

2.2 ดาต้าไมนิง

2.2.1 ดาต้าไมนิงคืออะไร

ดาต้าไมนิง (Data Mining) เป็นขั้นตอนในการขุดค้นและวิเคราะห์กลุ่มของข้อมูลต่างๆ หารูปแบบข้อมูลที่ยัง ไม่ถูกค้นพบและค้นหาความหมายจากข้อมูลด้วย ซึ่งจะเกี่ยวข้องกับการอธิบายแนวทางของข้อมูลในอดีตและทำนายความเป็นไปในอนาคต เพื่อให้เราทำการตัดสินใจในการดำเนินการทางธุรกิจ ตัวอย่างเช่น ในการค้าปลีกสามารถที่กำหนดการจัดตำแหน่งพื้นที่ในการวางขายสินค้าได้ เช่น ถ้าต้องการที่จะเพิ่มยอดขายอุปกรณ์ก่อสร้าง ก็จะมีการตรวจสอบรายการขายย้อนหลังในช่วงหลายๆปีที่ผ่านมา และได้ข้อสังเกตมาว่าลูกค้ามักจะซื้อรองเท้าสุภาพบุรุษพร้อมกับการซื้ออุปกรณ์ก่อสร้าง ทำให้ห้างสรรพสินค้าตัดสินใจวางตำแหน่งก่อสร้างกลับถัดจากแผนกรองเท้าสุภาพบุรุษเพื่อเป็นการเพิ่มโอกาสในการสร้างยอดขาย เป็นต้น

2.2.2 คำนิยาม (Seidman, Claude. 2001.)

Data Mining ขั้นตอนในการค้นหาความหมายของรูปแบบข้อมูลและความสัมพันธ์ระหว่างการวิเคราะห์และจัดแบ่งประเภทของข้อมูลที่ถูกจัดเก็บไว้

Data Warehouse เป็นศูนย์กลางในการจัดเก็บข้อมูล ที่นำมาจากข้อมูลปฏิบัติการในธุรกิจที่อยู่ในรูปแบบฐานข้อมูลออนไลน์ทรานแซกชัน (OLTP Database) ซึ่งได้ผ่านกระบวนการเพื่อให้ง่ายต่อการออกรายงานและรวดเร็วต่อการดึงข้อมูล

Mining Models โครงสร้างของกลุ่มข้อมูลที่ถูกคอมพิวเตอร์จากอัลกอริทึมแบบค้ำไ่มนึ่ง และใช้ในการทำนายความเป็นไปของข้อมูลต่อไป

Patterns กลุ่มของเหตุการณ์ที่มีการเกิดขึ้นที่เพียงพอต่อความสัมพันธ์กันในกลุ่มข้อมูลนั้นๆ ตัวอย่างเช่น เราอาจจะเรียนรู้จากกลุ่มข้อมูลที่แสดงว่าทุกครั้งที่ถูกค้าจะซื้อเบียร์และจะซื้อผ้าอ้อมด้วย ถ้าเหตุการณ์นี้มีความถี่ในการเกิดที่เพียงพอ ก็จะกำหนดเป็น Pattern ที่จะถูกจัดเก็บในโมเดล

Cases ทุกรายการของข้อมูลที่ถูกใช้เป็นทรัพยากรสำหรับทำเป็น โมเดล ตัวอย่างเช่น ถ้าโมเดลอธิบายว่าลูกค้าจะซื้อของจากร้านขายของชำ ถ้าการซื้อนั้นเป็นการซื้อที่มีจำนวนไม่มากนัก และ Case นี้ก็จะถูกใช้เป็นประสบการณ์ให้กับ โมเดลต่อไป

Data-Mining Algorithms อัลกอริทึมแบบคณิตศาสตร์และสถิติที่จะแปลง Case ไปสู่

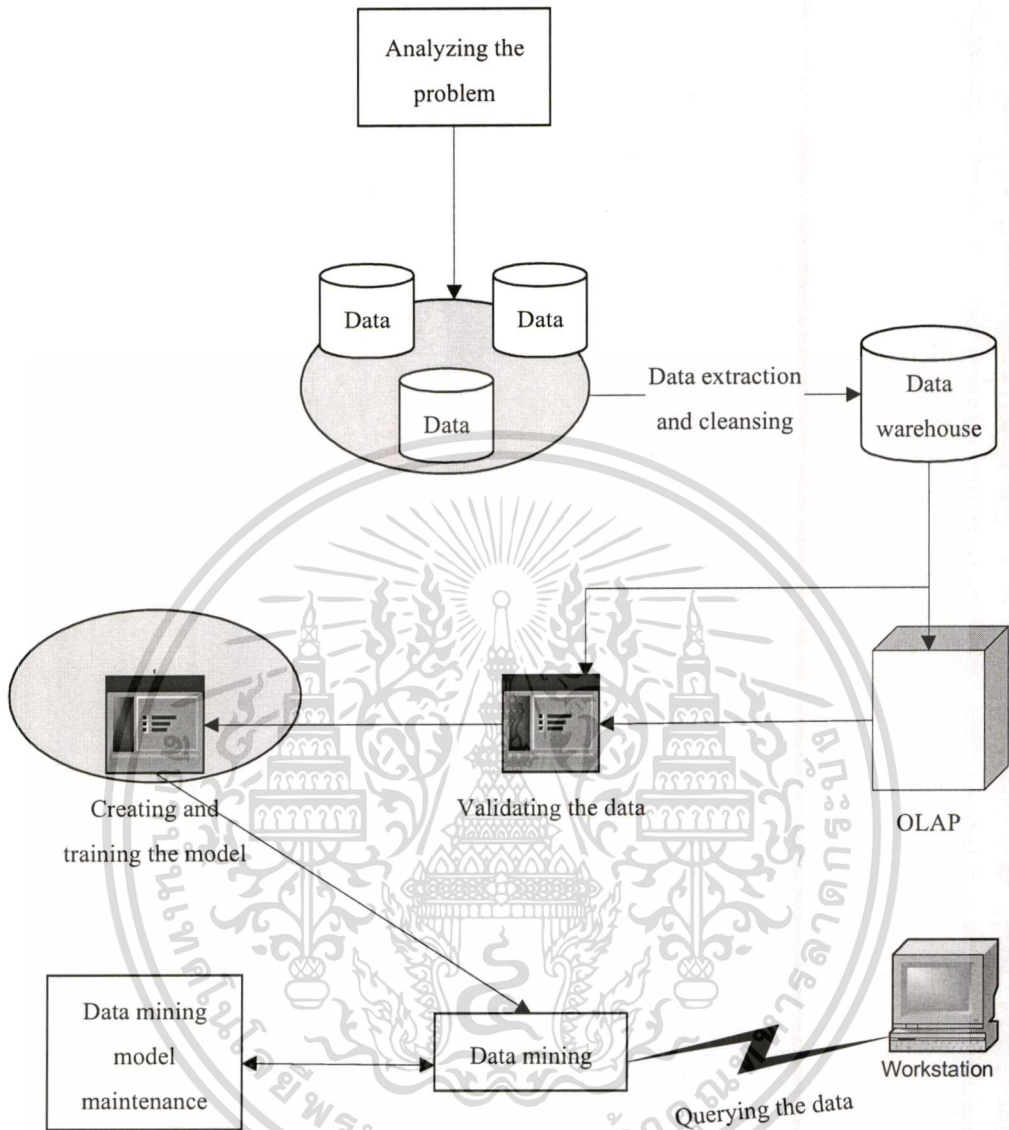
Mining Model

2.2.3 วิธีการในการทำค้ำไ่มนึ่ง

องค์ประกอบในแผนงานค้ำไ่มนึ่งแสดงดังรูปที่ 2.3 โดยมีรายละเอียดดังนี้ (Seidman, Claude. 2001.)

1. การวิเคราะห์ปัญหา (Analyzing the Problem)

ฐานข้อมูลที่จะถูกนำมาใช้จะต้องอยู่ภายใต้เงื่อนไขของการทำค้ำไ่มนึ่ง โดยมีคุณภาพและปริมาณของข้อมูลเป็นปัจจัยสำคัญ การคาดหวังผลลัพธ์จากการทำค้ำไ่มนึ่งก็จะต้องมีความเข้าใจด้วยว่าข้อมูลที่มีอยู่นั้นมีรูปแบบที่ถูกต้องของการที่จะเอาสารสนเทศออกมา



รูปที่ 2.3 วิธีการในการทำดาต้าไมนนิ่ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. การดึงข้อมูลและทำความสะอาดข้อมูล (Extracting and Cleansing the Data)

ข้อมูลที่จะถูกเอาออกมาจากแหล่งข้อมูลต้นกำเนิดนั้นอาจจะเป็นฐานข้อมูล OLTP เท็กซ์ไฟล์ ฐานข้อมูลเอ็กเซล (Microsoft Access) และสเปรดชีตต่างๆ ข้อมูลเหล่านี้จะถูกวางไว้ในดาต้าแวร์เฮาส์ที่มีโครงสร้างที่เหมาะสมกับดาต้าโมเดลและมีการใช้ดาต้าทรานสฟอร์มเมชันเซอร์วิส (Data Transformation Service) ที่ใช้ในการเอาข้อมูลออกมาและจัดข้อมูลที่ไม่ถูกต้องหรือไม่เหมาะสมออกไป

3. การตรวจสอบข้อมูล (Validating the Data)

หลังจากที่มีเอาข้อมูลออกมาและจัดข้อมูลที่ไม่ถูกต้องออกไปนั้น จะต้องมาตรวจสอบเพื่อให้แน่ใจว่า ข้อมูลทั้งหมดมีความถูกต้องและสมบูรณ์ ก่อนที่จะนำไปสร้างโมเดลต่อไป

4. การสร้างและฝึกสอนแบบจำลอง (Creating and Training the Model)

เมื่อมีการนำอัลกอริทึมมาใช้กับโมเดล ก็จะมีโครงสร้างต่างๆเกิดขึ้นมาและเป็นสิ่งสำคัญที่จะต้องดูว่า ข้อมูลที่เกิดขึ้นมานั้นมีความถูกต้องและเหมือนกับแหล่งข้อมูลเดิม

5. การสอบถามข้อมูลในแบบจำลองดาต้าไมนิ่ง (Querying the Data Mining Model Data)

เมื่อมีโมเดลและข้อมูลที่พร้อมในการสนับสนุนการตัดสินใจแล้ว (Decision support) ก็จะต้องมีการพัฒนาแอปพลิเคชันเพื่อดึงข้อมูลมาแสดงผล

6. การบำรุงรักษาความถูกต้องของแบบจำลองดาต้าไมนิ่ง (Maintaining the Validity of the Data-Mining Model)

เมื่อโมเดลมีการถูกใช้งานไปสักระยะหนึ่งคุณสมบัติของโมเดล เช่นความถูกต้องของข้อมูลอาจจะเปลี่ยนไป จึงอาจจะต้องมีการปรับเปลี่ยนคุณสมบัติของโมเดลใหม่เพื่อให้การทำงานมีความถูกต้อง

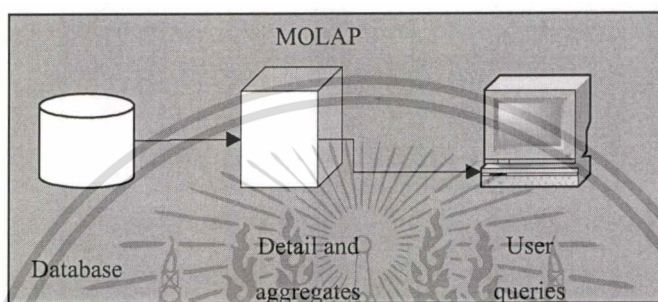
2.24 ดาต้าไมนิ่งในวิธีการของไมโครซอฟท์

ไมโครซอฟท์จะมี Microsoft SQL Server Analysis Services ในการให้บริการด้านการทำดาต้าไมนิ่งรวมทั้งดาต้าแวร์เฮาส์ โดยก่อนอื่นต้องมารู้จักกับเทคโนโลยี Online Analytical Processing (OLAP) ซึ่งจะเป็นแหล่งข้อมูลในการทำดาต้าไมนิ่ง โดยหน้าที่ของ OLAP นี้ใช้ในการวิเคราะห์ฐานข้อมูลขนาดใหญ่ ซึ่งฐานข้อมูลที่ใช้ในระบบ OLAP จะเก็บข้อมูลสถิติและ/หรือข้อมูลแบบสรุปรวมและอยู่ในรูปแบบมัลติไดเมนชัน (Multidimensional) ที่ถูกเรียกว่า คิวบ์ (Cube) ที่ประกอบด้วย ไดเมนชัน (Dimension) ตัวอย่างเช่น เขตการขาย และ เมเชอร์ (Measures) ตัวอย่าง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

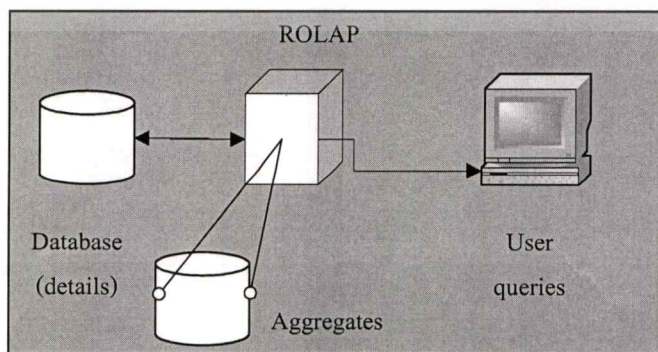
เช่น ยอดการขาย OLAP จะมีการจัดเก็บข้อมูลเพื่อสร้างคิวบ์ได้หลายแบบดังต่อไปนี้ (Seidman, Claude. 2001.)

1. **MOLAP** (Multidimensional Online Analytical Processing) ใช้สร้างคิวบ์จากข้อมูลที่เก็บอยู่ในดาต้าแวร์เฮาส์ วิธีนี้เหมาะกับข้อมูลขนาดใหญ่และการดึงข้อมูลจะมีประสิทธิภาพ โครงสร้างของ MOLAP แสดงดังรูปที่ 2.4



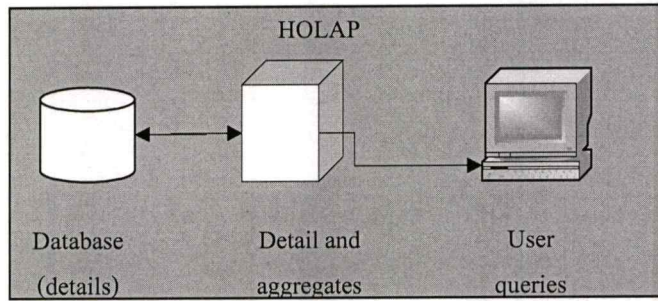
รูปที่ 2.4 โครงสร้างของ MOLAP

2. **ROLAP** (Relational Online Analytical Processing) จะเก็บข้อมูลในฐานะข้อมูลเดิมที่เป็น RDBMS แต่การดึงข้อมูลจะไม่มีประสิทธิภาพ โครงสร้างของ ROLAP แสดงดังรูปที่ 2.5
3. **HOLAP** (Hybrid Online Analytical Processing) รวมข้อดีของ MOLAP และ ROLAP ด้วยการเก็บผลรวมในระดับบนไว้ที่ MOLAP และผลรวมในระดับล่างและข้อมูลไว้ที่ฐานข้อมูล ส่วนข้อจำกัดของ ROLAP นั้นเกี่ยวกับตาราง (Table) ที่มีความซับซ้อนได้ถูกแก้ไขใน HOLAP โครงสร้างของ HOLAP แสดงดังรูปที่ 2.6



รูปที่ 2.5 โครงสร้างของ ROLAP

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.6 โครงสร้างของ HOLAP

2.2.5 การสร้างดาต้าแวร์เฮาส์สำหรับดาต้าไมน์นิ่ง (Seidman, Claude. 2001.)

2.2.5.1 รูปแบบในการจัดเก็บข้อมูล ดาต้าไมน์นิ่งจะใช้ดาต้าแวร์เฮาส์เป็นแหล่งข้อมูลเนื่องจากดาต้าแวร์เฮาส์จะรวบรวมข้อมูลแบบสรุปจากฐานข้อมูล OLTP มาเก็บไว้เป็นศูนย์กลางซึ่งมีผลทำให้ข้อมูลมีคุณภาพและความถูกต้อง

2.2.5.2 บुरณาภาพของข้อมูล (Data Integrity) โครงสร้างการจัดเก็บข้อมูลในฐานข้อมูลแบบ RDBMS นั้นจะมีการรักษาความบรณาภาพของข้อมูล ซึ่งทำให้ต้องมีการนอร์มอลไลซ์จนได้ตารางหลายๆตารางในฐานข้อมูล และต่อมาเมื่อมีการคิวรีข้อมูล ก็อาจจะทำให้คำสั่ง SQL มีความซับซ้อนมากตามจำนวนของตารางต่างๆ ที่มาเป็นเงื่อนไขในการดูข้อมูล แต่ในส่วนของการทำดาต้าแวร์เฮาส์นั้น เราจะไม่สนใจการรักษาบรณาภาพของข้อมูล คือจะมีการทำดินอร์มอลไลซ์กับโครงสร้างของดาต้าแวร์เฮาส์ เหตุผลในการทำแบบนี้ก็คือ ดาต้าแวร์เฮาส์เป็นฐานข้อมูลขนาดใหญ่ มีข้อมูลเป็นจำนวนมาก ซึ่งจะทำให้การคิวรีข้อมูลเป็นไปอย่างง่ายและรวดเร็ว แต่จะเป็นการเพิ่มความซ้ำซ้อนของข้อมูล (Data redundancy) และต้องการเนื้อที่ในดิสก์เพิ่มขึ้น รวมทั้งมีผลต่อประสิทธิภาพในการทำงานด้วย

2.2.5.3 สถาปัตยกรรมของดาต้าแวร์เฮาส์สำหรับการทำดาต้าไมน์นิ่ง ขั้นตอนในการเอาข้อมูลจากฐานข้อมูล OLTP มาทำดาต้าแวร์เฮาส์นั้นค่อนข้างซับซ้อนใช้เวลามาก และคุณภาพรวมทั้งประโยชน์ของข้อมูลที่น่ามาจะมีผลต่อเนื่องไปยังผลลัพธ์ของการทำนายในกระบวนการทำดาต้าไมน์นิ่ง ต่อไปนี้เป็นสิ่งที่ต้องระวังเมื่อมีการสร้างดาต้าแวร์เฮาส์ขึ้นมา

- โปรแกรมที่มีการเชื่อมต่อเพื่อดึงข้อมูลนั้นจะต้องไม่ผูกติดกับชื่อไฟล์หรือตาราง
- ความยาวของฟิลด์เดียวกันแต่อยู่ในแต่ละตารางจะต้องไม่แตกต่างกัน
- การจัดเก็บข้อมูลในฟิลด์จะต้องไม่ผิดพลาด เช่น ถ้าฟิลด์ที่เป็นตัวเลขเมื่อไม่มีค่าจะต้องเป็นศูนย์แทนที่จะเป็นค่า Null ซึ่งมีผลในการทำสรุปรวมข้อมูล
- ค่าที่จัดเก็บในฟิลด์เดียวกันแต่คนละตารางจะต้องสอดคล้องตรงกัน (Consistent values)

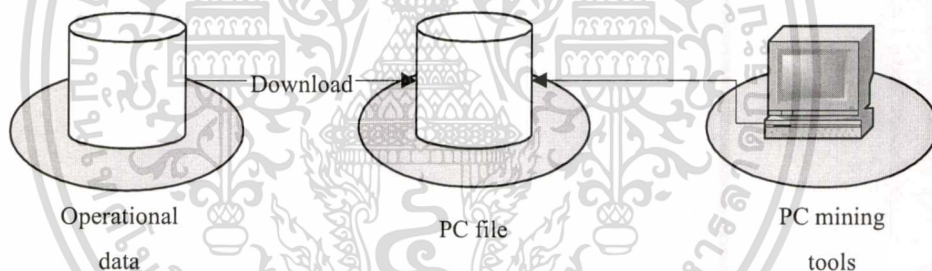
- ชนิดข้อมูลของฟิลด์เดียวกันแต่คนละตารางจะต้องสอดคล้องตรงกัน (Consistent values)

ข้อมูลที่ไม่ควรนำมาทำดาต้าแวร์เฮาส์ ข้อมูลเหล่านี้จะทำให้การทำงานผิดพลาดได้ ได้แก่

- ข้อมูลทดสอบ
- ข้อมูลที่มาจากเหตุการณ์ผิดปกติ เช่น ข้อมูลที่เพิ่มขึ้นจากยอดขายของอุปโภคบริโภค เมื่อเกิดภาวะเหตุการณ์แผ่นดินไหว เป็นต้น
- ข้อมูลที่ใช้ในการแก้ไขเพื่อให้ค่ารวมถูกต้อง เช่น ข้อมูลทางบัญชีที่ต้องใส่ค่าลบไป เพื่อให้ยอดมีความถูกต้องตรงกับเอกสารจริง เป็นต้น

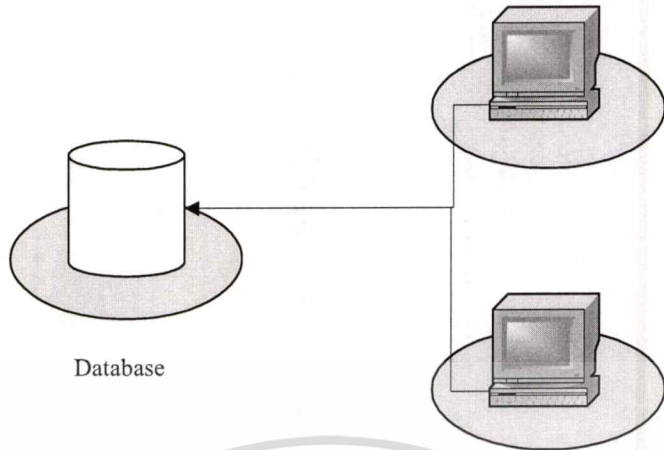
2.2.5.4 โครงสร้างดาต้าไมน์นึ่งทางกายภาพ

- สถาปัตยกรรมแบบ Single-Tier เหมาะสำหรับข้อมูลมีขนาดเล็ก และขบวนการในการดาวน์โหลดข้อมูลจากดาต้าแวร์เฮาส์มายังเครื่องคอมพิวเตอร์ที่ทำหน้าที่ในการประมวลผลดาต้าไมน์นึ่ง (PC mining) ไม่บ่อยครั้งนัก ดังแสดงรูปที่ 2.7



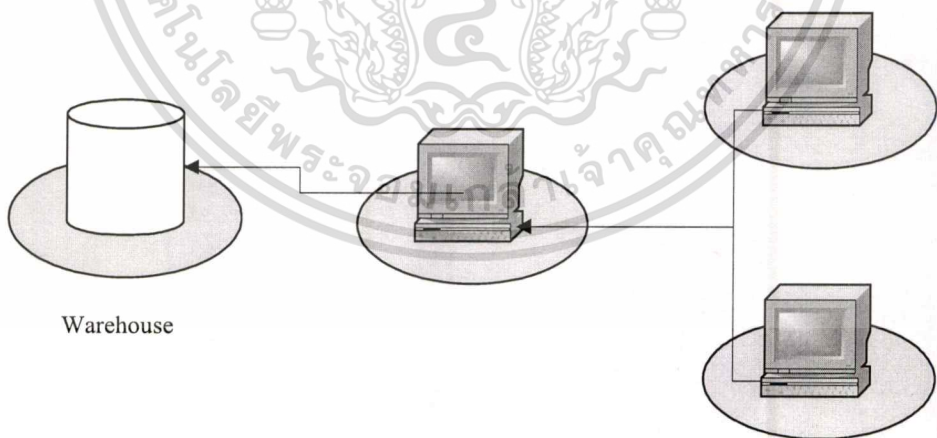
รูปที่ 2.7 สถาปัตยกรรมแบบ Single-Tier

- สถาปัตยกรรมแบบ Two-Tier เหมาะกับข้อมูลจำนวนมาก ที่ต้องการเซิร์ฟเวอร์ที่มีประสิทธิภาพสูง รองรับการทำงานของดาต้าแวร์เฮาส์และดาต้าไมน์นึ่ง โดยไคลเอนต์ทำการติดต่อเรียกใช้และรอผลลัพธ์จากเซิร์ฟเวอร์ดังแสดงรูปที่ 2.8



รูปที่ 2.8 สถาปัตยกรรมแบบ Two-Tier

- สถาปัตยกรรมแบบ Three-Tier เหมาะกับระบบที่มีความต้องการในการทำงานที่มีประสิทธิภาพสูง โดยดาต้าแวร์เฮาส์จะอยู่บนเครื่องเซิร์ฟเวอร์เครื่องหนึ่งและเครื่องเซิร์ฟเวอร์อีกเครื่องหนึ่งทำหน้าที่ไมน์นิ่งข้อมูลและส่งผลลัพธ์ไปยังเครื่องคอมพิวเตอร์ที่เป็นไคลเอนต์ดังแสดงรูปที่ 2.9



รูปที่ 2.9 สถาปัตยกรรมแบบ Three-Tier

2.2.6 วิธีการของดาต้าไมน์นิ่งและสถิติ (Seidman, Claude. 2001.)

2.2.6.1 ดาต้าไมน์นิ่งและสถิติ ดาต้าไมน์นิ่งจะเกี่ยวข้องกับหลักสถิติ ดังนั้นควรทำความเข้าใจ

เอกสารนี้เป็นใจกับคำดังต่อไปนี้ ทรัพยากรใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- Population คือกลุ่มตัวอย่างที่ประกอบด้วย Case (Case จะรวมเป็น โมเดลใน คาด้าไมน์นึ่ง) ลักษณะของกลุ่มจะต้องถูกนิยามไว้อย่างถูกต้องและมีจำนวน ข้อมูลที่เพียงพอจะสามารถแสดงลักษณะของกลุ่มได้อย่างถูกต้อง เช่น พฤติ กรรมของกลุ่มลูกค้าที่ซื้อสินค้า เป็นต้น
- Sample กลุ่มข้อมูลที่มีขนาดเล็ก หรือเลือกสุ่มมาจากข้อมูลขนาดใหญ่
- Range ประกอบด้วยช่วงข้อมูลจากกลุ่มข้อมูล
- Bias เมื่อมีการนำข้อมูลตัวอย่างมาวิเคราะห์ จะต้องมีการนำความเอนเอียงมา ใช้เพื่อป้องกันการที่ข้อมูลไม่ได้แสดงบางอย่างออกมา ซึ่งอาจทำให้วิเคราะห์ ผิดไป
- Mean คือค่าเฉลี่ยแต่จะทำให้มีการค่าเฉลี่ยของกลุ่มข้อมูลผิดไปถ้าข้อมูลใดข้อมูล หนึ่งมีค่าที่แตกต่างจากข้อมูลอื่นๆ อยู่มาก
- Median คือค่าที่อยู่ตรงกลางในกลุ่มข้อมูล
- Distribution ตำแหน่ง การจัดและความถี่ของความแตกต่างในกลุ่มข้อมูล
- Mode คือค่าที่เกิดขึ้นบ่อยๆ ใน Distribution
- Variance ผลรวมของความแตกต่างระหว่างทุกค่าในกลุ่มข้อมูลกับค่าเฉลี่ย (Mean) ใช้ในการดูความผันแปรจากค่าที่คาดไว้
- Standard Deviation เป็นสแควร์รูท (Square root) ของ Variance ใช้ในการ คำนวณหาการกระจายระหว่างค่าที่กำหนดและค่าเฉลี่ย (Mean)
- Correlation คือการที่ค่าข้อมูลหนึ่งเปลี่ยนไปและมีผลต่อค่าข้อมูลอีกค่าหนึ่ง
- Regression หลังจากที่ทำค่า Correlation แล้วจะมีการหารูปแบบของการ เปลี่ยนค่าระหว่างตัวแปรหนึ่งกับตัวแปรอื่นๆ เพื่อให้การทำนายค่ามีความถูกต้อง

2.2.6.2 การเรียนรู้จากประวัติข้อมูล คาด้าไมน์นึ่งจะทำนายค่าในอนาคตโดยเรียนรู้จากเหตุการณ์ในอดีต มีวิธีการ 6 วิธีที่คาด้าไมน์นึ่งใช้เรียนรู้เพื่อการทำนายค่าข้อมูล

- การวิเคราะห์อิทธิพล (Influence Analysis) จะกำหนดปัจจัยและตัวแปรที่มีผลกระทบต่อการวัดค่าต่างๆ
- การวิเคราะห์ความผันแปร (Variation Analysis) จะดูความผันแปรของกลุ่มข้อมูลและจะแยกปัจจัยที่อาจมีอิทธิพลต่อการวัดค่าข้อมูล
- การวิเคราะห์การเปรียบเทียบ (Comparison Analysis) ใช้ตัวอย่างก่อนหน้านี้ ในการเปรียบเทียบกับตัวอย่างปัจจุบัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- การวิเคราะห์สาเหตุและผลกระทบ (Cause and Effect Analysis) จะกำหนดผลกระทบท่อเหตุการณ์ต่างๆ เพื่อให้คำตอบใดหนึ่งสามารถรองรับการทำงานได้อย่างถูกต้อง
- การวิเคราะห์แนวทาง (Trend Analysis) ดูการเปลี่ยนแปลงของค่าข้อมูลในช่วงเวลาที่กำหนด
- การวิเคราะห์การบิดเบือน (Deviation Analysis) กำหนดข้อมูลที่ตกอยู่นอกขอบเขตของค่าที่คาดหวังไว้ การตรวจสอบการถือโง่งก็ใช้การวิเคราะห์การบิดเบือนเพื่อกำหนดจุดที่ข้อมูลมีความแตกต่างจากค่าปกติ

2.2.6.3 การทำนายค่าในอนาคต สามารถทำนายอนาคตโดยการใส่ค่าใดหนึ่งเพื่อเปรียบเทียบกลุ่มของอินพุตที่มีรูปแบบอยู่แล้วในระบบฐานข้อมูล โดยจะคาดเดาค่าสำหรับกลุ่มข้อมูลใหม่ โดยมีวิธีการดังต่อไปนี้

- การหาความน่าจะเป็น เช่น โมเดลที่มาจากฐานข้อมูลผู้แทนจำหน่ายรถยนต์จะแสดงว่าคนที่ซื้อสปอร์ต 2 ที่นั่งนั้นมักจะเป็นผู้ชายที่อายุต่ำกว่า 40 ปี เป็นต้น
- ซิมูเลชัน (Simulations) และ What-if การทำนายบนการซิมูเลชันกลุ่มข้อมูลนั้นเป็นวิธีที่ดีที่สุดในการตรวจสอบโมเดล เป็นการทดสอบใส่ปัจจัยต่างๆที่เป็นอินพุตของโมเดลเพื่อดูผลลัพธ์ที่เกิดขึ้น ส่วน What-if นั้นเป็นการกำหนดอินพุตที่ตายตัวลงไป

2.2.6.4 การเทรนดาต้าโมเดล ขบวนการในการวิเคราะห์ Case ที่เป็นอินพุตเพื่อให้ได้มาของรูปแบบของข้อมูลและเป็นส่วนหนึ่งในขบวนการ “Machine learning”

- การประเมิน โมเดลและหลีกเลี่ยงข้อผิดพลาดต่างๆ
- **Overfitting** คือการที่รูปแบบข้อมูลมีความผิดพลาดเนื่องจากกลุ่มข้อมูลมีขนาดเล็กเกินไป ทำให้ไม่ตรงกับความเป็นจริง ซึ่งการแก้ปัญหา Overfitting มีดังต่อไปนี้
 - (1) Pruning วิธีนี้ใช้อัลกอริทึม Decision trees ในการแก้ปัญหา Overfitting ด้วยการประเมินผลข้อมูลจากคุณลักษณะ
 - (2) การวิเคราะห์ Chi-Squared ใช้การกำหนดความเกี่ยวพันของคุณลักษณะ โดยมีสูตรสำหรับการวิเคราะห์ดังต่อไปนี้

$$c2 = \text{Sum}[(\text{Expected Value} - \text{Obtained Value})^2] / \text{Expected Value}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(3) Cross Validation ใช้เทคนิคที่ทดสอบกลุ่มย่อยของข้อมูลและดูการทำนายว่าถูกต้องหรือไม่

- **Underfitting** คือการที่ค่าค่าไม้นิ่งไม่สามารถกำหนดคุณลักษณะที่เพียงพอในการค้นหารูปแบบของข้อมูลที่สำคัญได้

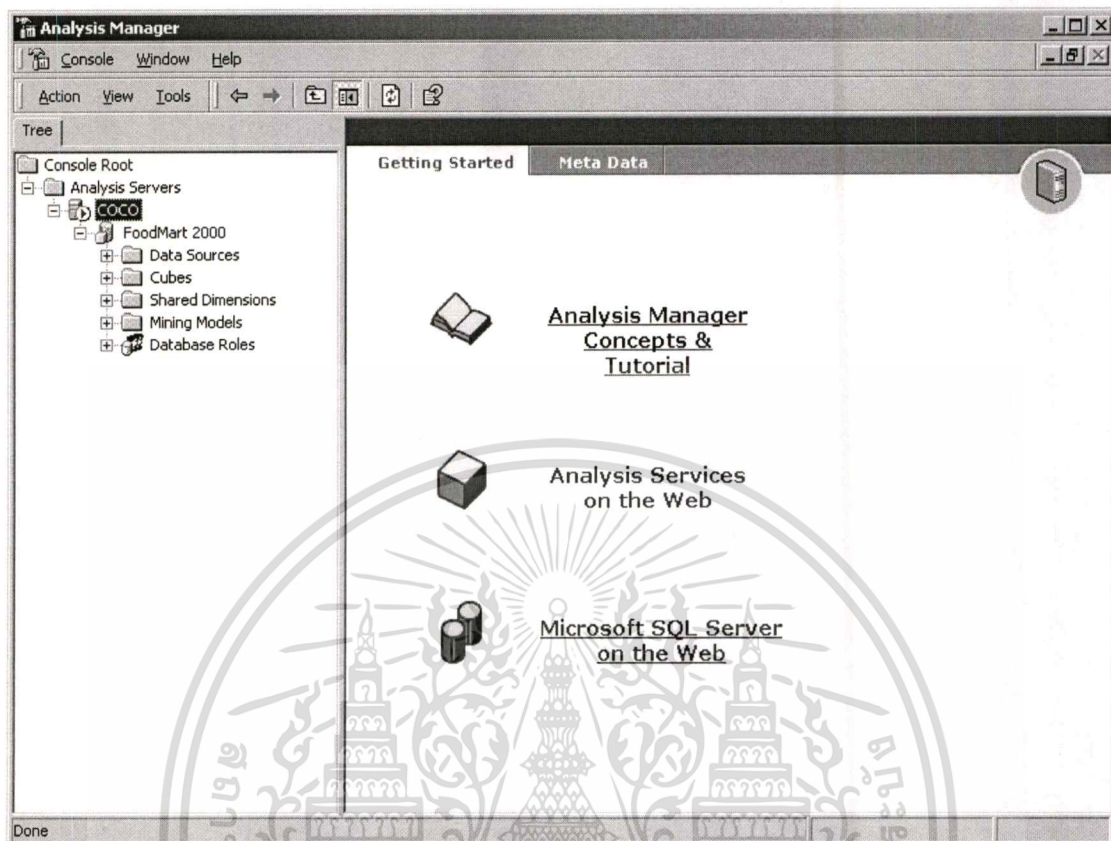
2.2.7 วิธีการทำค่าค่าไม้นิ่งของไมโครซอฟท์ที่ใช้ Decision Trees

Decision trees เป็นอัลกอริทึมที่นิยมใช้ในเครื่องมือที่ทำค่าค่าไม้นิ่ง โดยมีอัลกอริทึมที่อยู่ภายในได้แก่ Classification and Regression Trees (CART), Chi-Squared Automatic Interaction Detector (CHAID) และ C4.5 ซึ่งแต่ละอัลกอริทึมจะทำงานร่วมกันโดยมีผลลัพธ์เป็น โมเดลแบบต้นไม้ (Trees) โดยการทำนายข้อมูลจะเป็นลักษณะจัดแบ่งหรือแยกออกเป็นกลุ่มต่างๆ ตัวอย่างในการนำไปใช้งานจริงก็ได้แก่ ไคเร็คเมลล์ (Direct-mail) ในงานด้านการตลาด เป็นต้น

การสร้างโมเดล เป็นขั้นตอนแรกในการทำค่าค่าไม้นิ่ง โมเดลจะถูกสร้างจาก Case ที่นำมาจากแหล่งข้อมูลที่จะเป็นฐานข้อมูลเชิงสัมพันธ์หรือคิวบ์ จุดเริ่มต้นของการสร้างค่าค่าไม้นิ่งจะใช้ Analysis Manager ที่มากับ Microsoft SQL Server 2000 ดังแสดงในรูปที่ 2.10

ซึ่งจะมีส่วนประกอบต่างๆ ใน Analysis Manager ดังนี้ (Seidman, Claude. 2001.)

- **Databases** แสดงฐานข้อมูลที่อยู่ใน Analysis Manager ในที่นี้คือ FoodMart 2000
- **Data Sources** ประกอบด้วยแหล่งข้อมูลที่ถูกใช้เป็นฐานข้อมูลในค่าค่าไม้นิ่ง
- **Cubes** ประกอบด้วยคิวบ์ที่อยู่ในฐานข้อมูล โดยมี 3 แบบคือ Regular, Linked และ Virtual
- **Partitions** ประกอบด้วยพาร์ติชันที่อยู่ในคิวบ์ โดยมี 2 แบบคือ Local และ Remote
- **Cube Roles** ประกอบด้วยโรลที่อยู่ในคิวบ์



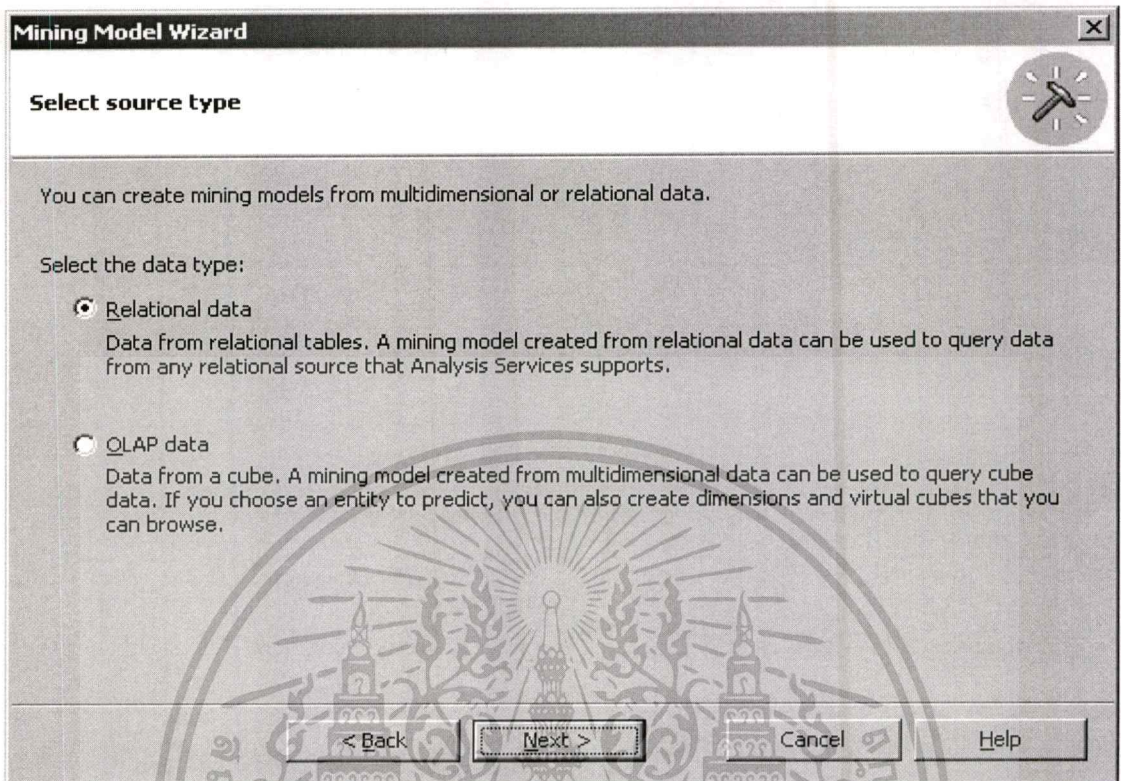
รูปที่ 2.10 Microsoft Analysis Manager

- **Shared Dimensions** ประกอบด้วยเซิร์ฟเวอร์โดเมนชั้นในฐานข้อมูล ซึ่งเซิร์ฟเวอร์โดเมนชั้นเหล่านี้สามารถถูกนำไปใช้กับคิวบ์ใดๆก็ได้ในฐานข้อมูล โดยมี 4 แบบคือ Regular, Virtual, Parent-Child และ Data-Mining
- **Mining Models** ประกอบด้วยโมเดลหนึ่งโมเดลที่อยู่ในฐานข้อมูล โดยมี 2 แบบคือแบบที่แหล่งข้อมูลมาจาก OLAP และแบบที่แหล่งข้อมูลมาจากฐานข้อมูลเชิงสัมพันธ์
- **Database Roles** แสดงโหนดทั้งหมดในฐานข้อมูล ซึ่งสามารถกำหนดให้กับคิวบ์หรือโมเดลใดๆก็ได้

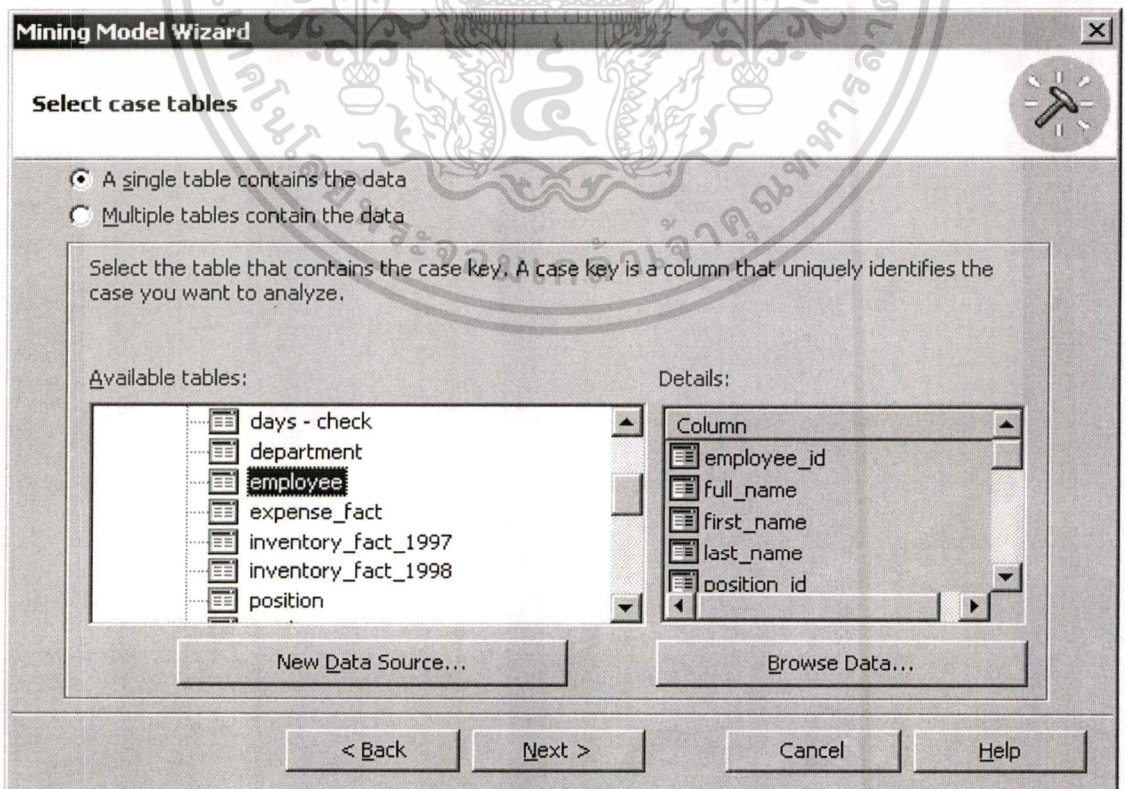
การสร้างโมเดลหนึ่งโมเดลด้วยการใช้คิวบ์จะมีขั้นตอนดังนี้

1. เลือกแหล่งข้อมูลในการทำคิวบ์โมเดลหนึ่ง ซึ่งอาจจะเป็นฐานข้อมูลเชิงสัมพันธ์หรือ OLAP ดังแสดงในรูปที่ 2.11 แต่ในที่นี้ให้เลือกเป็นฐานข้อมูลเชิงสัมพันธ์
2. เลือก Case หรือตารางสำหรับคิวบ์โมเดลหนึ่ง โมเดล ดังแสดงในรูปที่ 2.12

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้拿去ใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



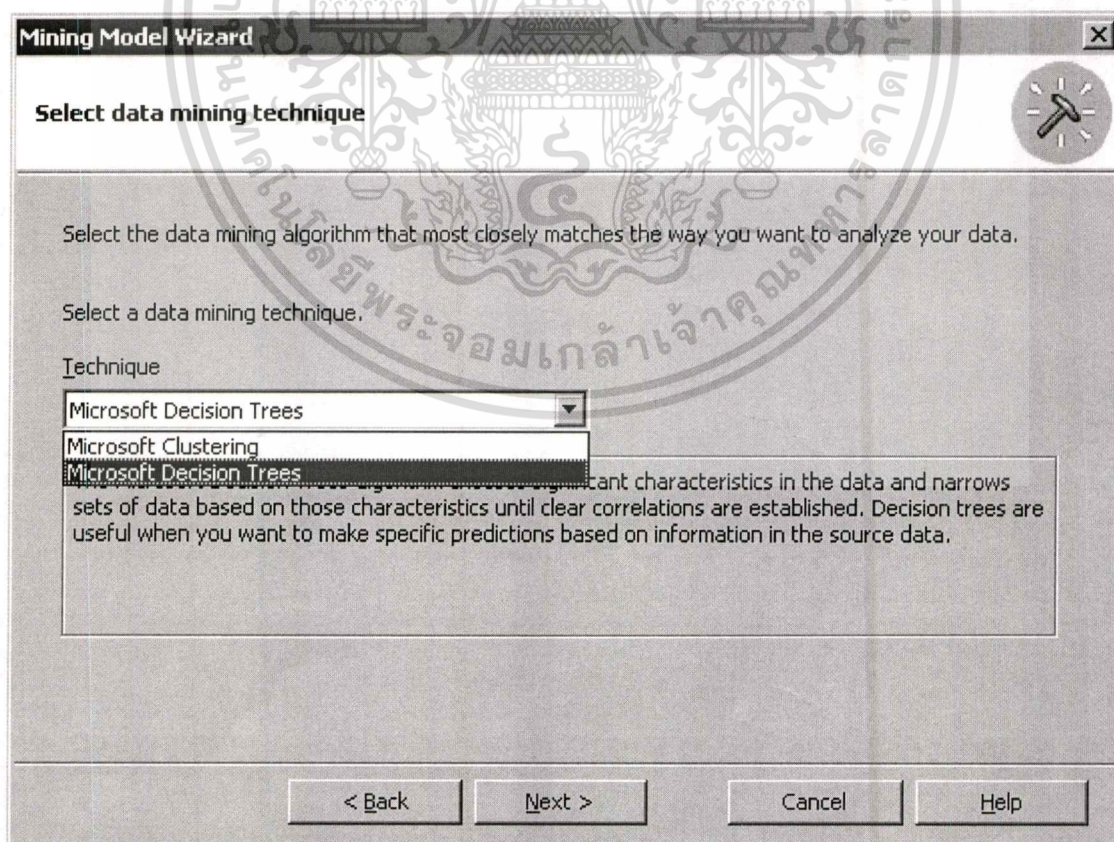
รูปที่ 2.11 การเลือกชนิดแหล่งข้อมูล



รูปที่ 2.12 การเลือกตารางข้อมูล

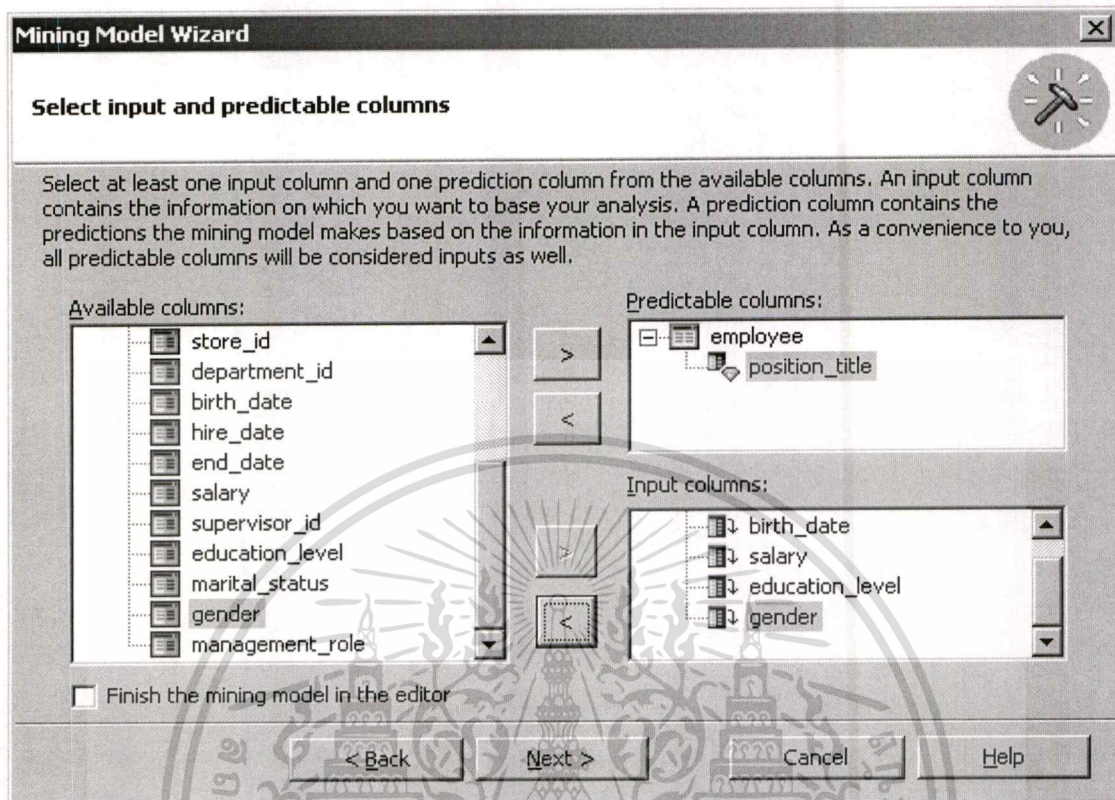
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. เลือกเทคนิคดาต้าไมน์นิ่ง (อัลกอริทึม) โดยวิศวาร์ดจะให้มา 2 วิธีคือ Microsoft Decision Trees และ Microsoft Clustering ดังแสดงในรูปที่ 2.13 ในที่นี้เลือกวิธี Microsoft Decision Trees
4. ทำการเชื่อมความสัมพันธ์ในกรณีที่ใช้หลายตารางจะเป็นการสร้างความสัมพันธ์ระหว่างตารางแม่และตารางลูกด้วยการใช้คอลัมน์ที่เป็นคีย์
5. เลือกคอลัมน์ที่จะเป็น Case Key เป็นสิ่งสำคัญที่มีผลต่อเอาท์พุทเพราะ Case Key จะใช้ในการกำหนด Case ต่างๆ
6. เลือกคอลัมน์ที่จะเป็นอินพุตและ Prediction โดยคอลัมน์อินพุตจะใช้ในการทำเทรนดาต้าไมน์นิ่งโมเดลส่วนคอลัมน์ Prediction ใช้ในการเป็นเอาท์พุทจากการทำนายจากดาต้าไมน์นิ่งโมเดลดังแสดงในรูปที่ 2.14
7. หลังจากนั้นดาต้าไมน์นิ่งจะถูกกำหนดขึ้นมาดังแสดงในรูปที่ 2.15 Relational Mining Model Editor โดยแสดงในส่วน Schema และแสดงในรูปที่ 2.16 Relational Mining Model Editor โดยแสดงในส่วน Content ตามลำดับ

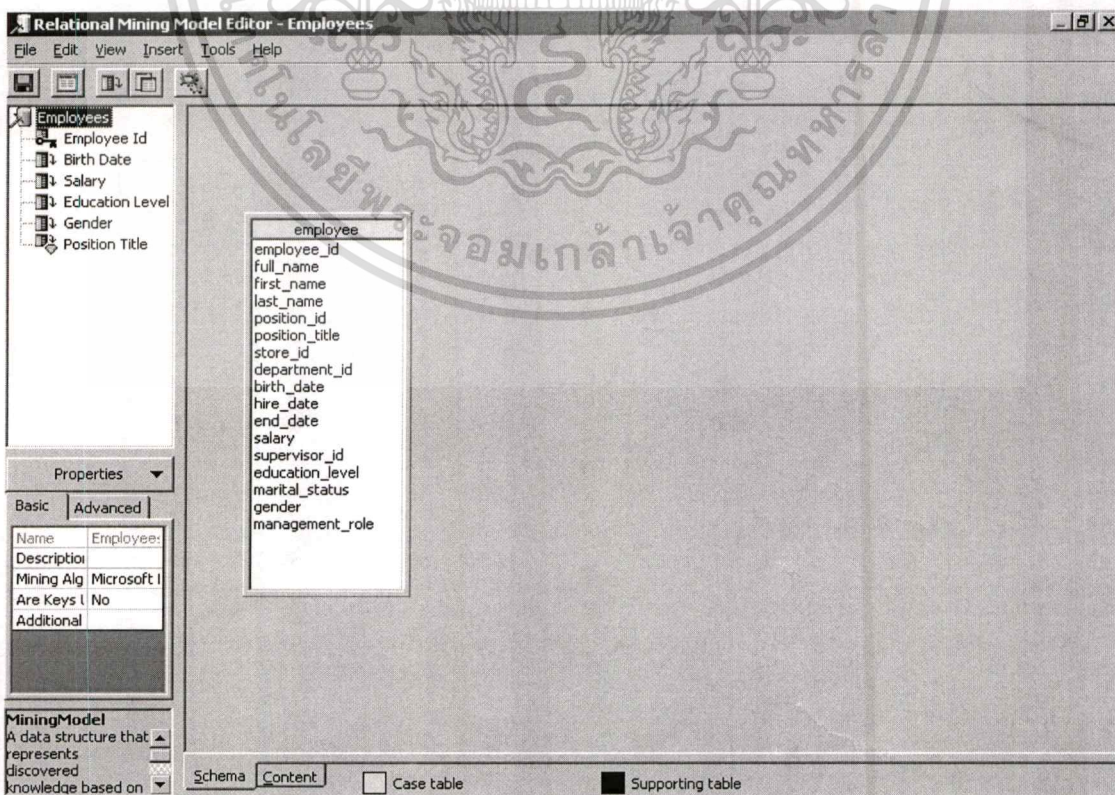


รูปที่ 2.13 การเลือกเทคนิคของดาต้าไมน์นิ่ง

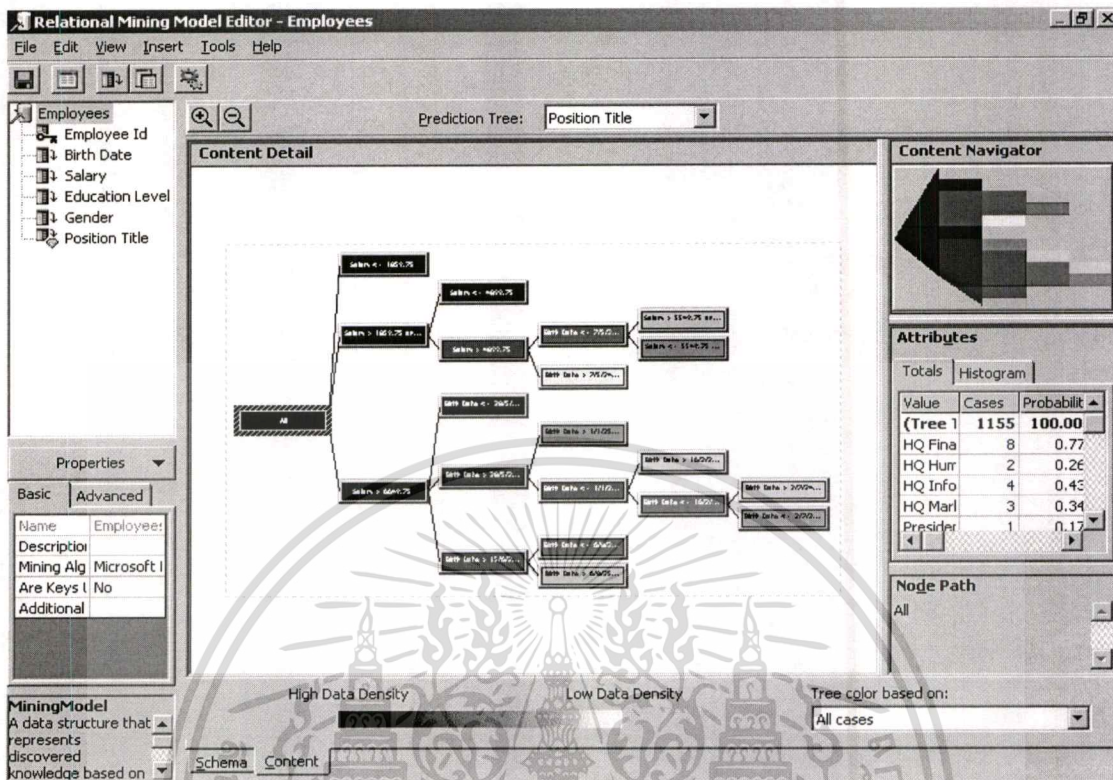
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.14 การเลือกคอลัมน์อินพุตและคอลัมน์ทำนายข้อมูล



เอกสารนี้เป็นเอกสารที่สงวนรูปที่ 2.15 Relational Mining Model Editor (Schema) ำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.16 Relational Mining Model Editor (Content)

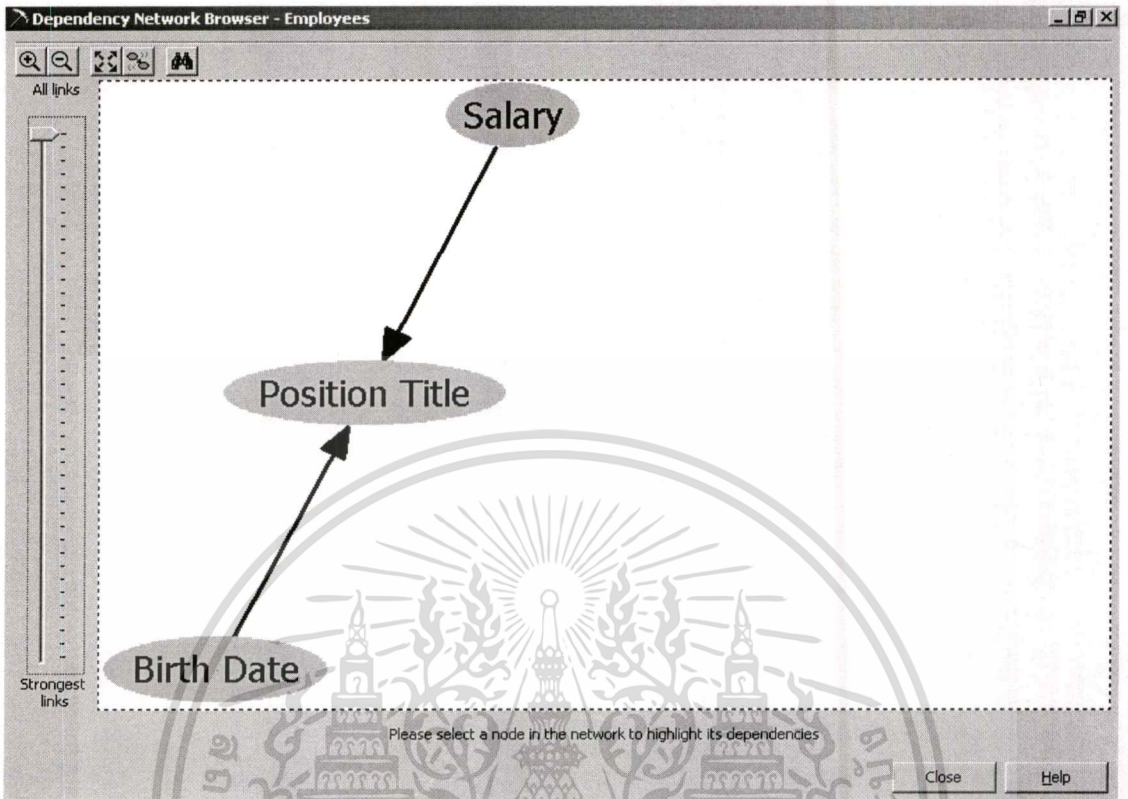
Data Mining Model Editor จะมีอยู่ 2 ส่วนที่สำคัญคือ Schema ที่ใช้สำหรับการเปลี่ยนแปลงโครงสร้างของโมเดลและ Content ใช้สำหรับแสดงข้อมูลที่ถูกจัดแบ่งภายใน Tree โดยสามารถเลือก Tree ได้ที่กรอบคานานิลิสต์บ็อก Prediction Tree และใช้กฎ IF->THEN ในการจัดแบ่งข้อมูล สีในแต่ละโหนดมีระดับความเข้มต่างกันตามความหนาแน่นของข้อมูลและเมื่อคลิกที่โหนดใดก็ตามที่ส่วนของ Attributes และ Node Path จะเปลี่ยนตาม โดย Node Path จะอธิบายกฎที่ใช้ในโหนดนั้นๆ

คอลัมน์ Prediction ข้อมูลจะถูกแบ่งเป็นกลุ่มๆ ที่สามารถใช้ในการทำนายได้ เพราะแต่ละโหนดหรือกลุ่มข้อมูลนั้นจะใช้พื้นฐานมาจากกฎของคุณลักษณะข้อมูลนั่นเอง

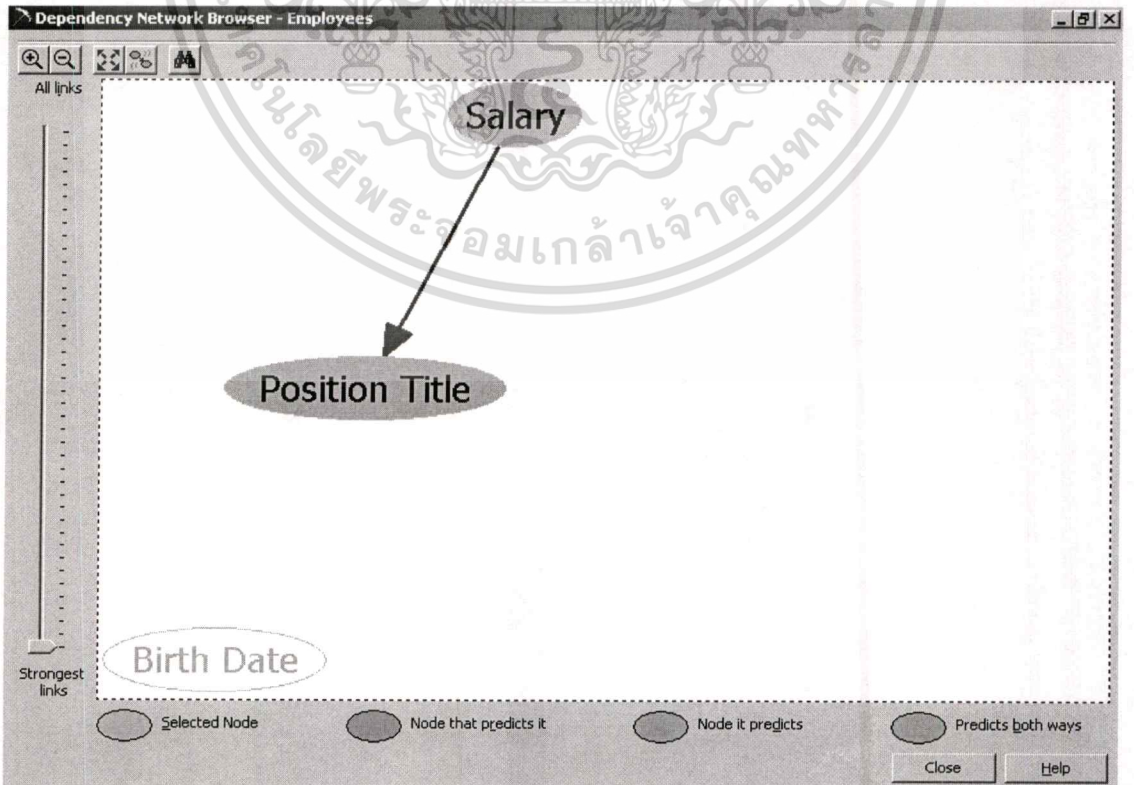
Dependency Network Browser เป็นเครื่องมือที่ใช้แสดงการขึ้นต่อกัน (Dependency) และความสัมพันธ์ระหว่างออบเจกต์ในดาต้าไมน์นิ่งโมเดล ดังแสดงรูปที่ 2.17

ใน Dependency Network Browser จะแสดงเครือข่ายของคุณลักษณะข้อมูล การขึ้นต่อกันจะแสดงด้วยลูกศรที่เชื่อมโยงถึงการทำนายข้อมูลระหว่างโหนดและถ้าเส้น สไลด์ “Strongest links” ก็จะเหลือโหนดที่มีความสัมพันธ์ต่อกันมากที่สุดนั่นเอง ดังแสดงในรูปที่ 2.18

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.17 Dependency Network Browser



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการ **รูปที่ 2.18** The links pointer มอนูญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่าการฉ้อโกงทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

อัลกอริทึมที่อยู่ภายใน Decision Tree ไมโครซอฟท์ได้ใช้อัลกอริทึมดังต่อไปนี้ใน Decision Tree (Seidman, Claude. 2001.)

- Classification an Regression Tress (CART) มีประสิทธิภาพมากสำหรับการจัดแบ่งข้อมูลโดยแยกออกเป็นสองโหนดหรือเป็นกิ่งสาขาของข้อมูลนั้นๆ จนได้เป็น Decision Tree
- Chi-Squared Automatic Interaction Detector (CHAID) ใช้ในการตรวจสอบ Tree
- C4.5 ใช้ในการเพิ่มประสิทธิภาพของ Decision Tree โดยก่อนหน้านั้นใช้ Iterative Dichotomizer version 3 (ID3) ที่ถูกพัฒนามา 20 กว่าปีแล้ว

เมื่อไหร่ควรจะใช้ Decision Trees (Seidman, Claude. 2001.)

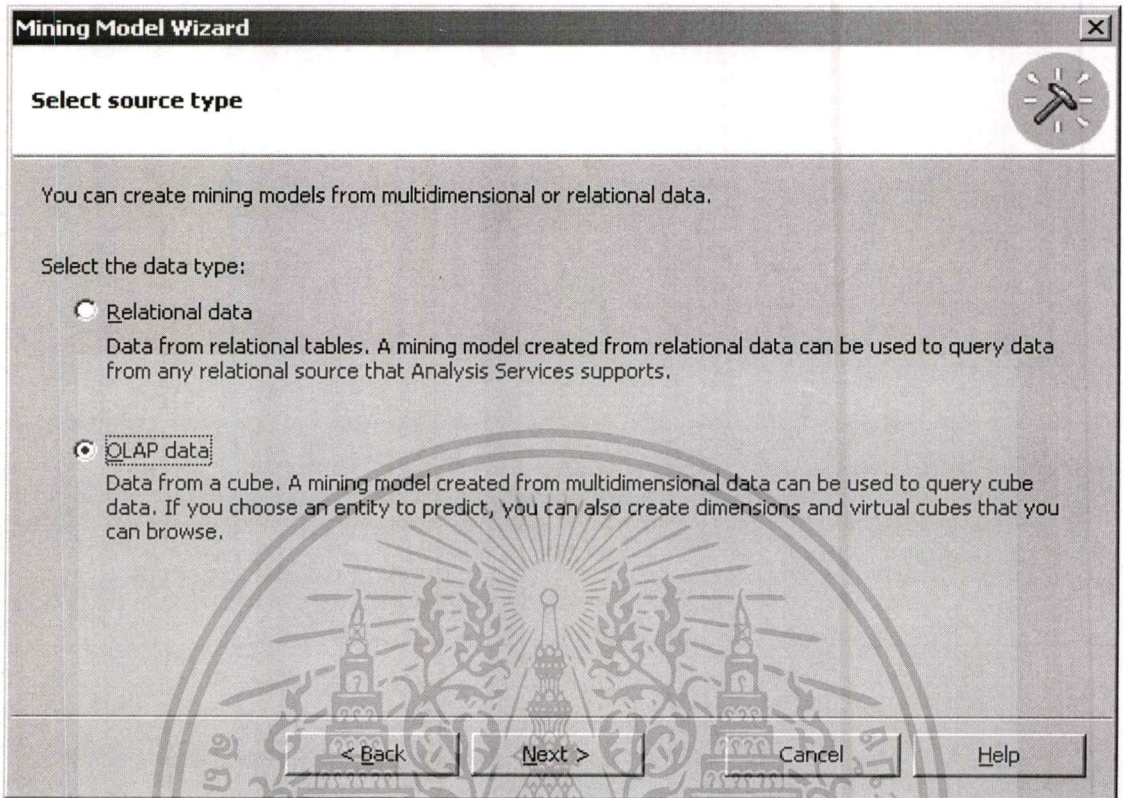
- เมื่อต้องการใช้โครงสร้างแบบจัดแบ่งกลุ่มข้อมูล
- เมื่อต้องการกำหนดความสัมพันธ์ที่เป็นไปได้ระหว่างตัวแปรที่เมื่อมีการเปลี่ยนแปลงจะส่งผลกระทบต่ออะไรบ้าง
- เมื่อต้องการนำเสนอภาพความสัมพันธ์ระหว่างตัวแปรที่ทำให้เข้าใจลักษณะของข้อมูลที่อยู่ในฐานข้อมูลนั้นๆ
- เมื่อต้องการประยุกต์ส่วนผสมที่หลากหลายของคุณลักษณะและประเภทต่างๆ เพื่อใช้ในการทำนายข้อมูล
- เมื่อต้องการสำรวจข้อมูลเพื่อกำหนดตัวแปรที่สำคัญในกลุ่มข้อมูลที่สามารถใช้เป็นเป้าหมายได้

2.2.8 การสร้าง Decision Trees ด้วย OLAP (Seidman, Claude. 2001.)

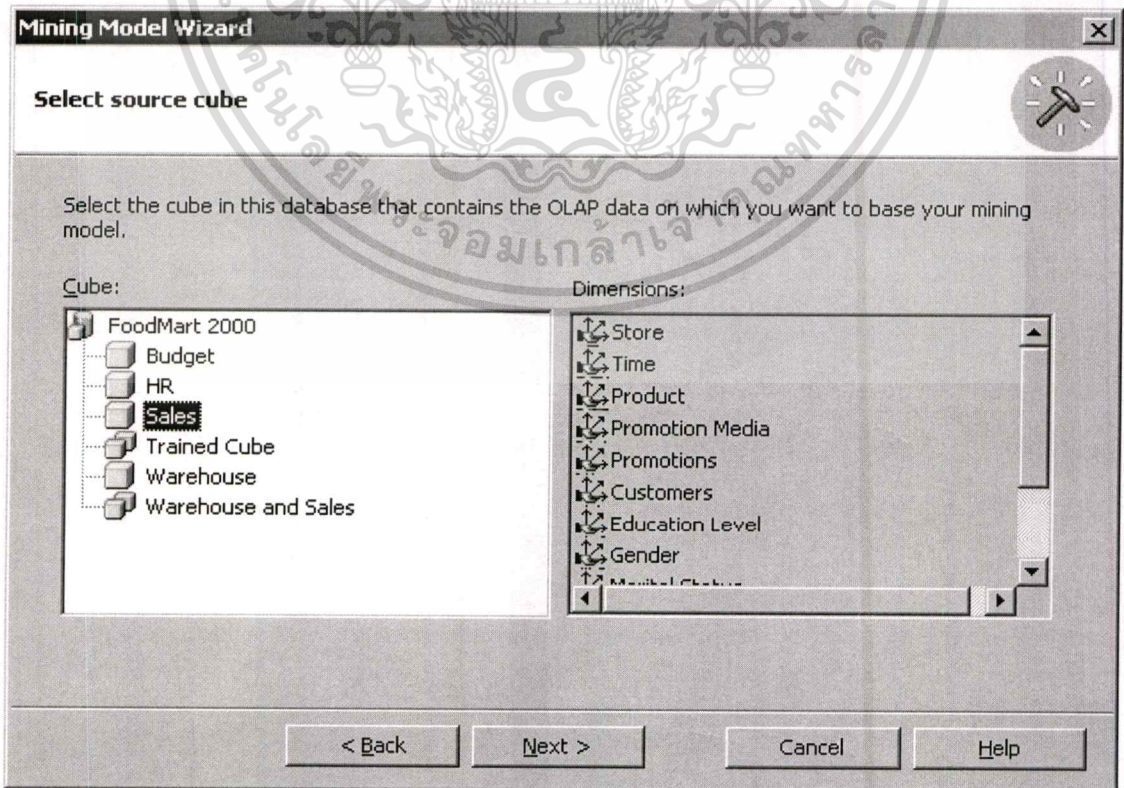
OLAP เป็นโครงสร้างรูปแบบที่ดีสำหรับการจัดเก็บข้อมูลแบบสรุปรวมคล้ายกับ GROUP BY ในคำสั่ง SQL นอกจากนั้นยังสะดวกเหมาะในการทำดาต้าไมนนิ่ง

2.2.8.1 การสร้างโมเดล

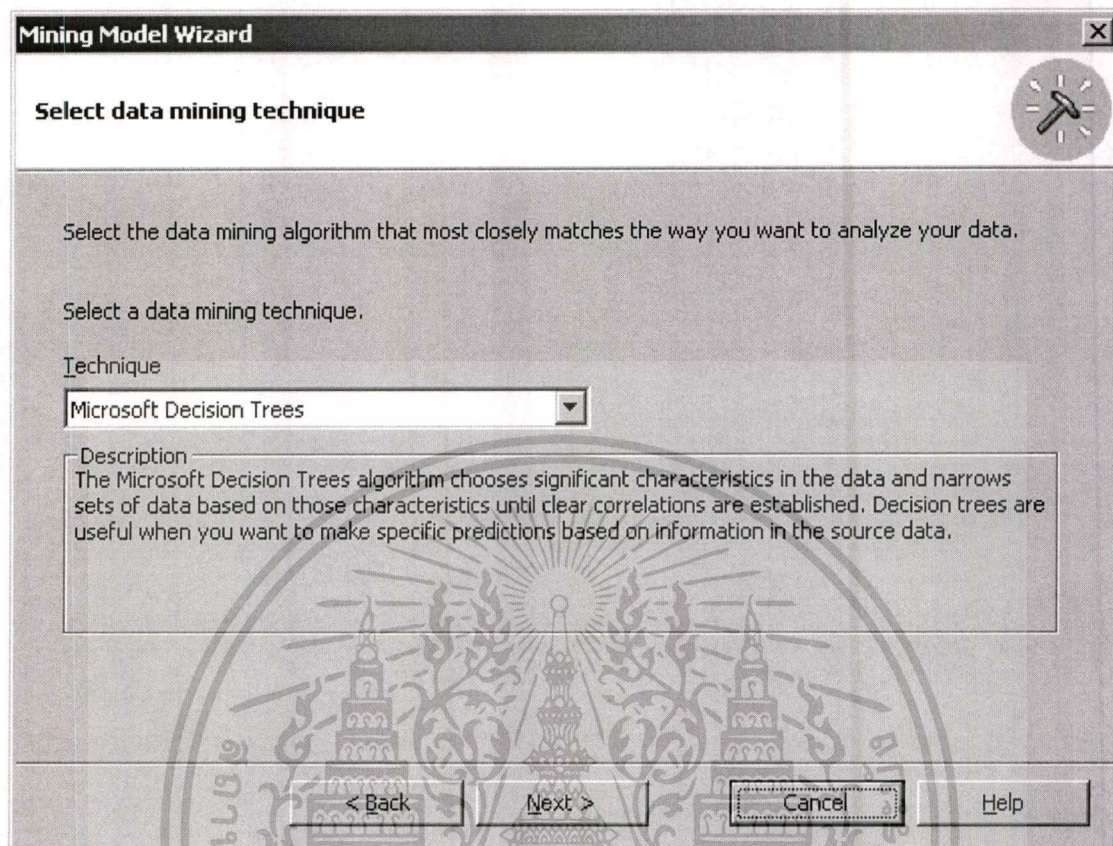
1. เลือกชนิดแหล่งข้อมูล ดังรูปที่ 2.19 ที่ให้มีการกำหนดชนิดแหล่งข้อมูลในที่นี้คือเลือกแบบ OLAP
2. เลือกแหล่งของคิวบ์และเทคนิคของดาต้าไมนนิ่ง ในรูปที่ 2.20 ให้เลือกคิวบ์ที่จะมี Case ที่จะต้องมีการเทรนโมเดล ต่อไปเลือกเทคนิคในการทำดาต้าไมนนิ่งดังแสดงในรูปที่ 2.21 ให้เลือก Microsoft Decision Trees



รูปที่ 2.19 การเลือกชนิดแหล่งข้อมูล



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับรูปที่ 2.20 การเลือกแหล่งข้อมูลโดยใช้คิวบ์ให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.21 Select data-mining technique

3. เลือก Case ในหน้าจอ “Select case” ให้เลือกโดเมนชั้น (Dimension) ที่เป็น Case ที่จะใช้ในการแทนค่าโมเดล ดังรูปที่ 2.22 นอกจากนั้นยังสามารถเลือกระดับ (Level) ได้อีกด้วย
4. เลือกเอนทิตีในการทำนาย ในหน้าจอ “Select predicted entity” ดังแสดงในรูป 2.23 โดยสามารถเลือกออกแบบชั้นได้ 3 แบบดังต่อไปนี้
 - Measure of the source cube ถ้าต้องการทำนายข้อมูลบนเมเชอร์ (Measures) หรือค่าที่เป็นตัวเลขในคิวบ์
 - Member property of the case level ทุกระดับของโดเมนชั้นใน OLAP จะประกอบด้วยเมมเบอร์พรอพเพอร์ตี้ (Member properties) ที่ใช้ช่วยอธิบายว่าในระดับนั้นๆประกอบด้วยเมมเบอร์พรอพเพอร์ตี้อะไรบ้าง
5. Members of another dimension ถ้ามีความสัมพันธ์ระหว่างโดเมนชั้นด้วยกันแล้วสามารถใช้โดเมนชั้นที่มีความสัมพันธ์มาทำนายข้อมูลได้

Mining Model Wizard

Select case

A case is the basic unit of analysis for mining models. In OLAP data, a dimension member represents a case. For example, in a customer credit application, the case is the customer. Select a dimension and the level that contains the set of cases you want to analyze. For a parent-child dimension, the lowest level is assumed to be the case level.

Dimension:
Customers

Level:
Name

< Back Next > Cancel Help

รูปที่ 2.22 การเลือก case

Mining Model Wizard

Select predicted entity

Select the measure, member property, or members you want to predict. For example, in a customer credit application, you can predict credit risk based on debt burden, education, or geographic location. You can use Mining Model Editor to specify more than one predictable entity.

A measure of the source cube
Measures:
Unit Sales

A member property of the case level
Member Properties:
Member Card

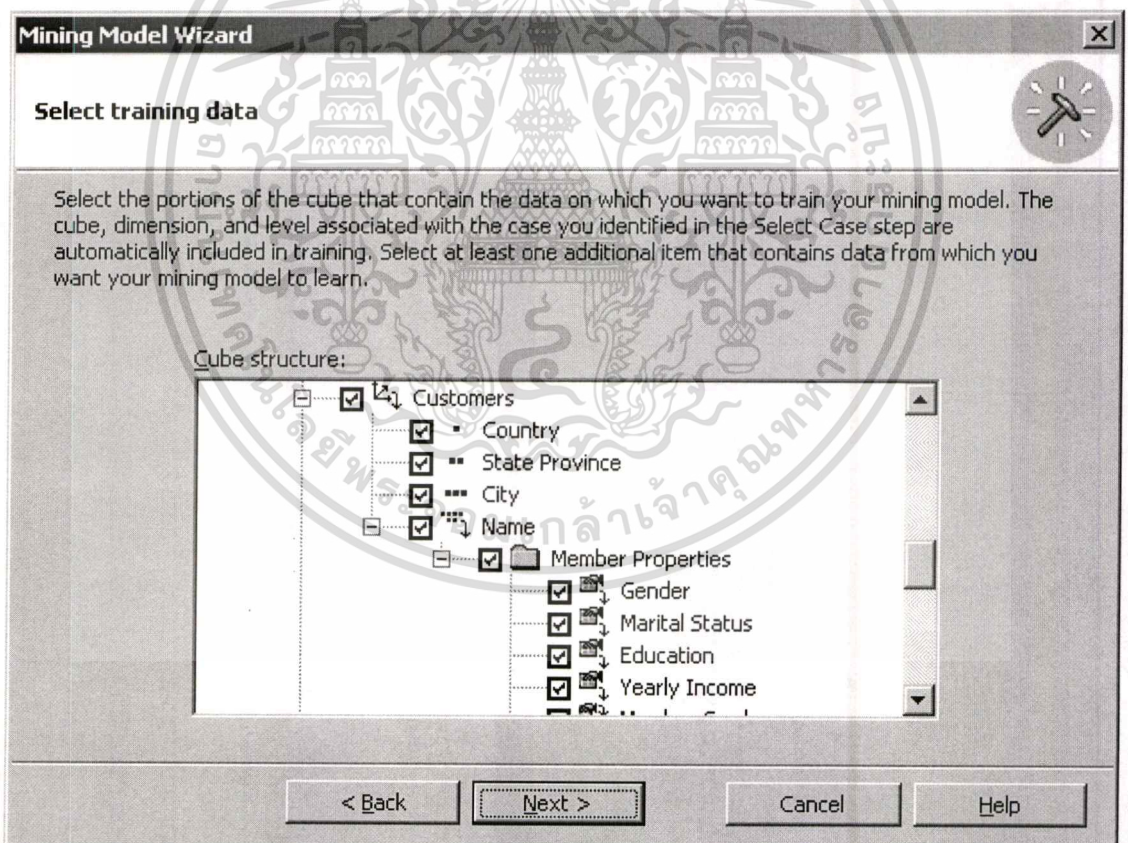
Members of another dimension
Dimensions:
Store

< Back Next > Cancel Help

รูปที่ 2.23 การเลือก “predicted entity”

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6. เลือกข้อมูลที่ใช้ในการทำเหมืองใดเมนชั้นที่ได้เลือกไปก่อนหน้านี้จะถูกเลือกโดยอัตโนมัติในหน้าจอ “Select training data” แต่ต้องเลือกใดเมนชั้นอย่างน้อยหนึ่งใดเมนชั้นดังในรูปที่ 2.24
7. สร้างใดเมนชั้นและเวอร์ชวลคิวบ์ (Virtual Cube) ชั้นตอนนี้เป็นแบบอบชั้นคือไม่เลือกก็ได้ แต่จะมีประโยชน์เป็นอย่างมากถ้าใช้แหล่งข้อมูลแบบ OLAP และอัลกอริทึมแบบ Microsoft Decision Trees ในรูปที่ 2.25 จะมีการถามว่าจะสร้างใดเมนชั้นและเวอร์ชวลคิวบ์เป็นเอทพุตหรือไม่ หลังจากนั้นในรูปที่ 2.26 จะแสดงขั้นตอนสุดท้ายในการสร้างคาค่าไมน์นิ่งโมเดลให้ใส่ชื่อโมเดลไปและคลิกปุ่ม “Finish” เพื่อทำการประมวลผลและสร้างคาค่าไมน์นิ่งโมเดล และได้ผลลัพธ์ดังรูปที่ 2.27



รูปที่ 2.24 การเลือกการทำ training ข้อมูล

Mining Model Wizard

Create a dimension and virtual cube (optional)

You can create a new dimension from the results of the mining model analysis of OLAP data. You can also optionally create a virtual cube that contains the created dimension and the original cube. You can browse this virtual cube to see the results of the data mining analysis.

Create a new dimension based on this mining model

Dimension name:

Create a new virtual cube

This cube will contain the source cube on which the mining model is based and the new dimension you specify in this wizard step.

Virtual cube name:

< Back Next > Cancel Help

รูปที่ 2.25 การสร้างเวอร์ชวลไคเมนชันและคิวบ์

Mining Model Wizard

Finish the **Mining Model Wizard**

Model name:

What do you want to do?

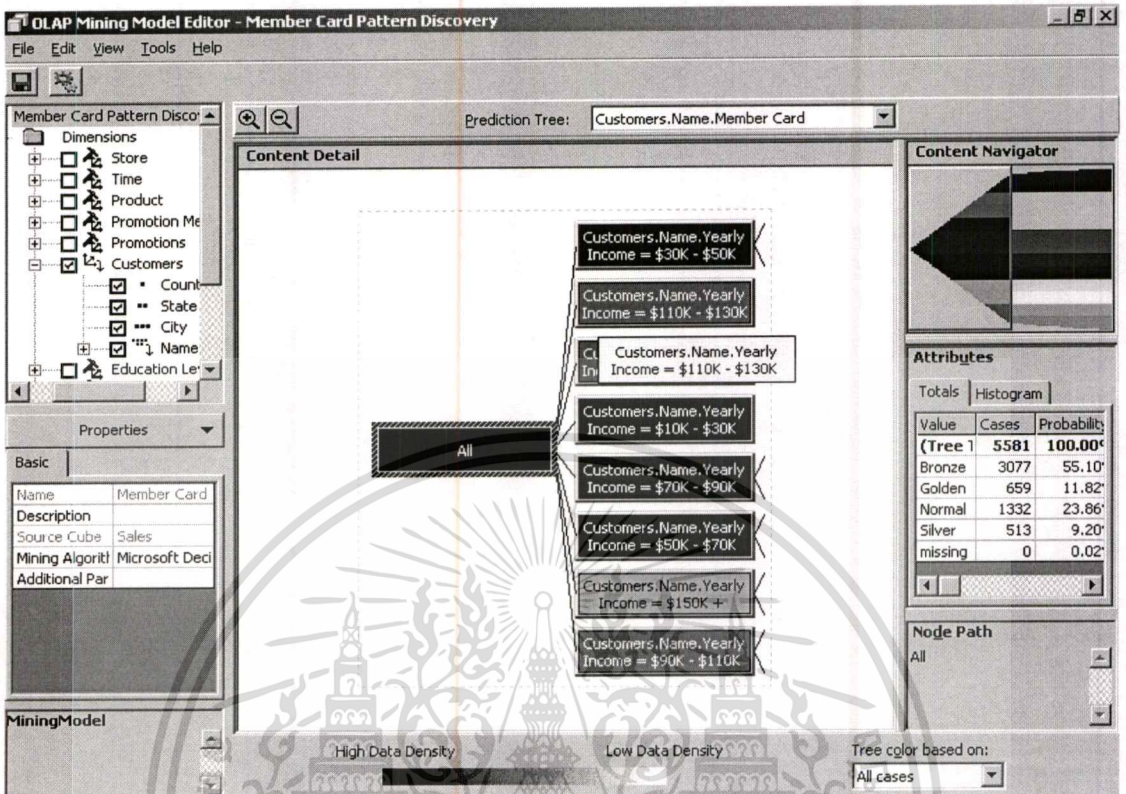
Save, but don't process now

Save and process now

Processing creates the model and trains it with data.

< Back Finish Cancel Help

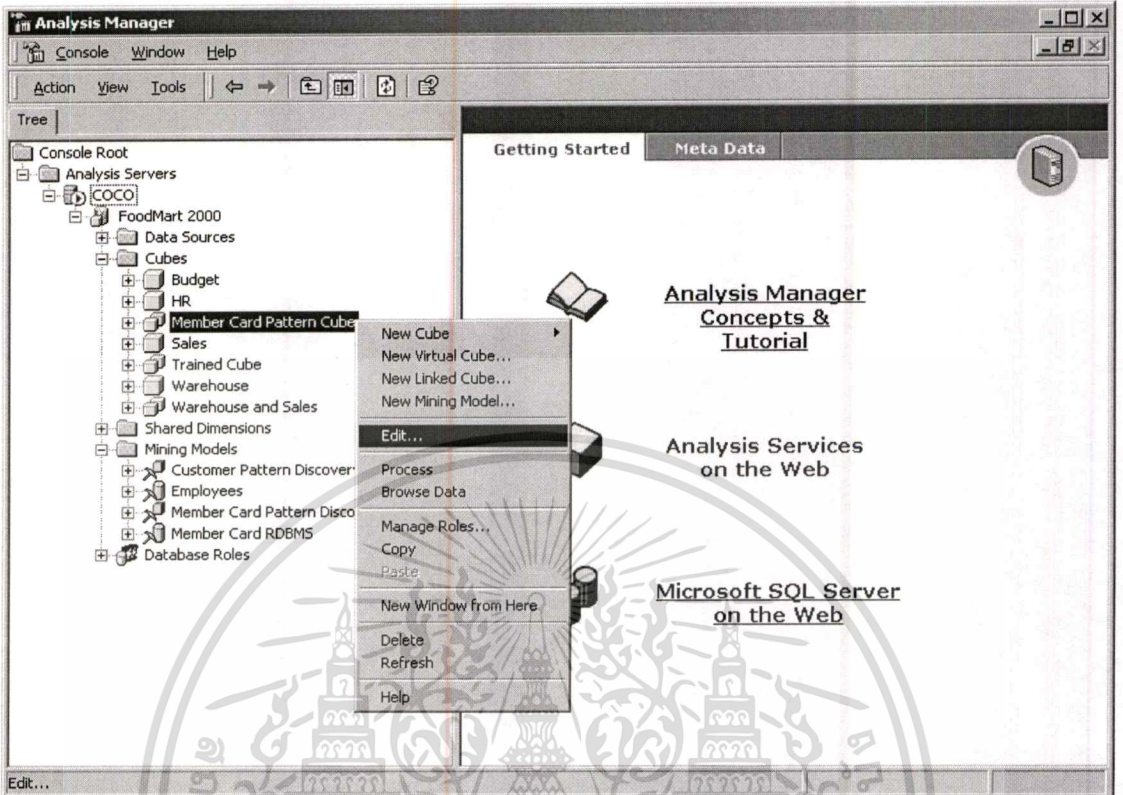
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับรูปที่ 2.26 การตั้งชื่อให้กับดาต้าไมน์นิ่งโมเดลให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



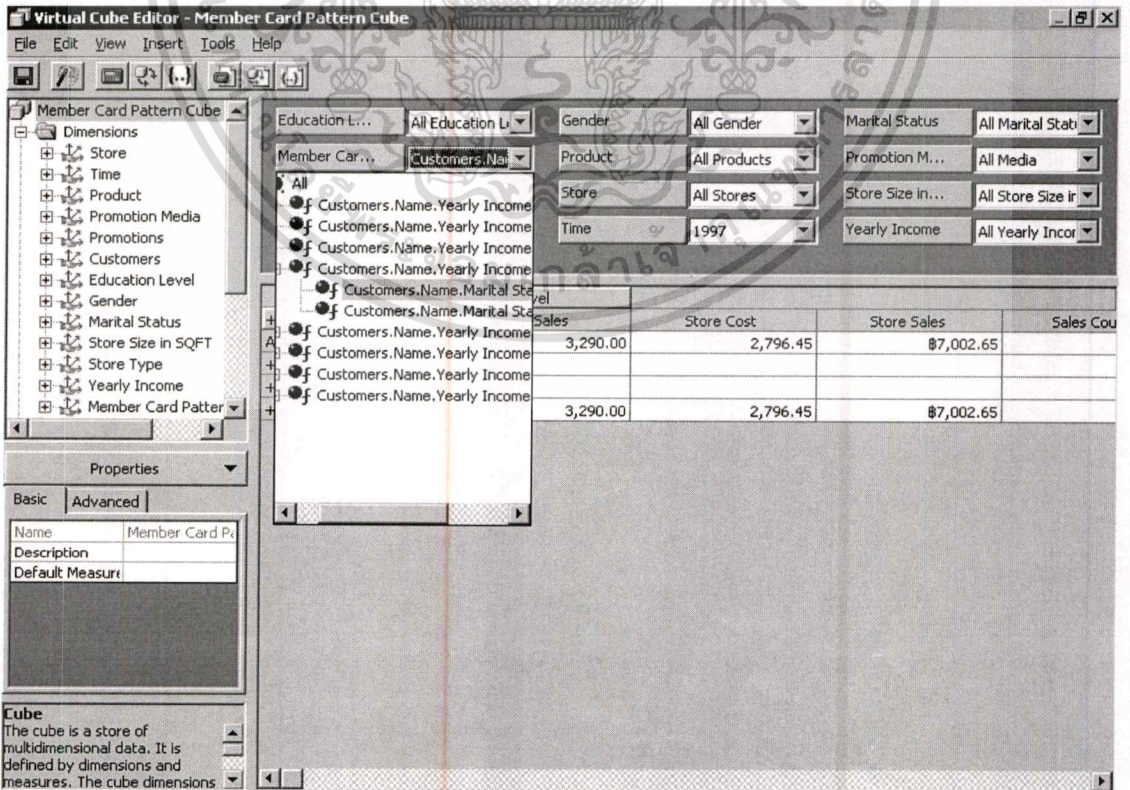
รูปที่ 2.27 OLAP Mining Model Editor

2.2.8.2 การวิเคราะห์ข้อมูลด้วย OLAP คาด้าไมน์นึ่งโมเดล ฐานข้อมูลเชิงสัมพันธ์จะใช้ดีที่สุดสำหรับการสนับสนุนการตัดสินใจข้อมูลแบบทรานแซกชัน เช่น ใบสั่งซื้อ แต่ฐานข้อมูลเชิงสัมพันธ์อาจจะไม่เหมาะกับรูปแบบบางอย่างจึงอาจทำให้ต้องแยกเพิ่มไปยังอีกคาด้าไมน์นึ่งโมเดลอื่นและเป็นการยากที่จะประยุกต์ข้อมูลในตารางต่างๆด้วย ส่วน OLAP นั้นถูกออกแบบมาสำหรับการรายงานและสนับสนุนการตัดสินใจ เพราะว่าได้รวมข้อมูลจากทรานแซกชันมาแล้ว

- การใช้เวอร์ชวลคิวบ์ ในขั้นตอนการสร้างคาด้าไมน์นึ่งที่ผ่านมา ได้มีการสร้างเวอร์ชวลคิวบ์เมื่อต้องการดูข้อมูลให้คลิกคิวบ์ คลิกขวา เลือก "Edit ..." ดังแสดงในรูปที่ 2.28 และจะแสดงข้อมูลนั้นดังแสดงในรูปที่ 2.29 โดยส่วนด้านล่างของหน้าจอจะแสดงข้อมูลที่เป็นตัวเลขที่สัมพันธ์กับรายการใดเมนชั้นที่เลือกในส่วนบนของหน้าจอ จะสังเกตว่าใดเมนชั้นจะประกอบด้วยระดับต่างๆ ที่คล้ายกับโครงสร้างโหนดใน Decision tree นั้นเอง



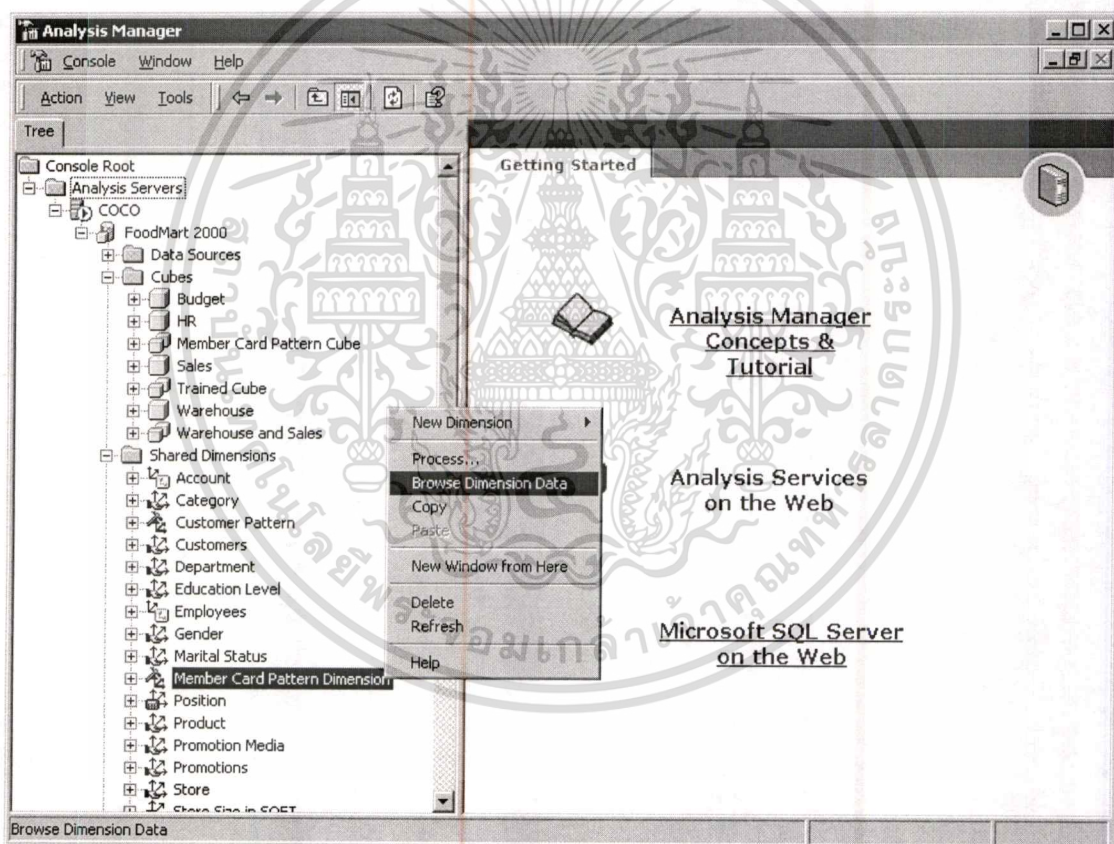
รูปที่ 2.28 วิธีการในการที่จะแสดงข้อมูลของเวอร์ชวลคิวบ์



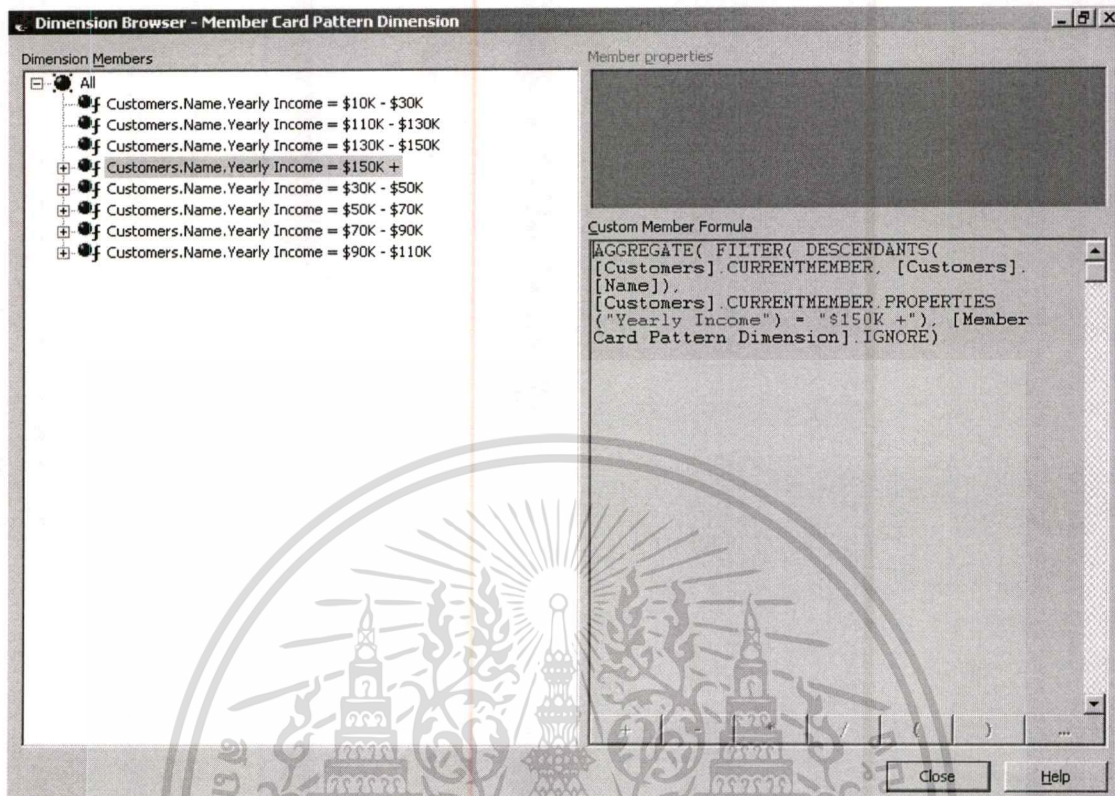
รูปที่ 2.29 การแสดงข้อมูลของเวอร์ชวลคิวบ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าการณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

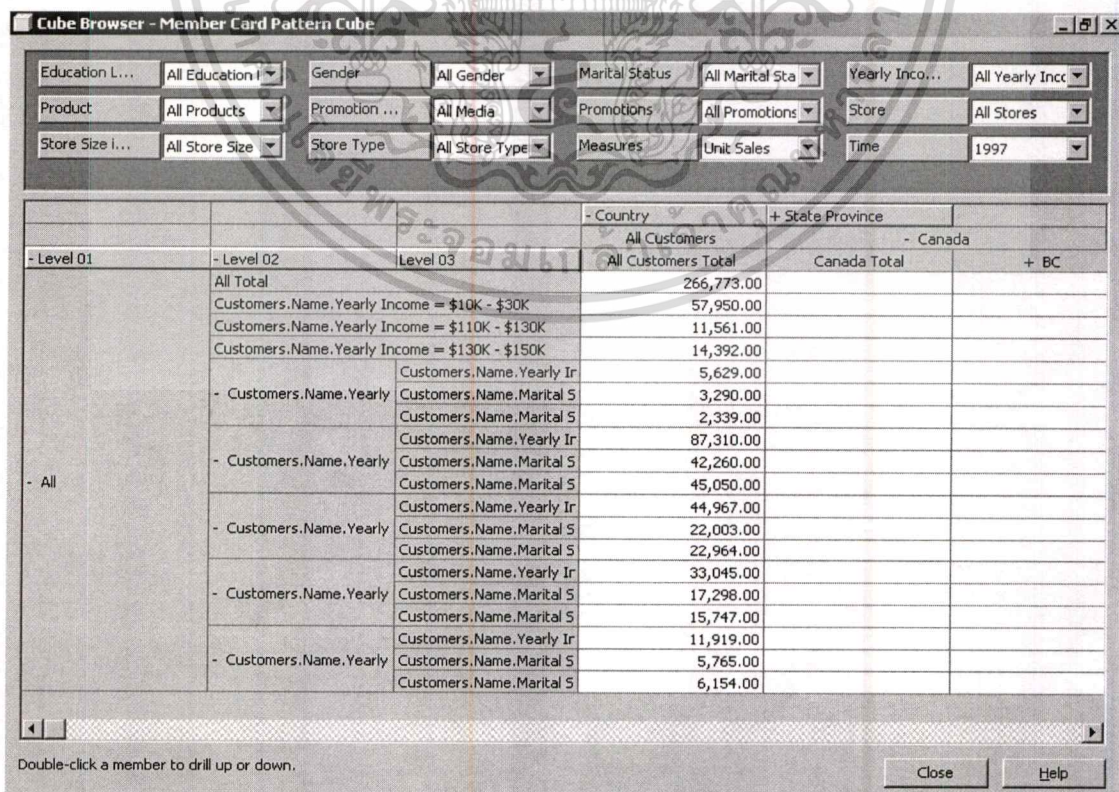
- การใช้โดเมนชัน โดเมนชันจะมีลักษณะเป็นลำดับชั้น การสร้างโดเมนชันด้วยวิธีอื่นนั้น คือการใช้ Multidimensional Expressions (MDX) ซึ่งสามารถสร้าง Calculated member ที่ตรงตามเงื่อนไขในโหนดของ Decision tree ได้ เพื่อให้เข้าใจการทำงานไปที่ไฟล์เตอร์ Shared Dimensions และคลิกขวาที่โดเมนชันเลือก “Browse Dimension Data” ดังรูปที่ 2.30 หลังจากนั้นจะปรากฏโครงสร้างของโดเมนชันและ MDX ดังในรูปที่ 2.31 สังเกตว่าที่ Customers.Name.Yearly Income = \$150K+ นั้นจะตรงกับในโหนดของ Decision tree โดเมนชันเป็นลักษณะที่สำคัญเพราะว่าสามารถแสดงข้อมูลในคิวบ์ที่ได้แสดงในค่าไบนารีหนึ่งโมเดลได้ดังแสดงในรูปที่ 2.32



รูปที่ 2.30 วิธีการในการที่จะแสดงข้อมูลของโดเมนชัน



รูปที่ 2.31 การแสดงข้อมูลของโดเมนชั้นในส่วน MDX



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับรูปที่ 2.32 การแสดงยอดการขายด้วยโดเมนชั้น ให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3

การวิเคราะห์ระบบ

3.1 ความต้องการของระบบ

บริษัท ไทยมาร์ท จำกัด (มหาชน) ได้ดำเนินธุรกิจค้าปลีกในนามของ ไทยมาร์ท (Thai Mart) ในลักษณะบริษัทข้ามชาติโดยจัดตั้งสาขาต่างๆ เพื่อดำเนินธุรกิจค้าปลีกของอุปโภคและบริโภคให้แก่ลูกค้าตามรัฐต่างๆ ในประเทศสหรัฐอเมริกา โดยมีฝ่ายที่เกี่ยวข้องกับระบบดังต่อไปนี้

- ฝ่ายขาย ได้ทำการจัดเก็บข้อมูลปฏิบัติการ (Data Operation) เช่น ยอดขาย ไว้ในระบบฐานข้อมูลและได้มีการสร้างดาต้าแวร์เฮาส์ขึ้นมา โดยได้จัดเก็บข้อมูลปฏิบัติการนั้นในรูปแบบสรุปและเป็นไคเมนชั้นต่างๆ เก็บไว้ในคิวบ์และส่งให้กับฝ่ายการตลาดนำไปทำการวิเคราะห์ต่อไป
- ฝ่ายการตลาด ได้จัดทำแผนการตลาดในส่วนบัตรสมาชิกให้แก่ลูกค้ามาก่อนแล้วโดยให้ลูกค้าเป็นผู้เลือกชนิดของบัตรสมาชิกนี้เอง แต่ต่อมาแผนกวิจัยการตลาดได้ทำการวิจัยและพบปัญหาตรงที่ว่า พฤติกรรมการซื้อและคุณสมบัติของลูกค้าไม่ว่าจะเป็น รายได้ สถานภาพ การแต่งงานและอื่นๆ อันเป็นลักษณะเฉพาะที่กระทบต่อพฤติกรรมผู้บริโภคไม่เหมาะสมชนิดของบัตร ซึ่งมีผลต่อเนื่องถึงการไม่บรรลุตามเป้าหมายของแผนการตลาดอื่นๆ ที่ได้จัดทำมา ณ ตอนนี้อย่างเห็นได้ชัด ฝ่ายการตลาดต้องการเพิ่มความพึงพอใจของลูกค้าและต้องการรักษาฐานลูกค้าไว้ให้เหนียวแน่นกับการให้บริการซื้อสินค้าจากไทยมาร์ทต่อไป โดยมี 2 สิ่งที่จะทำให้บรรลุวัตถุประสงค์ตามที่ตั้งไว้ได้คือ

1. แก้ไขในส่วนบัตรสมาชิกใหม่ ให้พฤติกรรมการซื้อและคุณสมบัติของลูกค้าเหมาะสมกับชนิดของบัตรสมาชิก เพื่อฝ่ายการตลาดจะได้บริการได้ดีและใกล้ชิดกับลูกค้ายิ่งขึ้น
2. จะมีการจัดทำแคตตาล็อกสินค้าแบ่งตามประเภทของบัตรสมาชิก และจัดส่งไปให้ลูกค้าในแต่ละสัปดาห์เพื่อเป็นการเพิ่มยอดขายให้กับไทยมาร์ทเพิ่มขึ้น

ฝ่ายการตลาดได้ทำการวิเคราะห์ข้อมูลจากคิวบ์แต่ประสบปัญหาในการวิเคราะห์ข้อมูลค่อนข้างมากเพราะว่าข้อมูลในคิวบ์มีอัตราการเพิ่มของข้อมูลที่รวดเร็วมก จนยากลำบากในการค้นหาและข้อสรุปของรูปแบบข้อมูลที่ซ่อนอยู่ในคิวบ์ได้ นอกจากนี้ยังเป็นการยากที่จะหาข้อมูลสารสนเทศโดยใช้คนมาจัดการได้อีกเช่นกัน

3.2 ขั้นตอนการทำดาต้าไมน์นิ่งของระบบ

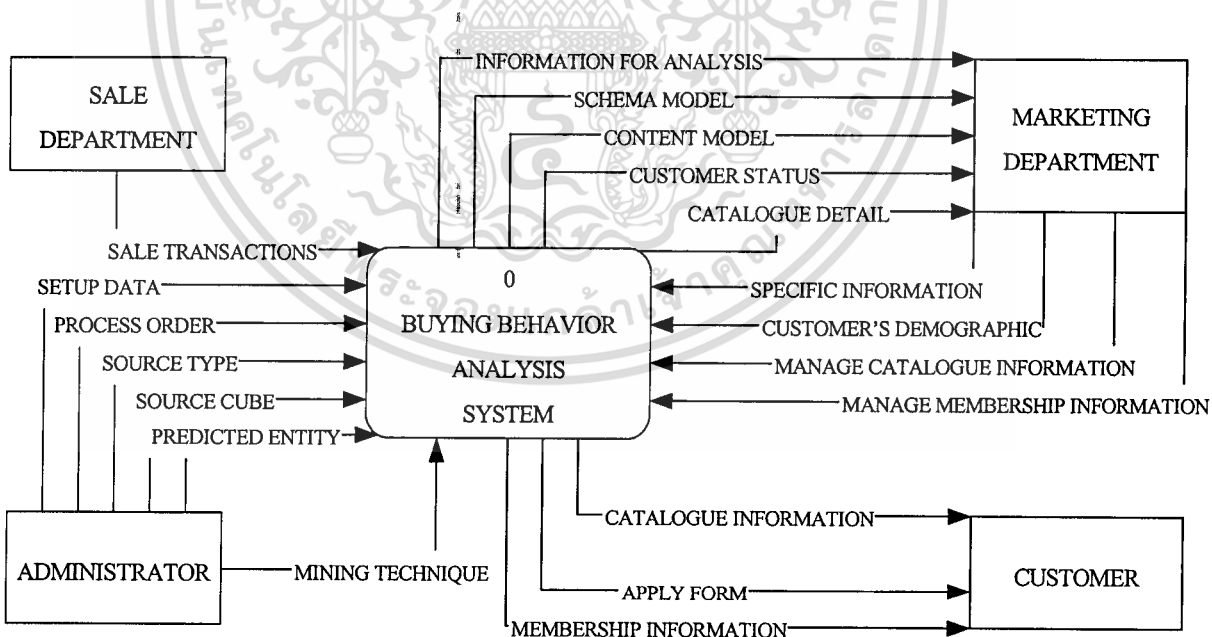
วิธีการในการจัดการกับปัญหาการวิเคราะห์ข้อมูลจากคิวบ์คือ การใช้เทคโนโลยีดาต้าไมน์นิ่ง ดาต้าไมน์นิ่งสามารถที่จะค้นหาและมีข้อสรุปเกี่ยวกับรูปแบบของข้อมูลที่ซ่อนอยู่ในคิวบ์ได้ นอกจากนี้ยังเป็นการง่ายที่จะหาข้อมูลสารสนเทศแบบอัตโนมัติด้วย เพราะดาต้าไมน์นิ่งได้จัดเตรียมอัลกอริทึมที่จะทำให้มีการค้นหารูปแบบของข้อมูลแบบอัตโนมัติ และมีเครื่องมือในการช่วยวิเคราะห์อีกด้วย โดยระบบจะประกอบด้วยผู้บริหารจัดการดาต้าไมน์นิ่งทำการสร้างและออกแบบดาต้าไมน์นิ่งโมเดลที่จะมีแอปพลิเคชันที่ถูกใช้ โดยผู้ใช้งานฝ่ายการตลาดมาทำการรันและวิเคราะห์ข้อมูลได้ โดยขั้นตอนวิธีการในการตอบสนองความต้องการของฝ่ายการตลาดมีดังนี้

1. การที่จะทำการจัดแบ่งประเภทบัตรสมาชิกใหม่นั้น ระบบจะใช้ข้อมูลจากยอดขายที่อยู่ในคิวบ์ และข้อมูลคุณสมบัติของลูกค้า เช่น สถานภาพการแต่งงาน รายได้ ประเภทบัตรสมาชิกที่ได้สมัครไป เป็นต้น มาทำการวิเคราะห์
2. สร้างดาต้าไมน์นิ่งโมเดลจากข้อมูลการขายที่จัดเก็บอยู่ในคิวบ์และใช้อัลกอริทึมแบบ Microsoft Decision Trees เพื่อทำการหารูปแบบของบัตรสมาชิกบนพื้นฐานของคุณสมบัติของลูกค้า
3. สร้างแอปพลิเคชันขึ้นมาเพื่อรองรับการใช้งานจากฝ่ายการตลาด ซึ่งจะเป็นลักษณะของการแสดงข้อมูลสารสนเทศจากดาต้าไมน์นิ่งโมเดลที่ได้สร้างมา
4. สร้างแอปพลิเคชันขึ้นมาเพื่อรองรับระบบบัตรสมาชิกใหม่ตามข้อมูลการวิเคราะห์จากดาต้าไมน์นิ่ง เพื่อให้ฝ่ายการตลาดได้กำหนดเงื่อนไขของประเภทสมาชิกโดยใช้การตัดสินใจ ประกอบร่วมกับข้อมูลการทำนายของดาต้าไมน์นิ่ง และมีการจัดส่งเอกสารแจ้งการปรับเปลี่ยนประเภทสมาชิก (สำหรับสมาชิกเก่า) นอกจากนี้ สามารถดูข้อมูลลูกค้าที่ประเภทบัตรสมาชิกยังไม่ตรงกับเงื่อนไขได้ การรับสมัครบัตรสมาชิกก็สามารถพิมพ์แบบฟอร์มพร้อมรายละเอียดของเงื่อนไขในแต่ละประเภทสมาชิกได้อีกเช่นกัน
5. สร้างแอปพลิเคชันขึ้นมาเพื่อรองรับการสร้างแคตตาล็อกสินค้าสำหรับลูกค้าแต่ละประเภทบัตรสมาชิกได้ นอกจากนี้ยังสามารถที่จะจัดส่งพร้อมทั้งคู่มือแคตตาล็อกที่เคยสร้างไปแล้วได้

3.3 แผนภาพกระแสข้อมูล

แผนภาพกระแสข้อมูล (DFD) เป็นเครื่องมือที่ช่วยในการอธิบายความคิดของระบบที่จะพัฒนาขึ้นมา โดยจะแสดงการไหลของข้อมูลผ่านเข้าไปในระบบ แต่ไม่ได้แสดงถึงขบวนการในการประมวลผลข้อมูล ระบบการวิเคราะห์พฤติกรรมการซื้อสินค้าในระดับ Context diagram นั้นแสดงในรูปที่ 3.1 ซึ่งแสดงว่าระบบประกอบด้วยเอนทิตีที่สำคัญได้แก่

- ฝ่ายขาย (Sale department) นำข้อมูลการขายเข้าสู่ระบบเพื่อนำไปประมวลผลเปลี่ยนแปลงคิวบ์
- ฝ่ายบริหารจัดการระบบ (Administrator) กำหนดข้อมูลต่างๆ เพื่อใช้ในการสร้างค่าต้นไม้โมเดลตามความต้องการของฝ่ายการตลาด
- ฝ่ายการตลาด (Marketing department) ให้ข้อมูลลูกค้าและข้อมูลเพื่อการใช้แอปพลิเคชันเข้าสู่ระบบและรับข้อมูลสารสนเทศเพื่อใช้ในการวิเคราะห์ต่อไป
- ลูกค้าหรือสมาชิกบัตร (Customer) จะรับข้อมูลสารสนเทศจากระบบในที่นี่คือ ใบแจ้งการเปลี่ยนประเภทสมาชิก แบบฟอร์มสมัครสมาชิกพร้อมเงื่อนไขการสมัครของแต่ละประเภทบัตรสมาชิกและแคตตาล็อกสินค้าที่จัดทำตามประเภทสมาชิก



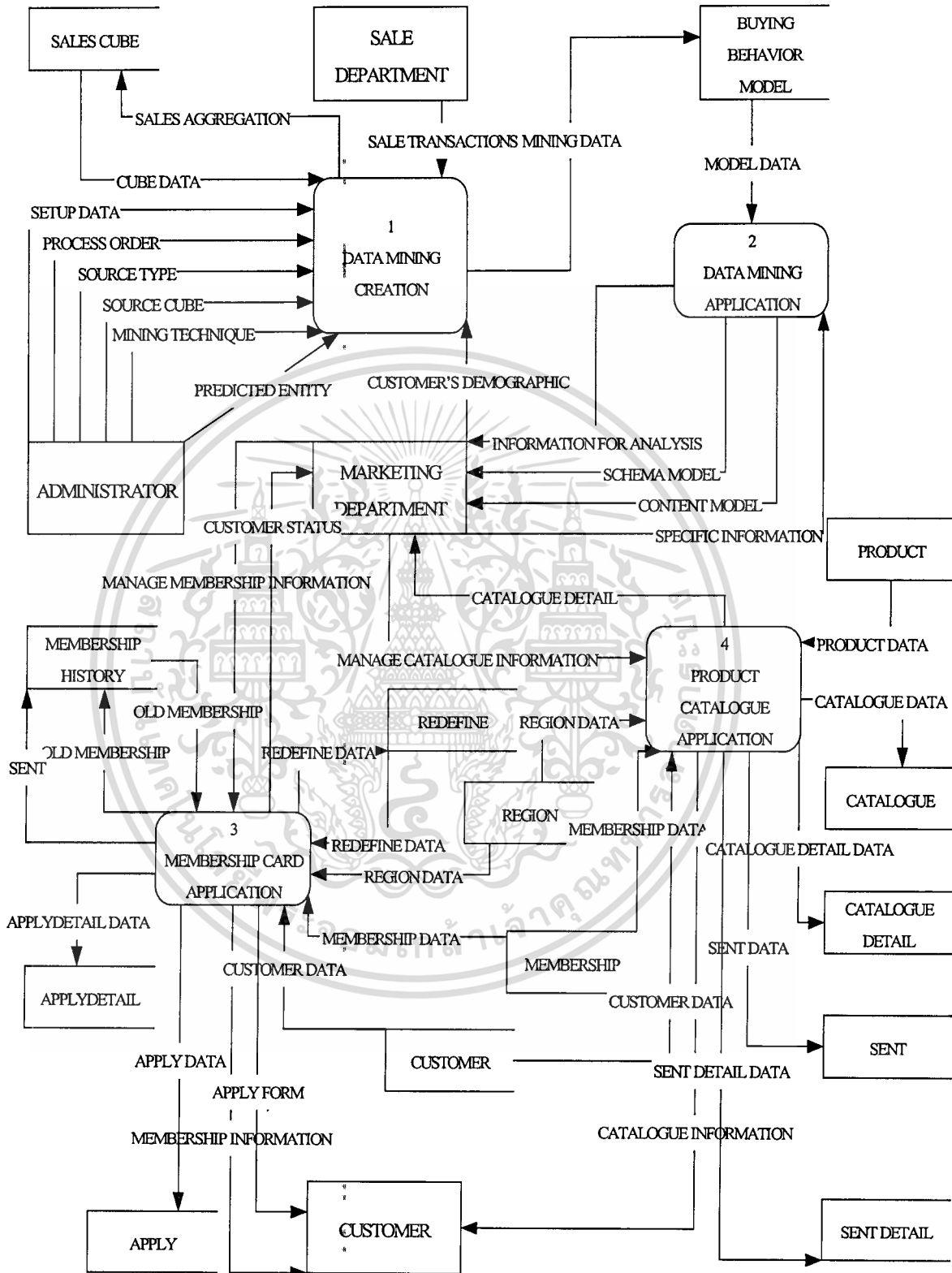
รูปที่ 3.1 DFD ระดับ Context diagram

ส่วนรูปที่ 3.2 แสดง DFD ในระดับ 0 ซึ่งประกอบด้วยโปรเซส (Process) ที่สำคัญดังนี้

- DATA MINING CREATION เป็นโปรเซสที่มีการสร้างดาต้าไมน์นิ่งขึ้นมาตามข้อมูลจากคิวบ์การขาย ยอดการขาย และข้อมูลลูกค้า
- DATA MINING APPLICATION เป็นโปรเซสที่มาสับสนุนการใช้งานดาต้าไมน์นิ่งที่ได้สร้างขึ้น
- MEMBERSHIP CARD APPLICATION เป็นโปรเซสที่ใช้ในการจัดการระบบบัตรสมาชิกใหม่ตามที่ได้ผลลัพธ์จากการทำนายข้อมูลจากดาต้าไมน์นิ่ง
- PRODUCT CATALOGUE APPLICATION เป็นโปรเซสที่ใช้ในการจัดการระบบแคตตาล็อกสินค้าตามประเภทบัตรสมาชิก

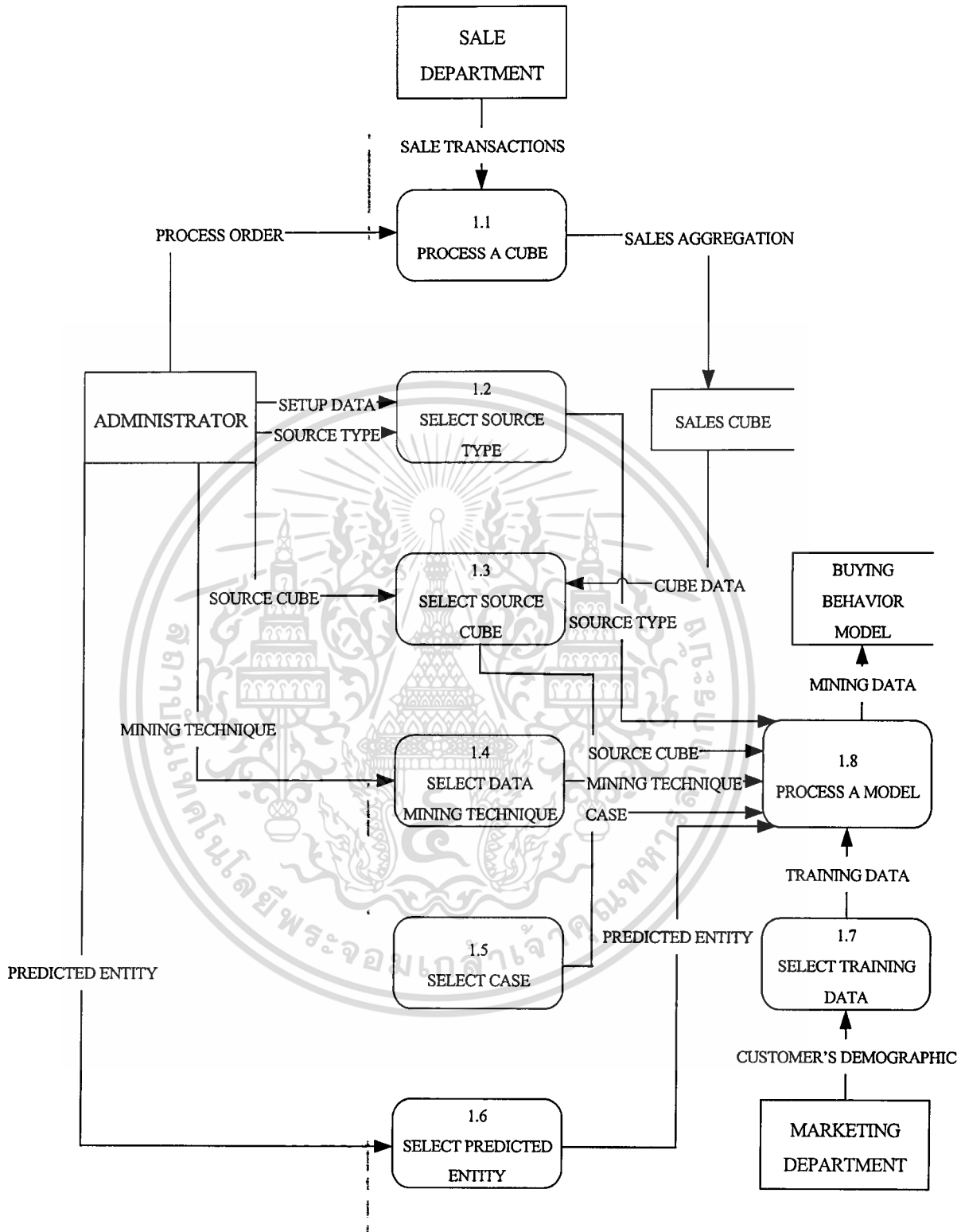
ส่วนรูปที่ 3.3 แสดงในระดับ 1 ซึ่งเป็นรายละเอียดของ DATA MINING CREATION ที่แสดงกระบวนการทั้งหมดในการสร้างดาต้าไมน์นิ่ง โดยประกอบด้วยโปรเซสที่สำคัญดังนี้

- PROCESS A CUBE เป็นโปรเซสที่ใช้ในการประมวลผลข้อมูลเพื่อปรับปรุงคิวบ์ที่ได้มาจากฝ่ายขายให้ข้อมูลมีทันสมัย โดยมีข้อมูลที่สำคัญมาจากข้อมูลปฏิบัติงานขาย
- SELECT SOURCE TYPE เป็นโปรเซสแรกที่ใช้ในการสร้างดาต้าไมน์นิ่งด้วยการกำหนดแหล่งชนิดของข้อมูล ในที่นี้เลือกแบบ OLAP
- SELECT SOURCE CUBE เป็นโปรเซสที่ใช้ในการเลือกแหล่งข้อมูลจากคิวบ์ ในที่นี้เลือกคิวบ์การขาย (Sales cube)



รูปที่ 3.2 DFD ระดับ 0

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.3 DFD ระดับ 1 ในรายละเอียดของ DATA MINING CREATION

- SELECT DATA MINING TECHNIQUE เป็นโปรเซสที่ใช้ในการเลือกเทคนิคอัลกอริทึมในการทำค้ำไม้หนึ่ง ในที่นี้เลือก Microsoft Decision Tree
- SELECT CASE เป็นโปรเซสที่ใช้ในการเลือกข้อมูลมาเป็นกลุ่มตัวอย่างในการทำนายข้อมูลต่อไป ในที่นี้เลือกไคเมนชัน Customers
- SELECT PREDICTED ENTITY เป็นโปรเซสที่ใช้ในการเลือกข้อมูลที่ต้องการทำนาย ในที่นี้เลือกเมมเบอร์พรอพเพอร์ตี้ Member Card
- SELECT TRAINING DATA เป็นโปรเซสที่ใช้ในการเลือกข้อมูลที่ต้องการทำการ Train ค้ำไม้หนึ่งโมเดล โดยในไคเมนชัน Customers นั้นเลือกเมมเบอร์พรอพเพอร์ตี้ดังนี้
 - Gender
 - Marital Status
 - Education
 - Yearly Income
 - Member Card
 - Birthdate
 - Num Children At Home
 - Occupation
 - Houseowner
 - Num Cars Owned

ต่อไปเลือกไคเมนชัน Products ระดับ (Level) Product Name และเมเชอร์ Unit Sales

- PROCESS A MODEL เป็นโปรเซสที่ใช้ในการประมวลผลโมเดลที่ได้กำหนดข้อมูลต่างๆ ที่จำเป็นต่อการประมวลผลในขั้นตอนที่ผ่านมา เพื่อให้ได้ค้ำไม้หนึ่งโมเดล

ส่วนรูปที่ 3.4 แสดงในระดับ 1 ซึ่งเป็นรายละเอียดของ DATA MINING APPLICATION ที่แสดงการจัดการ การแสดงข้อมูลของค้ำไม้หนึ่งเพื่อสนับสนุนการตัดสินใจของฝ่ายการตลาด โดยประกอบด้วยโปรเซสต่างๆ ดังนี้

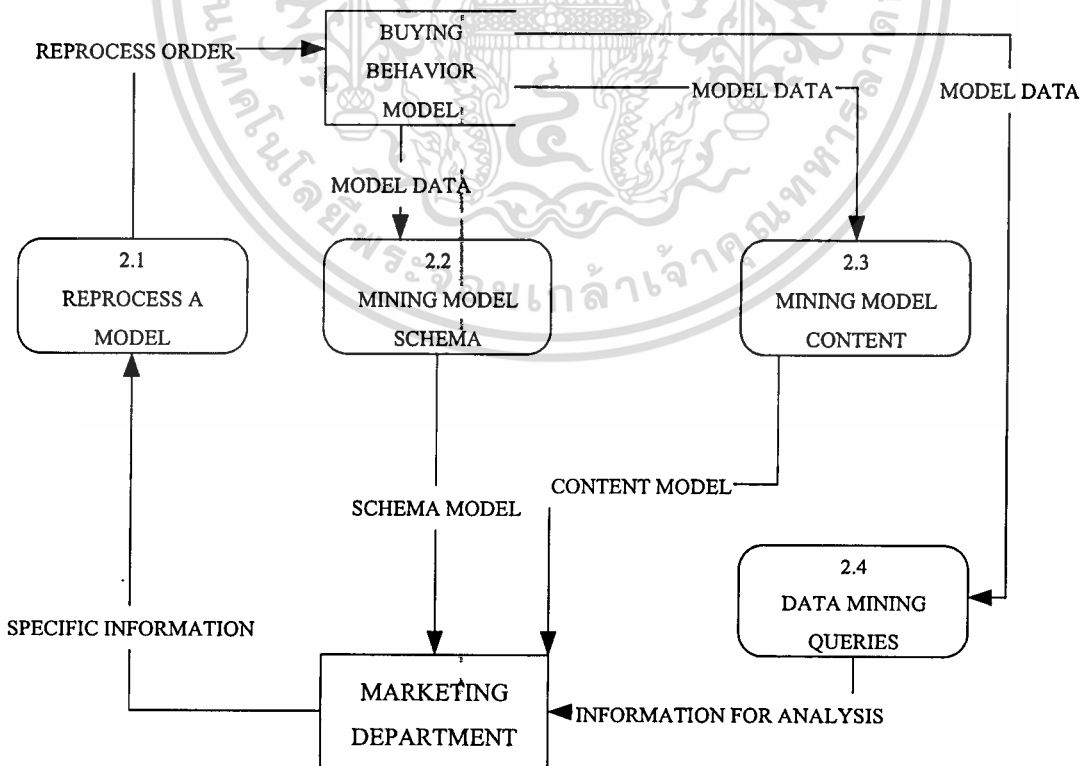
- REPROCESS A MODEL เป็นโปรเซสที่ใช้ในการประมวลผลค้ำไม้หนึ่งโมเดลเพื่อต้องการปรับปรุงโมเดลเมื่อฝ่ายขายได้มีการนำข้อมูลปฏิบัติการมาปรับปรุงเปลี่ยนแปลงคิวบ์การขายซึ่งเป็นแหล่งข้อมูลของค้ำไม้หนึ่งโมเดล
- MINING MODEL SCHEMA เป็นโปรเซสที่ใช้ในการแสดงโครงสร้างของค้ำไม้หนึ่งโมเดล

- MINING MODEL CONTENT เป็นโปรเซสที่ใช้ในการแสดงข้อมูลของโหนดของโครงสร้างของคาค้าไมน์นิ่งโมเดล
- DATA MINING QUERIES, เป็นโปรเซสที่ใช้ในการแสดงข้อมูลของคุณลักษณะของโหนดของโครงสร้างของคาค้าไมน์นิ่งโมเดล

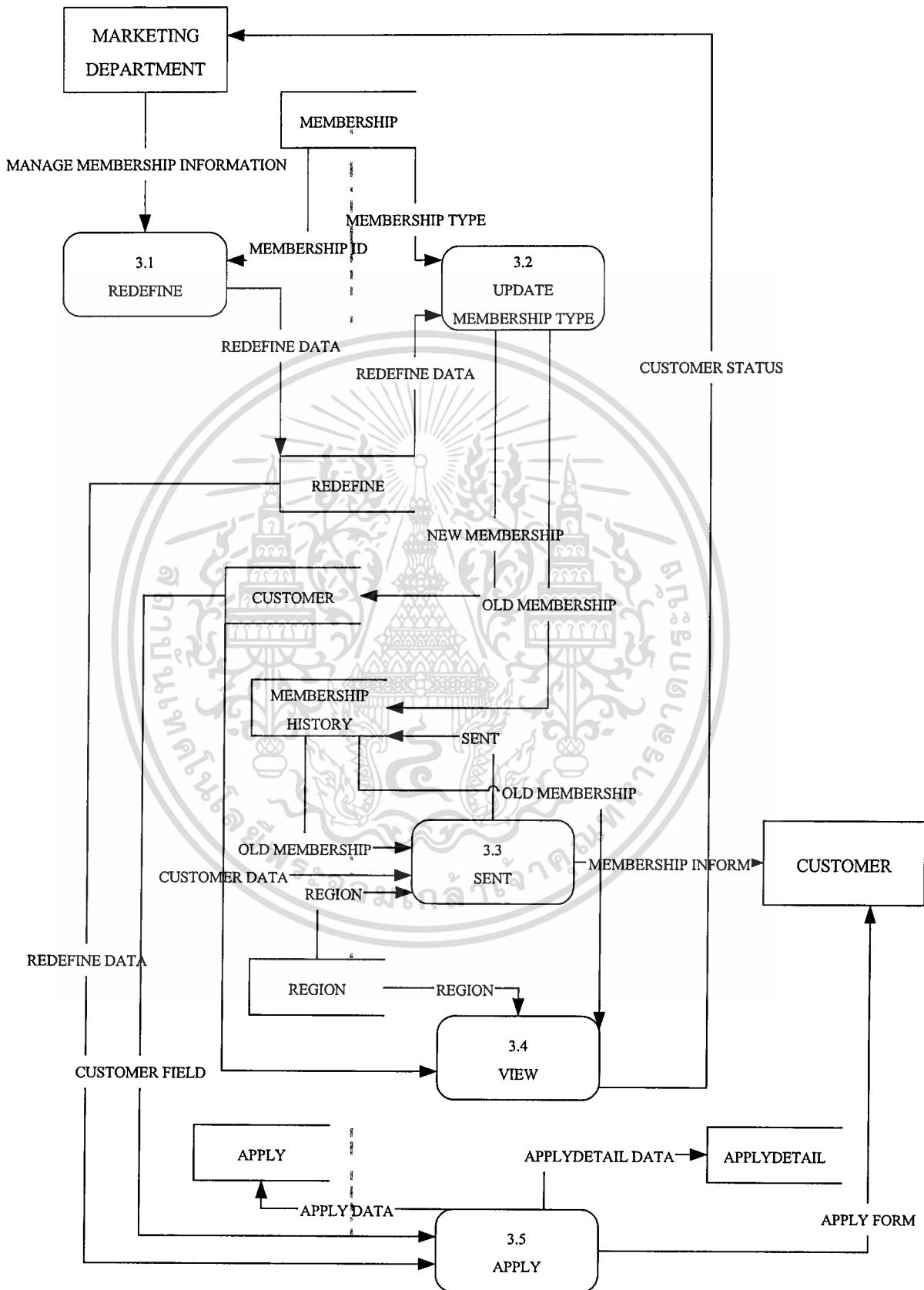
ส่วนรูปที่ 3.5 แสดงในระดับ 1 ซึ่งเป็นรายละเอียดของ MEMBERSHIP CARD

APPLICATION แสดงการจัดการระบบบัตรสมาชิกใหม่ตามความต้องการของฝ่ายการตลาดที่ให้คุณสมบัติของลูกค้าในด้านที่มีผลต่อพฤติกรรมการซื้อ ยอดขายและประเภทสินค้าได้ตรงตามความเหมาะสมโดยใช้ข้อมูลจากการทำนายของคาค้าไมน์นิ่งโมเดลเข้ามาช่วยทำการตัดสินใจ MEMBERSHIP CARD APPLICATION ประกอบด้วยโปรเซสต่างๆ ดังนี้

- REDEFINE เป็นโปรเซสที่ใช้ในการกำหนดเงื่อนไขของประเภทบัตรสมาชิก โดยใช้การตัดสินใจของฝ่ายการตลาดบนพื้นฐานการทำนายข้อมูลของคาค้าไมน์นิ่งโมเดล
- UPDATE MEMBERSHIP TYPE เป็นโปรเซสที่ใช้ในการแก้ไขประเภทบัตรสมาชิกของลูกค้าเมื่อมีคุณสมบัติไม่ตรงตามเงื่อนไขที่ได้กำหนดไว้พร้อมทั้งเก็บประวัติประเภทสมาชิกบัตรเดิมไว้ด้วย



รูปที่ 3.4 DFD ระดับ 1 ในรายละเอียดของ DATA MINING APPLICATION



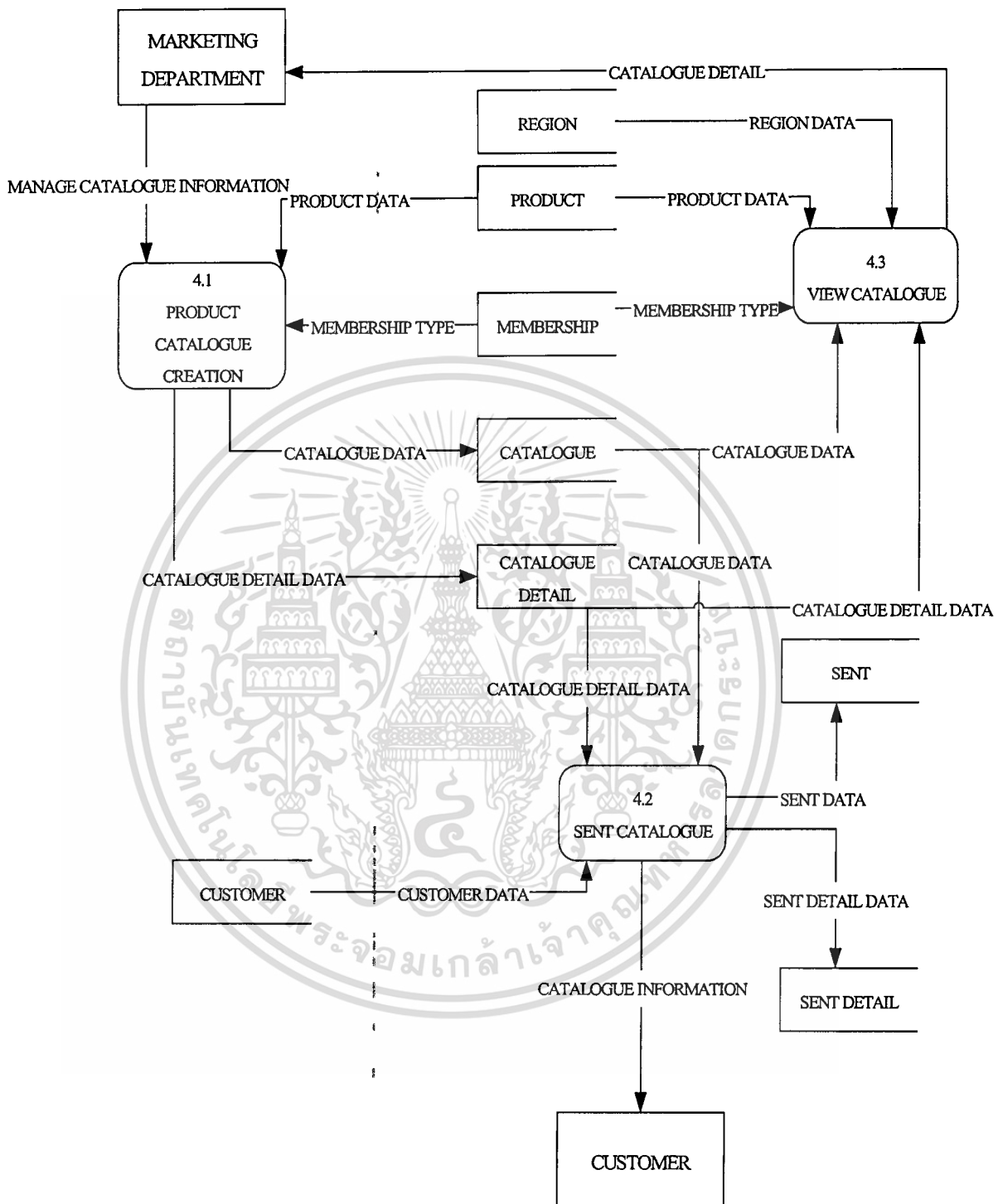
รูปที่ 3.5 DFD ระดับ 1 ในรายละเอียดของ MEMBERSHIP CARD APPLICATION

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- SENT เป็นโปรเซสที่ใช้ในการจัดส่งจดหมายการเปลี่ยนประเภทสมาชิกบัตรไปยังลูกค้า ซึ่งการจัดส่งสามารถเลือกส่งตามเขตที่อยู่ได้
 - VIEW เป็นโปรเซสที่ใช้ในการดูข้อมูลว่าลูกค้าใดที่ยังไม่เปลี่ยนประเภทสมาชิกบัตรหรือได้เปลี่ยนประเภทสมาชิกบัตรไปแล้ว
 - APPLY เป็นโปรเซสที่ใช้ในการจัดพิมพ์แบบฟอร์มการสมัคร โดยแสดงเงื่อนไขคุณสมบัติประเภทสมาชิกบัตรแต่ละประเภท พร้อมทั้งจัดเก็บประวัติในการจัดพิมพ์ด้วย
- ส่วนรูปที่ 3.6 แสดงในระดับ 1 ซึ่งเป็นรายละเอียดของ PRODUCT CATALOGUE

APPLICATION แสดงการจัดการระบบแคตตาล็อกสินค้า โดยใช้ข้อมูลจากการจัดแบ่งประเภทบัตรสมาชิกที่ได้จากการกำหนดเงื่อนไขของฝ่ายการตลาดบนพื้นฐานการทำงานของคาน่าไมนนิ่งโมเดลเข้ามาสร้างแคตตาล็อกสินค้าตามประเภทของบัตรสมาชิก PRODUCT CATALOGUE APPLICATION ประกอบด้วยโปรเซสต่างๆ ดังนี้

- PRODUCT CATALOGUE CREATION เป็นโปรเซสที่ใช้ในการสร้างแคตตาล็อกสินค้าตามประเภทบัตรสมาชิก โดยการกำหนดสินค้าที่แสดงนั้นขึ้นอยู่กับการตัดสินใจของฝ่ายการตลาดที่ต้องใช้แนวทางตามประเภทบัตรสมาชิกที่ลูกค้าแต่ละประเภทมีคุณสมบัติในด้านพฤติกรรมที่ซื้อที่แตกต่างกัน
- SENT CATALOGUE เป็นโปรเซสที่ใช้ในการจัดส่งแคตตาล็อกสินค้าฯ ไปยังลูกค้าพร้อมทั้งบันทึกประวัติในการจัดส่งด้วย
- VIEW CATALOGUE เป็นโปรเซสที่ใช้ในการดูแคตตาล็อกสินค้าที่เคยสร้างไปแล้วจัดแบ่งตามประเภทบัตรสมาชิก



รูปที่ 3.6 DFD ระดับ 1 ในรายละเอียดของ PRPRODUCT CATALOGUE APPLICATION

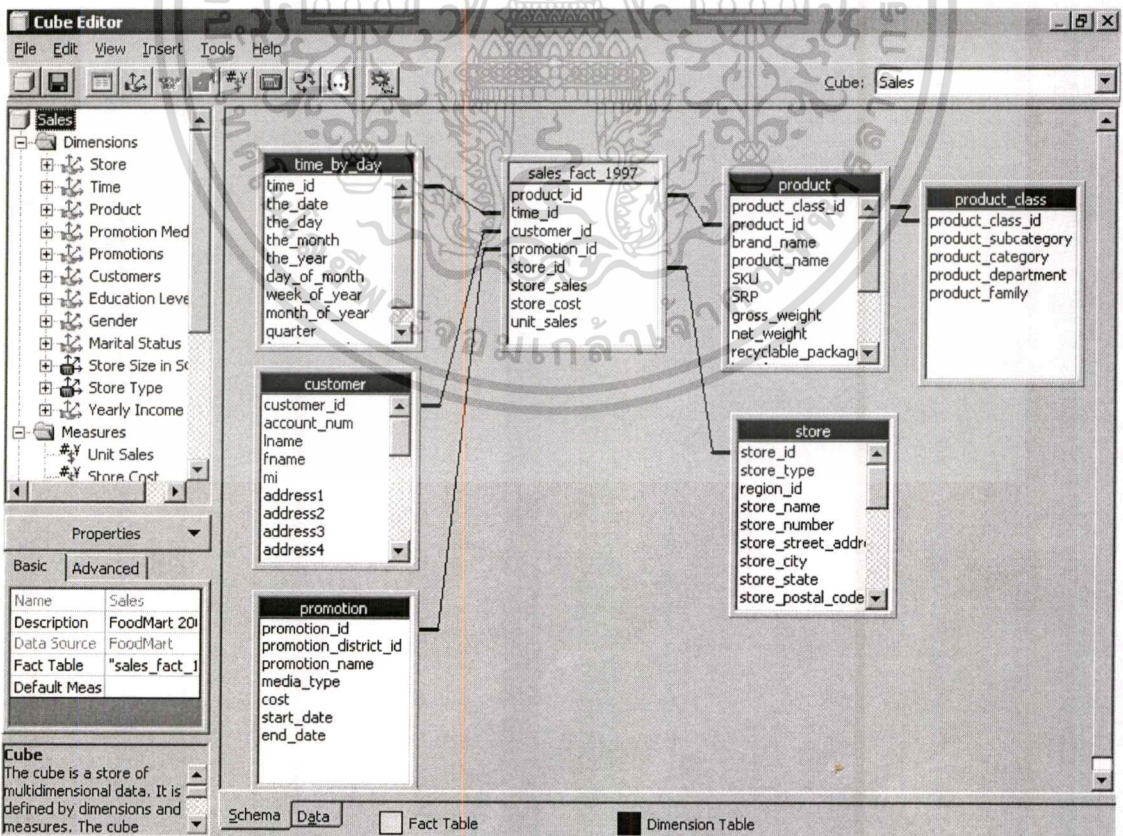
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

การออกแบบระบบ

4.1 โครงสร้างของคิวบ์ในดาต้าแวร์เฮาส์

คิวบ์การขาย (Sales Cube) เป็นรูปแบบของ Multidimensional โดยมาจากฐานข้อมูลเชิงสัมพันธ์ มีโครงสร้างแบบ Star schema ที่ประกอบด้วยตารางหลักที่เรียกว่าตาราง Fact ตาราง Fact ก็คือตารางที่ประกอบด้วยเมเชอร์ (Measure) หรือข้อมูลที่บอกในรูปของค่าปริมาณในฐานข้อมูลที่ต้องการวิเคราะห์ เช่น ยอดการขาย ต้นทุน และงบประมาณ เป็นต้น ทำการเชื่อมสัมพันธ์กับตารางที่เรียกว่า ตารางไดเมนชัน (Dimension) ที่ใช้อธิบายข้อมูลหรือเป็นกลุ่มของข้อมูลที่ใช้ในการวิเคราะห์ทางธุรกิจ เช่น เวลา ภูมิศาสตร์ เป็นต้น โดยสามารถแบ่งเป็นลำดับชั้นได้ เช่น ปี เดือน วัน เป็นต้น



รูปที่ 4.1 โครงสร้างของคิวบ์การขาย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โครงสร้างคิวบ์การขายแสดงดังรูปที่ 4.1 ส่วนรายละเอียดว่าประกอบด้วยตารางอะไรบ้าง แสดงดังตารางที่ 4.1

ตารางที่ 4.1 Data Dictionary ของคิวบ์การขาย

Table Name	Table Type	Column Name	Data Type	Length	Allow Nulls
Sales_fact_1997	Fact Table	Product_id	Int	4	Yes
		Time_id	Int	4	Yes
		Customer_id	Int	4	Yes
		Promotion_id	Int	4	Yes
		Store_id	Int	4	Yes
		Store_sales	Money	8	Yes
		Store_cost	Money	8	Yes
		unit_sales	Float	8	Yes
Time_by_day	Dimension Table	Time_id	Int	4	No
		the_date	Smalldatetime	4	Yes
		the_day	Nvarchar	15	Yes
		the_month	Nvarchar	15	Yes
		the_year	Smallint	2	Yes
		day_of_month	Smallint	2	Yes
		Week_of_year	Float	8	Yes
		Month_of_year	Smallint	2	Yes
		Quarter	Nvarchar	2	Yes
		Fiscal_period	Nvarchar	255	Yes
Customer	Dimension Table	Customer_id	Int	4	No
		Account_num	Float	8	Yes
		Lname	Nvarchar	100	Yes
		Fname	Nvarchar	50	Yes
		Mi	Nvarchar	20	Yes
		Address1	Nvarchar	100	Yes

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.1 Data Dictionary ของคิวบ์การขาย (ต่อ)

Table Name	Table Type	Column Name	Data Type	Length	Allow Nulls
		Address2	Nvarchar	100	Yes
		Address3	Nvarchar	100	Yes
		Address4	Nvarchar	100	Yes
		City	Nvarchar	50	Yes
		State_province	Nvarchar	50	Yes
		Postal_code	nvarchar	50	Yes
		Country	nvarchar	50	Yes
		Customer_region_id	int	4	Yes
		Phone1	nvarchar	50	Yes
		Phone2	nvarchar	50	Yes
		Birthdate	smalldatetime	4	Yes
		Marital_status	nvarchar	1	Yes
		Yearly_income	nvarchar	50	Yes
		Gender	nvarchar	1	Yes
		Total_children	smallint	2	Yes
		Num_children_at_home	smallint	2	Yes
		Education	nvarchar	30	Yes
		Date_acctn_opened	smalldatetime	4	Yes
		Member_card	nvarchar	50	Yes
		Occupation	nvarchar	50	Yes
		Houseowner	nvarchar	1	Yes
		Num_cars_owned	smallint	2	Yes
Promotion	Dimension Table	Promotion_id	int	4	No
		Promotion_district_id	int	4	Yes
		Promotion_name	nvarchar	255	Yes
		Media_type	nvarchar	255	Yes

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.1 Data Dictionary ของคิวบ์การขาย (ต่อ)

Table Name	Table Type	Column Name	Data Type	Length	Allow Nulls
		Cost	float	8	Yes
		Start_date	smalldatetime	4	Yes
		End_date	smalldatetime	4	Yes
Product	Dimension Table	Product_class_id	int	4	Yes
		Product_id	int	4	No
		Brand_name	nvarchar	255	Yes
		Product_name	nvarchar	255	Yes
		SKU	float	8	Yes
		SRP	money	8	Yes
		Gross_weight	real	4	Yes
		Net_weight	real	4	Yes
		Recyclable_package	bit	1	No
		Low_fat	bit	1	No
		Units_per_case	smallint	2	Yes
		Cases_per_pallet	smallint	2	Yes
		Shelf_width	real	4	Yes
		Shelf_height	real	4	Yes
		Shelf_depth	real	4	Yes
Store	Dimension Table	Store_id	int	4	No
		Store_type	nvarchar	255	Yes
		Region_id	int	4	Yes
		Store_name	nvarchar	255	Yes
		Store_number	float	8	Yes
		Store_street_address	nvarchar	255	Yes
		Store_city	nvarchar	50	Yes
		Store_state	nvarchar	50	Yes

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.1 Data Dictionary ของคิวบ์การขาย (ต่อ)

Table Name	Table Type	Column Name	Data Type	Length	Allow Nulls
		Store_postal_code	nvarchar	50	Yes
		Store_country	nvarchar	50	Yes
		Store_manager	nvarchar	255	Yes
		Store_phone	nvarchar	255	Yes
		Store_fax	nvarchar	255	Yes
		First_opened_date	smalldatetime	4	Yes
		Last_remodel_date	smalldatetime	4	Yes
		Lease_sqft	float	8	Yes
		Store_sqft	float	8	Yes
		Grocery_sqft	float	8	Yes
		Frozen_sqft	float	8	Yes
		meat_sqft	float	8	Yes
		Coffee_bar	bit	1	No
		Video_store	bit	1	No
		Salad_bar	bit	1	No
		Prepared_food	bit	1	No
		Florist	bit	1	No
Product_class	Dimension Table	Product_class_id	int	4	No
		Product_subcategory	nvarchar	50	Yes
		Product_category	nvarchar	50	Yes
		Product_department	nvarchar	50	Yes
		Product_family	nvarchar	50	Yes

4.2 การออกแบบดาต้าไมน์นึ่งโมเดล

ดาต้าไมน์นึ่งโมเดลคือโครงสร้างของข้อมูลที่จะเสนอฐานความรู้ (Knowledge based) ที่ได้มาจากการวิเคราะห์ข้อมูลด้วยอัลกอริทึมของการทำไมน์นึ่งที่จะใช้ในการทำนายข้อมูลได้รายละเอียดของโครงสร้างของดาต้าไมน์นึ่งโมเดลและข้อกำหนดอื่นๆแสดงในตารางที่ 4.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นอนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.2 โครงสร้างของดาต้าไมน์นิ่งโมเดล Buying Behavior

ชนิดของแหล่งข้อมูล :	แหล่งข้อมูลจาก OLAP ที่ข้อมูลมาจากคิวบ์	
แหล่งของคิวบ์ :	คิวบ์การขาย (Sales Cube)	
เทคนิคของดาต้าไมน์นิ่ง :	เทคนิคแบบ Microsoft Decision Trees	
Case ที่ใช้ :	Case หรือตัวอย่าง (ข้อมูล) เป็นหน่วยพื้นฐานของการวิเคราะห์ข้อมูลให้กับดาต้าไมน์นิ่งโมเดล ใน OLAP นั้น โดเมนชั้นจะเสมือนเป็น Case	
	Dimension :	Customers
	Level :	Name
เอนทิตีในการทำนายข้อมูล (Member property) :	Member Card	
ข้อมูลที่ใช้ในการ Train :	เลือกช่องทางของคิวบ์ที่ประกอบด้วยข้อมูลที่จะใช้ในการทำเทรนดาต้าไมน์นิ่งโมเดล ได้แก่ โดเมนชั้นและเมเชอร์ต่างๆ เพื่อให้ไมน์นิ่งโมเดลใช้ในการเรียนรู้	
	Dimension : Customers	Member property : Gender
	Level : Name	Member property : Marital Status
		Member property : Education
		Member property : Yearly Income
		Member property : Birthdate
		Member property : Num Children At Home
		Member property : Occupation
		Member property : Houseowner
		Member property : Num Cars Owned
Dimension : Product	Level : Product Name	
Measures :	Unit Sales	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากตารางที่ 4.2 จะเลือกเทคนิคในการทำค้ำไม้หนึ่งแบบ Microsoft Decision Tress ที่มีโครงสร้างลักษณะรูปแบบต้นไม้ที่ประกอบด้วยโหนดต่างๆ โดยโหนดนั้นจะเป็นการแตกออกเป็นสาขาต่างๆ ของต้นไม้ตัวเอง ตามเงื่อนไขลักษณะของข้อมูล และมีการบอกความน่าจะเป็นในการทำนายข้อมูลที่ต้องการด้วย ส่วน Case หรือกลุ่มตัวอย่างข้อมูลที่จะใช้ในการวิเคราะห์คือ Customers เพราะต้องการวิเคราะห์พฤติกรรมของลูกค้า และใช้ Member Card เป็นส่วนในการทำนายข้อมูล เพราะต้องการจัดทำระบบบัตรสมาชิกใหม่พร้อมแคตตาล็อกแยกตามประเภทสมาชิก การเทรดค้ำไม้หนึ่งเพื่อเป็นส่วนเพิ่มเติมในการวิเคราะห์ทำนายข้อมูล รวมทั้งใช้ในการเรียนรู้ที่ผู้ใช้ทฤษฎีการวิเคราะห์พฤติกรรมการซื้อขายสินค้าในส่วนลักษณะเฉพาะที่กระทบต่อพฤติกรรมผู้บริโภคปัจจัยส่วนบุคคลได้แก่

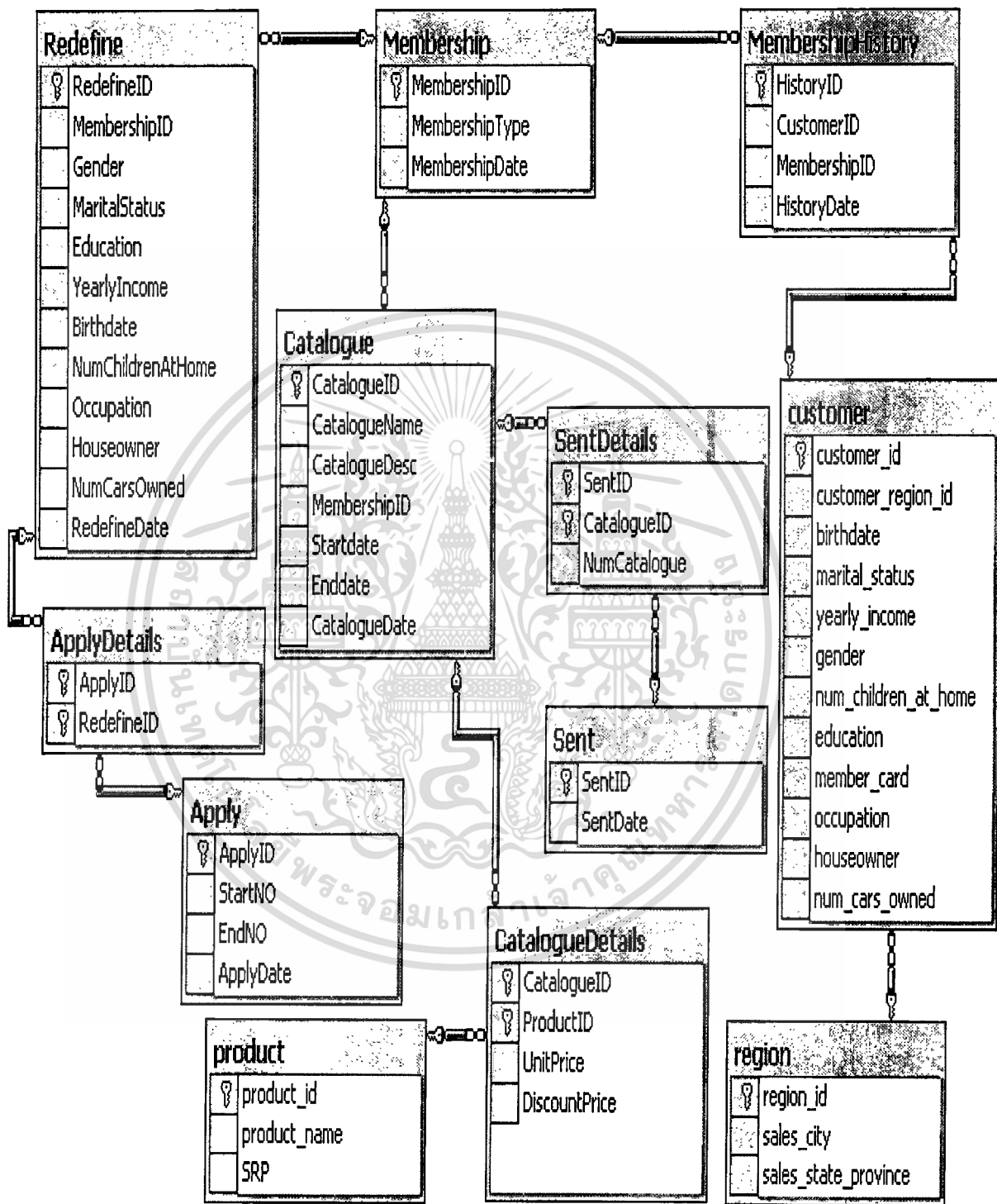
- อายุ จากตารางใช้ Member property : Birthdate
- ฐานะชีวิต จากตารางใช้ Member property : Marital Status, Gender, Num Children At Home
- อาชีพ จากตารางใช้ Member property : Occupation
- สถานะทางเศรษฐกิจ จากตารางใช้ Member property : Yearly Income, Houseowner, Num Cars Owned
- รูปแบบการดำเนินชีวิต (Lifestyle) จากตารางใช้ Member property : Education

นอกจากนั้น ยังนำข้อมูลจากการขายที่ประกอบด้วย ยอดการขายและสินค้าที่ขายนำมาทำค้ำไม้หนึ่งโมเดลเพื่อการวิเคราะห์ข้อมูลซึ่งประกอบด้วย

- ยอดการขาย จากตารางใช้ Level : Product Name
- สินค้าที่ขาย จากตารางใช้ Measures : Unit Sales

4.3 การออกแบบฐานข้อมูลเชิงสัมพันธ์

จากการวิเคราะห์ระบบ Buying Behavior Analysis System นั้นทำให้ต้องมีการออกแบบฐานข้อมูล Membership เพื่อรองรับการทำงานของระบบ ทั้งในส่วนของการจัดเก็บข้อมูลและการดึงข้อมูลมาใช้โดยสามารถแสดงความสัมพันธ์กันระหว่างตารางต่างๆ ในฐานข้อมูลด้วยการใช้รูปแบบ Database Diagram ของ Microsoft SQL Server 2000 ดังรูปที่ 4.2 และ Data Dictionary ของฐานข้อมูลดังตารางที่ 4.3 ส่วนตารางที่ 4.4 แสดง Data Dictionary ของฐานข้อมูล ThaiMart ซึ่งเป็นฐานข้อมูลที่รองรับการปฏิบัติงานของบริษัท ไทยมาร์ท จำกัด (มหาชน) โดยจะแสดงตารางและฟิลด์ข้อมูลที่เกี่ยวข้องกับระบบ Buying Behavior Analysis System เท่านั้น



รูปที่ 4.2 Database Diagram

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.3 Data Dictionary ของฐานข้อมูล Membership

ตารางที่ 1) Table 'Apply'						
Column Name	Data type	Length	Precision	Scale	Allow Nulls/PK	Description
1. ApplyID	Smallint	2	5	0	No/PK	รหัสชุดใบสมัคร บัตรสมาชิก (Identity Increment)
2. StartNO	Varchar	10	0	0	No	หมายเลขเริ่มต้น ของใบสมัครบัตร สมาชิกในรหัสชุด ใบสมัครบัตร สมาชิกนี้
3. EndNO	Varchar	10	0	0	No	หมายเลขสุดท้าย ของใบสมัครบัตร สมาชิกในรหัสชุด ใบสมัครบัตร สมาชิกนี้
4. ApplyDate	Smalldatetime	4	0	0	Yes	วัน เดือน ปี ที่ กำหนดรหัสชุดใบ สมัครบัตรสมาชิก

ตารางที่ 2) Table 'ApplyDetails'						
Column Name	Data type	Length	Precision	Scale	Allow Nulls/PK	Description
1. ApplyID	Smallint	2	5	0	No/PK	รหัสชุดใบสมัคร บัตรสมาชิก
2. RedefineID	Tinyint	1	3	0	No/PK	รหัสการกำหนด เงื่อนไขประเภท บัตรสมาชิก

ตารางที่ 4.3 Data Dictionary ของฐานข้อมูล Membership (ต่อ)

ตารางที่ 3) Table 'Catalogue'						
Column Name	Data type	Length	Precision	Scale	Allow Nulls/PK	Description
1. CatalogueID	Smallint	2	5	0	No/PK	รหัสแคตตาล็อกสินค้า (Identity Increment)
2. CatalogueName	Varchar	50	0	0	No	ชื่อแคตตาล็อกสินค้า
3. CatalogueDesc	Varchar	100	0	0	Yes	คำอธิบายแคตตาล็อกสินค้า
4. MembershipID	Tinyint	1	3	0	No	รหัสประเภทบัตรสมาชิก
5. Startdate	Smalldatetime	4	0	0	Yes	วันเดือนปีที่เริ่มใช้
6. Enddate	Smalldatetime	4	0	0	Yes	วันเดือนปีที่สิ้นสุดการใช้
7. CatalogueDate	Smalldatetime	4	0	0	Yes	วันเดือนปีที่สร้างแคตตาล็อกสินค้า

ตารางที่ 4) Table 'CatalogueDetails'						
Column Name	Data type	Length	Precision	Scale	Allow Nulls/PK	Description
1. CatalogueID	Smallint	2	5	0	No/PK	รหัสแคตตาล็อกสินค้า
2. ProductID	Int	4	10	0	No/PK	รหัสสินค้า
3. UnitPrice	Money	8	19	4	No	ราคาขาย
4. DiscountPrice	Money	8	19	4	Yes	ราคาขายแบบลด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.3 Data Dictionary ของฐานข้อมูล Membership (ต่อ)

ตารางที่ 5) Table 'Membership'						
Column Name	Data type	Length	Precision	Scale	Allow Nulls/PK	Description
1. MembershipID	Tinyint	1	3	0	No/PK	รหัสประเภทบัตรสมาชิก (Identity Increment)
2. MembershipType	Varchar	50	0	0	No	ประเภทบัตรสมาชิก
3. MembershipDate	Smalldatetime	4	0	0	Yes	วันเดือนปีที่กำหนดประเภทบัตรสมาชิก
ตารางที่ 6) Table 'MembershipHistory'						
Column Name	Data type	Length	Precision	Scale	Allow Nulls/PK	Description
1. HistoryID	Int	4	10	0	No/PK	รหัสประวัติประเภทบัตรสมาชิก (Identity Increment)
2. CustomerID	Int	4	10	0	No	รหัสลูกค้า
3. MembershipID	Tinyint	1	3	0	No	รหัสประเภทบัตรสมาชิกเดิม
4. Sent	Varchar	1	0	0	Yes	ส่งเอกสารการเปลี่ยนประเภทสมาชิกบัตร (Y=ส่งแล้ว)
5. HistoryDate	Smalldatetime	4	0	0	Yes	วันเดือนปีที่บันทึกประเภทบัตรสมาชิกเดิม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับกรใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้โดยไม่แจ้งให้ทราบ

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.3 Data Dictionary ของฐานข้อมูล Membership (ต่อ)

ตารางที่ 7) Table 'Redefine'						
Column Name	Data type	Length	Precision	Scale	Allow Nulls/ PK	Description
1. RedefineID	Tinyint	1	3	0	No/PK	รหัสการกำหนด เงื่อนไขประเภท บัตรสมาชิก (Identity Increment)
2. MembershipID	Tinyint	1	3	0	No	รหัสประเภทบัตร สมาชิก
3. Gender	Varchar	1	0	0	Yes	เพศ (N=ไม่ระบุ, F=ผู้หญิง, M=ผู้ ชาย)
4. MaritalStatus	Varchar	1	0	0	Yes	สถานะการแต่ง งาน (N=ไม่ระบุ, M=แต่งงานแล้ว, S=โสด)
5. Education	Varchar	30	0	0	Yes	ระดับการศึกษา (Not define หรือ Partial High School, Partial College, High School Degree, Bachelors Degree, ...)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.3 Data Dictionary ของฐานข้อมูล Membership (ต่อ)

ตารางที่ 7) Table 'Redefine' (ต่อ)						
6. YearlyIncome	Varchar	50	0	0	Yes	รายได้ต่อปี (\$10K - \$30K, \$30K - \$50K, \$50K - \$70K, \$70K - \$90K, \$90K - \$110K, \$110K - \$130K, \$130K - \$150K, \$150K + หรือไม่ ระบุหรือระบุตาม ช่วงรายได้)
7. Age	Varchar	10	0	0	Yes	ช่วงอายุ
8. NumChildrenAtHome	Tinyint	1	3	0	Yes	จำนวนบุตรที่พัก อาศัยด้วย
9. Occupation	Varchar	50	0	0	Yes	อาชีพ
10. Houseowner	Varchar	1	0	0	Yes	เป็นเจ้าของบ้าน (D=ไม่ระบุ, Y= เป็น, N=ไม่เป็น)
11. NumCarsOwned	Tinyint	1	3	0	Yes	จำนวนรถที่เป็น เจ้าของ
12. RedefineDate	Smalldatetime	4	0	0	Yes	วัน เดือน ปี ที่ กำหนดเงื่อนไข ประเภทบัตร สมาชิก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.3 Data Dictionary ของฐานข้อมูล Membership (ต่อ)

ตารางที่ 8) Table 'Sent'						
Column Name	Data type	Length	Precision	Scale	Allow Nulls/PK	Description
1. SentID	Smallint	2	5	0	No/PK	รหัสการส่งแคตตาล็อกสินค้า (Identity Increment)
2. SentDate	Smalldatetime	4	0	0	Yes	วันเดือนปีที่จะส่งแคตตาล็อกสินค้า
ตารางที่ 9) Table 'SentDetails'						
Column Name	Data type	Length	Precision	Scale	Allow Nulls/PK	Description
1. SentID	Smallint	2	5	0	No/PK	รหัสการส่งแคตตาล็อกสินค้า
2. CatalogueID	Smallint	2	5	0	No/PK	รหัสแคตตาล็อกสินค้า
3. NumCatalogue	Int	4	10	0	No	จำนวนแคตตาล็อกสินค้าที่จะส่ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.4 Data Dictionary ของฐานข้อมูล ThaiMart

ตารางที่ 1) Table 'customer'						
Column Name	Data type	Length	Precision	Scale	Allow Nulls/ PK	Description
1. customer_id	Int	4	10	0	No/PK	รหัสลูกค้า (Identity Increment)
2. customer_region_id	Int	4	10	0	Yes	รหัสเขตที่อยู่ลูกค้า
3. birthdate	Smalldatetime	4	0	0	Yes	วันเดือนปีเกิด
4. marital_status	Nvarchar	1	0	0	Yes	สถานะการแต่งงาน (M=แต่งงานแล้ว, S=โสด)
5. yearly_income	Nvarchar	50	0	0	Yes	รายได้ต่อปี (\$10K - \$30K, \$30K - \$50K, \$50K - \$70K, \$70K - \$90K, \$90K - \$110K, \$110K - \$130K, \$130K - \$150K, \$150K +)
6. gender	Nvarchar	1	0	0	Yes	เพศ (F=ผู้หญิง, M=ผู้ชาย)
7. num_children_at_home	Smallint	2	5	0	Yes	จำนวนบุตรที่พักอาศัยด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.4 Data Dictionary ของฐานข้อมูล ThaiMart (ต่อ)

ตารางที่ 1) Table 'customer' (ต่อ)						
8. education	Nvarchar	30	0	0	Yes	ระดับการศึกษา (Partial High School, Partial College, High School Degree, Bachelors Degree, Graduate Degree)
9. member_card	Nvarchar	50	0	0	Yes	ประเภทบัตรสมาชิก
10. occupation	Nvarchar	50	0	0	Yes	อาชีพ
11. houseowner	Nvarchar	1	0	0	Yes	เป็นเจ้าของบ้าน (Y=เป็น, N=ไม่เป็น)
12. num_cars_owned	Smallint	2	5	0	Yes	จำนวนรถที่เป็นเจ้าของ

ตารางที่ 2) Table 'product'						
Column Name	Data type	Length	Precision	Scale	Allow Nulls/PK	Description
1. product_id	Int	4	10	0	No/PK	รหัสสินค้า (Identity Increment)
2. product_name	Nvarchar	255	0	0	Yes	ชื่อสินค้า
3. SRP	Money	8	19	4	Yes	ราคาสินค้า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

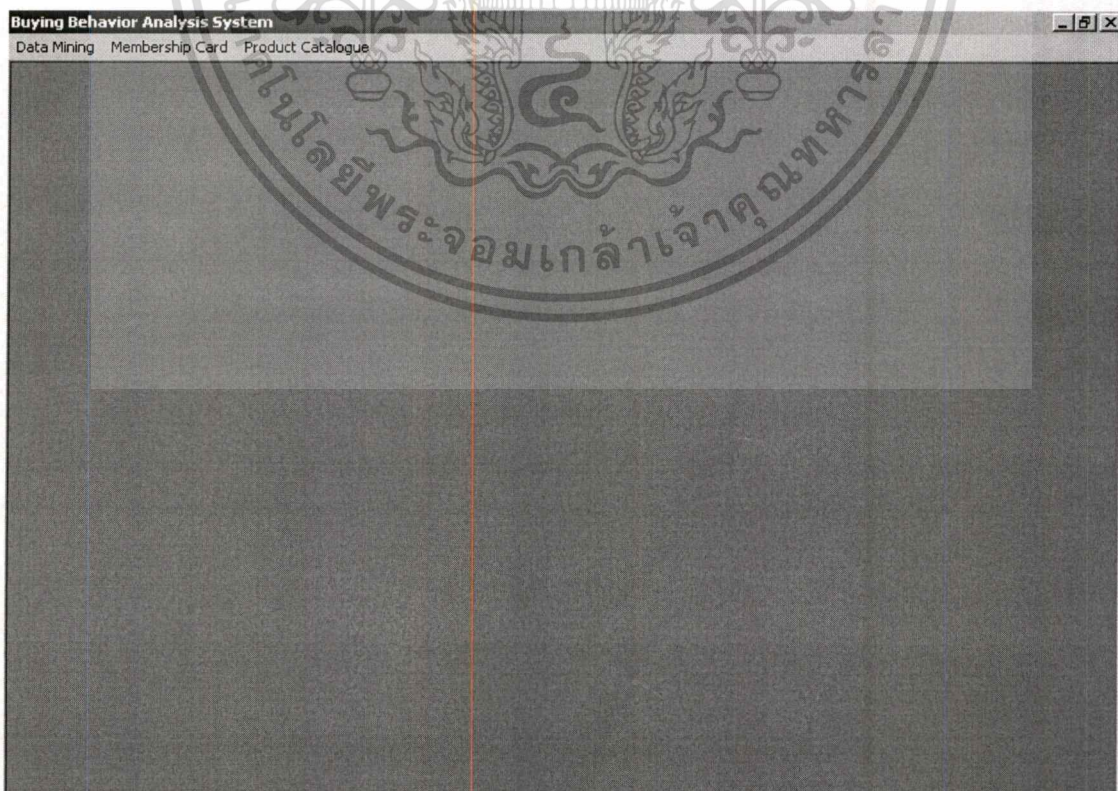
ตารางที่ 4.4 Data Dictionary ของฐานข้อมูล ThaiMart (ต่อ)

ตารางที่ 3) Table 'region'						
Column Name	Data type	Length	Precision	Scale	Allow Nulls/PK	Description
1. region_id	Int	4	10	0	No/PK	รหัสเขตที่อยู่ (Identity Increment)
2. sales_city	Nvarchar	50	0	0	Yes	เมือง
3. sales_state_province	Nvarchar	50	0	0	Yes	รัฐ

4.4 การออกแบบส่วนต่อประสานผู้ใช้

เพื่อให้ระบบ Buying Behavior Analysis System มีความสามารถในการรองรับการใช้งานจากผู้ใช้และครอบคลุมฟังก์ชันการทำงานจากการวิเคราะห์ห้ออกแบบระบบ จึงมีส่วนต่อประสานผู้ใช้หรือหน้าจอที่สำคัญดังนี้

1) หน้าจอเมนูหลัก (Buying Behavior Analysis System Menu)



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการรูปที่ 4.3 หน้าจอเมนูหลัก ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คำอธิบายการประมวลผล

เมื่อรันแอปพลิเคชัน Buying Behavior Analysis System จะปรากฏหน้าจอเมนูหลักดังรูปที่ 4.3 โดยส่วนประกอบของเมนูหลักมีดังต่อไปนี้

เมนู Data Mining ประกอบด้วยเมนูย่อย

- Reprocess... ใช้สำหรับการประมวลผลดาต้าไมน์นิ่งโมเดล
- Browse ใช้สำหรับดูข้อมูลของดาต้าไมน์นิ่งโมเดล
- Exit ใช้สำหรับออกจากระบบ

เมนู Membership Card ประกอบด้วยเมนูย่อย

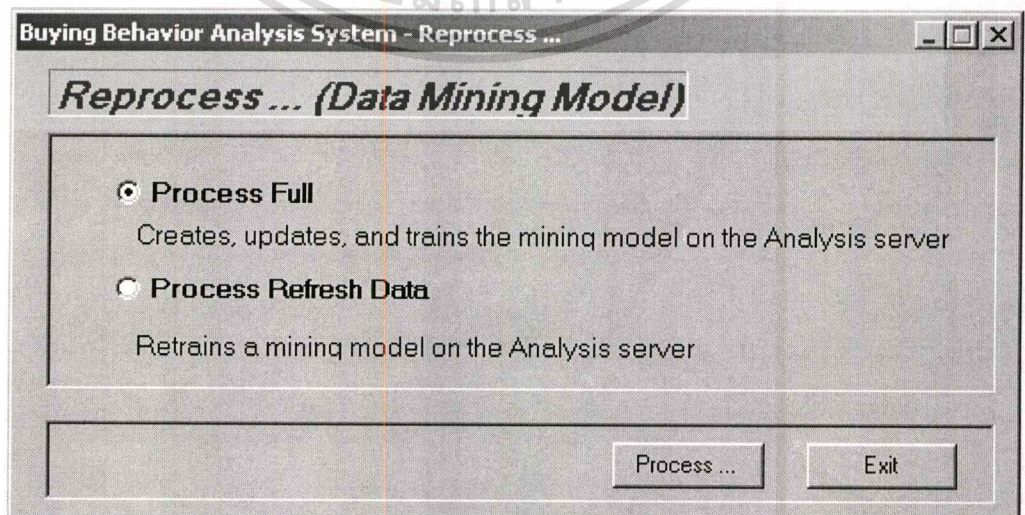
- Redefine ใช้สำหรับการกำหนดเงื่อนไขของบัตรสมาชิกประเภทต่างๆ
- Update Membership type ใช้เปลี่ยนแปลงประเภทบัตรสมาชิกของลูกค้า
- Sent ใช้ในการส่งประเภทบัตรสมาชิกใหม่ไปยังลูกค้า
- View ใช้ดูสถานะของลูกค้าในการเปลี่ยนประเภทบัตรสมาชิก
- Apply ใช้กำหนดจำนวนฟอร์มใบสมัครสมาชิกพร้อมเงื่อนไข

เมนู Product Catalogue ประกอบด้วยเมนูย่อย

- Create ใช้ในการสร้างแคตตาล็อกสินค้าสำหรับบัตรสมาชิกแต่ละประเภท
- Sent ใช้ในการจัดส่งแคตตาล็อกสินค้าไปยังลูกค้า
- View ใช้ในการดูแคตตาล็อกสินค้าที่ได้สร้างไปแล้ว

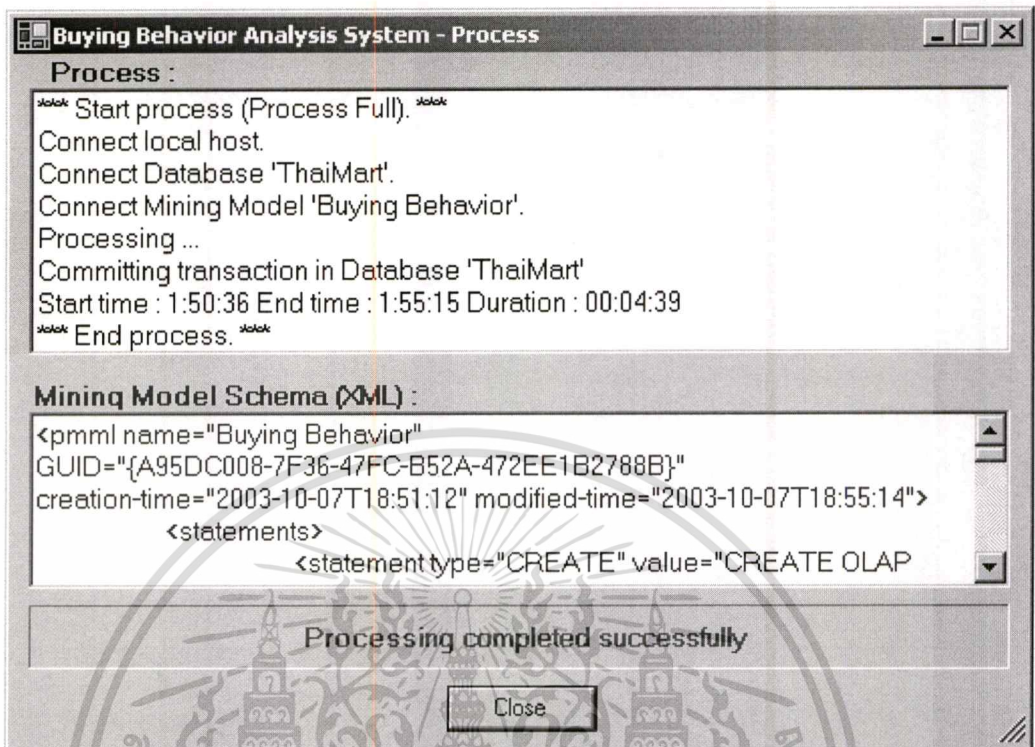
2) เมนู Data Mining (Data Mining Menu)

2.1) หน้าจอประมวลผลดาต้าไมน์นิ่งโมเดล (Reprocess)



รูปที่ 4.4 หน้าจอประมวลผลดาต้าไมน์นิ่งโมเดล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่อนักผู้ใดเห็นไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



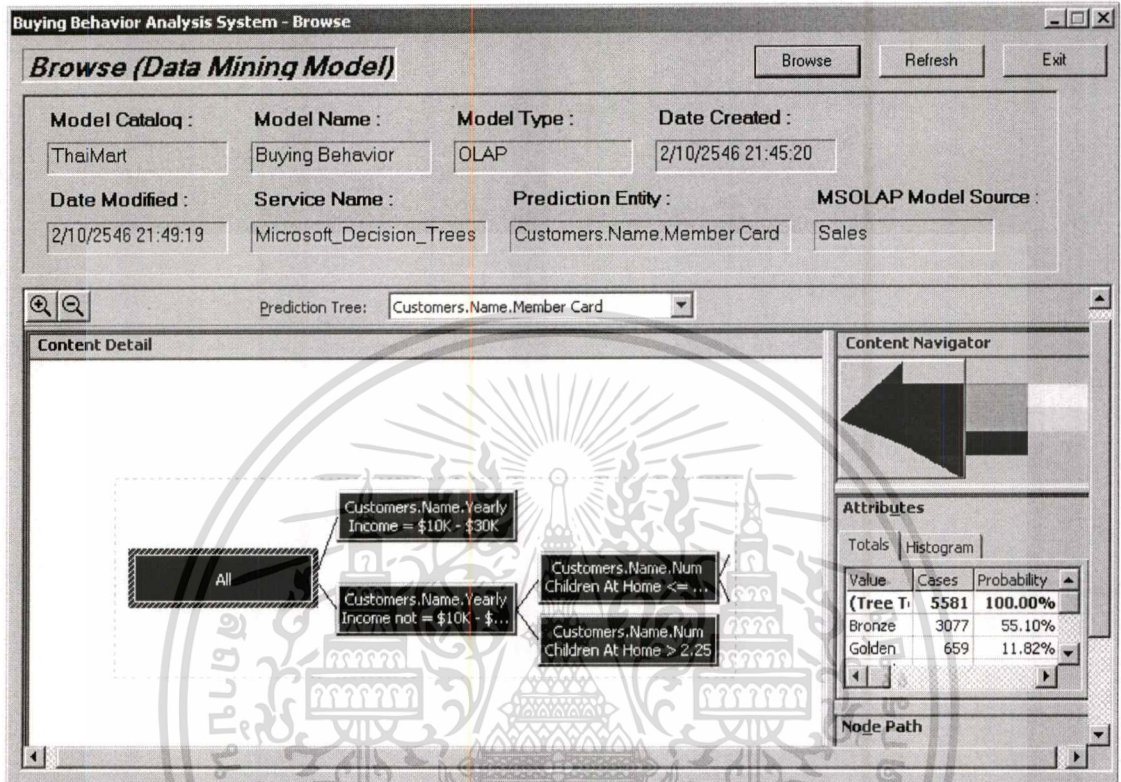
รูปที่ 4.5 หน้าจอแสดงขั้นตอนการประมวลผลดาต้าไมน์นิ่งโมเดล

คำอธิบายการประมวลผล

จากรูปที่ 4.4 การประมวลผลดาต้าไมน์นิ่งโมเดลนั้นจะมีการไปสร้าง แก๊ไข หรือ เทรนข้อมูลที่ Buying Behavior Model และจะใช้ก็ต่อเมื่อ

- มีการแก้ไขเปลี่ยนแปลง โครงสร้างของ Buying Behavior Model
 - ข้อมูลจากคิวบ์การขายมีการเปลี่ยนแปลง
- โดยการประมวลผลจะมีอยู่ 2 รูปแบบ คือ
- Process Full คือการประมวลผลที่มีการสร้าง แก๊ไขหรือเทรนข้อมูลใหม่ทั้งหมด
 - Process Refresh Data คือการเทรนข้อมูลเฉพาะที่มีการเปลี่ยนแปลงเท่านั้น
- จากนั้นจะมีการแสดงขั้นตอนการประมวลผลดาต้าไมน์นิ่งโมเดลตามรูปที่ 4.5 ซึ่ง จะแสดงโครงสร้างของดาต้าไมน์นิ่งโมเดลในรูปแบบ XML ด้วย
- ตารางข้อมูลที่เกี่ยวข้อง
- Buying Behavior Model (Update)

2.2) หน้าจอข้อมูลของดาต้าไมนนิ่งโมเดล (Browse)



รูปที่ 4.6 หน้าจอข้อมูลของดาต้าไมนนิ่งโมเดล

คำอธิบายการประมวลผล

จากรูปที่ 4.6 การดูข้อมูลของดาต้าไมนนิ่งโมเดลนั้น ส่วนบนจะเป็นข้อมูลทั่วไปของโมเดลซึ่งได้จากการใช้ MINING_MODELS Schema Rowset ที่เป็นการกำหนดคุณสมบัติในคอมโพเนนต์ที่ใช้ในการเข้าถึงโครงสร้างของโมเดล ส่วนด้านล่างของหน้าจอจะแสดง Content Detail Attributes และ Node Path

- ส่วนของ Content Detail จะแสดงโครงสร้างของโมเดลในลักษณะโครงสร้างต้นไม้
- ส่วนของ Attributes จะแสดงข้อมูลของความน่าจะเป็นในการทำนายประเภทบัตร โดยเมื่อคลิกที่โหนดของโครงสร้างต้นไม้ในส่วน Content Detail ข้อมูลจะแปรเปลี่ยนตาม
- ส่วนของ Node Path ก็เช่นกันแต่จะเป็นการแสดงเงื่อนไขข้อมูลของโหนด ตารางข้อมูลที่เกี่ยวข้อง

3) เมนู Membership Card (Membership Card Menu)

3.1) หน้าจอการกำหนดเงื่อนไขประเภทบัตรสมาชิก (Redefine)

Buying Behavior Analysis System - Redefine

Redefine

Membership ID : 1-Golden

Select Gender

Gender : Not Define Male Female

Select Marital Status

Marital Status : Not Define Single Marries

Education : Not define

Yearly Income : 6-\$110K 9-\$150K +

Age : Define 30 - 50

Number Children At Home : 0

Occupation :

Select House Owner

Houseowner : Not Define Have Not Have

Number Cars Owned : 0

Insert Exit

Ready ...

Data List

Redefine			
	Redefine ID	Membership ID	Gender
▶	1	1	N
	2	1	N
	3	1	N
	4	1	N
	5	1	M
	6	2	N
	7	2	N

รูปที่ 4.7 หน้าจอการกำหนดเงื่อนไขประเภทบัตรสมาชิก

คำอธิบายการประมวลผล

จากรูปที่ 4.7 การกำหนดเงื่อนไขประเภทบัตรสมาชิกนั้น จะเป็นการกำหนดข้อมูลที่เรียกว่า ข้อมูลลักษณะเฉพาะที่กระทบต่อพฤติกรรมผู้บริโภคในส่วนปัจจัยส่วนบุคคลโดยฝ่ายการตลาดเมื่อได้ข้อมูลจากการทำนายของคาด้าไมน์นิ่งแล้ว ก็จะนำมากำหนดโดยขึ้นอยู่กับวิจารณ์ญาณของฝ่ายการตลาดด้วย ส่วนด้านขวาที่เป็น Data List จะเป็นข้อกำหนดเงื่อนไขที่ได้มีการบันทึก (Insert) ลงฐานข้อมูลไปแล้ว

ตารางข้อมูลที่เกี่ยวข้อง

- Membership (Retrieve)
- Redefine (Insert, Retrieve)

3.2) หน้าจอการเปลี่ยนประเภทบัตรสมาชิก (Update Membership Type)

Buying Behavior Analysis System - Update Membership Type

Update Membership Type

Membership Type : 1-Golden

Gender : Male

Marital Status : Not define

Education : Not define

Yearly Income : \$150K - \$150K +

Age : Not define

Number Children At Home : 2

Occupation :

Houseowner : Have

Number Cars Owned : 0

11 Customer Updated Successfully

Update Exit

Customer Updated

รูปที่ 4.8 หน้าจอการเปลี่ยนประเภทบัตรสมาชิก

คำอธิบายการประมวลผล

จากรูปที่ 4.8 การเปลี่ยนประเภทบัตรสมาชิกจะต้องเลือกประเภทบัตรสมาชิกที่ต้องการเปลี่ยนโดยจะมีการดึงข้อกำหนดจากตารางข้อมูล Redefine ที่เป็นค่าที่กำหนดค่าสุดสำหรับแต่ละประเภทสมาชิคนั้นๆ แล้วไปทำการแก้ไขที่ตารางข้อมูล Customer ซึ่งจะแก้ไขลูกค้าที่ตรงตามเงื่อนไขแต่ประเภทบัตรไม่ตรงและมีการเก็บประวัติประเภทสมาชิกเดิมไว้ที่ตารางข้อมูล MemberHistory แต่ถ้าตรงก็จะเก็บบันทึกไว้ในตารางข้อมูล MemberHistory เช่นกัน สุดท้ายเมื่อเสร็จแล้วจะแสดงจำนวนลูกค้าที่ได้เปลี่ยนประเภทไป

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางข้อมูลที่เกี่ยวข้อง

- Membership (Retrieve)
- Redefine (Retrieve)
- Customer (Update)
- MembershipHistory (Insert)

3.3) หน้าจอการส่งเอกสารการเปลี่ยนประเภทบัตรสมาชิก (Sent Membership Inform)

Sent Membership Inform

Membership Type : 1-Golden

Region :

Region	Region Name
0	None
1	San Francisco
2	Mexico City
3	Los Angeles
4	Guadalajara

Sent Exit

Customer List :

756 , Vanderlinden , Kara , 8983 Haynes Court , Metchosis , BC , 91268 , Golden
 1608 , Edwards , Lucille , 7338 Green St. , Burbank , CA , 90503 , Golden , Gold
 1641 , Hesse , Tomas , 3937 Santa Monica , Orizaba , Veracruz , 39609 , Golde
 1837 , Martinez , Isoline , 8158 Pheasant Court , Novato , CA , 63366 , Golden ,
 5279 , Hayden , Robert , 867 Calico Way , Long Beach , CA , 40507 , Golden , C
 5995 , Roybal , Kathleen , 5586 P St. , Burlingame , CA , 27848 , Golden , Golde
 6843 , Mehaffy , Floyd , 282 Iowa Drive , Kirkland , WA , 38550 , Golden , Golde
 7639 , Ackerman , Anna , 10 Hillsborough Dr. , Albany , OR , 95420 , Golden , G
 8960 , Hull , Charles , 2796 Sudan Loop , Port Orchard , WA , 20442 , Golden , C
 9457 , Puckett , Carol , 6591 Cherokee Drive , Anacortes , WA , 69912 , Golden
 9494 , Pacheco , Kathleen , 1389 Walters Way , Long Beach , CA , 64044 , Gol

Sent 11 Customers.

รูปที่ 4.9 หน้าจอการส่งเอกสารการเปลี่ยนประเภทบัตรสมาชิก

คำอธิบายการประมวลผล

จากรูปที่ 4.9 การส่งเอกสารการเปลี่ยนประเภทบัตรสมาชิกจะต้องเลือกประเภทเอกสารนี้เป็นเอกสารบัตรสมาชิกที่จะส่ง จากนั้นเลือกเขตที่อยู่ลูกค้าโดยอาจเลือกทุกเขตก็ได้ สุดท้ายจะการคำนวณว่ากรณีใดทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แสดงข้อมูลลูกค้าที่จะต้องส่งเอกสารโดยประกอบด้วยข้อมูล รหัสลูกค้า ชื่อ ที่อยู่ ประเภทสมาชิกเดิมและประเภทสมาชิกใหม่ ซึ่งจะมีการบันทึกลงในไฟล์ข้อมูลเพื่อให้แผนกจัดส่งนำข้อมูลไปใช้รวมทั้งแก้ไขตารางข้อมูล MembershipHistory ที่ฟิลด์ Sent เพื่อเป็นการตรวจสอบได้ว่าส่งไปแล้ว

ตารางข้อมูลที่เกี่ยวข้อง

- Customer (Retrieve)
- Membership (Retrieve)
- MembershipHistory (Retrieve , Update)
- Region (Retrieve)

3.4) หน้าจอสถานะการเปลี่ยน ประเภทบัตรสมาชิก (View Customer Status)

The screenshot shows a window titled "Buying Behavior Analysis System - View Customer Status". Inside, there is a form with the following fields:

- Membership Type :** A dropdown menu currently showing "1-Golden".
- Customer Status :** A "Select Status" section with two radio buttons: "Changed" (selected) and "Not Changed".
- Region :** A list box showing a table of regions.

Region	Region Name
0	None
1	San Francisco
2	Mexico City
3	Los Angeles
4	Guadalaiaara

Below the form are "View" and "Exit" buttons. At the bottom, there is a "Customer List" table with 11 rows of data:

customer_id	Name	Address
756	Vanderlinden	8983 Haynes
1608	Edwards Luci	7338 Green S
1641	Hesse Tomas	3937 Santa M
1837	Martinez Isoli	8158 Pheasa
5279	Hayden Robe	867 Calico W
5995	Roybal Kathl	5586 P St. Bu
6843	Mehaffy Floy	282 Iowa Driv
7639	Ackerman An	10 Hillsborou
8960	Hull Charles	2796 Sudan L

At the bottom left of the window, it says "11 customers."

รูปที่ 4.10 หน้าจอสถานะการเปลี่ยนประเภทบัตรสมาชิก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นิยมนำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คำอธิบายการประมวลผล

จากรูปที่ 4.10 สถานะการเปลี่ยนประเภทบัตรสมาชิกนั้น ให้เลือกประเภทบัตรสมาชิก สถานะการเปลี่ยน และเขตที่อยู่ก็จะแสดงข้อมูลลูกค้าทั้งหมดที่ตรงตามเงื่อนไขที่ระบุ โดยข้อมูลที่แสดงนั้นก็คือ รหัสลูกค้า ชื่อ ที่อยู่

ตารางข้อมูลที่เกี่ยวข้อง

- Customer (Retrieve)
- Membership (Retrieve)
- MembershipHistory (Retrieve)
- Region (Retrieve)

3.5) หน้าจอใบสมัครบัตรสมาชิก (Apply Form)

Buying Behavior Analysis System - Apply Form

Apply Form

Total Form : 5000

Redefine :

Redefine ID	Membership ID	Gender	Marital Status	Education	Yearly Income
5	1	M	N	Not define	\$150K - \$150K
7	2	N	N	Not define	\$70K - \$130K

Condition :

Membership Type : Golden
 Gender : Male
 Yearly Income : \$150K - \$150K +
 Number Children At Home : 2
 House owner : Have
 Membership Type : Silver
 Yearly Income : \$70K - \$130K
 Number Children At Home : 1
 Start NO : 0000400001 End NO : 0000405000

Apply Record Added.

รูปที่ 4.11 หน้าจอใบสมัครบัตรสมาชิก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คำอธิบายการประมวลผล

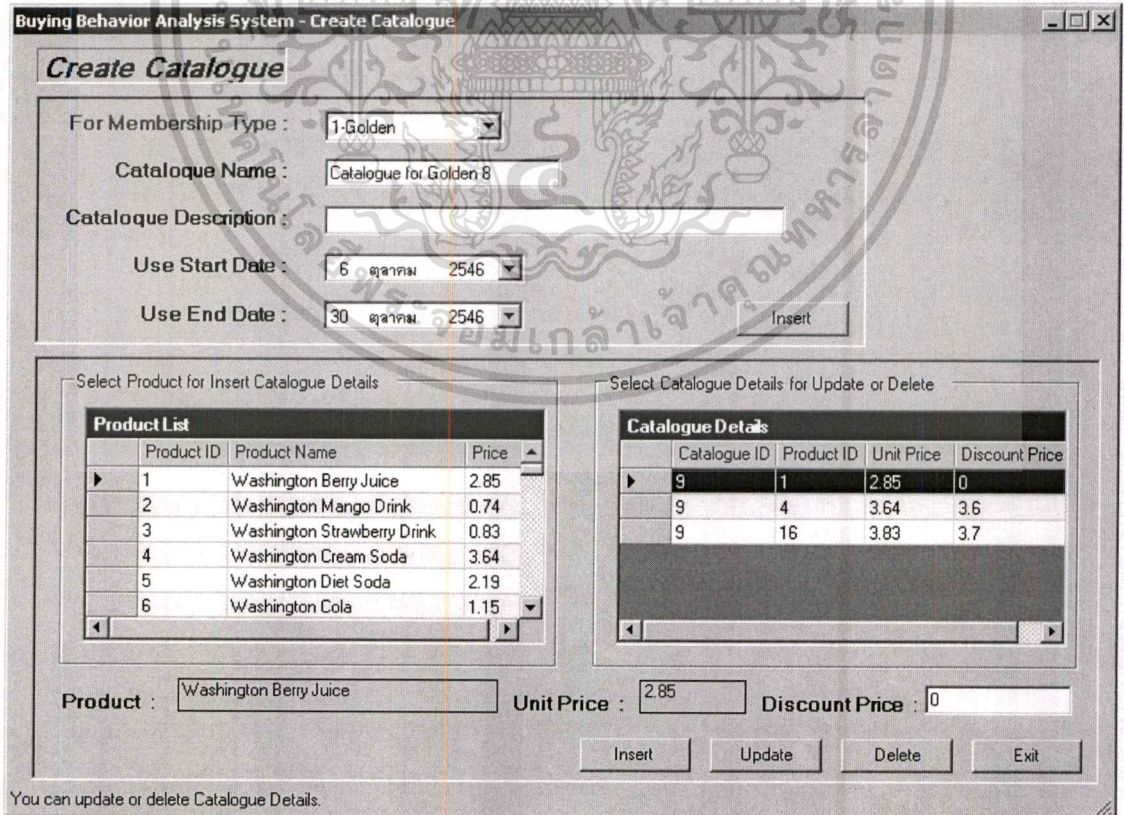
จากรูปที่ 4.11 ใบสมัครบัตรสมาชิกนั้นให้ใส่จำนวนใบสมัครที่ต้องการและจะมีการกำหนดรหัสหมายเลขเริ่มต้นและสิ้นสุดให้อัตโนมัติในรูปแบบตัวอย่าง เช่น จำนวนใบสมัครที่ต้องการ 10,000 ใบ รหัสชุดคือ 10 (ระบบกำหนดให้แบบ Identity Increment) จะได้หมายเลขดังนี้ 0001000001-0001010000 เพื่อใช้ในการตรวจสอบประวัติใบสมัครในภายหลังได้และจะมีการแสดงเงื่อนไขสำหรับบัตรทุกประเภท จากนั้นจะให้ข้อมูลนี้แก่แผนกจัดพิมพ์เพื่อจัดพิมพ์ต่อไป

ตารางข้อมูลที่เกี่ยวข้อง

- Apply (Insert)
- ApplyDetails (Insert)
- Redefine (Retrieve)

4) เมนู Product Catalogue (Product Catalogue Menu)

4.1) หน้าจอสร้างแคตตาล็อกสินค้า (Create Catalogue)



Buying Behavior Analysis System - Create Catalogue

Create Catalogue

For Membership Type : 1-Golden

Catalogue Name : Catalogue for Golden 8

Catalogue Description :

Use Start Date : 6 ตุลาคม 2546

Use End Date : 30 ตุลาคม 2546

Insert

Select Product for Insert Catalogue Details

Product ID	Product Name	Price
1	Washington Berry Juice	2.85
2	Washington Mango Drink	0.74
3	Washington Strawberry Drink	0.83
4	Washington Cream Soda	3.64
5	Washington Diet Soda	2.19
6	Washington Cola	1.15

Select Catalogue Details for Update or Delete

Catalogue ID	Product ID	Unit Price	Discount Price
9	1	2.85	0
9	4	3.64	3.6
9	16	3.83	3.7

Product : Washington Berry Juice Unit Price : 2.85 Discount Price : 0

Insert Update Delete Exit

You can update or delete Catalogue Details.

รูปที่ 4.12 หน้าจอสร้างแคตตาล็อกสินค้า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คำอธิบายการประมวลผล

จากรูปที่ 4.12 สร้างแคตตาล็อกสินค้าด้วยการกำหนดประเภทบัตรสมาชิกที่จะสร้างให้ กำหนดชื่อแคตตาล็อกสินค้า รายละเอียดเพิ่มเติม ช่วงระยะเวลาวันเดือนปีที่จะใช้โดยส่วนนี้เมื่อเพิ่ม (Insert) แล้วจะไม่สามารถเพิ่มได้อีกรวมทั้งไม่สามารถ แก้ไข (Update) และลบ (Delete) ได้ ต่อไปเมื่อต้องการเพิ่มรายการสินค้าลงในแคตตาล็อกให้คลิกเลือกรายการสินค้าได้ที่ Product List ถ้าต้องการให้มีส่วนลดให้ใส่ราคาลดได้ที่ Discount Priice การเพิ่มหรือลบรายการสินค้าออกจากแคตตาล็อกสินค้าให้เลือกรายการสินค้าที่ Catalogue Details โดยสามารถมีสิทธิทำการแก้ไขหรือลบได้ก่อน 3 วันที่จะใช้แคตตาล็อกสินค้านี้

ตารางข้อมูลที่เกี่ยวข้อง

- Catalogue (Insert)
- CatalogueDetails (Insert, Update, Delete)
- Membership (Retrieve)
- Product (Retrieve)

4.2) หน้าจอจัดส่งแคตตาล็อกสินค้า (Sent Catalogue)

Buying Behavior Analysis System - Sent Catalogue

Sent Catalogue

Catalogue For : 1-Golden

Catalogue Name : Catalogue for Golden 8

Catalogue Description :

Use Start Date : 6/10/2546 19:48:00

Use End Date : 30/10/2546 19:47:00

Catalogue Details :

Catalogue Details				
	Catalogue ID	Product ID	Unit Price	Discount Price
▶	9	1	2.85	0
	9	4	3.64	3.6
	9	16	3.83	3.7

Sent Date : 3 ตุลาคม 2546

Number of Catalogue : 2460

Sent Exit

Sent Record Added.

รูปที่ 4.13 หน้าจอจัดส่งแคตตาล็อกสินค้า

คำอธิบายการประมวลผล

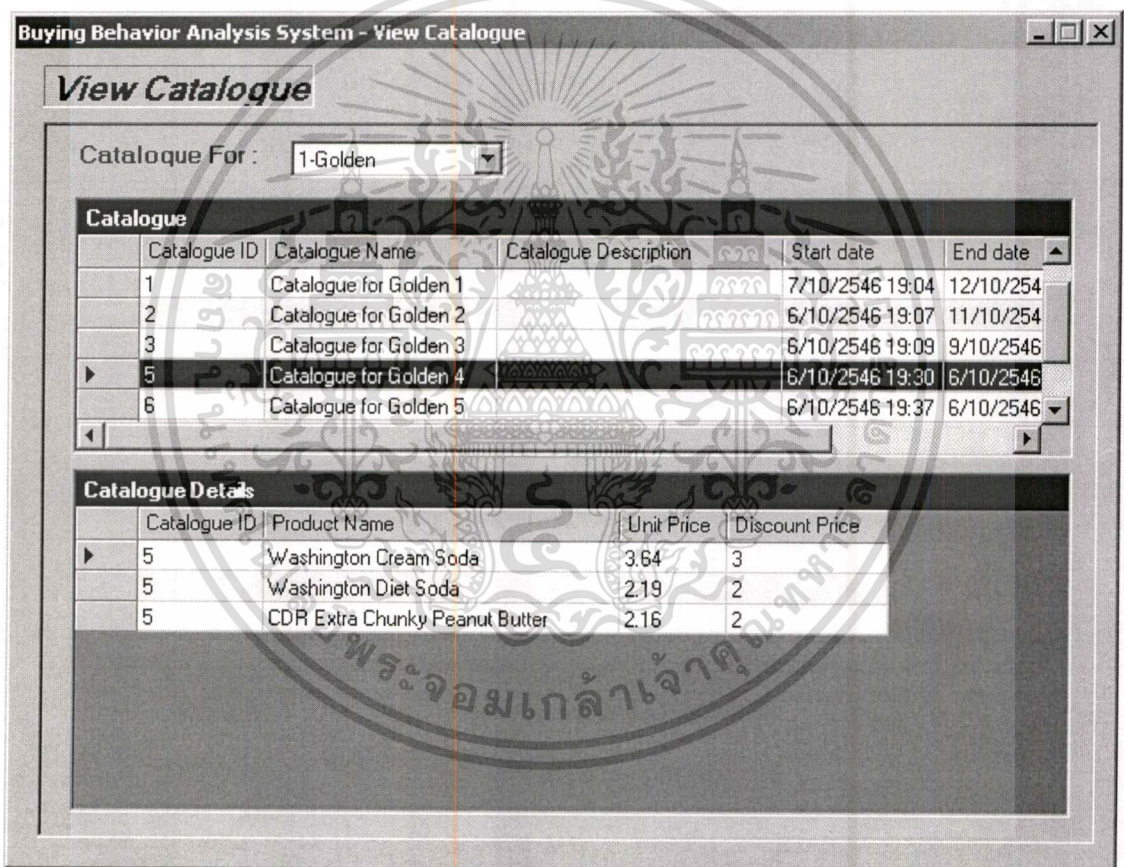
จากรูปที่ 4.13 จัดส่งแคตตาล็อกสินค้าให้เลือกระเภทบัตรสมาชิกและจะแสดงข้อมูลของแคตตาล็อกสินค้าที่เป็นแคตตาล็อกล่าสุดสำหรับประเภทบัตรสมาชิกที่ได้เลือกออกมา หลังจากนั้นจะกำหนดวันที่จะส่ง โดยมีเงื่อนไขว่าต้องส่งก่อนที่จะถึงวันที่ต้องใช้อย่างน้อย 2 วัน เมื่อทำการคลิกปุ่มส่ง (Sent) จะมีการไปนับจำนวนลูกค้าที่เป็นสมาชิกบัตรประเภทนั้นแล้วไปบันทึกลงที่ตารางข้อมูล SentDetails

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางข้อมูลที่เกี่ยวข้อง

- Catalogue (Retrieve)
- CatalogueDetails (Retrieve)
- Customer (Retrieve)
- Sent (Insert)
- SentDetails (Insert)

4.3) หน้าจอดูแคตตาล็อกสินค้า (View Catalogue)



รูปที่ 4.14 หน้าจอดูแคตตาล็อกสินค้า

คำอธิบายการประมวลผล

จากรูปที่ 4.14 ดูแคตตาล็อกสินค้าให้เลือกประเภทบัตรสมาชิกและจะแสดงข้อมูลของรายการแคตตาล็อกสินค้าออกมา เมื่อคลิกเลือกรายการใน Catalogue จะแสดงรายละเอียดของแคตตาล็อกสินค้าที่ Catalogue Details โดยประกอบด้วยข้อมูล รหัสสินค้า ชื่อสินค้า (ดึงมาจากตารางข้อมูล Product) ราคา และส่วนลด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นอนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางข้อมูลที่เกี่ยวข้อง

- Catalogue (Retrieve)
- CatalogueDetails (Retrieve)
- Membership (Retrieve)
- Product (Retrieve)



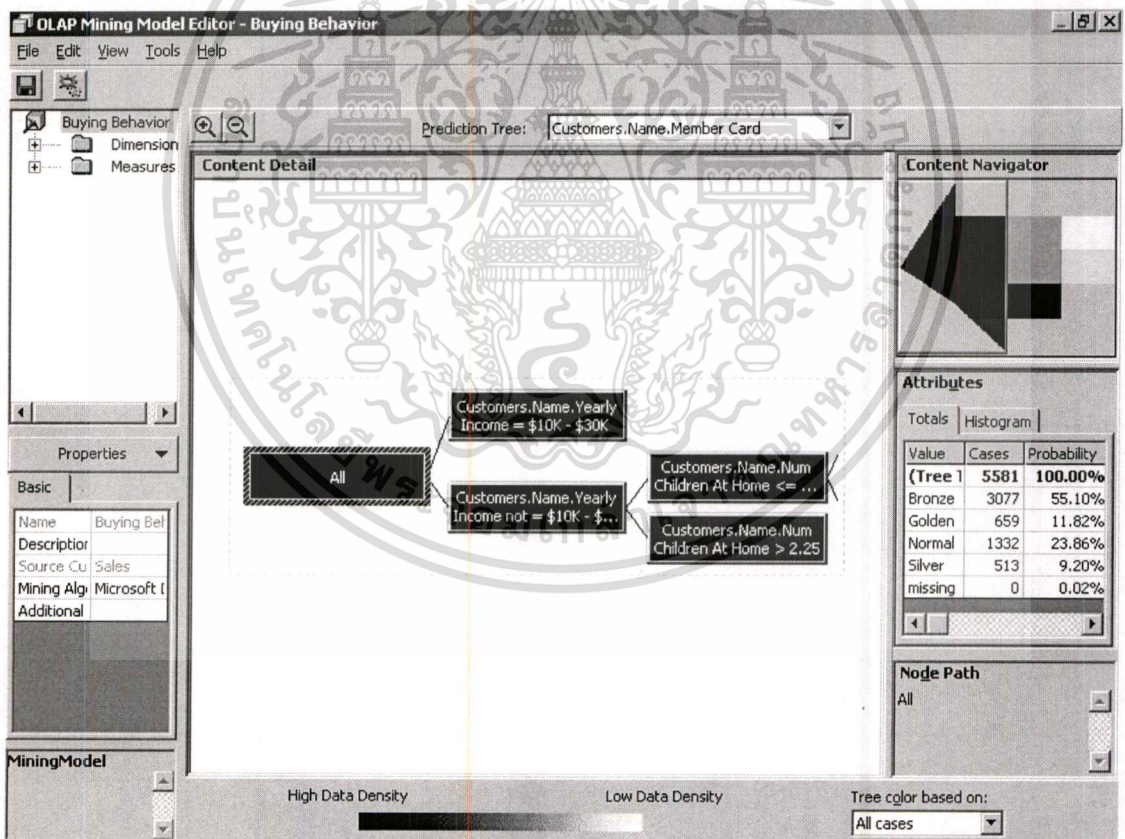
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

การพัฒนาระบบ

5.1 การพัฒนาดาต้าไมน์นิ่ง

ดาต้าไมน์นิ่งที่จะพัฒนานั้นใช้ผลิตภัณฑ์ของ Microsoft Analysis Services ซึ่งเป็นส่วนหนึ่งที่มาพร้อมกับ Microsoft SQL Server 2000 การสร้างโมเดลจะเลือก New Mining Model ... ที่ฐานข้อมูล ThaiMart แล้วกำหนดข้อมูลต่างๆ ตามที่ได้ออกแบบไว้ ขั้นสุดท้ายให้ตั้งชื่อโมเดลว่า Buying Behavior ดังรูปที่ 5.1

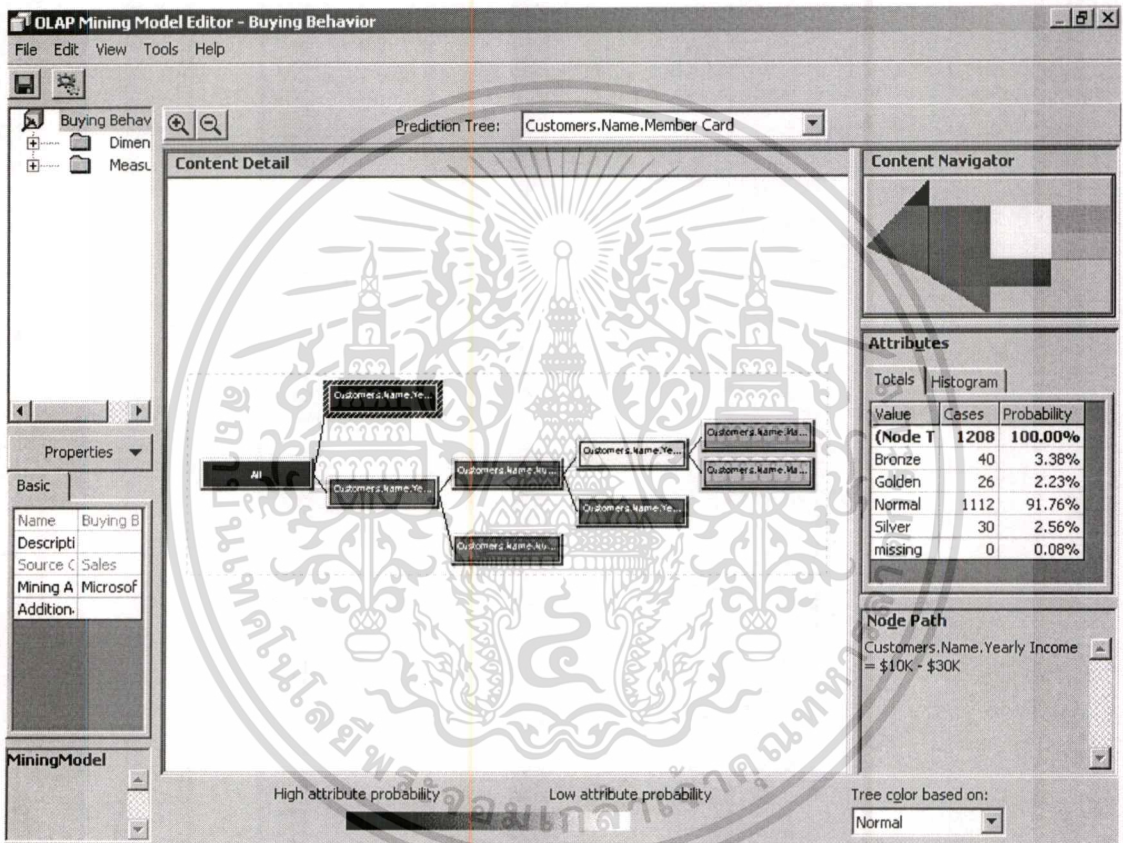


รูปที่ 5.1 Buying Behavior Model

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากคาด้าไมน์นี้โมเดล Buying Behavior นั้น

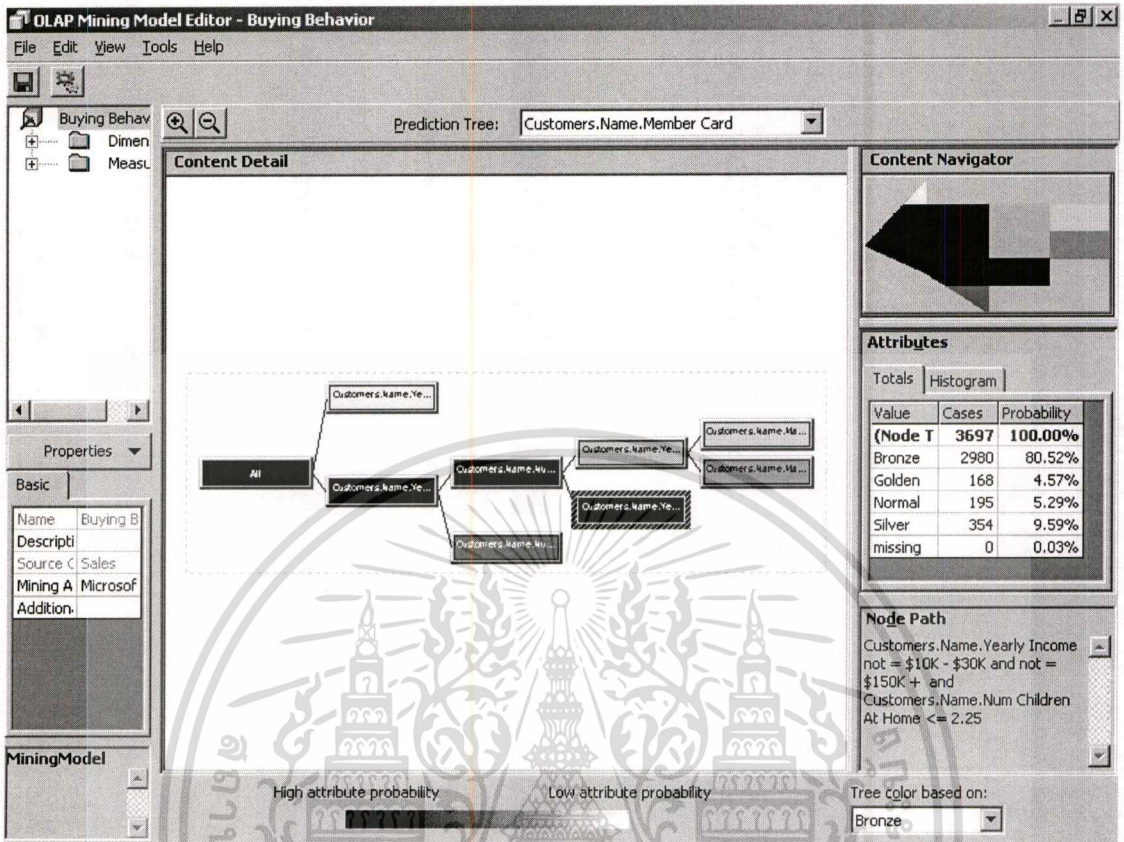
- เมื่อเลือก Tree color based on แบบบัตรสมาชิกประเภทธรรมดา (Normal) ตัวโหนดจะแสดงสีเข้มแล้วไปคลิกเลือกในระดับล่าง (Leaf) จะได้การทำนายความน่าจะเป็น 91.76 เปอร์เซ็นต์บนเงื่อนไขที่ว่า รายได้ต่อปีของลูกค้าอยู่ในช่วง \$10k และ \$30k (Customers.Name.Yearly Income = \$10K - \$30K) ดังรูปที่ 5.2



รูปที่ 5.2 Buying Behavior Model เมื่อเลือก Tree color based on แบบบัตรสมาชิกประเภทธรรมดา

- เมื่อเลือก Tree color based on แบบบัตรสมาชิกประเภททองแดง (Bronze) ตัวโหนดจะแสดงสีเข้มแล้วไปคลิกเลือกในระดับล่าง (Leaf) จะได้การทำนายความน่าจะเป็น 80.52 เปอร์เซ็นต์บนเงื่อนไขที่ว่า รายได้ต่อปีของลูกค้าอยู่ในช่วงมากกว่า \$30k และน้อยกว่า \$150k+ และจำนวนบุตรที่อาศัยอยู่ด้วยน้อยกว่าหรือเท่ากับ 2.25 คน (Customers.Name.Yearly Income not = \$10K - \$30K and not = \$150K + and Customers.Name.Num Children At Home <= 2.25) ดังรูปที่ 5.3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 5.3 Buying Behavior Model เมื่อเลือก Tree color based on แบบบัตรสมาชิกประเภททองแดง

- เมื่อเลือก Tree color based on แบบบัตรสมาชิกประเภทเงิน (Silver) ตัวโหนดจะแสดงสีเข้มแล้วไปคลิกเลือกในระดับล่าง (Leaf) จะได้การทำนายความน่าจะเป็น 86.67 เปอร์เซ็นต์บนเงื่อนไขที่ว่า รายได้ต่อปีของลูกค้าอยู่ในช่วงเท่ากับ \$150k+ และจำนวนบุตรที่อาศัยอยู่ด้วยน้อยกว่าหรือเท่ากับ 2.25 คน และสถานะการแต่งงานเป็นโสด (Customers.Name.Yearly Income = \$150K + and Customers.Name.Num Children At Home <= 2.25 and Customers.Name.Marital Status = S) ดังรูปที่ 5.4

OLAP Mining Model Editor - Buying Behavior

Prediction Tree: Customers.Name.Member Card

Content Detail

Content Navigator

Attributes

Value	Cases	Probability
(Node T	55	100.00%
Bronze	3	6.67%
Golden	1	3.33%
Normal	0	1.67%
Silver	51	86.67%
missing	0	1.67%

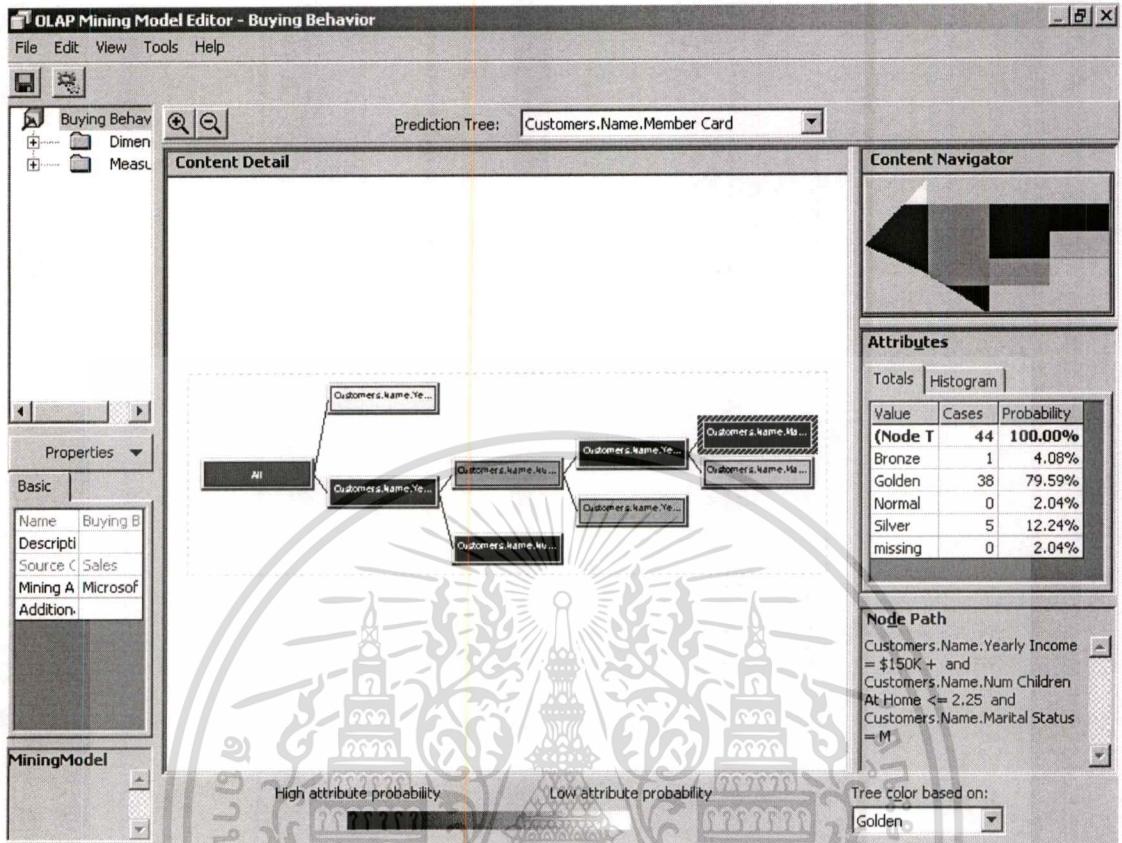
Node Path

Customers.Name.Yearly Income = \$150K + and Customers.Name.Num Children At Home <= 2.25 and Customers.Name.Marital Status = 5

Tree color based on: Silver

รูปที่ 5.4 Buying Behavior Model เมื่อเลือก Tree color based on แบบบัตรสมาชิกประเภทเงิน

- เมื่อเลือก Tree color based on แบบบัตรสมาชิกประเภททอง (Golden) ตัวโหนดจะแสดงสีเข้มแล้วไปคลิกเลือกในระดับล่าง (Leaf) จะได้การทำนายความน่าจะเป็น 79.59 เปอร์เซ็นต์บนเงื่อนไขที่ว่า รายได้ต่อปีของลูกค้าอยู่ในช่วงมากกว่า \$150k+ และจำนวนบุตรที่อาศัยอยู่ด้วยน้อยกว่าหรือเท่ากับ 2.25 คน และสถานะการแต่งงานเป็นแต่งงานแล้ว (Customers.Name.Yearly Income = \$150K + and Customers.Name.Num Children At Home <= 2.25 and Customers.Name.Marital Status = M) ดังรูปที่ 5.5



รูปที่ 5.5 Buying Behavior Model เมื่อเลือก Tree color based on แบบบัตรสมาชิกประเภททอง

สรุปผลการทำนายข้อมูลจากการพัฒนาตัวโมเดลหนึ่งที่ได้ใส่ข้อมูลที่ใช้ในการเทรนตัวโมเดลนั้นไม่ว่าจะเป็น เพศ (Gender) สถานะการแต่งงาน (Marital Status) การศึกษา (Education) รายได้ต่อปี (Yearly Income) อายุ (Birthdate) จำนวนบุตรที่พักอาศัยอยู่ด้วย (Num Children At Home) อาชีพ (Occupation) มีบ้านเป็นของตนเอง (Houseowner) และจำนวนรถที่เป็นเจ้าของ (Num Cars Owned) นั้น จะมีเฉพาะข้อมูล สถานะการแต่งงาน รายได้ต่อปี และจำนวนบุตรที่พักอาศัยอยู่ด้วยเท่านั้นที่มีอิทธิพลต่อการจัดแบ่งประเภทบัตรสมาชิก

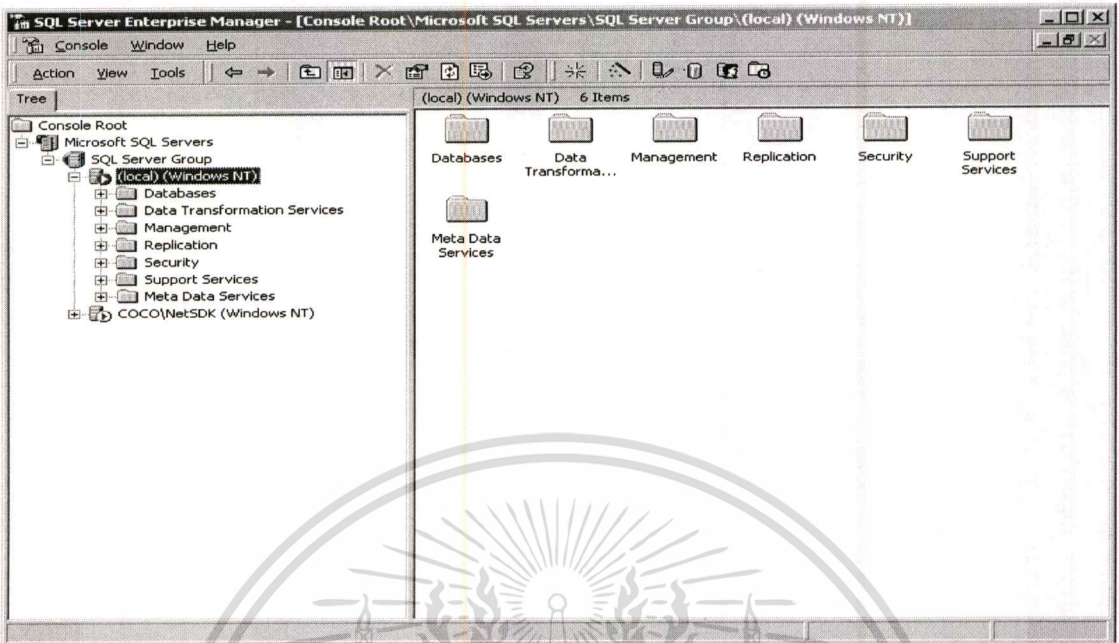
5.2 การพัฒนาระบบฐานข้อมูล

ระบบฐานข้อมูลที่จะรองรับการพัฒนาแอปพลิเคชันนั้นใช้ผลิตภัณฑ์ของ Microsoft SQL Server 2000 รุ่น Standard Edition ดังแสดงในรูปที่ 5.6 จากนั้นจะมีการสร้างฐานข้อมูลที่ชื่อว่า Membership ตามที่ได้ออกแบบไว้ดังรูปที่ 5.7 ในส่วน General ที่ได้มีการระบุชื่อตรงฟิลด์ Name ว่า Membership รูปที่ 5.8 ในส่วน Data Files นั้นเป็นไฟล์ที่สำหรับเก็บข้อมูลมีการกำหนดคุณสมบัติของไฟล์ข้อมูลทั้งในส่วน Logical และ Physical ดังนี้

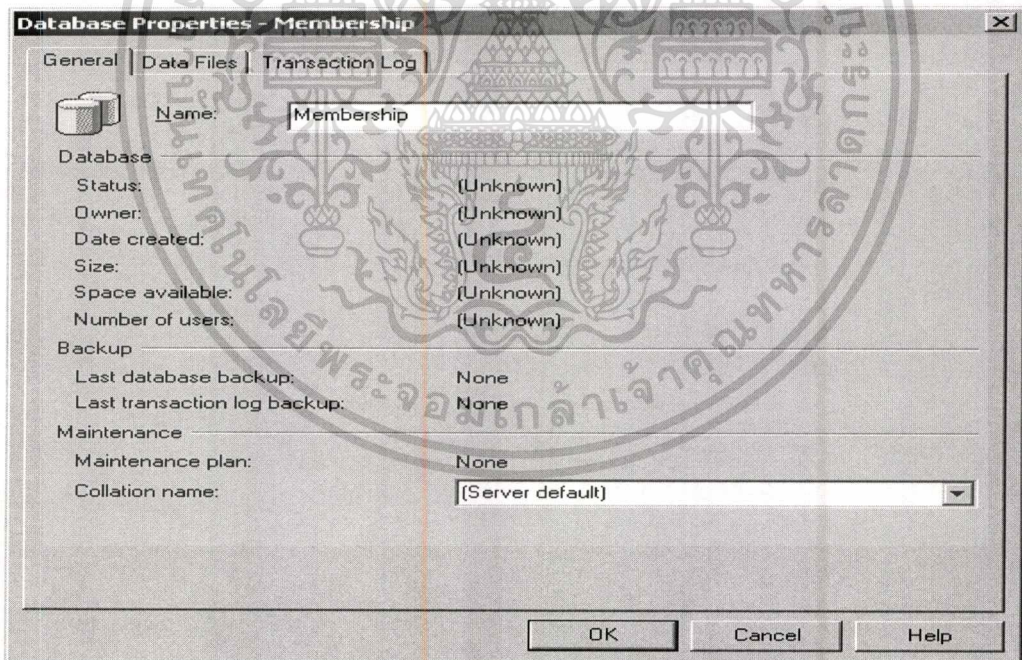
- File Name (ชื่อไฟล์ที่ไว้เก็บข้อมูล) : Membership_Data
- Location (ที่ที่เก็บไฟล์ข้อมูล) : D:\My WorkShop\MS.(IT)....
- Initial size (ขนาดเริ่มต้นของไฟล์ข้อมูล) : 1 MB
- File Group (กลุ่มของไฟล์ที่เก็บข้อมูล) : PRIMARY
- File Properties (คุณสมบัติของไฟล์ที่เก็บข้อมูล) : กำหนดให้มีการเพิ่มขนาดอัตโนมัติด้วยการเลือก Automatically grow file ในรูปแบบเพิ่มเปอร์เซ็นต์ครั้งละ 10 เปอร์เซ็นต์ ด้วยการเลือก By percent 10 และไม่จำกัดการเพิ่มขนาดของไฟล์ด้วยการเลือก Unrestricted file growth

รูปที่ 5.9 ในส่วน Transaction Log นั้นเป็นไฟล์ที่สำหรับเก็บล็อก (Log) มีการกำหนดคุณสมบัติของไฟล์ล็อกทั้งในส่วน Logical และ Physical ดังนี้

- File Name (ชื่อไฟล์ที่ไว้เก็บล็อก) : Membership_Log
- Location (ที่ที่เก็บไฟล์ล็อก) : D:\My WorkShop\MS.(IT)....
- Initial size (ขนาดเริ่มต้นของไฟล์ล็อก) : 1 MB
- File Properties (คุณสมบัติของไฟล์ที่เก็บล็อก) : กำหนดให้มีการเพิ่มขนาดอัตโนมัติด้วยการเลือก Automatically grow file ในรูปแบบเพิ่มเปอร์เซ็นต์ครั้งละ 10 เปอร์เซ็นต์ ด้วยการเลือก By percent 10 และไม่จำกัดการเพิ่มขนาดของไฟล์ด้วยการเลือก Unrestricted file growth

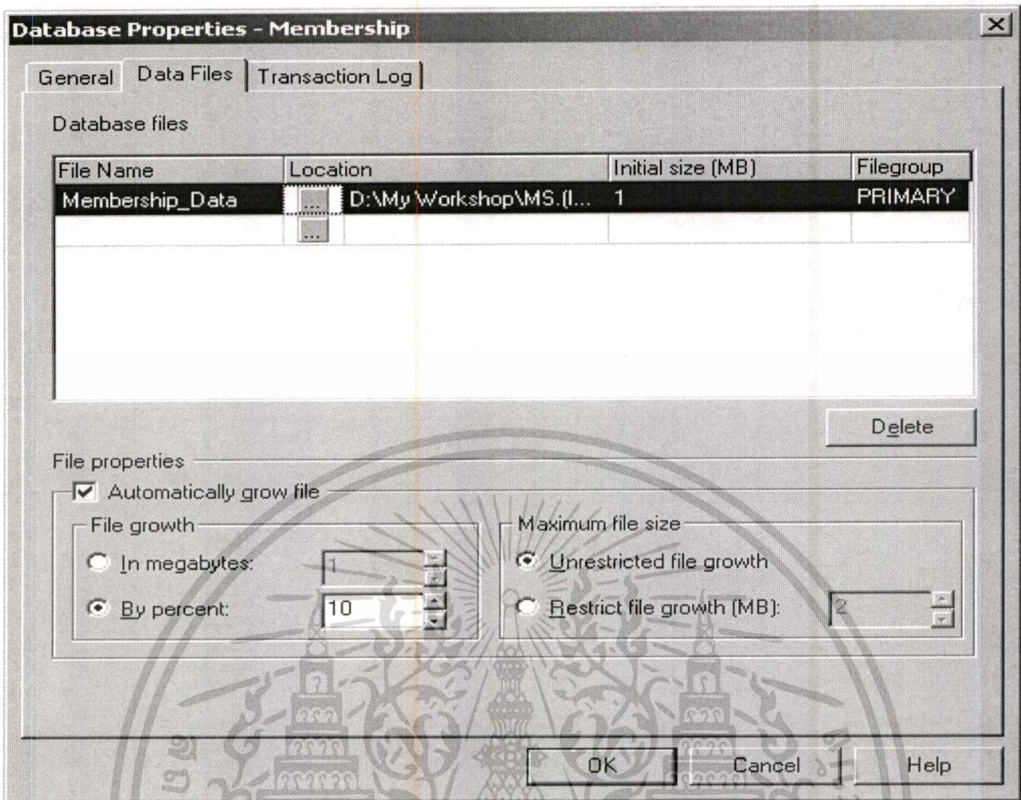


รูปที่ 5.6 Microsoft SQL Server 2000 รุ่น Standard Edition

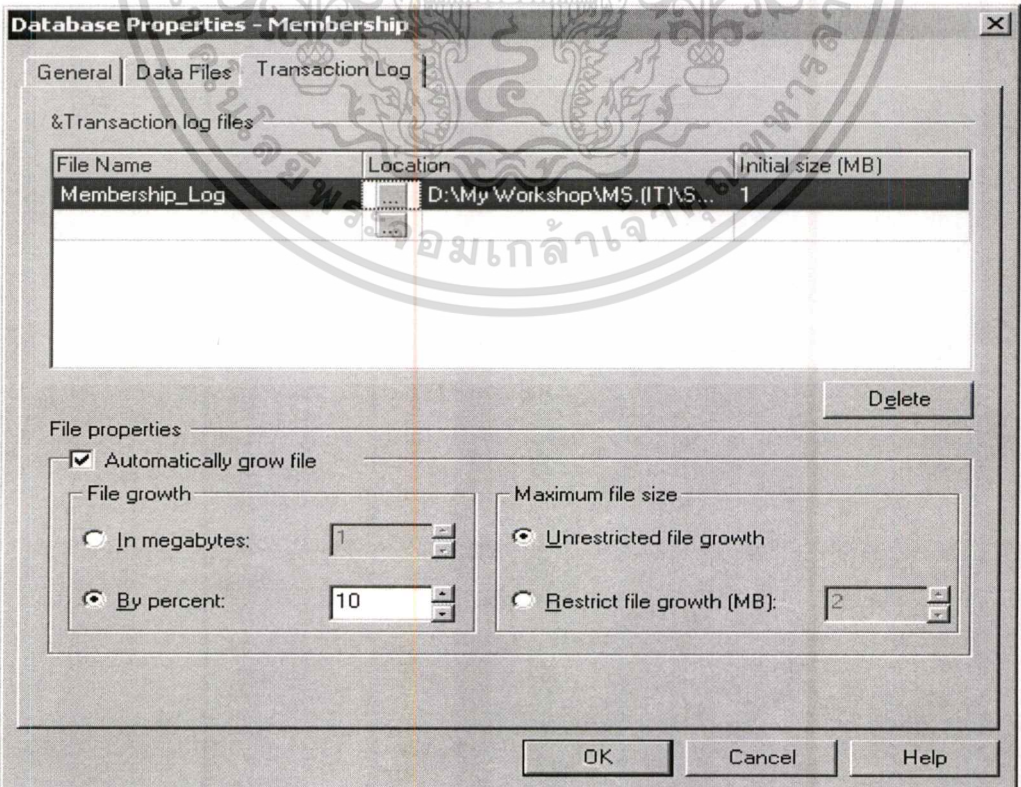


รูปที่ 5.7 การสร้างฐานข้อมูล Membership

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 5.8 การกำหนดคุณสมบัติของไฟล์ข้อมูล

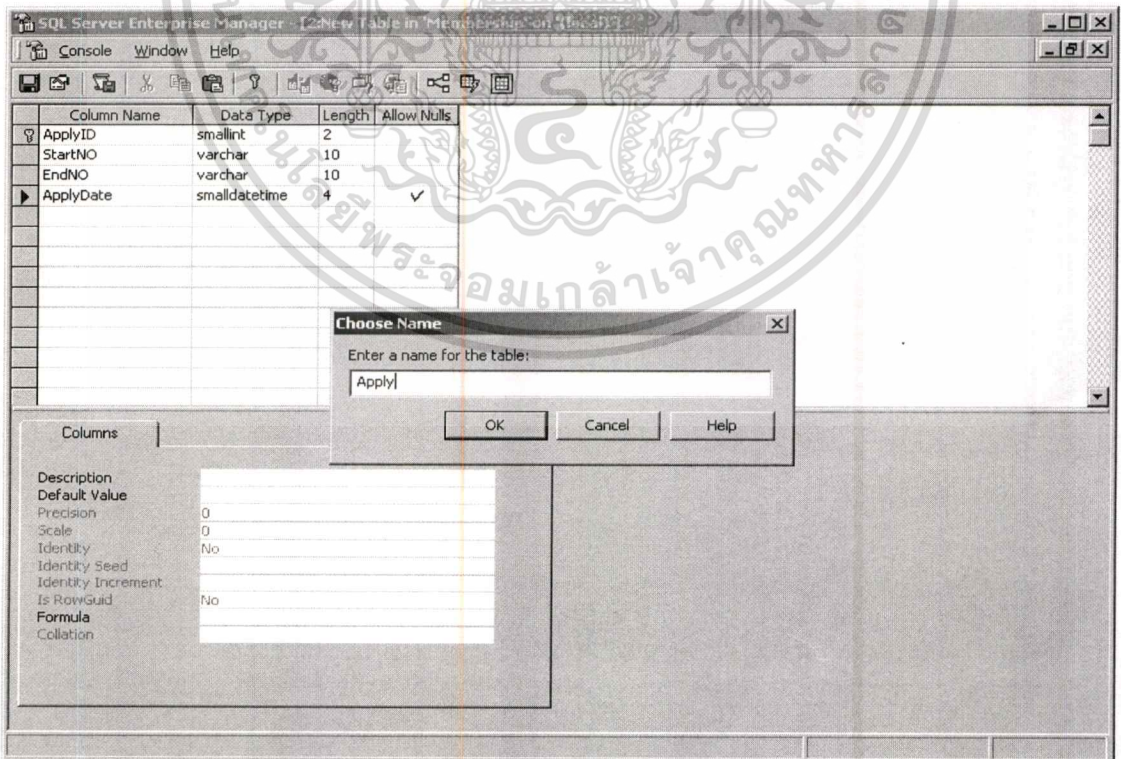


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับรูปที่ 5.9 การกำหนดคุณสมบัติของไฟล์ล็อก ตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การสร้างตารางข้อมูล (Table) ในฐานข้อมูล Membership นั้น จะเลือกเมนู New และ Table ... ที่ฐานข้อมูล Membership แล้วกำหนดชื่อคอลัมน์ (Column Name) ชนิดข้อมูล (Data type) และให้เป็นค่า Null และ/หรือเป็น Primary key (Allow Nulls/PK) ตาม Data Dictionary ที่ได้ออกแบบไว้สำหรับฐานข้อมูล Membership ดังแสดงในรูปที่ 5.10 ส่วนตารางข้อมูลที่ต้องสร้างได้แก่

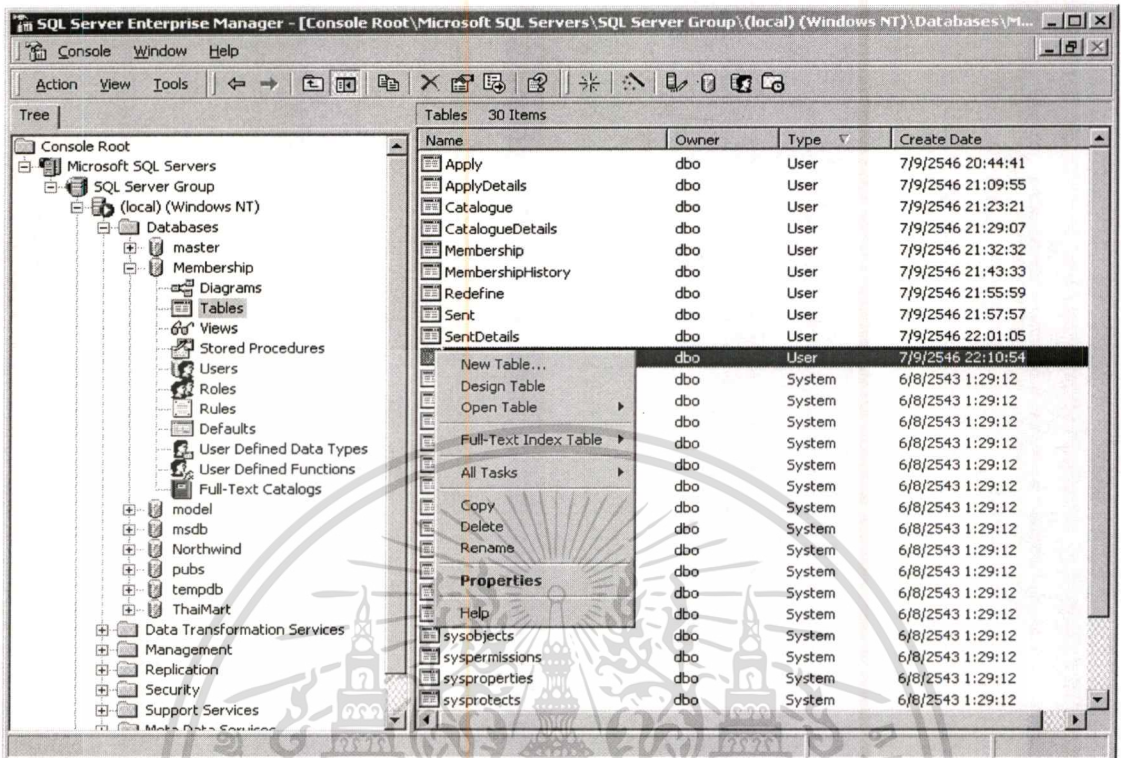
1. Apply
2. ApplyDetails
3. Catalogue
4. CatalogueDetails
5. Membership
6. MembershipHistory
7. Redefine
8. Sent
9. SentDetails

เมื่อสร้างตารางข้อมูลทั้งหมดเสร็จแล้วจะ ได้ดังรูปที่ 5.11



รูปที่ 5.10 การสร้างตารางข้อมูลในฐานข้อมูล Membership

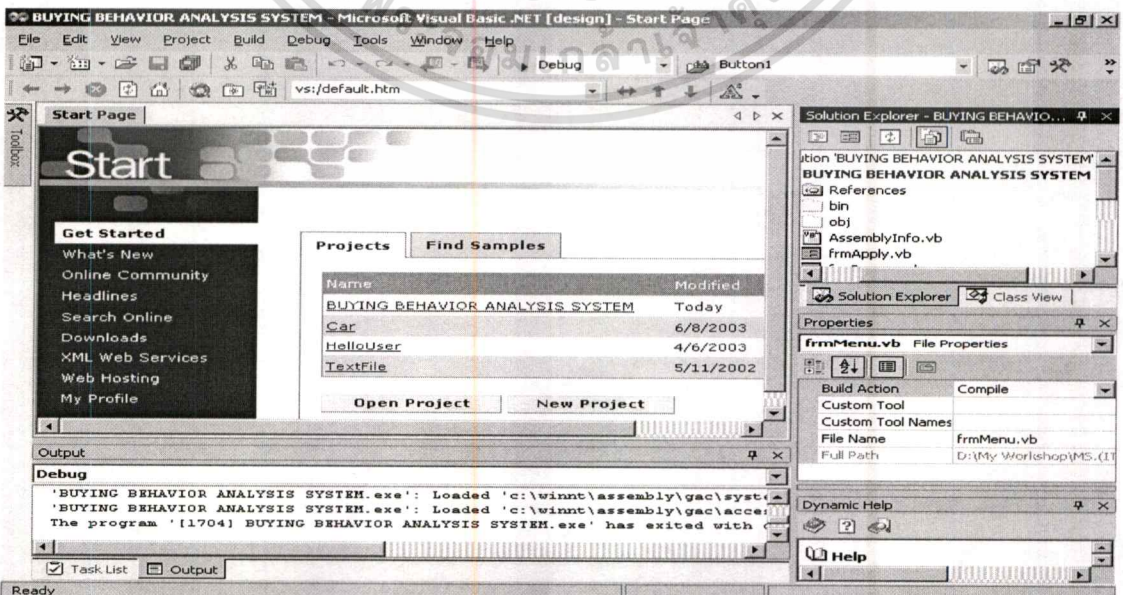
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 5.11 รายการตารางข้อมูลทั้งหมดในฐานข้อมูล Membership

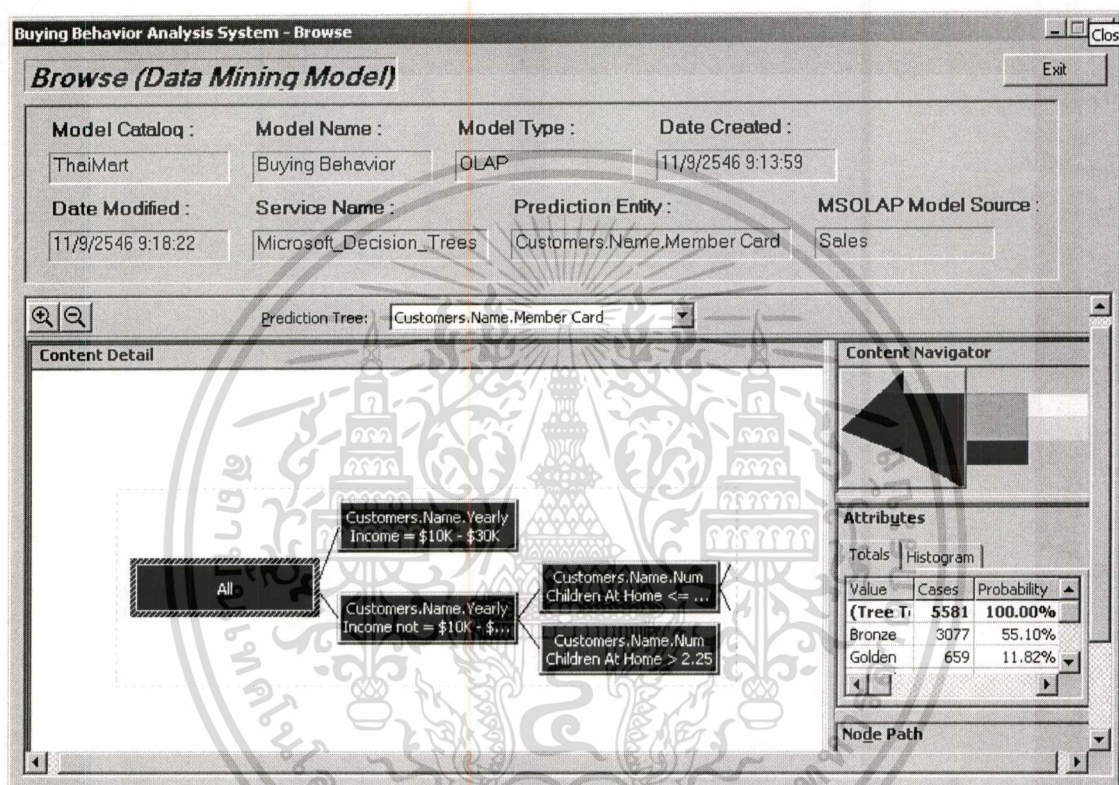
5.3 การพัฒนาแอปพลิเคชัน

เครื่องมือในการพัฒนาแอปพลิเคชันนั้นใช้ผลิตภัณฑ์ของ Microsoft Visual Studio .Net โดยใช้ภาษา Visual Basic .NET ดังแสดงในรูปที่ 5.12



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับรูปที่ 5.12 Microsoft Visual Studio .Net ญาติให้ผ่านไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อพัฒนาระบบตามที่ได้ออกแบบไว้ จะได้โปรแกรมต่างๆ เช่น โปรแกรมดูข้อมูลของ ค้าค้าไม้นิ่งโมเดล (Browse) ตามที่ได้ออกแบบโมเดลไว้จะแสดงดังรูปที่ 5.13 ซึ่งจะแสดงข้อมูล ทัวไปของคค้าไม้นิ่ง Buying Behavior ในด้านบน ส่วนของรายละเอียดใน Content Detail Attributes และ Node Path แสดงในส่วนล่าง สามารถคลิกเลือกเปลี่ยนได้ตามที่ต้องการ



รูปที่ 5.13 โปรแกรมดูข้อมูลของคค้าไม้นิ่งโมเดล

5.4 สรุปผลการพัฒนาระบบ

หลังจากที่ได้ผ่านการพัฒนาระบบในขั้นตอนต่างๆ เช่น การพัฒนาฐานข้อมูล การพัฒนา ค้าค้าไม้นิ่งโมเดล จนถึงการพัฒนาแอปพลิเคชัน ทำให้เกิดการเรียนรู้และทักษะในการใช้ Microsoft SQL Server 2000 ที่ใช้ในการรองรับการทำงานด้านระบบฐานข้อมูล Microsoft Analysis Services ที่ใช้ในการรองรับการทำงานด้านคค้าไม้นิ่ง และ Microsoft Visual Basic .NET ที่ใช้ในการสร้างส่วนต่อประสานผู้ใช้รวมทั้งการประมวลผลต่างๆ นอกจากนี้ทำให้มุมมองภาพโดยรวมของระบบที่การออกแบบระบบต้องออกแบบมาให้รองรับความต้องการของระบบและสามารถที่จะพัฒนาระบบได้ด้วย

บทที่ 6

บทสรุป

6.1 สรุปโครงการ

โครงการพัฒนาระบบงานนี้ประสบความสำเร็จตามความมุ่งหมายที่ได้ตั้งไว้ทั้งในส่วนทฤษฎีและปฏิบัติ คือ

- เป็นโครงการพัฒนาระบบงานที่ออกแบบและสร้างร้านค้าไม้นั่งโมเดลมาสนับสนุนการตัดสินใจในการจัดแบ่งประเภทบัตรสมาชิกของร้านค้าปลีก รวมทั้งมีแอปพลิเคชันที่รองรับการทำงานของผู้ใช้ในการจัดการร้านค้าไม้นั่งโมเดล การกำหนดเงื่อนไขของบัตรประเภทสมาชิก และการสร้างแคตตาล็อกสินค้าสำหรับบัตรสมาชิกแต่ละประเภท
- มีการศึกษาทฤษฎีของการวิเคราะห์พฤติกรรมการซื้อขายสินค้า ทฤษฎีและการใช้ร้านค้าแวร์ – เฮาส์ ร้านค้าไม้นั่ง คอมพิวเตอร์ ADO .NET คอมพิวเตอร์ Decision Support Objects (DSO) รูปแบบวิธีการเขียนโปรแกรมใน Visual Basic .NET การบริหารจัดการและเรียกใช้ Microsoft SQL Server 2000
- โครงการพัฒนาระบบงานนี้เริ่มด้วยการวิเคราะห์ระบบว่าระบบจะประกอบด้วยโมดูลอะไรบ้างรวมทั้งขั้นตอนการประมวลผลในแต่ละโมดูล การออกแบบระบบที่รวมถึงการออกแบบร้านค้าไม้นั่งโมเดล ฐานข้อมูล และส่วนประสานผู้ใช้ การพัฒนาระบบที่เริ่มจากการสร้างร้านค้าไม้นั่งโมเดล ฐานข้อมูล และการพัฒนาโปรแกรมเพื่อใช้งาน
- โครงการพัฒนาระบบงานนี้ร้านค้าไม้นั่งโมเดลสามารถให้เอาท์พุตการจัดแบ่งประเภทสมาชิกบัตรมาอย่างชัดเจนมีผลต่อการตัดสินใจได้ ส่วนแอปพลิเคชันสามารถทำงานได้อย่างถูกต้องตรงตามความต้องการและมีประสิทธิภาพ

6.2 ประโยชน์ที่ได้รับจากโครงการพัฒนาระบบงาน

จากการที่ได้ศึกษาค้นคว้าและวิเคราะห์ออกแบบรวมถึงการพัฒนาโครงการพัฒนาระบบงานนั้นได้ประโยชน์ดังนี้

- ได้รับความรู้ทางด้านการพัฒนาร้านค้าแวร์เฮาส์
- ได้รับความรู้ทางด้านการพัฒนาร้านค้าไม้นั่ง
- ได้รับความรู้ทางด้านการตลาดในส่วนการวิเคราะห์พฤติกรรมซื้อขายสินค้า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาค้นคว้าเท่านั้น เมื่ออนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ได้รับประสบการณ์ในการนำความรู้ทางสาขาทางเทคโนโลยีสารสนเทศและทางสาขาการตลาดมาประยุกต์ใช้ร่วมกัน
- ได้รับประสบการณ์ในการออกแบบระบบฐานข้อมูลที่ต้องออกแบบเพื่อรองรับการทำดาต้าแวร์เฮาส์และดาต้าไมน์นึ่ง ซึ่งส่วนหนึ่งนั้นมาจากการที่เข้าใจทฤษฎีการวิเคราะห์พฤติกรรมผู้บริโภคซื้อสินค้าทำให้รู้ว่าข้อมูลเกี่ยวกับลูกค้าข้อมูลไหนสามารถนำไปทำนายข้อมูลในดาต้าไมน์นึ่งได้และต้องมีในฐานข้อมูลด้วย
- ได้รับความรู้ในการพัฒนาแอปพลิเคชันด้วยการใช้ Microsoft Visual Basic .NET และคอมโพเนนต์ Decision Support Objects (DSO) รวมทั้งคอมโพเนนต์อื่นๆ

6.1 ปัญหา ข้อจำกัด และข้อเสนอแนะ

จากการพัฒนาโครงการพัฒนาระบบงานนั้นมีแนวทางที่ต้องพัฒนาหรือศึกษาต่อไปดังนี้

- ปรับปรุงพัฒนาแอปพลิเคชันให้รองรับการทำนายข้อมูลแบบอื่นๆ ต่อไปได้
- ศึกษาการทำดาต้าไมน์นึ่งในรูปแบบของการใช้อัลกอริทึม Clustering

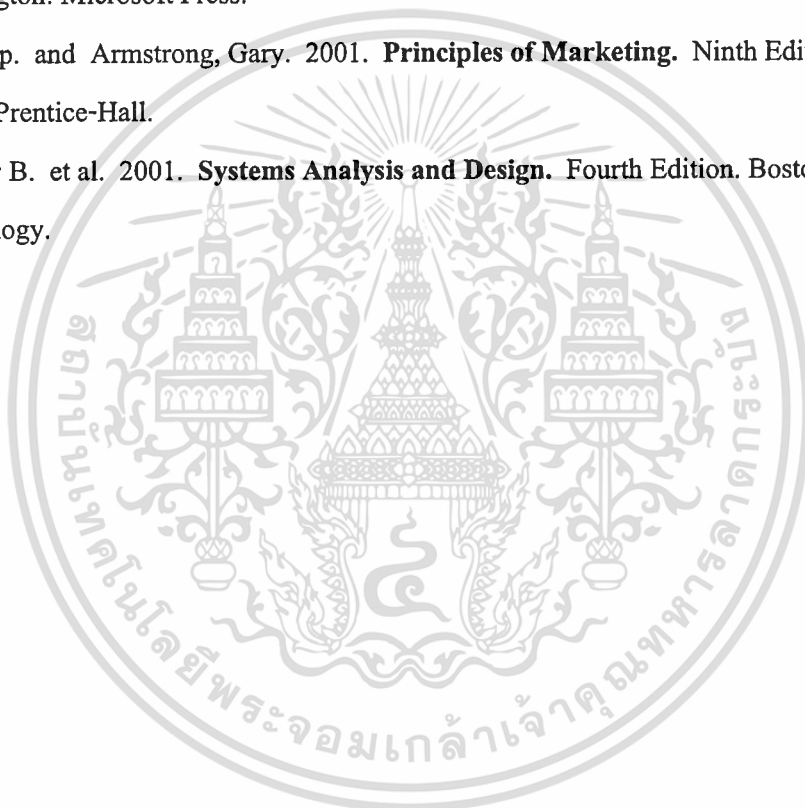
บรรณานุกรม

Reynolds, Matthew. et al. 2002. **Beginning Visual Basic .NET**. New York: Wrox Press.

Seidman, Claude. 2001. **Data Mining with Microsoft SQL Server 2000 Technical Reference**.
Washington: Microsoft Press.

Kotler, Philip. and Armstrong, Gary. 2001. **Principles of Marketing**. Ninth Edition. New
Jersey: Prentice-Hall.

Shelly, Gary B. et al. 2001. **Systems Analysis and Design**. Fourth Edition. Boston: Course
Technology.



ประวัติผู้เขียน

ชื่อ นาย ปรัชญา โกศล
รหัสนักศึกษา 44067453
ประวัติการศึกษา จบการศึกษา วิทยาศาสตร์บัณฑิต (วิทยาการคอมพิวเตอร์) จาก มหาวิทยาลัยรามคำแหง ปีการศึกษา 2539
ประวัติการทำงาน ปัจจุบันทำงานที่ บริษัท ลีอกซ์เลย์ จำกัด (มหาชน) ตำแหน่ง นักวิเคราะห์ระบบงาน (Systems Analyst)

