

ห้องสมุดคณะเทคโนโลยีสารสนเทศ พระจอมเกล้าลาดกระบัง

พัฒนาโปรแกรมวิเคราะห์การขายสินค้าในธุรกิจขายปูนซีเมนต์

DEVELOPMENT OF SALES ANALYSIS SYSTEM FOR  
CEMENT BUSINESS



ฉพ.

จ ๑๒๕ พ  
๒๕๔๙

รศ.ดร. วรพจน์ กรีสู่ระเดช

เลขหมู่.....  
เลขทะเบียน..... 04426  
วัน,เดือน,ปี..... - 5 ส.ย. 2551



b. 11๑ ๒๒๓๗๐  
i.....

รายงานนี้เป็นส่วนหนึ่งของวิชาโครงการพัฒนาระบบงาน  
หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ

คณะเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ภาคเรียนที่ ๒ ปีการศึกษา ๒๕๔๙

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นิยมนำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

**DEVELOPMENT OF SALES ANALYSIS SYSTEM FOR  
CEMENT BUSINESS**



**A SYSTEM DEVELOPMENT PROJECT  
OF THE REQUIREMENT FOR THE DEGREE OF  
MASTER OF SCIENCE PROGRAM IN INFORMATION TECHNOLOGY  
FACULTY OF INFORMATION TECNOLOGY  
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG**

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

2/ 2006

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



**COPYRIGHT 2007**

**FACULTY OF INFORMATION TECHNOLOGY**

**KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG**

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อ	พัฒนาโปรแกรมวิเคราะห์การขายสินค้าในธุรกิจขายปูนซีเมนต์
นักศึกษา	นายจักรพันธ์ ปลื้มเกษร
รหัสนักศึกษา	47066621
ปริญญา	วิทยาศาสตรมหาบัณฑิต
สาขาวิชา	เทคโนโลยีสารสนเทศ
แขนงวิชา	วิทยาการสารสนเทศ
ปีการศึกษา	2549
อาจารย์ที่ปรึกษา	รศ.ดร. วรพจน์ กรีสูระเดช

### บทคัดย่อ

เนื่องจากปัจจุบันการแข่งขันทางเศรษฐกิจที่รุนแรงขึ้นทุกวัน ทุกๆองค์กรจำเป็นต้องมีเครื่องมือที่ใช้ในการรวบรวมและวิเคราะห์ข้อมูลเพื่อให้ได้ข้อมูลที่ถูกต้องรวดเร็ว เพื่อใช้ในการทำแผนวางกลยุทธ์การเพิ่มยอดขายของตลาด

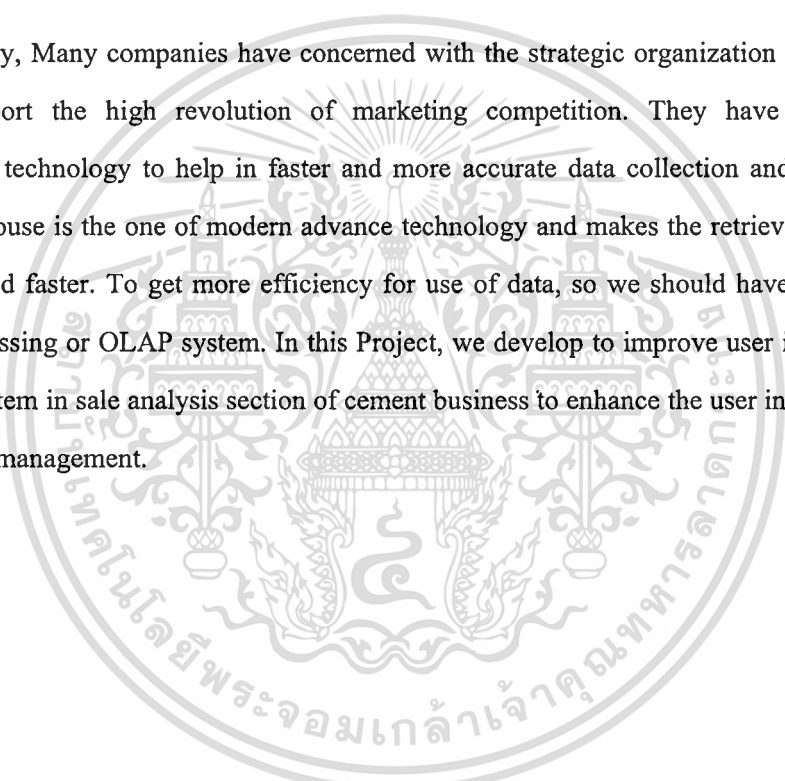
Data Warehouse นั้นเป็นกระบวนการหนึ่งที่ใช้ในการรวบรวมข้อมูลจากแหล่งต่างๆที่กระจัดกระจายมาจัดเก็บไว้ในที่เดียวกันเพื่อง่ายต่อการจัดการ และสามารถเรียกใช้ข้อมูลที่ต้องการได้อย่างสะดวก รวดเร็ว ในเพิ่มประสิทธิภาพของการนำข้อมูลมาใช้ให้มากยิ่งขึ้น โดยผ่านกระบวนการนำข้อมูลมาวิเคราะห์หรือที่เรียกว่า OLAP (Online Analytical Processing) มาช่วย

โดยโครงการนี้จะทำการพัฒนาโปรแกรมในส่วนของการนำ OLAP มาใช้งานของการวิเคราะห์ข้อมูลการขายของธุรกิจปูนซีเมนต์ ที่มีการผลิตและขาย เพื่อผู้ใช้งานและผู้บริหาร ได้นำข้อมูลไปวางแผนทางธุรกิจต่อไป

<b>Title</b>	Development of sales analysis system for Cement Business
<b>Student</b>	Mr. Jakrapan Pluemgasorn
<b>Student ID.</b>	47066621
<b>Degree</b>	Master of Science
<b>Programme</b>	Information Science
<b>Academic Year</b>	2006
<b>Advisor</b>	Assoc.Prof.Dr. Worapoj Kreesuradej

## ABSTRACT

Currently, Many companies have concerned with the strategic organization management system to support the high revolution of marketing competition. They have to use the advancement of technology to help in faster and more accurate data collection and processing. The Data warehouse is the one of modern advance technology and makes the retrieval process of data to easier and faster. To get more efficiency for use of data, so we should have the On-line Analytical processing or OLAP system. In this Project, we develop to improve user interface and using OLAP system in sale analysis section of cement business to enhance the user in the decision making or sales management.



# กิตติกรรมประกาศ

โครงการพัฒนาระบบงานฉบับนี้สำเร็จได้อย่างดี ด้วยคำแนะนำ และคำปรึกษาจากท่านอาจารย์ รศ.ดร. วรพงษ์ กรีสุระเดช ซึ่งเป็นอาจารย์ผู้ควบคุมโครงการนี้ ข้าพเจ้ารู้สึกทราบบ้างในความอนุเคราะห์จากท่านอาจารย์ และขอขอบพระคุณเป็นอย่างสูง

ขอกราบพระคุณคณาจารย์ คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ทุก ๆ ท่านที่ได้ประสิทธิ์ประสาทวิชาให้กับข้าพเจ้า

ขอขอบคุณบริษัท Microsoft Thailand และบริษัทต่างๆที่เกี่ยวข้อง ที่ได้สนับสนุนเครื่องมือตลอดจนข้อมูล และหนังสือต่างๆ ที่ใช้ในการทำโครงการ

ขอขอบคุณบัณฑิตศึกษาและบัณฑิตวิทยาลัย คณะเทคโนโลยีสารสนเทศที่ให้ความช่วยเหลือในเรื่องต่างๆ

สุดท้ายนี้ข้าพเจ้าขอกราบขอบพระคุณ บิดา มารดา และครอบครัวของข้าพเจ้าที่เป็นกำลังใจ และให้การสนับสนุนในทุกเรื่องๆ ทำให้ข้าพเจ้าสามารถทำโครงการฉบับนี้สำเร็จลุล่วงด้วยดี

คุณค่าและประโยชน์อันพึงมาจากโครงการฉบับนี้ ข้าพเจ้าขอบอบแด่ผู้มีพระคุณทุกท่าน

จักรพันธ์ ปลื้มเกษร

# สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	I
บทคัดย่อภาษาอังกฤษ.....	II
กิตติกรรมประกาศ.....	III
สารบัญ.....	IV
สารบัญตาราง.....	VIII
สารบัญรูป.....	IX
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาของปัญหา.....	1
1.2 วัตถุประสงค์ของโครงการ.....	2
1.3 ขอบเขตของโครงการ.....	2
1.4 ขั้นตอนและการวางแผนงานในการพัฒนา.....	3
1.5 เครื่องมือที่ใช้ในการพัฒนาระบบงาน.....	4
1.6 ประโยชน์ที่คาดว่าจะได้รับ.....	5
1.7 รายละเอียดบทต่างๆ.....	5
บทที่ 2 ทฤษฎีที่เกี่ยวข้องกับการพัฒนาระบบ.....	7
2.1 นิยามคลังข้อมูล.....	7
2.1.1 คลังข้อมูลคืออะไร.....	7
2.1.2 วัตถุประสงค์ของคลังข้อมูล.....	8
2.1.3 ประโยชน์ของระบบคลังข้อมูล.....	8
2.2 ความสัมพันธ์ระหว่างระบบคลังข้อมูลกับระบบฐานข้อมูลทั่วไป.....	9
2.2.1 การประมวลผลธุรกรรมออนไลน์ (On-Line Transaction Processing : OLTP).....	10
2.2.2 การประมวลผลเชิงวิเคราะห์ออนไลน์ (On-Line Analytical : OLAP).....	12
2.3 ความแตกต่างของคลังข้อมูลกับฐานข้อมูลทั่วไป.....	13
2.3.1 Consistency.....	13
2.3.2 Transaction.....	13
2.3.3 Time Dimension.....	13
2.4 กระบวนการคลังข้อมูล.....	14
2.4.1 Data Sources.....	14

เอกสารนี้เป็นเอกสารที่มอบให้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

# สารบัญ (ต่อ)

หน้า

2.4.2 Staging Area.....	14
2.4.3 Data Warehouse.....	14
2.4.4 Decision Support System.....	14
2.5 สถาปัตยกรรมคลังข้อมูล (Data Warehouse Architecture).....	15
2.5.1 รายละเอียดของฐานข้อมูลสำหรับคลังข้อมูล.....	15
2.5.2 สถาปัตยกรรมของคลังข้อมูล.....	16
2.6 กระบวนการรวบรวมข้อมูลและวิเคราะห์หารูปแบบข้อมูลในคลังข้อมูล.....	18
2.7 เทคนิคในการสร้างคลังข้อมูล.....	19
2.7.1 การเคลื่อนที่ของข้อมูลในคลังข้อมูล.....	20
2.7.2 การออกแบบคลังข้อมูล.....	20
2.7.3 วิธีการออกแบบฐานข้อมูลสำหรับคลังข้อมูล โดยใช้วิธี Nine-Step Methodology.....	21
2.7.4 การแปลงข้อมูลเข้าสู่ดาต้ามาร์ท.....	23
2.7.5 สิ่งที่ต้องพิจารณาก่อนสร้างคลังข้อมูล.....	24
2.8 สถาปัตยกรรมของการจัดการข้อมูลและการวิเคราะห์ข้อมูลของ MS-SQL SERVER 2005.....	25
2.8.1 สิ่งจำเป็นในการสร้างแอปพลิเคชัน.....	25
2.8.2 การติดต่อและการเชื่อมข้อมูล.....	25
2.8.3 การจัดการส่งต่อข้อมูล Data Transformation Service.....	26
2.8.4 การสร้าง OLAP และการรวบรวมให้เข้ากับคลังข้อมูล.....	27
2.8.5 การสร้างและการจัดการ Cube.....	28
2.8.6 การสร้างระบบงานที่รองรับ Business Intelligence.....	28
2.8.7 ภาษา MDX - Multidimensional Expressions Language.....	30
2.8.8 การทำ Cube Schema.....	30
2.8.9 การทำ Cube Data.....	31

บทที่ 3 การออกแบบและการจัดเก็บข้อมูลใน Data Warehouse.....	33
--	----

3.1 ขั้นตอนการเลือกสรรข้อมูลเพื่อใช้ในการวิเคราะห์.....	33
---	----

3.2 ขั้นตอนการ Transformation ข้อมูล.....	33
---	----

3.2.1 กระบวนการนำข้อมูลออกมาจากฐานข้อมูลหลัก.....	34
---	----

เอกสารนี้เป็นเอกสารที่รวบรวมขึ้นจากฐานข้อมูลก่อนนำมาใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

# สารบัญ (ต่อ)

หน้า

3.2.2	กระบวนการนำข้อมูลเข้า Staging Area.....	34
3.3	ขั้นตอนการออกแบบความสัมพันธ์ของข้อมูลใน Relation Database (Staging Area).....	40
3.4	พจนานุกรมข้อมูล (Data Dictionary).....	41
3.5	ทำการออกแบบตารางข้อมูลใน Data Warehouse.....	44
3.5.1	การวิเคราะห์ข้อมูลทางด้านลูกค้า.....	44
3.5.2	การวิเคราะห์ข้อมูลทางด้านสินค้า.....	45
3.5.3	การวิเคราะห์ข้อมูลทางด้านเวลาขาย.....	45
3.5.4	การวิเคราะห์ข้อมูลทางด้านเขตการขาย.....	46
3.5.5	การวิเคราะห์ข้อมูลตามประเภทของสินค้า.....	46
3.5.6	การวิเคราะห์ข้อมูลตามตารางแบ่งการบรรจุของสินค้า.....	46
3.5.7	การวิเคราะห์ข้อมูลตามตารางการขายสินค้า.....	47
3.6	ความสัมพันธ์และการสร้างข้อมูลใน Sales Cube.....	47
3.7	ทำการตรวจสอบข้อมูลของ Sales Cube ใน Analysis Services 2005.....	52
บทที่ 4	การพัฒนาและวิเคราะห์ระบบการขาย.....	53
4.1	สถาปัตยกรรมด้านการสื่อสารและฮาร์ดแวร์.....	53
4.2	สถาปัตยกรรมด้านซอฟต์แวร์.....	55
4.2.1	OS (Operating System).....	55
4.2.2	Database Server.....	55
4.2.3	Client Tool.....	55
4.2.4	Middleware.....	55
4.2.5	เครื่องมือในการพัฒนาระบบ.....	56
4.3	การออกแบบระบบสำหรับการวิเคราะห์ข้อมูลการขายบนซีเมนส์.....	56
4.3.1	รูปแบบการแสดงผลของโปรแกรม.....	56
4.3.2	รูปแบบการแสดงผลในรูปแบบรายงาน.....	58
4.3.3	ความสามารถต่างๆ และการทำงานของโปรแกรม.....	58
บทที่ 5	การใช้งานระบบและการบำรุงรักษาระบบ.....	60

เอกสารนี้เป็นเอกสารลิขสิทธิ์ของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ไม่ควรนำเอกสารนี้ไปเผยแพร่โดยไม่ได้รับอนุญาตจากทางมหาวิทยาลัย

5.1 การเข้าสู่การใช้งานของระบบเพื่อลดการศึกษานี้เท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ใด ๆ การค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## สารบัญ (ต่อ)

	หน้า
5.2 สิทธิในการเข้าใช้ข้อมูลของระบบ.....	61
5.3 การดูแลรักษาระบบฐานข้อมูล.....	64
บทที่ 6 สรุปผลการศึกษา และข้อเสนอแนะ.....	65
6.1 ผลที่ได้จากการพัฒนาระบบมีดังนี้.....	65
6.2 ข้อเสนอแนะ.....	65
6.3 สรุปผลดำเนินการพัฒนา.....	66
บรรณานุกรม.....	67
ภาคผนวก.....	68



# สารบัญตาราง

ตารางที่	หน้า
3.1 Customers Table.....	41
3.2 Product Table.....	41
3.3 ProdGroup Table.....	41
3.4 ProdLotype Table.....	42
3.5 Employee Table.....	42
3.6 Region Table.....	43
3.7 Username Table.....	43
3.8 Sales Table.....	43
3.9 Sex Table.....	43
3.10 Time.....	44
3.11 Customer Dimension Table.....	44
3.12 Product Dimension Table.....	45
3.13 Time Dimension Table.....	45
3.14 Region Dimension Table.....	46
3.15 ProdGroup Dimension Table.....	46
3.16 ProdLotype Dimension Table.....	46
3.17 Sales Fact Table.....	47

# สารบัญรูป

รูปที่	หน้า
2.1 ภาพความแตกต่างระหว่างคลังข้อมูลกับฐานข้อมูลที่ใช้ประจำวัน.....	14
2.2 ภาพกระบวนการคลังข้อมูล.....	15
2.3 ภาพสถาปัตยกรรมคลังข้อมูล.....	17
2.4 ภาพการสร้างคลังข้อมูล.....	19
3.1 การ Transformation ข้อมูล.....	34
3.2 การนำข้อมูลออกจากฐานข้อมูลหลัก.....	34
3.3 รูปแบบ SSIS Package สำหรับการ Transform ข้อมูล การขายเข้าสู่ Data Warehouse Server.....	36
3.4 รายละเอียดภายใน SSIS Package สำหรับการ Transform ข้อมูล.....	36
3.5 รูปแบบ SSIS Package สำหรับการ Transform ข้อมูล การขายเข้าสู่ OLAP Server.....	38
3.6 รายละเอียดภายใน SSIS Package สำหรับการ Transform ข้อมูลใน Data warehouse สู่ OLAP Server.....	38
3.7 รายละเอียดภายใน SSIS Package เป็นการลบข้อมูลใน Tables บน OLAP Server.....	39
3.8 E-R Diagram.....	40
3.9 Connection to MS-SQL Server 2005.....	48
3.10 การกำหนด Data Source View.....	48
3.11 ความสัมพันธ์ใน Sales Data Source View.....	49
3.12 การกำหนดค่าในการสร้าง Cube.....	50
3.13 โครงสร้าง Sale Cube.....	51
3.14 Analysis Services Management Studio.....	52
4.1 โครงสร้างและสถาปัตยกรรมของระบบ.....	54
4.2 หน้าจอของ โปรแกรม Cube Browser.....	57
4.3 รายงานต่างๆ ที่ใช้ Report Service.....	58
5.1 ขั้นตอนการใช้งานของระบบ.....	60
5.2 การ Log in เข้าสู่ระบบ.....	62
5.3 การจัดการสิทธิให้กับผู้เข้าใช้ระบบ.....	63

# บทที่ 1

## บทนำ

### 1.1 ความเป็นมาของปัญหา

เนื่องจากสภาพเศรษฐกิจปัจจุบันมีภาวะการแข่งขันที่รุนแรง ธุรกิจหลายประเภทจึงยังต้องการการวิเคราะห์ การวางแผน และการตัดสินใจอย่างถูกต้องรวดเร็วเพื่อช่วยให้ธุรกิจสามารถดำเนินไปได้และให้ผลกำไรอย่างสูงสุด ดังนั้นข้อมูลจึงเป็นปัจจัยที่สำคัญอย่างยิ่งต่อการดำเนินการนั้น การใช้ข้อมูลเป็นเครื่องมือสำคัญในการตัดสินใจลงทุนทางธุรกิจและวางแผนกลยุทธ์ทางการตลาดเพื่อแข่งขันกับคู่แข่งทางการค้า ฉะนั้นก็อาจกล่าวได้ว่าการมีข้อมูลมากทำให้มีโอกาสและมีชัยเหนือคู่แข่งในระดับหนึ่ง แต่ทว่าหากมองในทางกลับกันการมีข้อมูลจำนวนมากแต่ขาดการจัดเรียงให้เป็นระบบระเบียบยุ่งยากในการเข้าถึง และค้นคืนธุรกิจอาจต้องเสียค่าใช้จ่ายจำนวนมากมหาศาลในการเก็บรักษาข้อมูลเหล่านั้นไว้โดยไม่จำเป็น เพราะไม่ได้รับประโยชน์จากข้อมูลที่มี นอกจากนี้หากมีการนำข้อมูลมาวิเคราะห์อย่างผิดพลาดอาจจะก่อให้เกิดผลเสียหายได้ ซึ่งเป็นการสูญเสียโอกาสทางธุรกิจไป เพราะฉะนั้นในยุคที่ผู้บริหารมีความต้องการใช้ข้อมูลเพื่อการตัดสินใจมากขึ้นการจัดระบบระเบียบข้อมูลเพื่อนำเสนอข้อมูลที่มีคุณค่าและผ่านการกลั่นกรองแล้วแก่ผู้บริหารเพื่อใช้ในการตัดสินใจให้ทันต่อเหตุการณ์จึงเป็นสิ่งที่จะต้องทำเป็นอย่างยิ่ง ในหน่วยงานภายในองค์กรระบบงานขายก็เป็นอีกระบบหนึ่งที่มีความสำคัญ ที่ต้องใช้ความรวดเร็วในการวิเคราะห์ข้อมูลเพื่อนำมาวางแผนเพื่อเพิ่มยอดขายให้กับองค์กร แต่การนำข้อมูลการขายสินค้าในองค์กรหรือนำข้อมูลมาวิเคราะห์ยังคงนิยมแบบเดิม คือการรวบรวมข้อมูลที่จัดเก็บมาเรียงเรียงและทำให้อยู่ในเชิงของสถิติและสรุปในรูปแบบต่างๆมาให้กับผู้บริหารเพื่อนำไปใช้ในการวิเคราะห์ ซึ่งการรวบรวมข้อมูลแต่ละครั้งที่จะนำมาวิเคราะห์ต้องสิ้นเปลืองทรัพยากรต่างๆมากมาก เช่น เวลา , อุปกรณ์ทำเอกสาร , ทรัพยากรบุคคล ฯลฯ และทำให้เกิดความผิดพลาดได้ง่าย การนำเสนอข้อมูลรูปแบบเดิมๆไม่สามารถเปลี่ยนรูปแบบได้นี้ ในบางครั้งผู้บริหารที่ได้รับข้อมูลนั้นๆ อาจต้องการการปรับเปลี่ยนรูปแบบการนำเสนอข้อมูลเอง เพื่อให้ได้มุมมองหลายมิติเพื่อเข้าใจในข้อมูลมากยิ่งขึ้น นอกจากนี้หากแหล่งข้อมูลมีปริมาณมาก หรือ มีรูปแบบหลากหลายผู้ที่ทำการรวบรวมข้อมูลจะต้องใช้เวลามากยิ่งขึ้นในการรวบรวมข้อมูลเหล่านั้นมาทำรายงานให้กับผู้บริหารซึ่งทำให้การปรับกลยุทธ์หรือการวางแผนช้าไม่ทันกับความต้องการกับการตลาดที่มีการแข่งขันสูง

จากปัญหาที่กล่าวมานี้ได้มีแนวความคิดของการสร้างคลังข้อมูลจึงเกิดขึ้น เพื่อเป็นที่เก็บรวบรวมข้อมูลสำคัญและจำเป็นจากแหล่งต่างๆ ซึ่งเป็นประโยชน์ต่อการตัดสินใจของผู้บริหาร เพื่อให้ผู้บริหารสามารถเรียกใช้ข้อมูลที่ต้องการได้อย่างรวดเร็วและมีประสิทธิภาพมาก ซึ่งเรียกวิธีนี้ว่า คลังข้อมูล หรือ Data Warehouse เป็นวิธีการรวบรวมข้อมูลที่ต้องการมาวิเคราะห์หามาเก็บไว้ในที่เดียวกันเพื่อที่จะง่ายสะดวกในการจัดการต่อเรียกใช้ วิธีการนี้มีกระบวนการในการจัดเก็บไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รวบรวมข้อมูลจากแหล่งต่างๆ หลายๆ แห่ง หลายรูปแบบ มาไว้ในที่เดียวกันและรูปแบบเดียวกัน จะกล่าวถึงรายละเอียดในบทของทฤษฎีต่อไป

## 1.2 วัตถุประสงค์ของโครงการ

การพัฒนาโครงการนี้เกิดขึ้นจากวัตถุประสงค์ที่จะแก้ไขปัญหาลดความผิดพลาดของการนำเสนอข้อมูลการวิเคราะห์การขายสินค้าในธุรกิจปูนซีเมนต์ โดยบางครั้งการวางแผนการขายสินค้าได้ข้อมูลที่ไม่ครบถ้วนคลาดเคลื่อนจากความเป็นจริงและเสียเวลาในการรวบรวมข้อมูลเพื่อที่จะใช้ในการวิเคราะห์การตลาด เพื่อวางกลยุทธ์การขายสินค้าปูนซีเมนต์ในแต่ละพื้นที่ให้ตรงตามความต้องการของตลาด อีกประการหนึ่งเพื่อสามารถรองรับระบบงาน เกี่ยวกับการนำเสนอข้อมูลที่เป็นหลายมิติให้ตรงตามรูปแบบความต้องการของผู้บริหาร ทั้งยังช่วยแก้ปัญหาช่วยลดทรัพยากรต่างๆ ได้ โดยโปรแกรมวิเคราะห์การขายสินค้าในธุรกิจขายปูนซีเมนต์ที่มีจุดประสงค์ดังต่อไปนี้

1. เพิ่มความสามารถในการเข้าถึงและเรียกใช้ข้อมูล โดยให้ได้ข้อมูลที่รวดเร็ว ถูกต้อง และนำไปใช้ได้ทันต่อสถานการณ์
2. เพิ่มขีดความสามารถในการใช้ข้อมูลในการวิเคราะห์สำหรับการตัดสินใจของฝ่ายขาย และผู้บริหาร เพื่อเป็นแนวทางในการวางแผนการตลาดได้อย่างมีประสิทธิภาพ
3. เพิ่มประสิทธิภาพในการรวบรวมข้อมูลการขายสินค้าในฐานะข้อมูลต่างๆภายในองค์กร โดยการนำ Data Warehouse เข้ามาช่วยในการรวบรวมทำให้เกิดความสะดวก ถูกต้อง และรวดเร็วกว่าการจัดเก็บแบบเดิม
4. รวบรวมข้อมูลต่างๆ ให้อยู่ที่เดียวกันและนำมาวิเคราะห์
5. เพื่อให้ผู้บริหารสามารถวางแผนงานที่มีประสิทธิภาพและสามารถปรับเปลี่ยนแผนงานตามสถานการณ์ได้อย่างเหมาะสม

## 1.3 ขอบเขตของโครงการ

การทำการพัฒนาโปรแกรมดังกล่าวนี้ จะใช้เป็นเครื่องมือช่วยให้ผู้ใช้งานที่เกี่ยวข้องกับการขายสินค้าขององค์กร ได้ทำการเก็บรวบรวมข้อมูลที่กระจัดกระจายมาอยู่ที่เดียวกัน และนำข้อมูลเหล่านั้นมาวิเคราะห์ หลังจากนั้นนำผลลัพธ์ที่ได้มาวางแผนการขายในอนาคต โดยโปรแกรมที่ได้พัฒนานี้มุ่งเน้นในการข้อมูลที่อยู่ใน Data Warehouse ออกมาวิเคราะห์ หรือนำมาแสดงให้อยู่ในรูปแบบหลากหลาย ง่ายต่อการใช้งาน นอกจากนี้ในโครงการจะกล่าวในเรื่องการออกแบบในกระบวนการทำ OLAP เพื่อสอดคล้องกับการนำข้อมูลมาวิเคราะห์ ด้วยโปรแกรมที่จะพัฒนาขึ้นมาจะอยู่ในรูปแบบการนำเสนอข้อมูล และติดต่อกับผู้ใช้งานผ่านทางหน้าจอของ

โปรแกรมที่พัฒนาขึ้นบนระบบปฏิบัติการ Windows โดยรูปแบบการนำเสนอสามารถปรับเปลี่ยน

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ได้อย่างง่ายได้ตามความต้องการของผู้ใช้งาน โดยกระบวนการทั้งหมดจะครอบคลุมการทำงานดังต่อไปนี้

1. เก็บรวบรวมข้อมูลที่เกี่ยวข้องกับการขายสินค้าเพื่อให้เป็นรูปแบบเดียวกันทั้งองค์กร
2. จำแนกประเภท , รูปแบบของข้อมูลที่รวบรวมให้เป็นประเภทเดียวกัน
3. สร้างวิธีการหาความสัมพันธ์ของข้อมูล โดยให้มีรูปแบบสำหรับการวิเคราะห์
4. แสดงตัวเลขในรูปแบบตามความต้องการ เช่น รูปแบบของตัวเลขสรุป หรือแผนภูมิ
5. มีผลสรุปทางสถิติเพื่อใช้ในการตัดสินใจ

## 1.4 ขั้นตอนและการวางแผนงานในการพัฒนา

ขั้นตอนและการวางแผนงานในการพัฒนาระบบวิเคราะห์การขายปูนซีเมนต์ มีขั้นตอนที่สำคัญดังนี้

### 1.4.1 การวิเคราะห์การพัฒนา (Analysis Phase)

การวิเคราะห์การพัฒนารายละเอียดขั้นตอนมีดังนี้

- วิเคราะห์การทำงานในระบบงานขายปูนซีเมนต์ จากข้อมูลการขายในฐานะข้อมูลหลักและผู้ที่เกี่ยวข้อง
- เก็บรวบรวมความต้องการของผู้ใช้ และรับทราบปัญหาต่างๆ
- ศึกษาทรัพยากรต่างๆและความเป็นไปได้
- ศึกษาเครื่องมือและข้อมูลที่จะนำมาใช้ในระบบงาน
- สรุปผลที่ได้ทั้งหมดรวมถึงทางเลือกต่างๆที่ดีที่สุด
- กำหนดขอบเขตการทำงานของโปรแกรมให้ชัดเจน

### 1.4.2 การออกแบบ (Design Phase)

การออกแบบรายละเอียดขั้นตอนมีดังนี้

- ทำการวิเคราะห์และออกแบบข้อมูลทั้งหมด รวมทั้งจัดกลุ่มข้อมูลให้เป็นหมวดหมู่
- ออกแบบโครงสร้างข้อมูลต่างๆ
- กำหนดเครื่องมือและทรัพยากรต่างๆที่จะใช้ในการพัฒนา
- ออกแบบโครงสร้างโปรแกรมในส่วนอินพุทและเอาต์พุท รวมทั้งส่วนที่ติดต่อกับผู้ใช้งาน
- ออกแบบโครงสร้างของระบบ Network
- จัดรูปแบบข้อมูลที่จะนำเสนอเพื่อให้ตรงตามความต้องการของผู้ใช้
- ออกแบบและวิเคราะห์ระบบทั้งหมด ให้เหมาะสมและสอดคล้องกับองค์กร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 1.4.3 การทำการติดตั้งโปรแกรม (Implementation Phase)

การทำการติดตั้งโปรแกรมรายละเอียดขั้นตอนมีดังนี้

- พัฒนาโปรแกรมตามแผนงานและตามที่ได้ออกแบบไว้
- ทำการติดตั้งโปรแกรมที่ได้พัฒนาแล้ว
- กำหนดสิทธิต่างๆในการเข้าถึงข้อมูลของผู้ใช้งาน

### 1.4.4 การทดสอบโปรแกรม (Testing Phase)

การทดสอบโปรแกรมรายละเอียดขั้นตอนมีดังนี้

- ทำการทดสอบการทำงานต่างๆเช่น การทำงานของชุดคำสั่งต่าง , การรองรับความต้องการและสนับสนุนให้กับผู้ใช้งาน , ทดสอบการเข้าถึงข้อมูลว่าเป็นไปตามความต้องการหรือไม่ , ทดสอบการส่งข้อมูลและสร้างข้อมูลใน Data Warehouse
- ถ้าทดสอบแล้วพบข้อผิดพลาดหรือพบสิ่งที่ต้องทำการแก้ไข ให้กลับไปทำข้อ 1.4.2 ใหม่ จนกว่าจะได้โปรแกรมตรงตามความต้องการ

### 1.4.5 การดำเนินการทำเอกสารประกอบการทำงานของโปรแกรม

การดำเนินการทำเอกสารประกอบการทำงานของโปรแกรมรายละเอียดขั้นตอนมีดังนี้

- จัดทำเอกสารประกอบในการดำเนินการต่างๆตั้งแต่ต้น เช่นการออกแบบ
- ทำคู่มือเอกสารประกอบการใช้งานของโปรแกรม

### 1.4.6 การบำรุงรักษาโปรแกรม

การบำรุงรักษาโปรแกรมรายละเอียดขั้นตอนมีดังนี้

- ติดตามและประเมินผลการใช้งานของผู้ใช้
- ดูแลในเรื่องความปลอดภัยของการกำหนดสิทธิ ในการเข้าใช้งานข้อมูล

## 1.5 เครื่องมือที่ใช้ในการพัฒนาระบบงาน

ทรัพยากรทางด้านฮาร์ดแวร์และด้านซอฟต์แวร์ที่ใช้ในการพัฒนาระบบงาน มีดังนี้

### 1.5.1 รายละเอียดทรัพยากรทางด้านฮาร์ดแวร์

- Server 1 เครื่อง
- Client 1 เครื่อง
- Developer 1 เครื่อง

### 1.5.2 รายละเอียดทรัพยากรทางด้านซอฟต์แวร์

- Server
  - MS Windows 2003 server

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- OLAP Server ของ Ms SQL Server 2005 Analysis Service
- Client
  - MS Windows XP
  - MS Data Access Component 2.8 ใช้เชื่อมต่อข้อมูลกับ Server
  - MS Office
- Developer
  - MS Windows XP
  - MS Data Access Component 2.8 ใช้เชื่อมต่อข้อมูลกับ Server
  - MS Visual Basic.NET ใช้ในการพัฒนาโปรแกรม

## 1.6 ประโยชน์ที่คาดว่าจะได้รับ

เมื่อพัฒนาระบบงานใหม่ให้เหมาะสมกับองค์กรแล้ว ประโยชน์ที่จะได้รับมีดังนี้

### 1.6.1 ประโยชน์ในระดับองค์กร

- ประหยัดทรัพยากรต่างๆ เช่น เวลา , บุคลากร ฯลฯ
- สามารถปรับเปลี่ยนความต้องการ ได้ง่ายเพราะเป็น โปรแกรมที่พัฒนาในองค์กร
- มีระบบการจัดการข้อมูลมาใช้ในการวิเคราะห์ข้อมูลการขาย และสามารถนำข้อมูลมาใช้ได้อย่างมีประสิทธิภาพ
- สามารถใช้ในส่วนงานที่เกี่ยวข้องเพื่อใช้ในการวิเคราะห์ข้อมูล และวางกลยุทธ์ให้กับสินค้าที่จะวางขาย

### 1.6.2 ประโยชน์ต่อผู้ทำการศึกษาและพัฒนา

- นำความรู้ที่ได้ศึกษามาแล้วที่ได้ในบทเรียนและจากประสบการณ์มาใช้ในการปฏิบัติงานจริง กับงาน โครงการ
- สามารถทำการพัฒนาและแก้ไขปัญหาต่างๆด้วยตนเอง
- ใช้เป็นกรณีศึกษากับ โครงการอื่นๆต่อไป

## 1.7 รายละเอียดบทต่างๆ

ดังที่กล่าวถึงความสำคัญและขั้นตอนต่างๆในการพัฒนางาน โครงการนี้แล้ว ยังมีรายละเอียดที่จะกล่าวในบทอื่นๆอีกคือ

- บทที่ 2 กล่าวถึงทฤษฎีต่างๆที่เกี่ยวข้องกับการพัฒนาโครงการระบบทั้งหมด

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ของ บริษัท ข้อมูลคอมพิวเตอร์ จำกัด (มหาชน) หรือ Data Warehouse , OLAP , DTS , Cube , MDX ใช้ เครื่องมือที่จะใช้ในการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกการพัฒนา ฯลฯ ปลอดภัย และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- บทที่ 3 กล่าวถึงการออกแบบและการจัดเก็บข้อมูลใน Data Warehouse Server และ OLAP Server โดยแสดงข้อมูลที่ผ่านกระบวนการ Transform ใน ISS Package และแสดงรายละเอียดของฐานข้อมูลต่างๆที่เกี่ยวข้อง
- บทที่ 4 กล่าวถึงการพัฒนาและวิเคราะห์ระบบการขาย โดยแสดงโครงสร้างของสถาปัตยกรรมต่างๆเช่น ซอฟต์แวร์,การสื่อสารและฮาร์ดแวร์ ต่างๆในระบบ
- บทที่ 5 กล่าวถึงการใช้งานระบบและการบำรุงรักษาระบบ เช่น สิทธิต่างๆ สำหรับผู้ใช้งานของระบบ ละวิธีการดูแลฐานข้อมูลต่างๆ
- บทที่ 6 กล่าวถึงสรุปผลการศึกษา และข้อเสนอแนะ



## บทที่ 2

# ทฤษฎีที่เกี่ยวข้องกับการพัฒนาระบบ

ในการดำเนินธุรกิจปัจจุบัน เป็นยุคแห่งข้อมูลสารสนเทศที่ใช้ในการติดต่อสื่อสารถึงกันผ่านทางช่องทางได้หลายช่องทางไม่ว่าจะอยู่ในรูปของกระดาษ แฟ้มเอกสาร แฟ้มข้อมูล โดยใช้เครื่องมือสื่อสารเครื่องคอมพิวเตอร์ เพื่อประโยชน์กับหน่วยงานหรือองค์กรที่ต้องการนำข้อมูลสารสนเทศที่ได้ไปช่วยในการวิเคราะห์ และตัดสินใจในการดำเนินงานทางธุรกิจ จากอดีตที่ผ่านมาการนำข้อมูลมาใช้กับเครื่องคอมพิวเตอร์ ได้มีการพัฒนารูปแบบและลักษณะการจัดเก็บข้อมูลไว้หลายแบบเริ่มตั้งแต่จัดเก็บเป็นแฟ้มข้อมูล (File) , ฐานข้อมูล (Database) และคลังข้อมูล (Data Warehouse)

### 2.1 นิยามของคลังข้อมูล

ในปี 1990 จนถึงปัจจุบันการจัดเก็บข้อมูลสามารถทำได้หลายรูปแบบแตกต่างกันทั้งระบบปฏิบัติการหรือการจัดเก็บฐานข้อมูล การนำข้อมูลทั้งหมดมารวบรวมและจัดเก็บไว้ในที่เดียวกันและรูปแบบเดียวกันเรียกว่า คลังข้อมูล (Data Warehouse) ซึ่งเทคโนโลยีคลังข้อมูลนี้ จะรวมไปถึง Data Cleansing , Data Integration และ On-Line Analytical Processing (OLAP) ซึ่งเป็นเทคนิคการวิเคราะห์ข้อมูลในหลายๆ มิติ ซึ่งได้เกิดขึ้นมาตามลำดับ

#### 2.1.1 คลังข้อมูลคืออะไร

คลังข้อมูล คือ กระบวนการในการดึงข้อมูลจากระบบปฏิบัติการต่างๆ ในองค์กรมารวมกันในที่เก็บข้อมูลที่เดียวกัน (Single Repository) โดยข้อมูลเหล่านั้นจะถูกสรุปและผ่านการตรวจสอบ ตลอดจนการเปลี่ยนแปลงรูปเรียบร้อยแล้ว เพื่อให้ผู้ใช้สามารถใช้สารสนเทศต่างๆ ได้อย่างรวดเร็วและสามารถทำการวิเคราะห์ต่างๆ ได้ง่าย โดยเฉพาะอย่างยิ่งสารสนเทศเพื่อช่วยในการตัดสินใจสำหรับองค์กร

หรือ ระบบการจัดเก็บรวบรวมข้อมูลที่มีอยู่ในระบบปฏิบัติการต่างๆ ขององค์กร มักเป็นองค์กรขนาดกลาง ถึงขนาดใหญ่ โดยข้อมูลเหล่านั้นมักเป็นข้อมูลกระจัดกระจาย ให้มารวมไว้เป็นศูนย์กลางข้อมูล ซึ่งจะทำการแยกข้อมูลออกจากฐานข้อมูลที่ใช้ในประจำวัน (Operational Database) ขององค์กร และสามารถเก็บข้อมูลย้อนหลังได้หลายๆปี เพื่อใช้เป็นข้อมูลช่วยสนับสนุนการตัดสินใจ (Decision Support System) หรือใช้ในการวิเคราะห์ข้อมูลที่ต้องการ และมีประสิทธิภาพ โดยมากจะเป็นข้อมูลสรุป (Summary Data) โดยการวิเคราะห์ต้องทำได้แบบหลายมิติ (Multidimensional Analysis) ตลอดจนการวิเคราะห์ทางธุรกิจ เช่น การพยากรณ์ (Forecasting) , What-If Analysis , Data Mining เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษานั่น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 2.1.2 วัตถุประสงค์ของการสร้างคลังข้อมูล

การสร้างคลังข้อมูลคือ การแยกกลุ่มข้อมูลสารสนเทศที่ใช้ในการวิเคราะห์ทางธุรกิจออกจากฐานข้อมูล Operational Database มาเก็บอยู่ใน Relational Database Management Systems (RDBMS) ทำให้การเรียกใช้ข้อมูลชุดนี้ทำได้อย่างมีประสิทธิภาพมากยิ่งขึ้น สามารถเรียกดูจากเครื่องทั่วไป โดยมีเครื่องมือที่ใช้ในการตัดสินใจเป็นตัวนำเสนอ ช่วยประหยัดเวลาที่นำเสนอ (Response Time) รวดเร็วยิ่งขึ้นอย่างมากและผู้บริหารสามารถเรียกข้อมูลรายละเอียดที่จำเป็น ที่ถูกเก็บมาก่อนหน้านี้ มาช่วยในการตัดสินใจทางธุรกิจแม่นยำขึ้น โดยเป้าหมายในการสร้างคลังข้อมูลมีดังนี้

1. คลังข้อมูลทำให้สามารถเข้าถึงข้อมูลขององค์กรได้ ผู้ใช้งานสามารถต่อเข้าไปยังคลังข้อมูลจากเครื่องคอมพิวเตอร์ของตนได้ ซึ่งการเชื่อมต่อสามารถทำได้ทันทีตามความต้องการ เครื่องมือที่มีให้กับผู้ใช้งานสามารถใช้งานและทำรายงานได้สะดวก
2. ข้อมูลในคลังข้อมูลมีความถูกต้องตรงกันหมด
3. ข้อมูลสามารถนำมาวิเคราะห์ตามความต้องการของเงื่อนไขทางธุรกิจ โดยสามารถแบ่งข้อมูลหรือรวมข้อมูลมาวิเคราะห์ตามความต้องการได้ทันที
4. ข้อมูลที่นำมาใช้ ไม่เพียงแต่ถูกเก็บไว้ที่ศูนย์กลางอย่างเดียว จากแหล่งข้อมูลหลายๆแห่งนอกองค์กรด้วย แล้วมาปรับปรุงให้เหมาะกับการใช้งานนั้นๆ

## 2.1.3 ประโยชน์ของระบบคลังข้อมูล

โดยทั่วไปแล้วข้อมูลที่ใช้ในชีวิตประจำวัน จะถูกเก็บข้อมูลในรูปแบบ Transaction Systems เมื่อมีความต้องการข้อมูลที่จะนำมาใช้ช่วยในการตัดสินใจก็จะประสบปัญหาต่างๆ เช่น

1. ททรัพยากรบุคคลทางด้าน Information Systems จำเป็นต้องเรียกข้อมูลจากฐานข้อมูลขนาดใหญ่ ซึ่งมีข้อมูลมากเกินไปเกินความต้องการ ส่งผลให้ประสิทธิภาพของ Transaction Operational Database ทำงานได้ช้าลง
2. ข้อมูลจะเป็นข้อมูลแบบตารางเท่านั้น
3. ข้อมูลจะถูกนำเสนอในรูปแบบข้อมูลที่ตายตัว ไม่สามารถเปลี่ยนแปลงตามความต้องการของผู้ใช้
4. ไม่ตอบสนองความต้องการของการตัดสินใจ เพราะข้อมูลสำหรับการตัดสินใจมีความสลับซับซ้อนสูง มีการรวมตัวกันของข้อมูลจากตารางต่างๆ หลายตารางข้อมูล
5. ไม่ตอบสนองการสอบถามข้อมูล (Data Queries) สำหรับผู้ใช้
6. มีข้อมูลย้อนหลังน้อย (Historical Data)
7. ข้อมูลถูกจัดเก็บกระจายตามที่ตั้งต่างๆ ซึ่งยากต่อการเรียกใช้หรือขาดความสัมพันธ์ทางธุรกิจ อันอาจจะต้องเสียเวลาในการทำให้สอดคล้อง หรือเกิดความซ้ำซ้อนของข้อมูลได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากอุปสรรคที่กล่าวมาข้างต้นคลังข้อมูลได้ถูกออกแบบมาเพื่อตอบสนองงานในรูปแบบ การตัดสินใจโดยการแยกฐานข้อมูลออกจาก Operational Database และเก็บข้อมูลในรูปแบบ ข้อมูลสรุป (Summary Data) ซึ่งข้อมูลสรุปนี้จะเลือกแต่เฉพาะข้อมูลที่มีความสำคัญต่อการตัดสินใจหรือเพื่อใช้ในการบริหารไปจนถึงการกำหนดแผนงานในอนาคต

ในระบบฐานข้อมูล ข้อมูลที่ซับซ้อนจะถูกรวบรวมหรือเปลี่ยนแปลงให้ง่ายต่อการจัดเก็บ และสามารถเรียกกลับมาใช้ได้อย่างรวดเร็วและถูกต้อง โดยข้อมูลต่างๆเหล่านี้จะถูกนำมาใช้ สำหรับการวิเคราะห์และช่วยในเรื่องการตัดสินใจ โดยอาศัยเครื่องมือที่อยู่ในเครื่องคอมพิวเตอร์ที่เป็นซอฟต์แวร์มาใช้ในการจัดการทำรายงาน และเพิ่มประสิทธิภาพสำหรับการตัดสินใจให้รวดเร็วยิ่งขึ้น โดยผู้บริหาร นักวางแผน และนักวิเคราะห์ข้อมูลสามารถเรียกหาข้อมูล เพื่อให้ได้รับคำตอบในรูปแบบตาราง หรือรายงาน กราฟ เพื่อมาทำการวิเคราะห์ข้อมูลด้วยตนเอง เช่น

1. การเปรียบเทียบยอดขายระหว่างช่วงเวลาในอดีตกับปัจจุบัน ไปจนถึงการทำพยากรณ์ยอดขายในอดีต (Forecasting)
2. การหายอดขายสูงสุดหรือต่ำสุด
3. การเปรียบเทียบยอดขาย ต้นทุน กำไร ในรูปแบบตาราง รายงาน หรือรายงานกราฟ

ซึ่งเครื่องมือนี้ถือได้ว่าเป็นสิ่งสำคัญในอันที่จะนำองค์กร ไปสู่ความสำเร็จในการตัดสินใจการวางแผนต่างๆ ในปัจจุบันเครื่องมือที่ช่วยตอบสนองงานเพื่อช่วยผู้บริหารสำหรับการตัดสินใจมีอยู่มากมายในตลาด ทั้งนี้ก็เป็นทางเลือกของผู้ใช้ในการที่จะเลือกเครื่องมือที่มีประสิทธิภาพสูงสุดเพื่อตอบสนองงานของผู้บริหารในกระบวนการตัดสินใจต่อไป

## 2.2 ความสัมพันธ์ระหว่างระบบคลังข้อมูลกับระบบฐานข้อมูลทั่วไป

ในปัจจุบันมีการใช้ฐานข้อมูลอย่างกว้างขวางในระบบงานทั่วไป จึงมีการวิจัยและพัฒนาวิธีเก็บข้อมูลจำนวนมาก รวมถึงการค้นหาและนำข้อมูลที่ต้องการออกมาจากระบบฐานข้อมูลด้วย แต่เนื่องจากระบบฐานข้อมูลทั่วไป (Operational Database) ที่นิยมใช้อยู่ในปัจจุบันมีหลักในการเก็บข้อมูลที่เน้นในเรื่องการลดความซ้ำซ้อน (Redundancy) รักษาความถูกต้อง (Integrity) ลดการสูญหายของข้อมูล (Information lost) และลดความผิดพลาดที่เกิดขึ้นจากการแก้ไขข้อมูล (Update Anomalies)

เนื่องจากฐานข้อมูลทั่วไป มีความสามารถเพียงแค่การเรียกใช้ข้อมูลที่มีอยู่ แต่ไม่สามารถจะนำมาช่วยในการสนับสนุนการตัดสินใจได้ เพราะเมื่อมีการเรียกใช้ข้อมูลจะต้องเรียกใช้ฐานข้อมูลขนาดใหญ่ ซึ่งมีข้อมูลจำนวนมากและมีการแยกตารางที่นอร์มัลไลซ์ (Normalized Table) แล้วออกเป็นหลายตาราง จึงไม่รองรับเงื่อนไขทางธุรกิจที่ต้องการจะนำมาใช้ช่วยในการสนับสนุนการตัดสินใจ (Decision Support Queries) มีการรวม (Join) กันของตารางต่างๆ ที่ซับซ้อน ซึ่งจะทำให้มีประสิทธิภาพของการค้นหาข้อมูลจากฐานข้อมูลน้อยลง และทำงานช้า ไม่สามารถเรียกใช้ข้อมูลที่ต้องการ ได้ทั้งหมด จึงมีความสามารถในการค้นหาข้อมูลแบบที่ไม่ซ้ำซ้อน ไม่วารณใดๆทั้งสิ้น อีกทั้งยังมีเทคนิคแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เท่านั้น นอกจากนี้การเก็บข้อมูลในระบบฐานข้อมูลทั่วไป (Operational Routine) ยังไม่มีการเก็บข้อมูลย้อนหลัง เพื่อใช้ช่วยในการคาดคะเนแนวโน้มที่คาดว่าจะเป็นไปได้ในอนาคต

ระบบฐานข้อมูลคลังข้อมูล จะแยกกลุ่มข้อมูลสารสนเทศที่ใช้ในการวิเคราะห์ทางธุรกิจ ออกจากฐานข้อมูลที่ใช้ประจำวัน (Operational Database) มาเก็บอยู่ในระบบจัดการฐานข้อมูล (Relational Database Management Systems) ประสิทธิภาพสูง และทำให้การเรียกใช้ข้อมูลชุดนี้ทำได้ยืดหยุ่น จากเครื่องมือที่อยู่บนเครื่องคอมพิวเตอร์ทั่วไป โดยลดการทำงานให้กับระบบ (Off-Loading) เพิ่มกลไกการช่วยตัดสินใจ ปรับปรุงเวลาที่ตอบสนอง (Response Time) รวดเร็วขึ้นอย่างมากและผู้บริหารสามารถเรียกข้อมูลรายละเอียดที่จำเป็นที่ถูกเก็บมาก่อนหน้านี้ (Historical Data) มาใช้ช่วยในการตัดสินใจทางธุรกิจแม่นยำขึ้น

ความแตกต่างอีกประการหนึ่งก็คือ ผู้ใช้คลังข้อมูลมักจะต้องการจัดกลุ่มข้อมูลด้วยตนเองมากกว่าผู้ใช้ในระบบฐานข้อมูลทั่วไป ในกรณีต่างๆ เหล่านี้ผู้ใช้ต้องการที่จะเลือกจัดกลุ่มข้อมูลได้ตามใจชอบ นอกจากการนำข้อมูลเข้ามารวมกันแล้ว ผู้ใช้ยังอาจต้องการที่จะแยกแยะข้อมูลในแบบที่ตนเองต้องการได้ โดยปกติแล้วการจัดทำฐานข้อมูลให้สามารถวิเคราะห์แยกแยะข้อมูลในแบบนี้ด้วยนั้นเป็นเรื่องไม่ยาก แต่ในการออกแบบคลังข้อมูลนั้นจำเป็นต้องเพื่อให้ผู้ใช้หลายคนสามารถแยกแยะข้อมูลตามความต้องการที่แตกต่างกันได้ด้วย ผู้ใช้จำนวนมากในปัจจุบันนี้อาจใช้ซอฟต์แวร์หลากหลายประเภทสำหรับเครื่องคอมพิวเตอร์ส่วนบุคคล ผู้ใช้บางคนอาจจะใช้โปรแกรมสเปรดชีตในการวิเคราะห์ข้อมูล และผู้ใช้บางคนอาจต้องการใช้โปรแกรมวิเคราะห์สถิติอื่นๆ ดังนั้นผู้ใช้เหล่านี้จึงอาจมีความต้องการในการนำเข้าข้อมูลจากคลังข้อมูลมาไว้ในแฟ้มข้อมูลที่มีรูปแบบตรงกับโปรแกรมที่ตนต้องการใช้ ความต้องการด้านนี้นับว่าสำคัญมากที่สุดในการจัดทำคลังข้อมูล

ในบางครั้งที่นิยมใช้ฐานข้อมูลกันมากก็คือ งานบันทึกข้อมูลทางธุรกิจเอาไว้เพื่อประมวลผล ข้อมูลเหล่านี้ได้แก่ ข้อมูลการสั่งซื้อสินค้าของลูกค้า แต่เดิมนั้นการบันทึกข้อมูลเริ่มต้นด้วยการใช้กระดาษแบบฟอร์มสำหรับให้ลูกค้ากรอกข้อมูล จากนั้นจึงนำแบบฟอร์มมาบันทึกข้อมูลลงในฐานข้อมูลของระบบคอมพิวเตอร์ในแบบแบตช์ (Batch) ปัจจุบันนี้การบันทึกข้อมูลธุรกรรมได้เปลี่ยนไปเป็นระบบออนไลน์ (Online) เป็นส่วนใหญ่ในระบบแบบนี้กระบวนการบันทึกข้อมูลมีลักษณะอัตโนมัติมากขึ้น และใช้อุปกรณ์บันทึกข้อมูลที่สามารถเก็บข้อมูลลงในฐานข้อมูลของระบบคอมพิวเตอร์ได้ทันที เช่น ในห้างสรรพสินค้าก็มีการใช้เครื่องบริการ ณ จุดขาย (Point of Sale : POS) สำหรับอ่านรหัสแท่งแสดงราคาสินค้า แล้วบันทึกข้อมูลการขายไปเก็บไว้ในฐานข้อมูลได้ทันที การดำเนินการในลักษณะนี้เรียกกันว่า การประมวลผลธุรกรรมออนไลน์ (On-Line Transaction Processing : OLTP)

### 2.2.1 การประมวลผลธุรกรรมออนไลน์ (On-Line Transaction Processing : OLTP)

ระบบ OLTP โดยทั่วไปจะต้องสามารถดำเนินการกับข้อมูลธุรกรรม ได้อย่างมี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปเผยแพร่โดยไม่ได้รับอนุญาต

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประสิทธิภาพ งานที่ทำกับข้อมูลได้แก่การปรับค่าของข้อมูลให้เป็นปัจจุบันและการเพิ่มข้อมูลลงไปในฐานะข้อมูล ข้อมูลเหล่านี้อาจจะมีจำนวนมากและเพิ่มขึ้นตลอดเวลา ในบางครั้งอาจจะมีการประมวลผลข้อมูลจำนวนนับแสนเรคคอร์ดก็ได้ การออกแบบระบบ OLTP แบบนี้จำเป็นต้องหาทางให้ระบบสามารถทำงานได้อย่างถูกต้องรวดเร็วตลอดเวลา เอื้ออำนวยให้ผู้ใช้จำนวนมากสามารถใช้ระบบได้พร้อมกัน อีกทั้งยังต้องสามารถแก้ไข ปัญหาในการกู้คืนระบบให้กลับคืนดั้งเดิมได้หากเกิดความขัดข้องเสียหาย

การที่จะจัดทำระบบ OLTP ให้มีความสามารถในแบบนี้ได้ต้องคำนึงถึงปัจจัยต่อไปนี้

- 1) ขนาดและตำแหน่งของ Rollback Segment
- 2) ดัชนี การจัดกลุ่ม และการคำนวณตำแหน่งที่อยู่ (Hashing)
- 3) การออกแบบข้อมูลธุรกรรมให้เหมาะกับงานประยุกต์
- 4) หน่วยเก็บและเนื้อที่ว่างสำหรับการเก็บข้อมูลใหม่
- 5) ความเข้าใจลักษณะงานประยุกต์และการเขียนคำสั่งสำหรับค้นคืนข้อมูล
- 6) การปรับปรุงสมรรถนะของระบบอย่างต่อเนื่อง

ระบบ OLTP ที่พัฒนาขึ้นโดยใช้เทคนิคด้านฐานข้อมูลตามปกติมักจะไม่สามารถรับกับปริมาณข้อมูลที่เพิ่มขึ้นอย่างมากมาเป็นประจำทุกวันได้ การนำระบบเช่นนี้มาใช้จึงมีความเสี่ยงที่จะเกิดความผิดพลาดเสียหายขึ้น วิธีการแก้ไขก็คือการแยกฐานข้อมูลออกมาเป็นส่วน ๆ ให้เหมาะสมกับการใช้งาน

งานที่เกี่ยวข้องกับฐานข้อมูลอีกอย่างหนึ่งก็คืองานที่เรียกว่า การประมวลผลเชิงวิเคราะห์ออนไลน์ (On-Line Analytical : OLAP) ระบบ OLTP ที่กล่าวไปแล้วนั้นเน้นที่การบันทึกเก็บข้อมูลใหม่ๆ เพิ่มเข้าไปในฐานข้อมูล ส่วนระบบ OLAP นั้นเน้นที่การค้นคืนข้อมูลที่มีอยู่แล้วจากฐานข้อมูลเพื่อนำมาวิเคราะห์อย่างละเอียด ผู้ใช้ระบบ OLAP ส่วนใหญ่คือผู้บริหาร หรือผู้ใช้อื่นๆ ดังนั้นปัจจัยสำคัญสำหรับความสำเร็จของระบบ OLAP ก็คือระบบจะต้องทำงานได้รวดเร็วสามารถค้นหาข้อมูลจากฐานข้อมูลขนาดใหญ่มาคำนวณได้อย่างครบถ้วนไม่ตกหล่น ในขณะที่เดียวกันระบบก็จะต้องมีความมั่นคง ไม่ผิดพลาดได้ง่ายระหว่างการใช้งาน ปัจจัยที่จะทำได้ตามที่กล่าวนี้มีอยู่ 3 ข้อ คือ

1. จะต้องมียระบบจัดคำสั่งค้นคืนข้อมูลให้ทำงานได้รวดเร็วที่สุด (Query Optimization)
2. การจัดดัชนี จัดกลุ่มข้อมูลและการคำนวณตำแหน่งที่อยู่ข้อมูล
3. การประมวลผลคำสั่งค้นคืนในแบบขนาน

แม้ว่าระบบ OLTP และ OLAP นี้จะเกี่ยวข้องกับข้อมูลธุรกรรมเหมือนกันแต่ก็มีความแตกต่างกันมากในกระบวนการทำงานที่เกี่ยวข้องกับข้อมูล หากพบว่าการอ่านข้อมูลจากฐานข้อมูลมาประมวลผลมีช่วงเวลาได้ตอบ (Response Time) ช้ามากและต้องการปรับการเก็บโดยการจัดดัชนีเพิ่มเติมให้การค้นคืนข้อมูลได้สะดวกขึ้น ก็จะส่งผลให้การบันทึกข้อมูลกลับต้องช้าลง

เพราะต้องเสียเวลาดำเนินการกับดัชนีมากขึ้นกว่าระบบเดิม ด้วยเหตุนี้จึงเป็นเรื่องยากที่เราจะปรับระบบทั้งสองให้มีสมรรถนะดีมากขึ้นพร้อมกัน

ปัจจุบันนี้แนวทางแก้ไขปัญหาลักษณะข้างต้นก็คือการแยกระบบ OLTP และระบบ OLAP ออกจากกันให้เป็นคนละระบบ โดยให้ระบบ OLTP สามารถจัดเก็บข้อมูลจำนวนมากได้อย่างรวดเร็ว มีประสิทธิภาพ และระบบ OLAP ก็สามารถค้นคืนและวิเคราะห์ข้อมูลตามความต้องการของผู้ใช้ได้อย่างรวดเร็ว ระบบ OLTP นั้นปกติยังคงปล่อยให้มันเป็นแบบเดิม หากใช้คอมพิวเตอร์ขนาดใหญ่ เช่น เครื่องเมนเฟรมและใช้ระบบจัดการฐานข้อมูลขนาดใหญ่อยู่แล้วก็เพียงแค่ปรับให้สามารถบันทึกจัดเก็บข้อมูลให้เร็วขึ้น จากนั้นก็จัดทำระบบขึ้นใหม่ให้แยกข้อมูลพื้นฐานออกจากฐานข้อมูลในระบบเดิมแล้วนำข้อมูลมาจัดทำดัชนีใหม่เพื่อให้ผู้บริหารวิเคราะห์ อย่างไรก็ตามทั้งระบบ OLTP และ ระบบ OLAP ก็อาจจะยังไม่เหมาะที่เราจะนำมาใช้ในการวิเคราะห์ทางธุรกิจ หรือช่วยผู้บริหารสำหรับการตัดสินใจ (Decision Support System) ทางธุรกิจได้เพราะต้องใช้เวลาในการประมวลผลที่นานพอสมควร และส่งผลกระทบต่อระบบการทำงานของเครื่องที่ใช้งานประจำวัน

เราจะมีวิธีการอย่างไรเพื่อที่จะทำให้ข้อมูลที่เราที่มีอยู่สามารถนำมาใช้ตอบสนองความต้องการทางธุรกิจได้อย่างรวดเร็ว และมีประสิทธิภาพ ดังนั้นจึงได้นำเอาแนวความคิดระบบ “คลังข้อมูล (Data Warehouse)” มาใช้ร่วมกันเพื่อตอบสนองงานในรูปแบบของคลังเก็บข้อมูลสำหรับการบริหาร และ หากองค์กรใดสามารถที่จะนำข้อมูลที่มีอยู่มาใช้ได้อย่างมีประสิทธิภาพย่อมที่จะทำให้องค์กรประสบความสำเร็จเหนือคู่แข่ง

ข้อมูลส่วนมากที่จัดเก็บในคลังข้อมูลนั้นปกติจะมีน้อยกว่าข้อมูลในฐานข้อมูลของระบบ OLTP เพราะเป็นข้อมูลที่ได้นำมาจัดกลุ่มให้เหมาะสมแก่การค้นคืนแล้ว ข้อมูลเหล่านี้จะมีลักษณะ Consistent กล่าวคือ ข้อมูลทุกรายการที่แสดงเรื่องเดียวกันจะต้องเขียนให้เหมือนกัน สะกดแบบเดียวกันหรือมีรหัสเดียวกัน หากข้อมูลมีลักษณะแตกต่างกันแล้วจะวิเคราะห์ข้อมูล ได้ยากหรืออาจทำให้ได้ผลลัพธ์ที่ไม่ถูกต้อง ในหน่วยงานและบริษัทขนาดใหญ่ นั้นโอกาสที่ข้อมูลทั้งหมดจะ “สะอาด” นั้นเป็นเรื่องที่ยาก ดังนั้นจึงจำเป็นจะต้องมีผู้ทำหน้าที่กั่นกรองและควบคุมคุณภาพของข้อมูลด้วย

### 2.2.2 การประมวลผลเชิงวิเคราะห์ออนไลน์ (On-Line Analytical : OLAP)

OLAP ย่อมาจาก Online Analytical Processing เป็นเทคโนโลยีที่ใช้ข้อมูลจากคลังข้อมูลเพื่อนำไปใช้ในการวิเคราะห์ และตัดสินใจทางด้านธุรกิจหรือทางการบริหารอย่างมีประสิทธิภาพ สามารถค้นหาคำตอบที่ต้องการ และแก้ปัญหาที่มีความซับซ้อนได้โดยใช้ระยะเวลาเพียงไม่นาน OLAP จะจัดการกับข้อมูลโดยคำนวณหายอดรวมที่มีการใช้งานบ่อยครั้งไว้ล่วงหน้า ดังนั้นเมื่อต้องการวิเคราะห์หรือหาผลรวมของข้อมูลอีกก็จะสามารถเรียกได้ง่าย และรวดเร็ว

## 2.3 ความแตกต่างของคลังข้อมูลกับฐานข้อมูลทั่วไปได้ดังนี้

สามารถสรุปความแตกต่างแต่ละหัวข้อได้ดังนี้

### 2.3.1 Consistency

ทั้ง OLTP และคลังข้อมูลต่างก็ให้ความสำคัญในเรื่องข้อมูลควรจะมีคุณสมบัติเหมือนกัน สำหรับ OLTP ซึ่งมีการทำ Transaction จำนวนมากๆ สิ่งที่ต้องการคือการทำ Transaction ให้ครบ ไม่มีการสูญหาย ดังนั้นจึงมีความจำเป็นว่า ผู้ส่งและผู้รับจะต้องรับรู้และตรวจสอบอยู่ตลอดเวลาว่าขณะนี้มีการทำ Transaction เกิดขึ้นหรือไม่ สำหรับคลังข้อมูลจะไม่สนใจทำการทำ Transaction แต่ครั้งแต่ละจะสนใจว่าการ Load Data ใหม่เข้ามานั้นทำสำเร็จหรือยังและการ Load Data เข้ามาทั้งหมดนั้นถูกต้องหรือไม่

### 2.3.2 Transaction

สำหรับระบบ OLTP นั้น ในแต่ละวันอาจมีการทำ Transaction มากมายซึ่งการทำ Transaction แต่ละครั้งจะใช้ข้อมูลเพียงแค่อเล็กน้ยเท่านั้น สำหรับคลังข้อมูล แต่ละวันจะทำแค่เพียง 1 Transaction ซึ่ง Transaction นี้อาจต้องใช้ข้อมูลเป็นจำนวนมากมาย

### 2.3.3 Time Dimension

สำหรับ OLTP นั้นจะทำงานอย่างรวดเร็วและทำ Transaction อย่างสม่ำเสมอ สถานะของข้อมูลต่างๆ มีการเปลี่ยนแปลงอยู่ตลอดเวลา และความสัมพันธ์ระหว่างเอนติตี้ต่างๆก็เปลี่ยนแปลงไปด้วย สำหรับระบบคลังข้อมูลมักจะเก็บข้อมูลในอดีตเพื่อใช้ในการวิเคราะห์ ดังนั้นข้อมูลจะไม่ค่อยมีการเปลี่ยนแปลงตลอดวัน เป็นที่น่าสังเกตว่าคลังข้อมูลไม่ต้องทำการ Normalization เหมือนกับฐานข้อมูลประจำวันที่ต้องทำการ Normalization ทั้งนี้เพราะในฐานข้อมูลประจำวันข้อมูลจำนวนมากมีการเปลี่ยนแปลงทำให้ทันสมัยอยู่ตลอดเวลา ดังนั้นประเด็นสำคัญจึงอยู่ที่การเปลี่ยนแปลงทำให้ทันสมัย การออกแบบฐานข้อมูลประจำวันจึงต้องทำให้มีความซ้ำซ้อนหรือ Redundancy น้อยที่สุด วิธีการที่จะทำได้ตามจุดประสงค์คือการทำ Normalization แต่สำหรับข้อมูลในคลังข้อมูลเป็นข้อมูลที่มีการกลั่นกรองมาแล้ว ใช้ในการวิเคราะห์ตอบคำถามของผู้บริหารประเด็นสำคัญจึงไม่อยู่ที่การทำให้ข้อมูลมีความทันสมัย ทำให้ข้อมูลในคลังข้อมูลสามารถมีความซ้ำซ้อนได้ เพราะความซ้ำซ้อนมีข้อดีคือ การตอบคำถามและการออกรายงานสามารถทำได้รวดเร็ว เนื่องจากไม่ต้อง Join หลายตาราง ดังนั้นในคลังข้อมูลจึงไม่มีความจำเป็นต้องทำการ Normalization

เราสามารถสรุปความแตกต่างของคลังข้อมูลกับฐานข้อมูลที่ใช้ประจำวันหรือระบบการประมวลผลธุรกรรมออนไลน์ (On-Line Transaction Processing : OLTP)

<i>OLTP systems</i>	<i>Data warehousing systems</i>
Holds current data	Holds historic data
Stores detailed data	Stores detailed, lightly, and highly summarized data
Data is dynamic	Data is largely static
Repetitive processing	<i>Ad hoc</i> , unstructured, and heuristic processing
High level of transaction throughput	Medium to low level of transaction throughput
Predictable pattern of usage	Unpredictable pattern of usage
Transaction driven	Analysis driven
Application oriented	Subject oriented
Supports day-to-day decisions	Supports strategic decisions
Serves large number of clerical/operational users	Serves relatively lower number of managerial users

รูปที่ 2.1 ภาพความแตกต่างระหว่างคลังข้อมูลกับฐานข้อมูลที่ใช้ประจำวัน

## 2.4 กระบวนการคลังข้อมูล

กระบวนการคลังข้อมูลจะมีส่วนในการจัดการทั้งหมด 4 ส่วนดังนี้

### 2.4.1 Data Sources

เป็นส่วนของข้อมูลจากหลายๆ แหล่งหรือจากฐานข้อมูลจากที่ต่างๆ เช่น ฐานข้อมูลที่เป็น Application , Packaged Application/ERP Data , Desktop Data , External Data , Web-Based Data

### 2.4.2 Staging Area

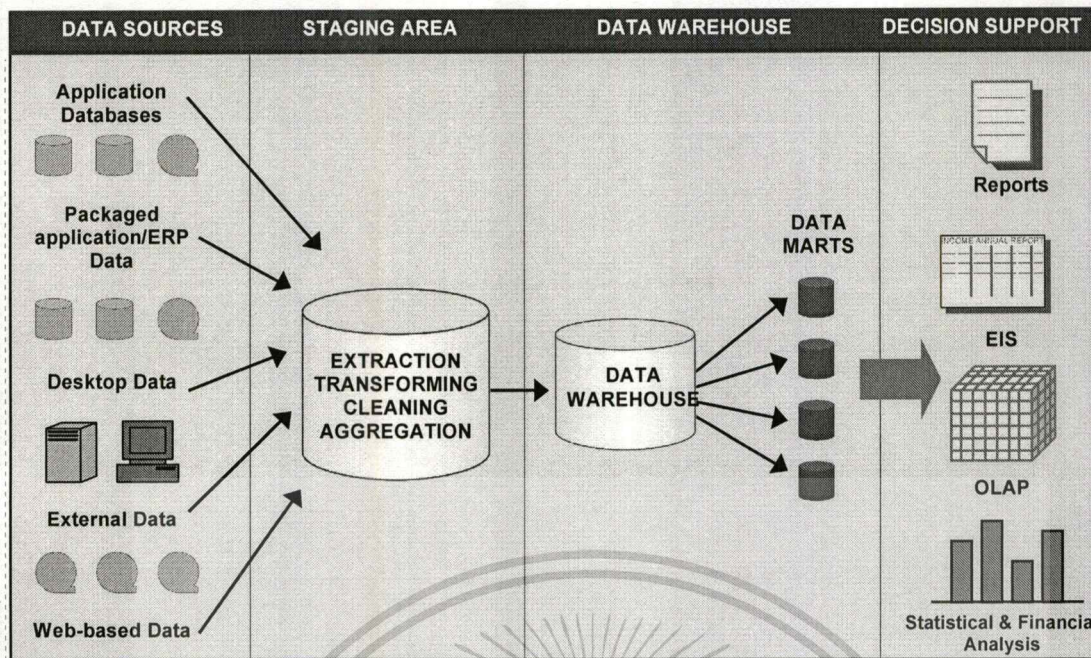
เป็นวิธีการที่คลังข้อมูลจะปรับข้อมูลเพื่อลดความซ้ำซ้อน (Cleaning) และการเลือกเฉพาะข้อมูลที่เป็นประโยชน์ (Filtering) เพื่อนำมาเก็บไว้ในคลังข้อมูลเพื่อใช้งานต่อไป

### 2.4.3 Data Warehouse

เป็นการจัดเก็บข้อมูลหรือการนำเอาข้อมูลที่ผ่านมาจากการ Cleaning และ Filtering แล้วมาเก็บบันทึกลงในคลังข้อมูล ซึ่งส่วนที่ทำหน้าที่จัดเก็บข้อมูลนั้นเรียกว่า “Data Warehouse Database”

### 2.4.4 Decision Support System

การเตรียมข้อมูลเพื่อใช้งาน หรือ เป็นการนำเอาข้อมูลที่มีอยู่ในคลังข้อมูลมาประมวลผลเพื่อให้ได้สารสนเทศที่ต้องการเพื่อการใช้งานต่อไป



รูปที่ 2.2 ภาพกระบวนการคลังข้อมูล

## 2.5 สถาปัตยกรรมคลังข้อมูล (Data Warehouse Architecture)

คลังข้อมูลอาจจะมีข้อมูลเป็นจำนวนมากมายมหาศาล ดังนั้นจึงจำเป็นต้องมีฐานข้อมูลของตนเองในการเก็บและประมวลผลข้อมูล หน่วยงานต้องมีโปรแกรมจัดการฐานข้อมูลและโปรแกรมอื่นๆ สำหรับช่วยในการเปลี่ยนรูปแบบข้อมูล, จัดกลุ่มข้อมูล, รวมข้อมูล และ โยกย้ายข้อมูลจากฐานข้อมูลหนึ่งไปยังฐานข้อมูลอื่นๆ โปรแกรมเหล่านี้ต้องทำงานได้ทั้งกับข้อมูลที่เป็นจำนวน, ข้อมูลกราฟิกข้อมูลภาพลักษณ์ และ ข้อมูลแบบมัลติมีเดีย โปรแกรมเหล่านี้จะต้องสามารถแปลงข้อมูลให้เหมาะสมที่จะนำไปวิเคราะห์และจัดทำรายงานในรูปแบบต่างๆ ได้

โดยที่คลังข้อมูลมีบริการสำคัญหลายอย่างให้แก่ผู้ใช้ ซึ่งอาจจะไม่ชำนาญในด้านคอมพิวเตอร์ ดังนั้นการจัดคลังข้อมูลจึงจำเป็นต้องจัดการฝึกอบรมให้แก่ผู้ใช้ด้วย นอกจากนี้ยังอาจจะต้องจัดระบบอธิบายการใช้เอาไว้ในระบบด้วย เพื่อให้ผู้ใช้สามารถเรียกคำอธิบายมาใช้เมื่อต้องการได้ ส่วนประกอบสำคัญอีกอย่างหนึ่งก็คือที่ปรึกษาเกี่ยวกับคลังข้อมูลเพื่อช่วยเหลือผู้ใช้ให้สามารถใช้ระบบได้อย่างมีประสิทธิภาพมากขึ้น

### 2.5.1 รายละเอียดของฐานข้อมูลสำหรับคลังข้อมูล

ในที่นี้จะขอกล่าวรายละเอียดของฐานข้อมูลสำหรับคลังข้อมูลซึ่งประกอบด้วย

1. Dimensional Modeling เป็นชื่อเรียกของเทคนิคในการทำให้ฐานข้อมูลง่ายต่อการทำความเข้าใจโดยการมองภาพของฐานข้อมูลเป็นลูกบาศก์ที่มี 3 , 4 , 5 มิติ หรือมากกว่านั้น ทำให้สามารถจินตนาการการหัน หรือ แบ่งลูกบาศก์ที่มีลักษณะเหมือนลูกเต๋านี้ได้ นั่นคือสามารถตัดข้อมูล มาวิเคราะห์ดูในช่วงใดก็ได้ และหมุนข้อมูลดูได้จากทุกๆ ด้านของลูกเต๋า ตัวอย่างเช่น เราขายสินค้า (Product) ในหลายๆ ที่ (Market) และในช่วงเวลาต่างๆ กัน (Time) เราสามารถสร้าง

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Dimensional Modeling ได้โดยให้ Label คือ Product , Market และ Time อยู่บนแต่ละด้านของลูกบาศก์ที่เป็น 3 มิติ แต่ละจุดภายในลูกบาศก์เกิดจากการตัดของ Coordinate ซึ่งมี Label อยู่ที่ขอบของลูกบาศก์ ดังนั้นจุดต่างๆภายในลูกบาศก์คือผลลัพธ์ทางด้านธุรกิจที่พิจารณาจากทั้ง 3 เรื่องคือ สินค้า , ที่ขายสินค้า และเวลา พร้อมๆกัน

2. Star Join Schema เป็นชื่อหนึ่งของ Dimensional Model ซึ่งเป็นชื่อที่ใช้กันเนื่องจาก Diagram มีรูปร่างคล้ายดาว ซึ่งมีตารางใหญ่ 1 ตารางอยู่ตรงกลางซึ่งเรียกว่า Fact Table และมีตารางเล็กๆ ที่มีความสัมพันธ์กับตารางหลักนั้นอยู่รอบๆ เรียกว่า Dimensional Table ซึ่งตารางหลักนี้เป็นตารางเดียวที่ใช้ Multiple Join เพื่อเชื่อมต่อกับตารางอื่นๆ แต่ตารางอื่นๆ ที่อยู่รอบๆ จะมีเพียงแค่ Single Join เพื่อเชื่อมเข้ากับตารางหลักเท่านั้น

2.1 Fact Table ข้อมูลที่เก็บอยู่ใน Fact Table เรียกว่า “Grain of fact Table” แต่ละเรคคอร์ดใน Fact Table จะแสดงถึง ยอดขายทั้งหมดของสินค้าหนึ่งที่ยขายได้ในสถานที่ที่แห่งหนึ่งในหนึ่งวัน Fact Table จะเป็นที่เก็บข้อมูลที่วัดได้ของธุรกิจหนึ่งๆ เช่น ข้อมูลที่เป็นตัวเลข ข้อมูลที่วัดได้แต่ละตัวนี้มาจากการ Intersection กันของทุกๆ Dimensions

2.2 Dimensional Table เก็บคำอธิบายของแต่ละ Dimension ของธุรกิจเอาไว้ ซึ่งคำอธิบายเหล่านี้จะช่วยในการอธิบายถึงสมาชิกในทุกๆ Dimension และใน Dimension Table จะประกอบด้วยหลายๆ Attribute ซึ่ง Attribute ที่จำเป็นจะต้องเป็นตัวอักษรและแต่ละ Attribute ต้องแยกออกจากกัน

### 2.5.2 สถาปัตยกรรมของคลังข้อมูล

สถาปัตยกรรมคลังข้อมูลเป็น โครงสร้างมาตรฐานที่ใช้อธิบาย เพื่อให้เข้าใจแนวคิดและกระบวนการของคลังข้อมูลนั้นๆ ซึ่ง โดยทั่วไปแล้วคลังข้อมูลแต่ละระบบอาจจะมีรูปแบบที่ไม่เหมือนกันได้ เพื่อให้เหมาะสมกับองค์กรนั้นๆ ทั้งนี้ส่วนประกอบต่างๆ ภายใน Data Warehouse Architecture ที่สำคัญ ได้แก่

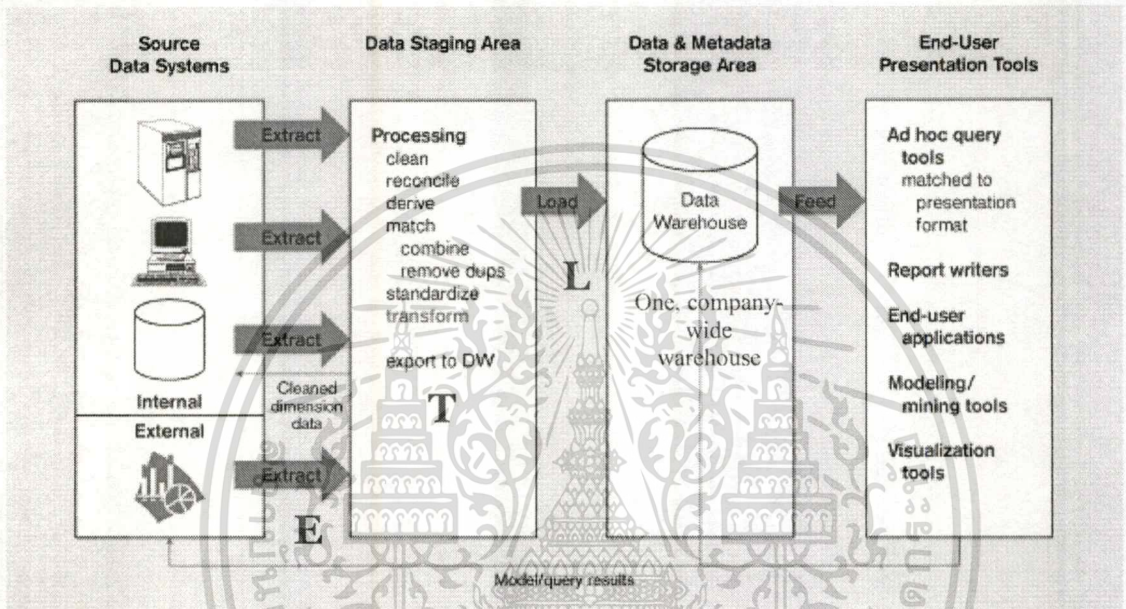
1. Operational Database หรือ External Database Layer ทำหน้าที่จัดการกับข้อมูลในระบบงานปฏิบัติการหรือแหล่งข้อมูลภายนอกองค์กร

2. Information Access Layer เป็นส่วนที่ผู้ใช้ปลายทางติดต่อผ่านโดยตรง ประกอบด้วย ฮาร์ดแวร์และซอฟต์แวร์ที่ใช้ในการแสดงผลเพื่อการวิเคราะห์ โดยมีเครื่องมือช่วยเป็นตัวกลางที่ผู้ใช้ใช้ติดต่อกับคลังข้อมูล โดยในปัจจุบันเครื่องมือที่ได้รับความนิยมเพิ่มขึ้นอย่างรวดเร็ว นั่นคือ Online Analytical Processing Tool หรือ OLAP Tool ซึ่งเป็นเครื่องมือที่มีความสามารถในการวิเคราะห์ที่ซับซ้อนและแสดงข้อมูลในรูปแบบหลายมิติ

3. Data Access Layer เป็นส่วนต่อประสานระหว่าง Information Access Layer กับ Operational Layer

4. Data Director (Metadata) Layer เพื่อให้เข้าถึงข้อมูลได้ง่ายขึ้นและเป็นการเพิ่มความ

5. Process Management Layer ทำหน้าที่จัดการกระบวนการทำงานทั้งหมด
6. Application Messaging Layer เป็นMiddleware ทำหน้าที่ในการส่งข้อมูลภายในองค์กรผ่านทางเครือข่าย
7. Data Warehouse (Physical) Layer เป็นแหล่งเก็บข้อมูลของทั้ง Information Data และ External Data ในรูปแบบที่ง่ายแก่การเข้าถึงและยืดหยุ่นได้
8. Data Staging Layer เป็นกระบวนการแก้ไขและดึงข้อมูลจาก External Database



รูปที่ 2.3 ภาพสถาปัตยกรรมคลังข้อมูล

**Source Data System**

เป็นข้อมูลจากภายนอก ซึ่งในที่นี้หมายถึงรวมทั้งข้อมูลที่มาจากภายนอกองค์กรและภายในองค์กร (แต่อยู่นอกระบบคลังข้อมูล) ข้อมูลจะได้รับการตรวจสอบความถูกต้องในขั้นต้น ก่อนที่จะส่งไปยังส่วนอื่นๆ ของคลังข้อมูลต่อไป

การจะรับข้อมูลภายนอกนั้นอาจจะได้มาโดยผ่านทางระบบเครือข่ายคอมพิวเตอร์ ระบบอินเทอร์เน็ต หรือแม้แต่ได้มาจากการใช้โปรแกรมคอมพิวเตอร์เพื่อการป้อนข้อมูลด้วยเจ้าหน้าที่ก็ได้เช่นกัน

**Data String Area**

Data String Area จะทำหน้าที่เป็นเสมือนด่านสุดท้ายของคลังข้อมูล ที่จะทำหน้าที่เป็นที่พักและตรวจตราข้อมูลในรายละเอียด เมื่อข้อมูลเข้ามาถึงยัง Data String Area แล้ว ข้อมูลที่พักอยู่ที่นี่จะถูกดำเนินการโดยกระบวนการหลายๆ อย่างเพื่อทำให้ข้อมูลนั้นพร้อมสำหรับการนำไปเก็บไว้ในฐานข้อมูลคลังข้อมูล

## Data Warehouse Database

Data Warehouse Database ถูกใช้เพื่อเก็บบันทึกข้อมูลต่างๆ ที่จำเป็นสำหรับการวิเคราะห์ข้อมูลขององค์กร กระบวนการส่วนใหญ่ที่เกิดขึ้นในการพัฒนาระบบคลังข้อมูล จะมีจุดประสงค์เพื่อการออกแบบข้อมูลภายใน Data Warehouse Database

## End-User Presentation Tool

เป็นเครื่องมือที่ใช้ในการดึงข้อมูลจากคลังข้อมูลมาแสดงผลในรูปแบบต่างๆ เช่น Ac hoc Query Tool , Report Writer , End-user Applications , Modeling/mining tool , Visualization Tool

กระบวนการต่างๆที่ทำใน Data String Area ถือเป็นกระบวนการที่ทำให้ข้อมูลสามารถเก็บบันทึกใน Data Warehouse Database ได้อย่างถูกต้อง ในขณะที่ผลลัพธ์ต่างๆที่ต้องการสามารถดึงได้จาก Data Warehouse Database เช่นเดียวกัน

## 2.6 กระบวนการการรวบรวมข้อมูลและวิเคราะห์หารูปแบบข้อมูลในคลังข้อมูล

เนื่องจากข้อมูลพื้นฐานของฐานข้อมูลในคลังข้อมูลประกอบด้วยเรคคอร์ดจำนวนมาก จะต้องมีการออกแบบคลังข้อมูลเพื่อรวบรวมข้อมูลและวิเคราะห์หารูปแบบข้อมูลในคลังข้อมูล ขึ้นตอนการออกแบบประกอบด้วย 4 ขั้นตอน

ขั้นตอน 1 เลือก Business Process ที่ต้องการสร้าง ซึ่งเป็นขบวนการหลักที่ต้องการทำในองค์กร ซึ่งขบวนการนั้นมีระบบเดิมสนับสนุนอยู่ ข้อมูลในระบบนั้นสามารถนำมารวบรวมเพื่อทำเป็นคลังข้อมูลได้ เช่น ใบส่งของ (Order) , ใบส่งของ (Invoices) , รายการสินค้า (Inventory) , ยอดขาย (Sales) Business Process ที่เลือกมานั้นต้องเป็นหัวข้อธุรกิจที่สนใจ เพื่อจะได้สามารถทำการออกแบบคลังข้อมูลให้เกี่ยวข้องกับเฉพาะหัวข้อธุรกิจที่สนใจเท่านั้น ส่วนข้อมูลที่ไม่ได้เกี่ยวข้องกับหัวข้อธุรกิจก็จะไม่นำมารวมในคลังข้อมูล

ขั้นตอน 2 เลือก Grain ของ Business Process ข้อมูลที่เป็นข้อมูลพื้นฐานที่เก็บอยู่ใน Fact Table เรียกว่า Grain ใน Business Process นี้มี Grain ที่มีอยู่ทั่วไป เช่น ข้อมูลของการทำ Transaction ในแต่ละครั้ง (Individual Transaction) , ข้อมูลของการทำงานในแต่ละวัน , สรุปในแต่ละวัน (Individual Daily Snapshots) , ข้อมูลจากการสรุปการทำงานในแต่ละเดือน (Individual Monthly Snapshots)

ขั้นตอน 3 เลือก Dimension ที่จะถูกนำมาใช้กับแต่ละ Record ของ Fact Table นั้น Dimension ที่มีอยู่ทั่วไป เช่น เวลา (Time) , สินค้า (Product) , ลูกค้า (Customer) ซึ่งแต่ละ Dimension จะถูกอธิบายแยกกัน ในลักษณะของ Dimensional Attribute ซึ่งอธิบายแต่ละ Dimension เป็นตัวหนังสือ ซึ่ง Attributes เหล่านี้จะถูกเก็บอยู่ในแต่ละตาราง Dimension

ขั้นตอน 4 เลือก Measured Fact (ข้อมูลที่มีการวัด , การประมวลผล หรือการคำนวณไว้แล้ว) ที่จะเก็บอยู่ในแต่ละเรคคอร์ดของ Fact Table ปริมาณต่างๆหรือ Measured Fact ที่เพิ่มเข้าไป

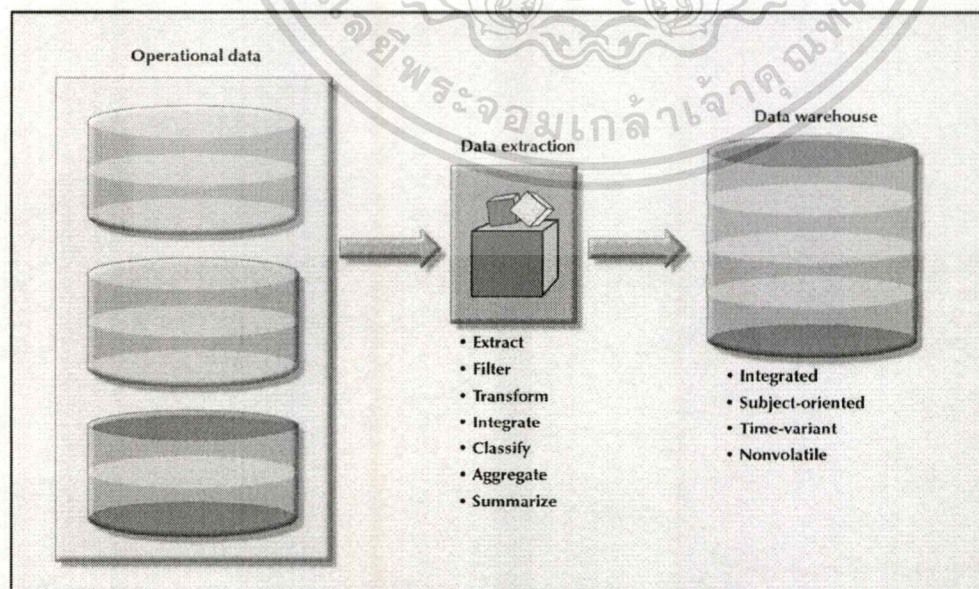
ซึ่งจะมีลักษณะเป็นตัวเลขได้แก่ ปริมาณที่ขายได้ (Quality Sold) และ จำนวนเงินที่ได้รับจากการขาย (Dollars Sold) ถ้านักวิเคราะห์ต้องการสอบถามข้อมูล (Query) โดยไม่มีการกำหนดเงื่อนไขให้กับบางตาราง Dimension แล้วในคลังข้อมูลมีแต่เฉพาะข้อมูลพื้นฐานเท่านั้น นั่นหมายความว่า การสอบถาม ข้อมูล (Query) จะต้องทำการรวมข้อมูลภายในเรคคอร์ดจำนวนมากมาย ถ้าเป็นเช่นนี้จะทำให้การทำ Query แต่ละครั้งเกิดการสิ้นเปลืองมาก ดังนั้นจึงต้องทำการรวบรวมหรือข้อสรุป (Aggregate) ข้อมูลไว้ล่วงหน้าเพื่อเร่งให้การทำ Query สามารถทำได้เร็วขึ้นมี ประสิทธิภาพในการทำงานมากขึ้น การรวบรวมหรือข้อสรุป (Aggregate) สามารถถูกสร้างได้มากมายตามที่ต้องการในเฉพาะส่วนที่ต้องการเท่านั้น

ในการคิวรีข้อมูลในคลังข้อมูลจะใช้ Structured Query Language (SQL) เป็นมาตรฐาน สำหรับการสอบถามข้อมูลทั้งหมดในคลังข้อมูลที่เกี่ยวข้องกับ Fact Table การสอบถามข้อมูลจะต้องมีการใช้ Dimension Table ในการทำการสอบถามข้อมูลเสมอ เพราะใน Dimension Table จะเก็บชื่อและคำอธิบายที่การสอบถามข้อมูลต้องการใช้อาไว้ การสอบถามข้อมูลจะประกอบไปด้วย Where Clause 2 ส่วนคือ

- การ Join ระหว่าง Fact Table และ Dimension Table
- เซตของข้อกำหนดหรือเงื่อนไข (Criteria) สำหรับคอลัมน์ที่อยู่ภายใน Dimension Table

## 2.7 เทคนิคในการสร้างคลังข้อมูล

ในการดำเนินการสร้างคลังข้อมูลนั้นต้องคำนึงถึงแบบแผนและหน้าที่ต่างๆของแต่ละขั้นตอนสิ่งที่สำคัญที่สุดคือการเตรียมข้อมูลในแต่ละขั้นตอนให้ถูกต้องโดยขบวนการสร้างจะมีดังนี้



รูปที่ 2.4 ภาพการสร้างคลังข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## Operation data

ในเมื่อรูปแบบของข้อมูลมีความแตกต่างกันในแต่ละข้อมูลที่ได้รับเข้ามาสู่กระบวนการ จึงเกิดเป็นกระบวนการและวิธีการที่คลังข้อมูลจะปรับข้อมูลเพื่อลดความซ้ำซ้อนและเลือกเฉพาะข้อมูลที่เป็นประโยชน์เพื่อนำมาเก็บไว้ในคลังข้อมูลเพื่อใช้งานต่อไป

## Data extraction

เป็นการดึงข้อมูลออกจากฐานข้อมูลที่จัดเตรียมไว้สำหรับการทำคลังข้อมูล โดยมีกระบวนการเพื่อให้ได้ข้อมูลที่ต้องการ เช่น Transform, Filter, Extract, Integrate เป็นต้น

## Data warehouse

การนำเอาโครงสร้างของข้อมูลที่สร้างไว้ใน Physical Data Base มาสร้างให้เกิดเป็นจริง Operation data เพื่อพร้อมสำหรับการเก็บข้อมูลที่จะเข้ามาจากภายนอก

### 2.7.1 การเคลื่อนที่ของข้อมูลในคลังข้อมูล

ข้อมูลที่จะจัดเก็บภายในคลังข้อมูลมีการเคลื่อนที่ของข้อมูล (Information Flow) 5 ประเภท ดังนี้

1. Inflow คือการนำข้อมูลจากฐานข้อมูลอื่นเข้าสู่คลังข้อมูลทั้งฐานข้อมูลภายในและภายนอกองค์กร โดยในขั้นนี้อาจมีการเปลี่ยนแปลงโครงสร้างข้อมูล การทำ Denormalize การลบหรือเพิ่มฟิลด์เพื่อให้ข้อมูลทั้งหมดอยู่ในเนื้อหาที่สนใจเดียวกัน ในขั้นตอนนี้อาจใช้เครื่องมือที่เรียกว่า Data Warehouse Tool

2. Upflow เมื่อข้อมูลที่เราต้องการอยู่ในคลังข้อมูลแล้วในบางครั้งอาจต้องมีการเพิ่มคุณค่าให้กับข้อมูลด้วย เพื่อให้ข้อมูลอยู่ในรูปแบบที่เป็นประโยชน์มากที่สุดต่อการนำเครื่องมือมาใช้ ซึ่งได้แก่การจัดกลุ่มข้อมูลหาค่าทางสถิติที่ซับซ้อน จัดข้อมูลให้อยู่ในรูปแบบ หรือ เทมเพลตมาตรฐาน

3. Downflow เป็นขั้นตอนของการปรับปรุงเปลี่ยนแปลงข้อมูลเก่าและไม่อยู่ในเนื้อหาขององค์กรสนใจออกไปจากคลังข้อมูลขององค์กร

4. Outflow เป็นขั้นที่ผู้ใช้เรียกใช้ข้อมูลในคลังข้อมูลผ่านเครื่องมือต่างๆ โดยการเรียกใช้ อาจมีเพียงขอเรียกเป็นครั้งคราวเป็นประจำทุกวัน/เดือน หรือแม้กระทั่งต้องการแบบทันที

5. Metaflow ข้อมูลที่จัดเก็บในคลังข้อมูลจะถูกทำข้อมูลไว้อีกชุดหนึ่งเป็นแหล่งที่มาของข้อมูลนั้น หรือแม้กระทั่งที่อยู่ของข้อมูลนั้นในคลังข้อมูลและข้อมูลอื่นที่เกี่ยวข้อง

### 2.7.2 การออกแบบคลังข้อมูล

การออกแบบคลังข้อมูลต้องประกอบไปด้วย

#### 1. กำหนดความต้องการ

ซึ่งประกอบด้วยความต้องการทางธุรกิจของผู้ใช้ และความต้องการทางเทคนิคกำหนด

ขอบเขตงาน และกำหนดแหล่งที่มาของข้อมูลที่ต้องใช้ โดยศึกษาจากงานที่ใช้อยู่เดิม ความ

## 2. ออกแบบและสร้างฐานข้อมูล

โครงสร้างของฐานข้อมูลสำหรับงานคลังข้อมูล จะแตกต่างจากการออกแบบงานที่เป็นฐานข้อมูลเชิงสัมพันธ์ สำหรับคลังข้อมูลนั้นจะไม่เน้นการออกแบบให้ نرمัลไลซ์ คือ ไม่จำเป็นต้องเป็นไปตามกฎของการ نرمัลไลซ์ เพื่อให้การคิวรีได้ผลลัพธ์รวดเร็วที่สุด จึงยอมให้เกิดการซ้ำซ้อนของข้อมูลในแต่ละตารางได้ การซ้ำซ้อนของข้อมูลนี้จะผิดกฎของการ نرمัลไลซ์ แต่จะอยู่ในหลักของการทำดีน نرمัลไลซ์ข้อมูล ซึ่งในการออกแบบฐานข้อมูลของคลังข้อมูล ประกอบด้วยตารางที่สำคัญ 2 ประเภท คือ

1. ตารางหลัก เป็นตารางที่เก็บข้อมูลโดยรวมทั้งหมดและสามารถตอบคำถามที่ต้องการได้เพียงพอ เช่น ถ้าเก็บข้อมูลยอดขายรวมการขายและมีคอลัมน์ของรหัสลูกค้ารหัสสินค้า เวลายอดขาย เป็นต้น การออกแบบตารางหลักจะเลือกเฉพาะคอลัมน์ที่ต้องการใช้งานเท่านั้น และพยายามลดขนาดของคอลัมน์ที่มีความยาวมากเกินไปโดยไม่จำเป็น ข้อมูลจะไม่มีการแก้ไขหรือเปลี่ยนแปลงยกเว้นแต่การเพิ่มข้อมูลใหม่เข้าไปในตาราง

2. ตารางไคเมนชัน เป็นตารางที่เก็บความหมายของรหัสที่ใช้ในตารางหลัก มีประโยชน์เพื่อช่วยให้การสอบถามข้อมูลแสดงคำอธิบายได้ชัดเจนขึ้นกว่า การแสดงโดยใช้รหัสเพียงอย่างเดียว เช่น ระบบการขาย จะมีตารางไคเมนชันเป็น ตาราง Project เพื่อเก็บรหัสสินค้า ชื่อสินค้า และรายละเอียดของสินค้า ตาราง Customer เก็บชื่อรหัสสินค้า และที่อยู่ลูกค้า เป็นต้น นอกจากนี้ ตารางไคเมนชันยังอาจจัดข้อมูลเป็นหลายระดับ คือ มีระดับใหญ่ ระดับรอง ได้ เช่น ถ้าเลือกไคเมนชันของเวลา สามารถจัดให้ระดับไคเมนชันใหญ่สุดคือปี ระดับต่อมาคือไตรมาส และระดับถัดไปคือเดือน ดังนั้นในการดึงข้อมูลของไคเมนชันเวลาจะเลือกดูได้ตั้งแต่ระดับปี ไตรมาส และเดือน ซึ่งทำได้โดยใช้เครื่องมือของ OLAP เพื่อทำการ Drill down หรือ Drill up ที่ไคเมนชันนั้นได้

### 2.7.3 วิธีการออกแบบฐานข้อมูลสำหรับคลังข้อมูลโดยใช้วิธี Nine-Step Methodology

วิธีการนี้ถูกเสนอโดย Kimball ในปี 1996 เรียกว่าระเบียบวิธี 9 ขั้นหรือ Nine-Step Methodology โดยวิธีการนี้เริ่มจากการออกแบบจากส่วนย่อย ที่แสดงถึงแต่ละระบบงานขององค์กร หรือเรียกอีกอย่างหนึ่งว่าดาต้ามาร์ท (Data Mart)

ดาต้ามาร์ท คือ คลังข้อมูลขนาดเล็กนั่นเอง ข้อมูลในดาต้ามาร์ทเป็นสับเซตของคลังข้อมูลซึ่งลักษณะดังนี้

1. กำหนดดาต้ามาร์ท คือ การเลือกที่จะสร้างดาต้ามาร์ทของระบบงานใดบ้าง และระบบงานใดเป็นระบบงานแรก

2. กำหนดตารางหลักของดาต้ามาร์ท คือ กำหนดเนื้อหาหลักที่ควรจะเป็นของดาต้ามาร์ท โดยการเลือกเอนทิตีหลักและกระบวนการที่เกี่ยวกับเอนทิตีนั้นๆ ออกมาจาก E-R Model ขององค์กร นั้นหมายถึงจะทำให้เราทราบถึง Dimension Table ที่ควรจะมีด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่อนำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. กำหนดแอตทริบิวต์ที่จำเป็นในแต่ละ Dimension Table คือ การกำหนดแอตทริบิวต์ที่บอกหรืออธิบายรายละเอียดของ Dimension

4. กำหนดแอตทริบิวต์ที่จำเป็นใน Fact Table โดยแอตทริบิวต์หลักใน Fact Table จะมาจาก Primary Key ในแต่ละ Dimension Table นอกจากนี้แล้ว ยังสามารถมีแอตทริบิวต์ที่จำเป็นอื่นๆ ประกอบอยู่ด้วย

5. จัดเก็บค่าการคำนวณเบื้องต้นใน Fact Table คือ การจัดเก็บที่ได้จากการคำนวณให้เป็นแอตทริบิวต์ใน Fact Table ถึงแม้ว่าจะสามารถหาค่าได้จากแอตทริบิวต์อื่นๆ ก็ตาม ทั้งนี้เพื่อให้การสอบถามมีประสิทธิภาพมากขึ้น สามารถทำงานด้วยความเร็วที่เพิ่มขึ้น เนื่องจากไม่ต้องคำนวณค่าใหม่ทั้งหมด ถึงแม้ว่าจะเกิดความซ้ำซ้อนของข้อมูลในการจัดเก็บบ้างก็ตาม

6. เขียนคำอธิบาย Dimension Table ทั้งนี้ก็เพื่อให้ผู้ใช้สามารถใช้งานดาต้ามาร์ทได้อย่างมีประสิทธิภาพเพราะเกิดความเข้าใจอย่างดีในส่วนต่างๆ

7. กำหนดระยะเวลาในการจัดเก็บข้อมูลในฐานข้อมูล โดยอาจจะเป็นการจัดเก็บเพียงช่วงระยะเวลา 1-2 ปี หรือนานกว่านั้น ขึ้นอยู่กับความต้องการขององค์กร

8. การติดตามปัญหาการเปลี่ยนแปลงของ Dimension อย่างซ้ำๆ คือ การเปลี่ยนเอาแอตทริบิวต์ของ Dimension Table เก่ามาใช้แล้วส่งผลกระทบต่อข้อมูลปัจจุบันของ Dimension Table

9. กำหนดคิวรี เป็นการออกแบบด้านกานภาพเพื่อให้ผู้ใช้เกิดความสะดวกในการใช้งานและสามารถทำงานได้อย่างมีประสิทธิภาพเมื่อดำเนินการทั้ง 9 ขั้นตอนสำหรับแต่ละดาต้ามาร์ทเสร็จแล้ว จึงจะนำทั้งหมดมารวมกันเป็นภาพของคลังข้อมูลขององค์กรต่อไป

เมื่อดำเนินการทั้ง 9 ขั้นตอนสำหรับแต่ละดาต้ามาร์ทเสร็จแล้ว จึงจะนำทั้งหมดมารวมกันเป็นคลังข้อมูลขององค์กรต่อไป

1. การเตรียมข้อมูล คือการเตรียมแผนในการดึงข้อมูลจากระบบ OLTP มาสร้างเป็นข้อมูลของระบบคลังข้อมูลและเตรียมหาข้อมูลจาก OLTP ที่ต้องใช้ย้อนหลังสำหรับการสร้างคลังข้อมูล

2. ออกแบบการเพิ่มหรือปรับปรุงข้อมูลในคลังข้อมูล การเพิ่มข้อมูลในคลังข้อมูลเป็นสิ่งปกติและสำคัญ เนื่องจากจะต้องมีการนำข้อมูลจากระบบ OLTP มาเพิ่มที่คลังข้อมูล เพื่อให้มีข้อมูลที่ทันสมัย นอกจากนี้ยังอาจจะมีการปรับเปลี่ยนโครงสร้างที่เก็บข้อมูลเพราะเปลี่ยนแนวความคิดหรือเพิ่มวิธีในการวิเคราะห์ข้อมูล เป็นต้น การเพิ่มและปรับปรุงข้อมูลทำได้ 3 วิธีคือ

2.1 Incremental Update คือ การเพิ่มข้อมูลต่อท้ายข้อมูลที่มีอยู่เดิม โดยไม่ได้เปลี่ยนแปลง โครงสร้างของข้อมูล และข้อมูลส่วนนี้เป็นข้อมูลที่ยังไม่เคยนำเข้าไปในคลังข้อมูล

2.2 Refresh Data คือ การทำให้ข้อมูลทั้งหมดมีความถูกต้องทันสมัย โดยทำการประมวลผลข้อมูลใหม่ทั้งหมด มักเกิดจากการเปลี่ยนแปลงโครงสร้างของฐานข้อมูล เช่น การเพิ่มไคเมนชัน หรือการเพิ่มค่าชี้วัดเชิงปริมาณ

2.3 Rebuild the Dimension Structure คือ การปรับปรุงโครงสร้างที่ใช้สำหรับการวิเคราะห์ เช่น การเพิ่มไคเมนชัน หรือการเพิ่มค่าชี้วัดเชิงปริมาณ

3. การบำรุงรักษาระบบฐานข้อมูล การเตรียมแผน และสำรองข้อมูลจากระบบคลังข้อมูลอย่างสม่ำเสมอรวมทั้งการทดสอบ และวางแผนการกู้ระบบเมื่อมีความเสียหายเกิดขึ้น

#### 2.7.4 การแปลงข้อมูลเข้าสู่ดาต้ามาร์ท

เมื่อเราออกแบบฐานข้อมูลสำหรับแต่ละดาต้ามาร์ทเสร็จแล้ว ขั้นตอนต่อไปที่สำคัญยิ่งก็คือการนำข้อมูลจากแหล่งข้อมูลไปแปลงให้อยู่ในแพลตฟอร์มของฐานข้อมูลที่ได้ออกแบบไว้ นั่นก็คือการแปลงข้อมูล หรือ Extraction Transformation and Loading (ETL) นั่นเอง โดยที่คุณภาพของการแปลงข้อมูลเป็นสิ่งสำคัญมากสำหรับการสร้างคลังข้อมูล จะแตกต่างกันไปตามคลังข้อมูลที่แต่ละองค์กรต้องการ โดยที่การแปลงข้อมูลหมายรวมถึงแต่การวิเคราะห์แหล่งข้อมูล กำหนดการส่งข้อมูลรวบรวมหรือสร้างข้อมูลภายนอก วางแผนและสร้างรูทีนของการแปลงข้อมูล และตรวจสอบความถูกต้องของข้อมูลที่ได้สามารถสรุปเป็นขั้นตอนได้ดังนี้

1. วิเคราะห์แหล่งข้อมูล เช่น ปริมาณของข้อมูล จำนวน และ ชนิดของการเข้าถึงแหล่งข้อมูล แพลตฟอร์มและภาษาโปรแกรมที่ใช้ เป็นต้น

2. ย้ายข้อมูลที่ต้องการจากระบบเดิมมาไว้ในบริเวณที่ใช้ปรับแต่งข้อมูล หรือเรียกบริเวณนี้ว่า Staging Area เพื่อนำมาเลือกเฉพาะส่วนที่ต้องการแปลงข้อมูลและตรวจสอบความถูกต้องหรือการทำความสะอาดข้อมูล

3. กำหนด Primary Key ของ Fact Table และ Dimension Table และกำหนด Foreign Key ระหว่าง Fact Table กับ Dimension Table

4. ย้ายข้อมูลที่ทำความสะอาดแล้วจาก Staging Area ลงสู่เซิร์ฟเวอร์ของดาต้ามาร์ท

5. สร้าง Metadata ของแต่ละดาต้ามาร์ท โดยเก็บรายละเอียดของข้อมูลการอัปเดตและส่งออกไว้ในดาต้ามาร์ท

6. ตรวจสอบความถูกต้องของข้อมูลซึ่งจะต้องกระทำตลอดทั้งกระบวนการแปลงข้อมูลทำได้ดังนี้

6.1 ตรวจสอบผลรวมทั้งหมดของจำนวนข้อมูลที่ดึงมาจากแหล่งข้อมูลที่เพิ่มเข้าไป

6.2 ตรวจสอบแก้ไขข้อมูลในระบบเดิมของแหล่งข้อมูล หรือในรูทีนของการแปลง ซึ่งควรจะต้องเก็บ

6.3 ข้อมูลในการตรวจแก้ไขไว้ใน Metadata ของการแปลงข้อมูลด้วย

6.4 ตรวจสอบค่าของข้อมูลให้ถูกต้องในกระบวนการรวบรวมข้อมูล

## 6.5 ตรวจสอบผลรวมของข้อมูลหลังจากย้ายข้อมูลลงสู่ตลาดมืดแล้ว

### 2.7.5 สิ่งที่ต้องพิจารณา ก่อนสร้างคลังข้อมูล

เนื่องจากการลงทุนสร้างคลังข้อมูลขึ้นมาใช้ เพื่อสนับสนุนการทำงานขององค์กรนั้น จำเป็นต้องมีค่าใช้จ่ายในการลงทุนมหาศาลทั้งที่สามารถวัดออกมาเป็นตัวเงินได้ เช่น ค่าใช้จ่ายด้านฮาร์ดแวร์ ซอฟต์แวร์ และ Infrastructure อื่นๆ ที่จำเป็นต้องใช้ ส่วนค่าใช้จ่ายที่ไม่เป็นตัวเงิน แต่มีความสำคัญอย่างมากได้แก่ กำลังแรงงานที่จะเสียไปของทรัพยากรบุคคลขององค์กรและเวลาที่ใช้ไปกับการพัฒนา ดังนั้น เมื่อองค์กรตัดสินใจสร้างคลังข้อมูลขึ้นแล้ว ควรจะประสบความสำเร็จด้วย โดยมีสิ่งที่จะต้องให้ความสนใจ โดยมีรายละเอียดดังนี้

1. ควรมีความชัดเจนในเป้าหมายร่วมของการสร้างระบบนี้ของคนในองค์กรเหมือนการตอบคำถามว่าทำไมคุณถึงคิดจะสร้างคลังข้อมูล? ซึ่งคำตอบขององค์กร ที่จะได้คือเป้าหมายที่ต้องการ โดยควรจะเขียนเป้าหมายนี้ ออกมาเป็นลายลักษณ์อักษรที่ชัดเจนเพื่อให้ทีมพัฒนาได้เข้าใจเป้าหมายร่วมกัน

2. ทำความเข้าใจสถาปัตยกรรมของระบบ เพื่อให้ทีมพัฒนาเข้าใจตรงกัน ในที่นี้หมายถึง Blueprint ที่แสดง E-R Model รวมของระบบความเข้าใจที่ตรงกันทำให้งานเดินไปได้เร็วขึ้น

3. เทคโนโลยีที่ใช้ควรอยู่ในวิสัยที่เหมาะสมทั้งด้านของตัวเงินและความยากง่ายในการเรียนรู้ ทั้งนี้หมายรวมทั้งฮาร์ดแวร์ ซอฟต์แวร์ และเครือข่าย อาจต้องมีการทดสอบและฝึกอบรมก่อนการใช้งานจริง

4. ทีมพัฒนาต้องมีวิสัยทัศน์เชิงบวกในการทำงาน เนื่องจากทีมพัฒนามักมาจากส่วนงานด้านเทคโนโลยีสารสนเทศแต่ในเนื้องานจริงๆ แล้วผู้ใช้ชั้นปลายเป็นส่วนงานอื่นๆขององค์กร ดังนั้นจึงจำเป็นอย่างยั้งที่จะให้ผู้ใช้ ชั้นปลายที่เป็นเจ้าของงานเข้ามาร่วมทำงานด้วยตั้งแต่ต้นโครงการ

5. ต้องมั่นใจได้ว่าทีมพัฒนาเข้าใจเป็นอย่างดีถึงความแตกต่างกันระหว่างฐานข้อมูลปฏิบัติการและฐานข้อมูลสนับสนุนการตัดสินใจ

6. จัดให้มีการฝึกอบรมโดยควรเป็นการฝึกอบรมก่อนเริ่มโครงการ โดยเฉพาะอย่างยิ่งการฝึกอบรมเกี่ยวกับเครื่องมือที่องค์กรจะใช้ในการพัฒนา ทั้งนี้อาจเป็นการฝึกอบรมจากบริษัทผู้ขาย

7. ควรหาบุคลากรที่มีประสบการณ์ในการพัฒนากลังข้อมูล เพื่อทำหน้าที่เป็นผู้จัดการโครงการ หรือถ้าในองค์กรไม่เคยมีประสบการณ์เลยอาจจ้างที่ปรึกษาที่มีความเชี่ยวชาญและมีประสบการณ์ด้านนี้โดยเฉพาะมาช่วยทีมพัฒนา

8. โปรแกรมที่จะใช้นำเสนอข้อมูลในคลังข้อมูล ต้องสามารถเรียนรู้ได้ง่ายและผู้ใช้สามารถใช้งานได้อย่างมีประสิทธิภาพ

## 2.8 สถาปัตยกรรมของการจัดการข้อมูลและการวิเคราะห์ข้อมูลของ MS SQL SERVER 2005

จากแนวคิดที่บริษัท ไมโครซอฟท์ ต้องการสร้างฐานข้อมูลและเครื่องมือที่มีคุณสมบัติ การวิเคราะห์ การผสานข้อมูล และการแปลงข้อมูลของเพื่อช่วยให้บริษัทต่างๆ ได้ปรับปรุงให้แอปพลิเคชันที่มีอยู่เดิมมีคุณค่ามากขึ้น ซึ่งมีระบบบริหารข้อมูลระดับเอนเตอร์ไพรส์ พร้อมกับมีเครื่องมือระบบธุรกิจอัจฉริยะ (Business Intelligence -BI) ในตัว กลไกดาต้าเบสของ SQL Server 2005 ช่วยให้จัดเก็บข้อมูลรีเลชันแนลและข้อมูลที่มีโครงสร้างได้อย่างปลอดภัยมากขึ้นและมีเสถียรภาพมากขึ้น รวมทั้งช่วยให้การสร้างและบริหารแอปพลิเคชันข้อมูลประสิทธิภาพสูงและพร้อมที่จะให้บริการตลอดเวลา รวมทั้งใน SQL Server 2005 ยัง ผสมผสานระบบวิเคราะห์ ระบบทำรายงาน ระบบผสานข้อมูล และระบบแจ้งเตือนที่ดีที่สุดเข้าไว้ด้วยกัน วิธีการนี้จะช่วยให้ธุรกิจต่างๆ สามารถสร้างและติดตั้งโซลูชัน BI ที่คุ้มค่าที่ช่วยในการจัดสรรข้อมูลไปยังทุกจุดภายในองค์กรได้ผ่านระบบข้อมูลสำหรับผู้บริหาร เว็บเซอร์วิส และอุปกรณ์โมบายล์ต่างๆ ได้

### 2.8.1 สิ่งจำเป็นในการสร้างแอปพลิเคชัน

ในโครงการนี้ได้นำ SQL Server 2005 มาใช้ในการจัดการคลังข้อมูลและสร้างแอปพลิเคชันต่างๆ ในระบบธุรกิจอัจฉริยะ (BI) ซึ่งจะมีส่วนประกอบด้วย

- การจัดการฐานข้อมูลเชิงสัมพันธ์สำหรับจัดเก็บข้อมูลเชิงสัมพันธ์หรือ คลังข้อมูล ทำงานโดยใช้ SQL Server นั่นเอง
- การนำ Analysis Server มาใช้ (OLAP)
- DTS (Data Transformation Service) เป็น Tool ใน SQL Sever ที่ช่วยการการโอนย้ายข้อมูลจาก Database หลากๆตัว ซึ่งจะคัดแยกแปลงรูปแบบและโหลดข้อมูล (ETL) จากระบบปฏิบัติการสู่หน่วยจัดเก็บข้อมูลเชิงสัมพันธ์
- ซอฟแวร์ในการบริหารจัดการระบบ ใช้ในการจัดการทั้งฐานข้อมูลและจัดการใน OLAP
- สามารถทำการวิเคราะห์ในการจัดทำ Data Mining
- แสดงผลในรูปของกราฟิกและสนับสนุนการเชื่อมต่อได้อย่างกว้างขวาง
- สามารถบริการข้อมูลต่างๆ ได้ เช่น Metadata

### 2.8.2 การติดต่อและการเชื่อมข้อมูล

การติดต่อและเชื่อมโยงข้อมูลใน Business Intelligence ของไมโครซอฟต์นั้นใช้ในอินเทอร์เน็ต และมาตรฐานแบบเปิดสำหรับการทำงานกับส่วนประกอบอื่นๆ

ทั้ง MS\_SQL Server และ Analysis Sever มีออปเจ็ทโมเดล เปิดโอกาสให้เชื่อมต่อผ่านทางโปรแกรมโดยตรง โดยไม่ต้องผ่านชั้นตอนผ่านอินเทอร์เน็ตในการจัดการของไมโครซอฟต์ เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- SQL Distributed Management Object (SQL-DMO) เป็นออปเจ็คโมเดลสำหรับ SQL Server
- ActiveX Data Object Multidimensional (ADO MD) เป็นออปเจ็คโมเดลสำหรับ PivotTable Service ซึ่งเป็นส่วนประกอบด้าน Client ของ Analysis service
- DSO (Decision Support Object) เป็นออปเจ็คโมเดลสำหรับ Analysis Sever  
ในบางครั้งการเชื่อมต่อในระดับต่ำไปยังฐานข้อมูลสามารถทำได้ทั้งองค์กร โดยใช้ OLE DB ซึ่งเป็นการติดต่อแบบที่ได้รับการออกแบบให้เรียกใช้ได้ทั้งแหล่งข้อมูลแบบสัมพันธ์และไม่ใช่แบบสัมพันธ์ แต่ ODBC มีความแตกต่างคือได้รับการออกแบบให้ครอบคลุมฐานข้อมูลเชิงสัมพันธ์ทั้งหมด โดยมาตรฐานของ OLE DB เป็นข้อกำหนดแบบเปิดที่ได้รับการพัฒนาาร่วมกันจากหลายองค์กร และยังเปิดให้นักพัฒนาทั่วโลกได้ใช้งานอย่างทั่วถึง

ในการติดต่อของ Analysis Service ได้ทำการกำหนดเพิ่มมาตรฐานใน OLE DB ด้วยข้อกำหนดเพิ่มอีก 3 อย่างด้วยกัน คือ

- OLE DB สำหรับ OLAP  
เหมาะสมกับการใช้งานในสำหรับการเรียกค้นและจัดทำรายงาน โดยในบางครั้งการค้นและทำรายงานอาจใช้มาตรฐานมากกว่าหนึ่งแบบ เพราะมีศักยภาพสามารถเชื่อมต่อข้อมูลได้หลายประเภท และติดต่อได้หลายมิติ
- OLE DB สำหรับ Data mining  
มีความสามารถในการใช้งานร่วมกับ Data Mining ที่มีอยู่กับ Analysis Service โดยทำงานร่วมกันกับอัลกอริธึมพิเศษหรือแอปพลิเคชันที่เป็นส่วนหนึ่งของ Analysis Service หรือใช้ Analysis Service cube
- XML สำหรับ Analysis  
สนับสนุนการทำงานแบบไม่ผูกติดและเชื่อมโยงแบบไม่มีสถานะในสภาพแวดล้อมของเว็บจึงมีความสามารถในการทำงานในรูปการวิเคราะห์ทุกรูปแบบผ่านเว็บ

### 2.8.3 การจัดการส่งต่อข้อมูล Data Transformation Service

เป็นขั้นตอนแรกๆของการทำ Business Intelligence เป็นการสร้างจัดเก็บข้อมูลเชิงสัมพันธ์ด้วยข้อมูลที่เหมาะสมสำหรับการวิเคราะห์ ซึ่งอาจจะส่งต่อง่ายๆหรือ ซับซ้อนก็ได้เพราะผู้ใช้งานไม่เพียงต้องการคัดเลือกข้อมูลจากแหล่งข้อมูลต่างๆเท่านั้น ในบางที่ที่ต้องการเชื่อมโยงข้อมูลและแปลงข้อมูลด้วย เพื่อให้มั่นใจถึงความถูกต้องและแน่นอนของข้อมูลที่ได้มาจากหลายระบบ จะสังเกตได้ว่าเป็นการบริการการจัดการที่ยืดหยุ่น สามารถปรับเปลี่ยนได้อย่างมากและสามารถจัดการกับข้อมูลที่มีปริมาณมากได้อย่างมีประสิทธิภาพด้วย

เอเจนซีเหล่านี้มีแนวโน้มที่จะอนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การใช้งาน DTS นั้นทำได้โดยการสร้างแอปพลิเคชัน DTS เป็น โมดูล โดยสร้างคอมโพเนนท์ที่นำกลับมาใช้ใหม่ได้ โดยโครงสร้างพื้นฐานของของ DTS นั้นจะทำการบันทึกการทำงานผลที่ได้ และจำนวนข้อมูลที่น่าเข้าในแต่ละขั้นตอนโดยอัตโนมัติและสามารถแจ้งให้ผู้ดูแลฐานข้อมูลให้ทราบได้ตลอดเวลาโดยผ่านช่องทางที่ได้ตั้งไว้ ในขณะที่เดียวกัน DTS ทำให้การรวบรวมข้อมูลจากแหล่งข้อมูลต่างๆง่ายขึ้นเป็นอย่างมากโดยทำการติดต่อผ่าน OLE DB หรือ ODBC ก็จะสามารถแปลงเข้าและรวมกันได้โดยใช้ DTS ตั้งเวลาอัตโนมัติเพื่อจัดส่งข้อมูลในระบบของ Business Intelligence ทั้งหมด ซึ่งงานพิเศษของ DTS ที่เกี่ยวกับ Business Intelligence ได้แก่

- การประมวลผล Cube จะสามารถประมวลผลแบบเพิ่มเติมหรือแบบสมบูรณ์ของมิติและ Cube ได้
- การคาดการณ์โดยใช้ Data Mining เมื่อพัฒนาและทดสอบโดยโมดูลของ Data Mining เรียบร้อยแล้วผู้ใช้สามารถใช้ DTS Data Mining Prediction เพื่อใช้โมเดลดังกล่าวกับข้อมูลชุดใหม่ได้

#### 2.8.4 การสร้าง OLAP และการรวบรวมให้เข้ากับคลังข้อมูล

เราสามารถใช้งาน Analysis Service ให้มีประสิทธิภาพสูงสุด เพราะ Analysis Service ง่ายต่อการสร้างแอปพลิเคชันสำหรับงานวิเคราะห์ที่มีความเปลี่ยนแปลงและซับซ้อนอย่างมาก สามารถยกตัวอย่างได้คือ

1. OLAP Cube สามารถ เรียกค้นข้อมูลที่มีความซับซ้อนเมื่อเทียบกับฐานข้อมูลเชิงสัมพันธ์ทั่วไปวิธีหนึ่งที่ทำให้ OLAP ประสบความสำเร็จคือ ใช้การสรุปผลที่คำนวณไว้ล่วงหน้า ด้วยการเตรียมผลสรุปจากการเชื่อมโยงข้อมูลหลายมิติและการวัดผลไว้ล่วงหน้า จึงทำให้ได้ผลลัพธ์ที่รวดเร็วและถูกต้อง โดยไม่ต้องคำนวณข้อมูลจากฐานข้อมูลใหม่
2. การเรียกค้นใน OLAP Cube ทำให้ผู้ใช้งานทางธุรกิจสามารถใช้ข้อมูลได้มากขึ้น โดยเตรียมโครงสร้างข้อมูลที่เกี่ยวข้องง่ายควบคู่กับการเรียกค้นแบบโต้ตอบ
3. ฐานข้อมูลทั่วไปจัดเก็บข้อมูลอย่างเฉียด แต่ OLAP Cube สามารถ มีได้หลายมิติ ซึ่งเข้ากันกับวิธีคิดทางธุรกิจ คือพิจารณาในหลายๆมุมมอง เช่น ลูกค้า เวลา สถานที่ เป็นต้น

ภาษาในการจัดการของ OLAP นั้นมีดังนี้

- Structure Query Language (SQL) เป็นภาษาที่ใช้ในการค้นฐานข้อมูลเชิงสัมพันธ์เหมาะสำหรับโครงสร้างข้อมูลที่เป็นตารางที่ผ่านการ Normalize มาแล้ว แต่ถ้านำมาใช้ในการวิเคราะห์ทางธุรกิจแล้วเป็นการยากที่จะค้นหาข้อมูลโดยใช้ SQL

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- Analysis Service ใช้ในการเรียกข้อมูลในหลายมิติ และ MDX เป็นภาษาที่ดีที่สุดสำหรับการทำงานเป็นแบบลำดับชั้นจึงง่ายต่อการค้นหาข้อมูลทางธุรกิจ

### 2.8.5 การสร้างและการจัดการ Cube

การที่จะให้ OLAP ประสบความสำเร็จสูงสุดในเรื่องของประสิทธิภาพการเรียกค้นที่ซับซ้อนก็คือ การออกแบบกลุ่มข้อมูลไว้ล่วงหน้าจำนวนหนึ่งซึ่งใช้ในการปรับแต่งประสิทธิภาพในการเรียกค้นจำเป็นต้องจัดการกับข้อมูลเหล่านี้ก่อน การจัดข้อมูลนั้นที่แท้จริงเป็นการจัดเพิ่มข้อมูลลงไป Cube ซึ่งอาจไปถึงจุดที่ กลุ่มข้อมูลมีปริมาณมากกว่าข้อมูลเริ่มต้น ปัญหาที่เรียกว่า Data Explosion

Analysis Manager มีการจัดการข้อมูลที่ชาญฉลาด โดยทำให้กระบวนการจัดการระบบ OLAP ง่ายขึ้น ลดการจัดเก็บกลุ่มข้อมูลมีดังนี้

- ในการเลือกเฉพาะกลุ่มที่มีความสำคัญมากที่สุด ให้ทำการจัดทำข้อมูลก่อนล่วงหน้า
- ทำการติดตามเรียกค้น ซึ่งทำให้ผู้ดูแลฐานข้อมูลสามารถปรับแต่งจัดการข้อมูลได้ตรงตามการเรียกค้นที่ขึ้นกับ Cube อย่างแท้จริง
- สามารถเปรียบเทียบผลดีและผลเสียของปริมาณข้อมูลในเนื้อที่จัดเก็บกลุ่มข้อมูลกับประสิทธิภาพที่ได้รับ และยอมให้ผู้ดูแลฐานข้อมูลได้ปรับแต่งให้สมดุล

### 2.8.6 การสร้างระบบงานที่รองรับ Business Intelligence

ในการวิเคราะห์ข้อมูลทางธุรกิจ ผู้ใช้จะไม่รู้ว่าตัวเองต้องการอะไรจากแอปพลิเคชัน จนกระทั่งได้ผลการวิเคราะห์แล้ว ดังนั้นแพลตฟอร์มสำหรับการวิเคราะห์ที่ดีต้องมีการยืดหยุ่นและปรับตัวได้ รวมทั้งสามารถขยายระบบรองรับจำนวนของผู้ใช้, ข้อมูล ที่จะเพิ่มมากขึ้น

นอกเหนือจากนี้ Analysis Service ยังมีความสามารถด้านการวิเคราะห์ที่ซับซ้อนและยังสามารถตอบสนองความต้องการใหม่ๆทางธุรกิจได้ดังนี้

#### 1. การวิเคราะห์ขั้นสูงของ Analysis Service

- การคำนวณ สามารถเรียกใช้ได้มากกว่า 130 ฟังก์ชัน รวมทั้งสามารถใช้งานใน Excel และ Visual Basic ได้ หรือเขียนสูตรคำนวณขึ้นมาเอง
- โดย Virture Cube ทำหน้าที่คล้ายกับ View ในฐานข้อมูลเชิงสัมพันธ์ จะมีการเชื่อมโยงข้อมูลจากหลายๆ Cube ที่มีมิติรวมกัน
- สามารถกำหนดสูตรคำนวณลงในแต่ละเซลล์ได้

#### 2. การติดตั้งและวิเคราะห์ในเชิงธุรกิจ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Analysis Service มีคุณสมบัติที่เรียกว่า Actions ซึ่งสามารถเชื่อมโยงและวิเคราะห์เข้ากับการทำงานในธุรกิจจึงทำให้สามารถตอบสนองความต้องการทางธุรกิจได้

### 3. การรักษาความปลอดภัย

มีการรักษาความปลอดภัยได้หลายระดับที่ยืดหยุ่นต่อความต้องการ โดยแบ่งได้ดังนี้

- ระดับ Server ฐานข้อมูล Cube และการทำโมเดลการวิเคราะห์ข้อมูล โดยผู้บริการสามารถกำหนดสิทธิ์เข้าใช้ให้กับผู้ใช้ได้
- ระดับ มิติ สามารถกำหนดได้ว่า มิติใดสามารถเข้าใช้งานได้บ้าง
- ระดับ เซลล์ กำหนดสิทธิ์ละเอียดขึ้นกว่าเดิม โดยจำกัดผู้ใช้เข้าได้เพียงแค่เซลล์ที่ระบุไว้ในระดับต่างๆ

### 4. การจัดการกับงานที่มีจำนวนข้อมูลมากและซับซ้อน

SQL Server และ Analysis Services มีการให้บริการสำหรับผู้ดูแลระบบหรือสำหรับจัดการฐานข้อมูล

- สามารถทำการบีบอัดเก็บข้อมูลให้เล็กลงที่ใช้เก็บ Cube ของ Analysis Services ลงเหลือเพียง 25-40 %
- สามารถสร้าง logical Cube เพื่อทำเป็นบทสรุป การประมวลผลต่างๆ หรือจัดเก็บส่วนของมูล แยกออกจาก Cube
- สามารถทำการจัดเก็บแบบ ROLAP มีประโยชน์สำหรับมิติขนาดใหญ่เนื่องจากแหล่งข้อมูล กลุ่มข้อมูล และ มิติ ถูกเก็บอยู่ในฐานข้อมูลเชิงสัมพันธ์
- สามารถจัดเก็บแบบ HOLAP มีประโยชน์สำหรับข้อมูลจำนวนมาก ซึ่งมีรายละเอียดถูกเรียกค้นไม่บ่อยนัก
- สามารถปรับค่าตามการใช้งาน และให้ผู้ดูแลระบบ ทำการสร้างข้อมูลชุดเล็ก ให้ตรงตามความต้องการการเรียกค้นของผู้ใช้

สรุป SQL Server2005 และ Analysis Services เป็นเครื่องมือที่เหมาะสมกับการที่จะนำมาใช้ในการจัดการคลังข้อมูล และสามารถใช้ประโยชน์ในการวิเคราะห์ข้อมูลที่ซับซ้อนได้อย่างมีประสิทธิภาพ ในขณะที่รูปแบบการใช้งานแต่ละอย่างสามารถยืดหยุ่นต่อผู้ใช้งานข้อมูล รวมทั้งสามารถเลือกแอปพลิเคชันที่จะนำข้อมูลจาก Analysis Services มาทำการวิเคราะห์ได้ ประโยชน์ด้านการคำนวณว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 2.8.7 ภาษา MDX - Multidimensional Expressions Language

MDX เป็นคำสั่ง Query Language ซึ่งเป็นส่วนหนึ่งใน MS SQL 2000 Analysis Services โดยมีโครงสร้างของคำสั่งและหน้าที่การทำงานคล้ายกับภาษา SQL แต่มีส่วนที่ต่างกันคือ MDX สามารถดึงข้อมูลจาก Multidimensional Database หรือ OLAP cube ได้ทันที เพราะมีโครงสร้างภาษาที่ออกแบบทำหน้าที่เฉพาะอยู่แล้ว และ ข้อมูลที่อาจยังคงอยู่ในสภาพที่เป็นลักษณะหลายมิติก็ได้ แต่ข้อมูลที่ได้อาจจากการเรียกในคำสั่ง SQL จะแสดงในรูปสองมิติคือ Column และ Row เท่านั้น ดังนั้น MDX จึงมีการระบุ Dimension ให้กับข้อมูลที่รับมาด้วย

รูปแบบคำสั่ง MDX มีโครงสร้างคล้ายกับ SQL โดยมีลักษณะดังนี้

```
SELECT [ < axis_specification > [, axis_specification > ..]]
```

```
FROM [ <cube_specification > ]
```

```
[WHERE[< slicer_specification >]]
```

โดยที่

Axis\_Specification เป็นการเลือก member หรือ เซตของ member ซึ่งตามที่กล่าวไปแล้ว ต้องระบุ Dimension ให้กับ Member ที่เลือกด้วย ดังนั้น รูปแบบของ Axis\_Specification จะต้องกำหนดด้วยว่า on row หรือ on column ไว้ด้วยเสมอ ดังนั้น axis\_specification จะมีรูปแบบคือ <set> on < axis\_name > ซึ่งเซตก็คือ เซตของ member และ axis\_name สามารถระบุชื่อ ได้แก่ columns , row , page , section , chapters หรือระบุเป็น Index เช่น axis(0) , axis(1) , axis(2) ไปเรื่อยๆก็ได้ ซึ่ง column คือ axis(0) และ row คือ axis(1) และ page คือ axis(2) นั่นเอง ใน MDX สามารถมีได้ 64 axis

Cube\_Specification คือ ชื่อ cube ที่ต้องการเหมือนระบุชื่อ table

Slicer\_Specification เป็นการระบุอีกชื่อหรือเงื่อนไขที่ มิติหรืออีกด้านหนึ่งซึ่งไม่ซ้ำกับ axis\_specification , cube\_specification หรือต้องเป็น Dimension ที่ต่างกัน

คำสั่ง MDX มีความสำคัญในการพัฒนาโปรแกรม สำหรับการวิเคราะห์การขาย ปูนซีเมนต์อย่างมากสำหรับโครงการนี้ เนื่องจาก เนื้อหาในโปรแกรมจะมีการนำสิ่งต่างๆ ที่ผู้ใช้ต้องการมาสร้างเป็นคำสั่ง MDX เพื่อส่งไปขอข้อมูลจาก OLAP Server แล้วนำข้อมูลมาเสนอในโปรแกรมต่อไป

## 2.8.8 การทำ Cube Schema

Cube Schema คือ โครงสร้างของ Cube โดยก่อนที่จะทำการสร้าง Cube จะต้องมีการกำหนดโครงสร้างเหล่านี้ขึ้นมาก่อน แล้วจึงทำการสร้าง (Process Cube) ขึ้นมา ซึ่งโครงสร้างเหล่านี้ประกอบด้วย

- Dimension คือ โครงสร้างหลักของ Cube ถ้ามองอย่างง่าย ถ้า Cube เป็นสี่เหลี่ยมลูก

บาศก์ Dimension ก็คือด้านทั้งสามของ Cube นั่นเอง แต่ในความเป็นจริง Dimension อาจมีได้มากกว่า 3 Dimension ก็ได้

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- Hierarchy คือความสัมพันธ์แบบ Parent – Child ภายใน Dimension เดียวกัน เนื่องจากในแต่ละ Dimension นั้น อาจมีการจัดหมวดหมู่ของข้อมูล ตัวอย่างเช่น Geographic Dimension มีการแบ่งเป็นภาคและจังหวัด ซึ่งจังหวัดหนึ่งๆจะต้องขึ้นอยู่กับภาคใดภาคหนึ่งเสมอ ในกรณีนี้ภาคจะเป็น Parent และ จังหวัดที่อยู่ภายในภาคก็จะเป็น Child ของภาคนั้นๆ เป็นต้น

- Level เป็นการแบ่งลำดับชั้น ภายใน Dimension จากตัวอย่างเดิม คือ ภาคและจังหวัด จะมี Level ที่เป็นภาค และ Level ที่เป็นจังหวัด เป็นต้น

- Member คือข้อมูลที่อยู่ภายในแต่ละ Dimension ซึ่งก็คือสมาชิกทั้งหมดภายใน Dimension นั้นๆนั่นเอง

ใน ADO-MD จะมี Object กลุ่มที่ทำงานกับ Cube Schema โดยเฉพาะ ซึ่งก็คือ Catalog , CubeDef , Dimension , Hierarchy และ Level โดยทำงานสอดคล้องกับ Cube Schema โดยตรง

### 2.8.9 การทำ Cube Data

Cube Data หรือข้อมูลภายใน Cube บางที่เรียกว่า Measurement ซึ่งใน ADO-MD จะสามารถส่งคำสั่งไปขอข้อมูลใน Cube โดยใช้คำสั่ง MDX และผลของคำสั่ง MDX นี้จะกลับมามีอยู่ในรูปแบบ Table ธรรมดาที่มีสองมิติ หรือให้อยู่ในรูปแบบที่มากกว่าสองมิติก็ได้ โดยเรียกกลุ่มข้อมูลที่ได้กลับมาจากกระบวนการนี้ว่า CellSet ซึ่งใน ADO-MD จะมี Object ที่ทำหน้าที่ในการจัดการผลของ CellSet ประกอบด้วย

- Axis ก็คือ แกนของข้อมูลที่ได้กลับมา ถ้าเปรียบเทียบกับฐานข้อมูลแบบธรรมดา ผลของการ Query จะได้เป็นสองแกน คือ แนวตั้งและแนวนอน แต่ใน CellSet นี้สามารถมีได้มากกว่าสองแกน ซึ่ง axis นี้มักสอดคล้องกับ Dimension ที่อยู่ใน Cube ซึ่งส่วนมากมักจะ Query โดยระบุให้ axis เป็น Dimension หนึ่งใน Cube แต่ก็มีบางกรณี ที่สามารถระบุให้ axis ประกอบด้วย Dimension มากกว่าหนึ่ง Dimension ก็ได้ ซึ่งจะเรียก axis นั้นว่าเป็น Nested Dimension

- Position คือตำแหน่งต่างๆ หรือสมาชิกใน axis เช่น ถ้าระบุให้ Query ดึงข้อมูลหนึ่ง Dimension ลงมาใน axis ผลที่ได้ก็คือ Position ก็คือ Member ของ Dimension นั้นนั่นเอง แต่หาก axis นั้นเป็น Nested Dimension จะได้ว่า Position ก็คือคู่อันดับของ Member ของทั้งสอง Dimension นั้น

- Cell คือที่เก็บข้อมูลจริงๆ หรือมองอีกนัยหนึ่ง ก็คือ ข้อมูลที่เก็บอยู่ที่เกิดจากการระบุ Position จากทุกๆ axis โดยที่ Cell หนึ่ง Cell จะเก็บข้อมูลตัวเลขได้เพียงหนึ่งชุดเท่านั้น

ในบางกรณีคำสั่ง MDX ที่ส่งไปขอข้อมูลนั้นอาจระบุให้ Measurement ใน Cube กลับมาเป็น axis ก็ได้ โดยลักษณะนี้จะมอง Measurement เป็นเหมือนกับ Dimension และ Position ของ axis ก็คือค่าต่างๆที่เป็น Measure นั่นเอง

ADO-MD Object Library มีความสำคัญต่อการพัฒนาโปรแกรมสำหรับวิเคราะห์ข้อมูลการวิเคราะห์การขายปูนซีเมนต์สำหรับโครงการนี้ เนื่องจากโปรแกรมดังกล่าวจะต้องมีการเรียกใช้ ADO-MD เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Object เพื่อติดต่อกับ OLAP Cube ในการทำงานกับโครงสร้างของ Cube หรือขอข้อมูลที่ถูกรวบรวมอยู่ภายใน Cube ดังกล่าว



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทที่ 3

# การออกแบบและการจัดเก็บข้อมูลใน Data Warehouse

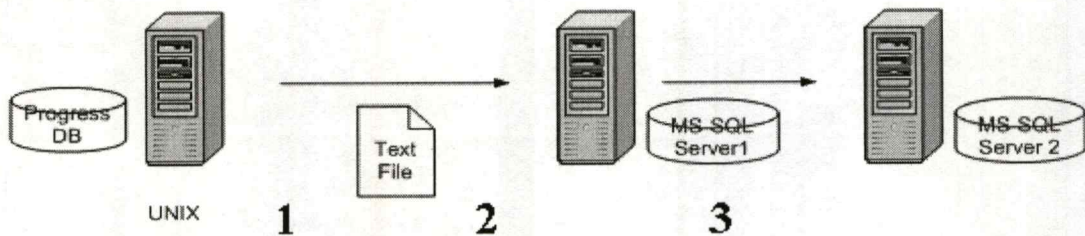
ในการพัฒนาลคลังข้อมูล สำหรับโปรแกรมวิเคราะห์การขายสินค้า ในธุรกิจขายปูนซีเมนต์นี้ มีวัตถุประสงค์ต้องการที่จะทำการวิเคราะห์ข้อมูลการขายปูนซีเมนต์อย่างมีประสิทธิภาพ จึงต้องมีการศึกษากระบวนการที่เกี่ยวข้องกับข้อมูลการขายปูนซีเมนต์ที่มีอยู่ในระบบเดิม และออกแบบจัดเตรียมข้อมูลการขายแบบเดิมนี้เสียใหม่ โดยข้อมูลเดิมถูกจัดเก็บอยู่ในฐานข้อมูลหลักขององค์กร (Operational Database) ยกต่อการค้นหาและเรียกใช้ข้อมูลเมื่อต้องการวิเคราะห์ในเชิงธุรกิจ เพราะฉะนั้นจึงต้องมีการจัดการเลือกสรรข้อมูลที่ต้องการนำมาวิเคราะห์และจัดเก็บไว้ในฐานข้อมูลของ Data Warehouse โดยเน้นความสะดวก ความถูกต้องในการนำมาใช้เป็นหลักด้วย ข้อมูลที่ถูกจัดเตรียมควรอยู่ในรูปแบบที่เหมาะสมพร้อมกับการใช้งาน สิ่งสำคัญที่ต้องคำนึงถึงคือความสะดวกในการใช้งานผ่านระบบ โครงการวิเคราะห์การขายที่พัฒนาขึ้นนี้ด้วย เพื่อที่จะนำข้อมูลดังกล่าวมาใช้ได้ง่ายขึ้น

### 3.1 ขั้นตอนการเลือกสรรข้อมูลเพื่อใช้ในการวิเคราะห์

เนื่องจากฐานข้อมูลหลักขององค์กรได้ถูกจัดเก็บไว้ในฐานข้อมูลโดยใช้ Progress เป็น RDBMS ซึ่งทำงานบนระบบปฏิบัติการ UNIX แบบเก่า โดยฐานข้อมูลนี้จะเก็บข้อมูลการขายสินค้าที่ขายได้ในทุกวัน ภายในฐานข้อมูลนี้ยังการจัดเก็บข้อมูลของ รายละเอียดของลูกค้า, สินค้า, ประเภทสินค้า, ประเภทการบรรจุของสินค้า, เขตการขาย และ อื่นๆอีกด้วย ดังนั้นการเลือกข้อมูลที่ใช้ในการวิเคราะห์การขายนี้เราได้คัดเลือกข้อมูลที่ใช้ในการวิเคราะห์ ได้แก่ ข้อมูลการขาย, รายละเอียดข้อมูลของลูกค้า, สินค้า, ประเภทสินค้า, ประเภทการบรรจุของสินค้า, เขตการขาย มาเก็บไว้ในฐานข้อมูล Data Warehouse โดยจะนำข้อมูลดังกล่าวมาใช้ในการวิเคราะห์ผ่านระบบโครงการที่พัฒนาขึ้น

### 3.2 ขั้นตอนการ Transformation ข้อมูล

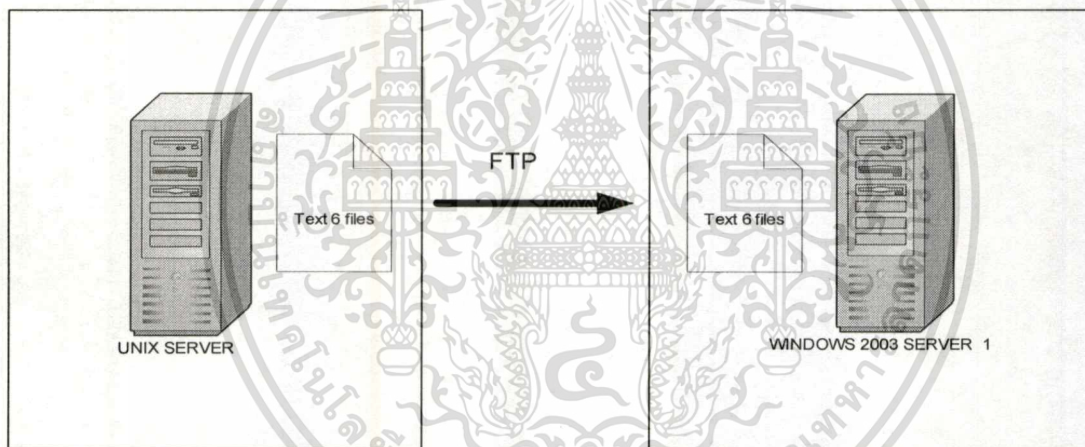
ในการออกแบบและการวางแผนที่จะนำข้อมูล หลักขององค์กรมาใช้งานในระบบ ฐานข้อมูล Data Warehouse นั้นเราได้ทำออกแบบ การ Transformation ข้อมูลเป็น 3 ขั้นตอน ดังรูปที่ 3.1 คือ



รูปที่ 3.1 การ Transformation ข้อมูล

### 3.2.1 กระบวนการนำข้อมูลออกจากฐานข้อมูลหลัก

การนำข้อมูลออกมานั้น โดยใช้วิธีการ Export ข้อมูลที่ต้องการออกมาในลักษณะของ Excel Files ทั้งหมด 6 files โดยมีรายละเอียดของข้อมูลตามที่ได้กล่าวมาแล้ว หลังจากนั้น ทำการส่ง Excel Files โดยใช้วิธีการ FTP มาเก็บในเครื่อง Windows Server ดังรูปที่ 3.2



รูปที่ 3.2 การนำข้อมูลออกจากฐานข้อมูลหลัก

### 3.2.2 กระบวนการนำข้อมูลเข้า Staging Area

เมื่อ Excel ทั้งหมด 6 files ได้ถูกส่งมาครบถ้วนแล้ว จากระบบฐานข้อมูลการขายสินค้าขององค์กร โดยมีข้อมูลที่เกี่ยวข้องดังนี้คือ ข้อมูลการขาย , รายละเอียดข้อมูลของลูกค้า , สินค้า , ประเภทสินค้า , ประเภทการบรรจุของสินค้า , เขตการขาย โดยข้อมูลทั้งหมดสามารถนำเข้าสู่ Staging Area ที่เป็นฐานข้อมูล Microsoft SQL Server 2005 โดยผ่านกระบวนการ ที่เรียกว่า ETL (Extract Transform Load) โดย Microsoft SQL Server 2005 มีเครื่องมือในการสร้าง ETL คือ SSIS (SQL Server Integration Service) ซึ่งการออกแบบ SSIS นั้นถูกรวมอยู่ใน SQL Server Business Intelligence Development Studio ซึ่งใช้ในงานส่วน Business Intelligence ของ Microsoft SQL Server 2005 สามารถสร้างข้อมูลในแต่ละตัว แยกออกจากกันเป็น Package และ SSIS นี้ยังสามารถกำหนดค่าให้แต่ละ Package ที่สร้างไว้ทำงานตามเวลาได้บ้างแล้วแต่กำหนด

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เนื่องจากกระบวนการเหล่านี้ทำหลังจากข้อมูลการซื้อขายในข้อมูลหลักปิดขายในแต่ละวันแล้ว เพื่อให้ได้ข้อมูลคงที่ซึ่งจะสามารถวิเคราะห์ได้อย่างมีประสิทธิภาพ ในที่นี้จะกล่าวเรื่องการ Transform ข้อมูลแต่ละ Package แบบคร่าวๆ จะไม่กล่าวถึงรายละเอียดในการเขียน SSIS แต่จะอธิบายขั้นตอนในการทำงานแต่ละ Package เท่านั้น

สาเหตุที่มีเครื่อง Server MS- SQL Server 2005 ถึง 2 เครื่อง เนื่องจากได้ทำการออกแบบเป็น Staging Area และจัดเก็บฐานข้อมูลในการเข้าใช้เช่น รหัส, ข้อมูลพนักงาน ฯลฯ ร่วมกันไว้ในที่เครื่องเดียวกัน ทั้งนี้ ในการจัดการการแก้ไขข้อมูลโดยการผ่านเว็บ จะทำการแก้ไขข้อมูลได้ก่อนก่อนการสร้าง Cube ส่วนอีกเครื่องหนึ่งจำทำการสร้าง เป็น Data warehouse อย่างเดียว

### 3.2.2.1 การทำการ Transform ของข้อมูลลูกค้า, สินค้า, ประเภทสินค้า, ประเภทการบรรจุของสินค้า และ เขตการขาย

กระบวนการ Transform ของข้อมูลต่างๆที่ได้กล่าวมาแล้ว เราสามารถรวม SSIS เป็น Package ได้เลยเพราะกระบวนการ Transform ข้อมูล ไม่ต่างกัน และข้อมูลเป็นข้อมูลที่คงที่มีการเคลื่อนไหวน้อยมาก ดังนั้นในการ สร้าง SSIS Package จึงมีขั้นตอนการดำเนินการดังนี้

1. ทำการ Backup ข้อมูลในแต่ละ Tables โดย export เป็น Textfile.
2. ลบข้อมูลที่อยู่ในแต่ละ Tables ทั้งหมด โดยใช้คำสั่ง SQL ดังนี้

```
TRUNCATE TABLE [FrontSale].[dbo].[Customers]
```

```
TRUNCATE TABLE [FrontSale].[dbo].[ProdGroup]
```

```
TRUNCATE TABLE [FrontSale].[dbo].[ProdLotType]
```

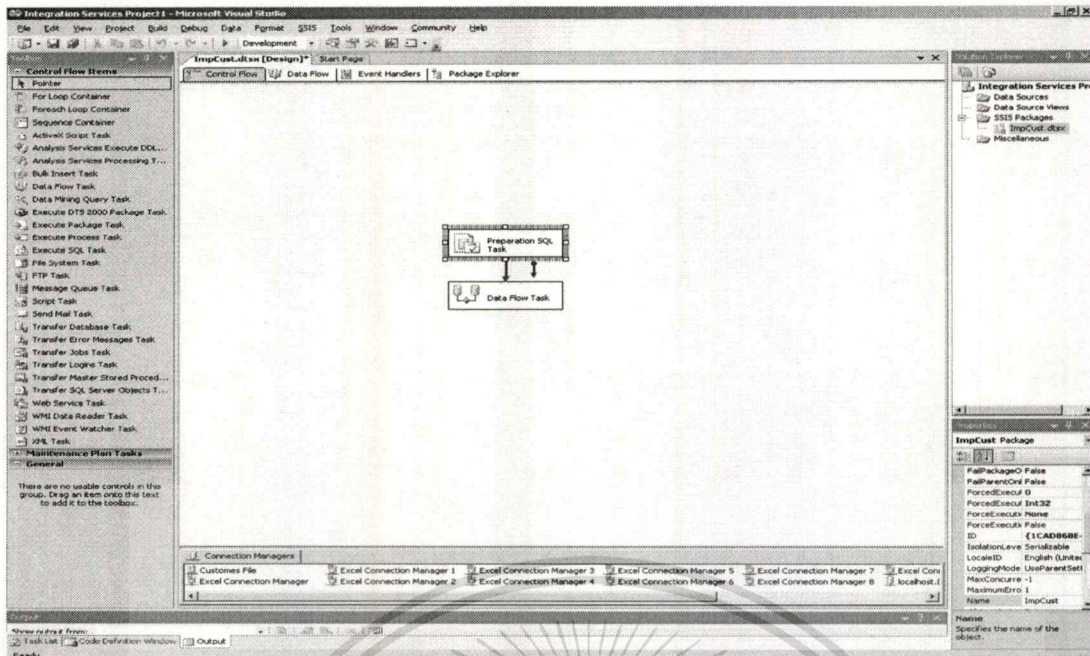
```
TRUNCATE TABLE [FrontSale].[dbo].[Products]
```

```
TRUNCATE TABLE [FrontSale].[dbo].[Region]
```

```
TRUNCATE TABLE [FrontSale].[dbo].[Sales]
```

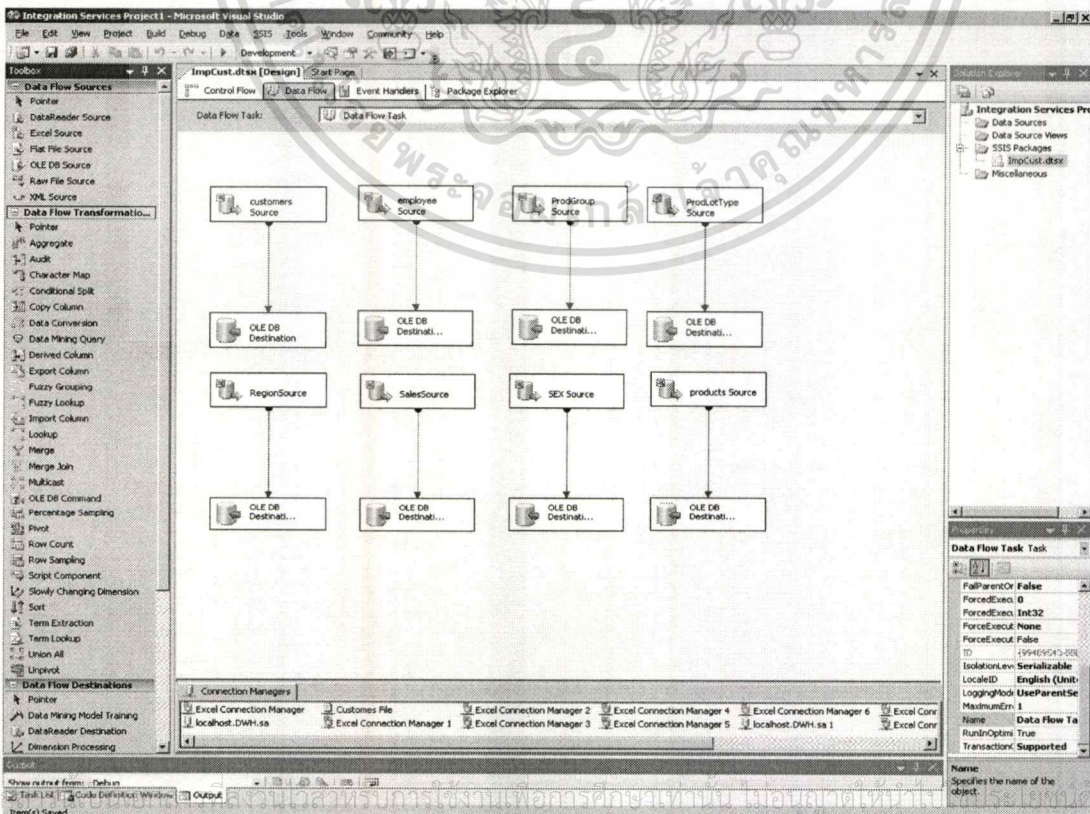
3. ทำการ Transform ข้อมูลจาก Excels File ลงสู่ Tables ต่างๆให้ตรงตามชื่อไฟล์ ดังรูปที่ 3.4

4. นำข้อมูลจากการทำ Backup File มีรายละเอียดที่ไม่มีในฐานข้อมูลที่ไม่ใช่ข้อมูลปัจจุบันเพิ่มเข้าไปใน ฐานข้อมูล



รูปที่ 3.3 รูปแบบ SSIS Package สำหรับการ Transform ข้อมูล กระจายเข้าสู่ Data Warehouse Server

เมื่อทำการคลิกเข้าไปดูในโครงสร้างภายในของ Data Flow Task จะเป็นกระบวนการ Transform ข้อมูลจาก Excel File ลงสู่ ฐานข้อมูล Data Warehouse โดยข้อมูลที่ต้องการ จะถูกสร้างลงในตารางที่ออกแบบไว้ในขั้นต้น จากรูปที่ 3.4



ไม่ว่าการรูปที่ 3.4 รายละเอียดภายใน SSIS Package สำหรับการ Transform ข้อมูล ทุกครั้งที่มีการนำไปใช้

### 3.2.2.2 การทำการ Transform ของข้อมูลการขายใน OLAP Server

ในกระบวนการนี้เป็นขั้นตอนการ Transform ข้อมูลจากเครื่อง Data Warehouse Server เข้าสู่เครื่อง OLAP Server มีกระบวนการคล้ายกับการนำข้อมูลจาก Excel Files เข้าสู่ Data Warehouse Server ดังนั้นรูปแบบสร้าง SSIS Package จึงเป็นรูปแบบของ การ Transform ระหว่าง Server ผู้ Server โดยมีขั้นตอนดังนี้

1. ลบข้อมูลที่อยู่ในแต่ละ Dimension Tables ที่เกี่ยวข้องบนเครื่อง OLAP Server โดยใช้คำสั่ง SQL ดังนี้

```
TRUNCATE TABLE [DWH].[dbo].[Customers]
```

```
TRUNCATE TABLE [DWH].[dbo].[ProdGroup]
```

```
TRUNCATE TABLE [DWH].[dbo].[ProdLotType]
```

```
TRUNCATE TABLE [DWH].[dbo].[Products]
```

```
TRUNCATE TABLE [DWH].[dbo].[Region]
```

```
TRUNCATE TABLE [DWH].[dbo].[Sales]
```

2. เลือกข้อมูลที่ได้จัดเตรียมไว้แล้วใน Tables ต่างๆบนเครื่อง Data Warehouse Server และทำการ Transform ข้อมูลที่ต้องการสู่ OLAP Server โดยใช้คำสั่ง SQL ดังนี้

```
Use DWH
```

```
Select * From [FrontSale].[dbo].[Customers]
```

```
Select * From [FrontSale].[dbo].[ProdGroup]
```

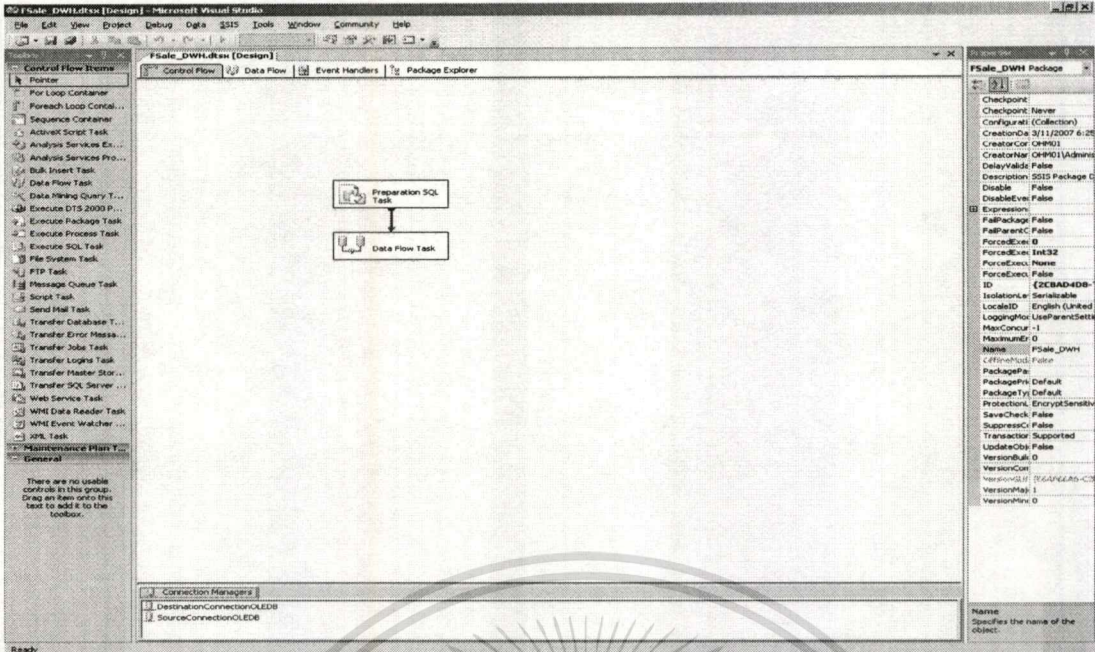
```
Select * From [FrontSale].[dbo].[ProdLotType]
```

```
Select * From [FrontSale].[dbo].[Products]
```

```
Select * From [FrontSale].[dbo].[Region]
```

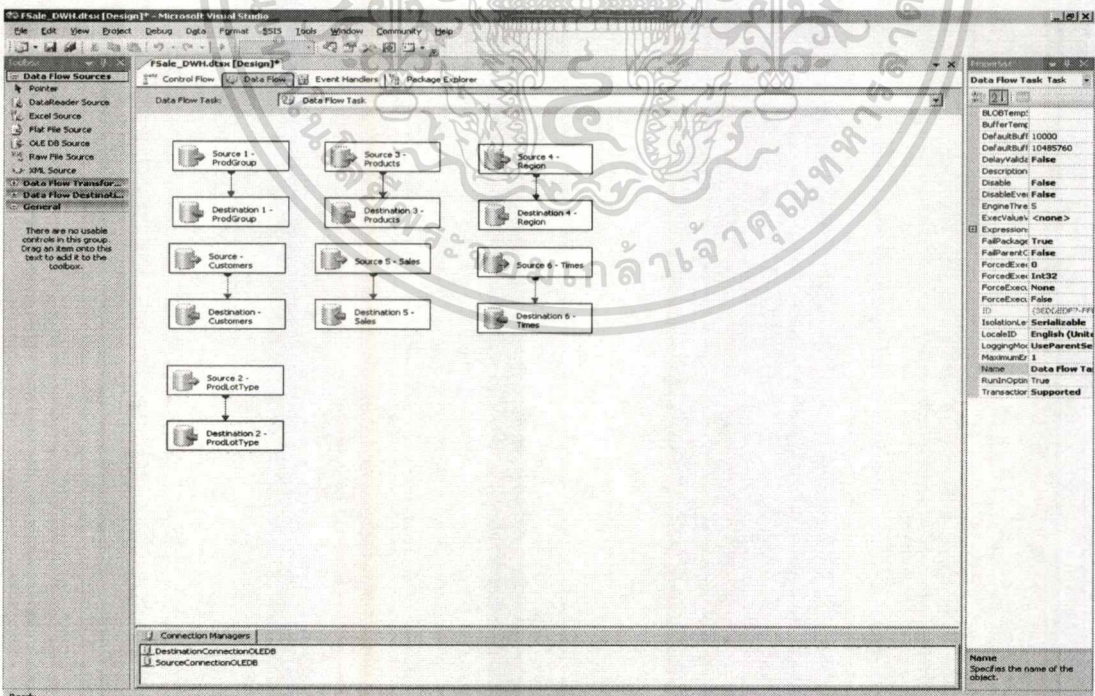
```
Select * From [FrontSale].[dbo].[Sales]
```

```
Select * From [FrontSale].[dbo].[ProdGroup]
```



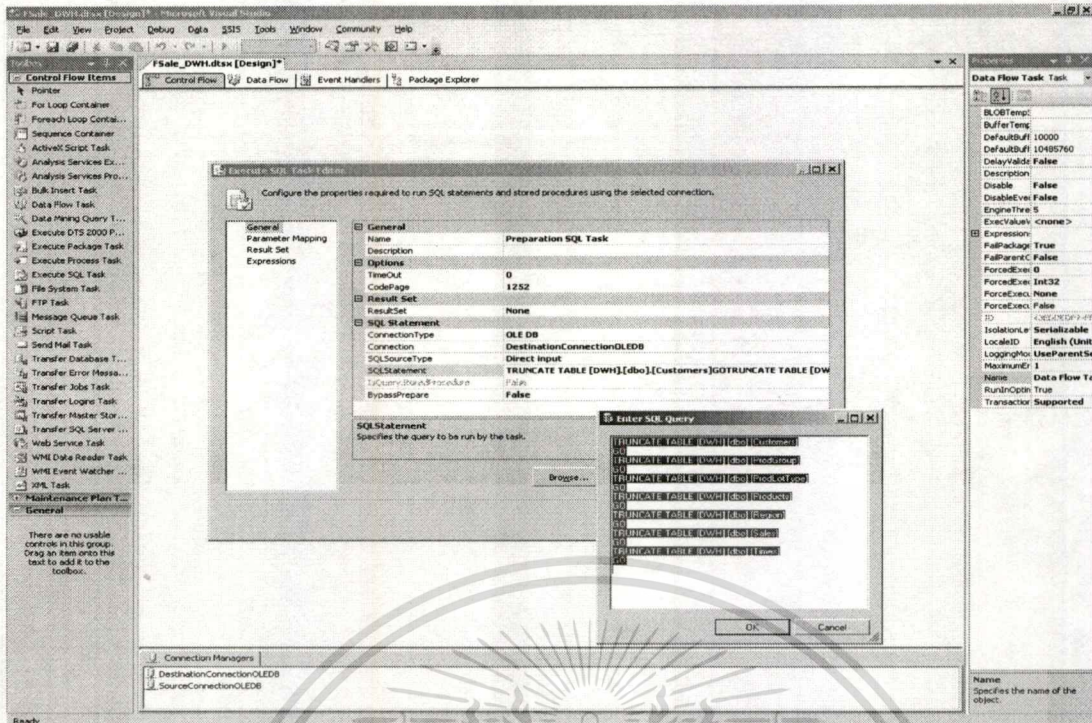
รูปที่ 3.5 รูปแบบ SSIS Package สำหรับการ Transform ข้อมูล การขายเข้าสู่ OLAP Server

เมื่อทำการคลิกเข้าไปดูในโครงสร้างภายในของ Data Flow Task จะเป็นกระบวนการ Transform ข้อมูลจาก Data Warehouse Server ลงสู่ฐานข้อมูล OLAP Server โดยข้อมูลที่ต้องการจะถูกสร้างลงในตารางที่ออกแบบไว้ในขั้นตอน จากรูปที่ 3.6



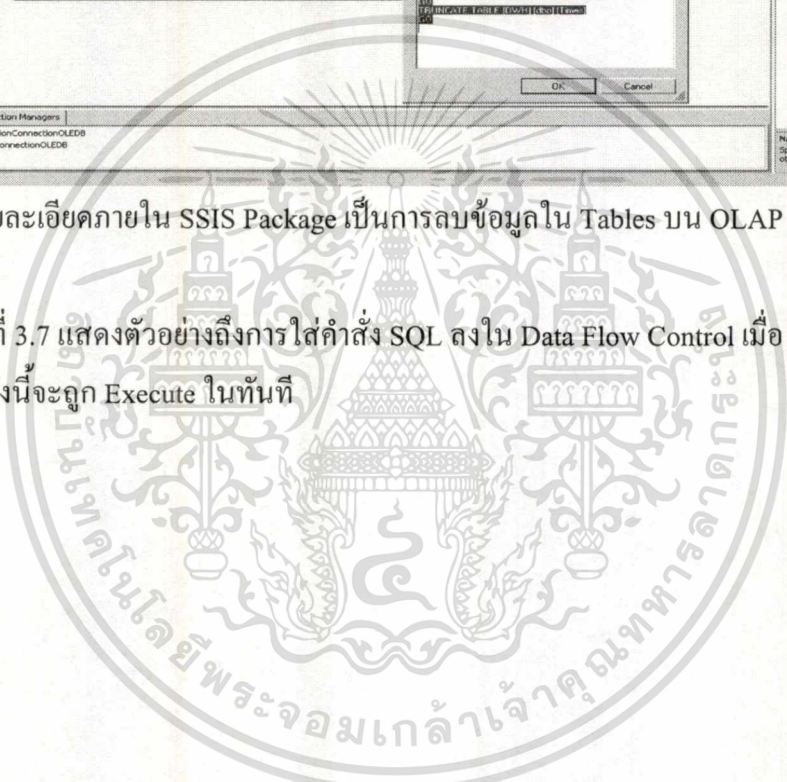
รูปที่ 3.6 รายละเอียดภายใน SSIS Package สำหรับการ Transform ข้อมูลใน Data warehouse OLAP Server

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



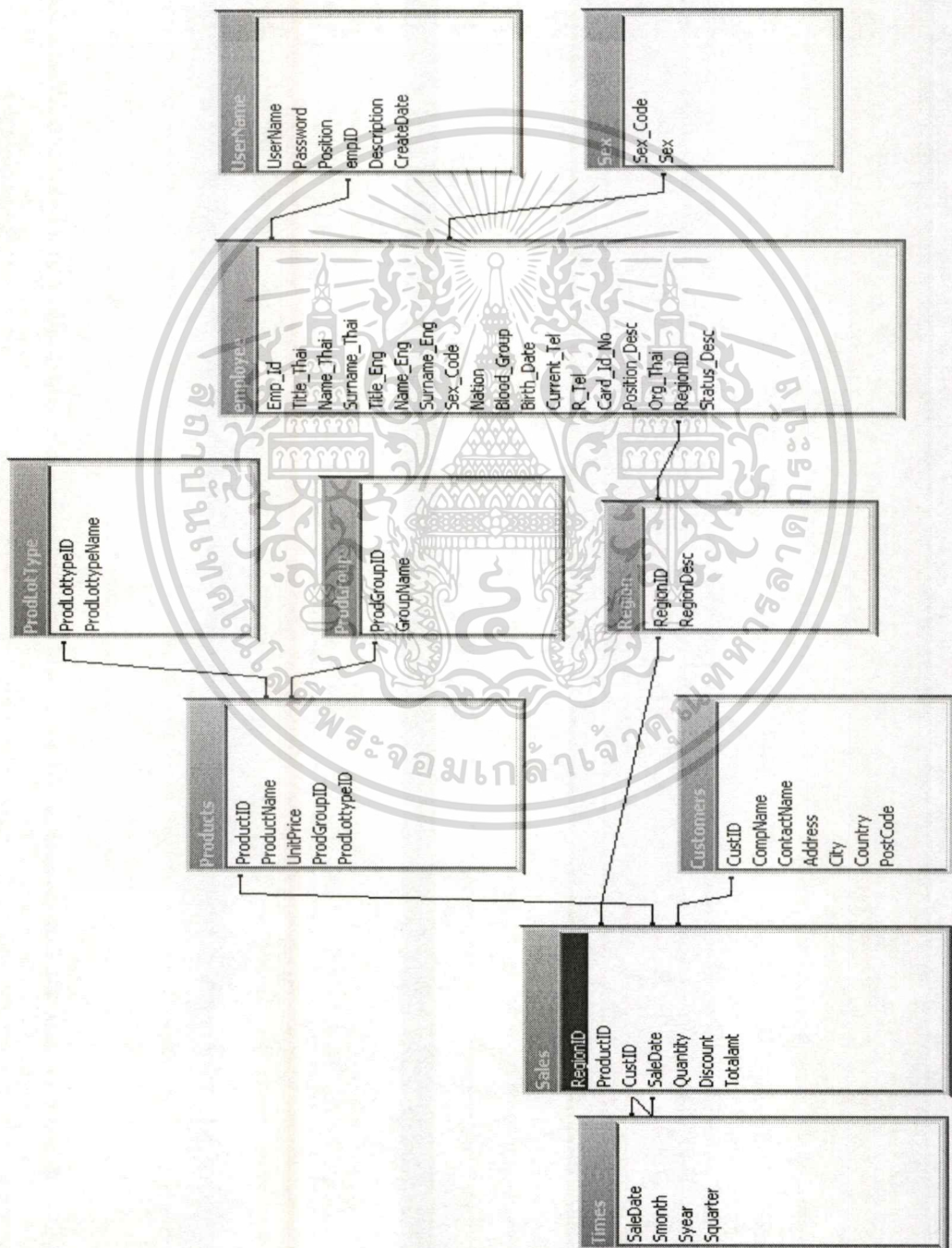
รูปที่ 3.7 รายละเอียดภายใน SSIS Package เป็นการลบข้อมูลใน Tables บน OLAP Server

จากรูปที่ 3.7 แสดงตัวอย่างถึงการใส่คำสั่ง SQL ลงใน Data Flow Control เมื่อ Package ทำการ Run คำสั่งนี้จะถูก Execute ในทันที



### 3.3 ขั้นตอนการออกแบบความสัมพันธ์ของข้อมูลใน Relation Database (Staging Area)

ในการออกแบบข้อมูลในที่นี้เราใช้ฐานข้อมูล Microsoft SQL Server 2005 เป็นตัวจัดการ โดยมีตารางทั้งหมด 10 ตาราง มีความสัมพันธ์ ดังรูปที่ 3.8



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้ภายในเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
 รูปที่ 3.8 E-R Diagrams  
 ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 3.4 พจนานุกรมข้อมูล (Data Dictionary)

พจนานุกรมข้อมูลเป็นการแสดงรายละเอียดฟิลด์ต่างๆที่เกี่ยวข้องกับ โปรแกรมวิเคราะห์ การขายสินค้าในธุรกิจขายปูนซีเมนต์ว่าประกอบด้วยตารางอะไรบ้าง ความยาวของข้อมูลมีขนาดเท่าไร รายละเอียดของพจนานุกรมข้อมูลแสดงดังนี้

ตารางที่ 3.1 Customers table

Field No.	Field Name	Type(length)	Pk	Fk	Description
1	CustID	varchar(10)	Y		รหัสลูกค้า
2	CompName	varchar(50)			ชื่อลูกค้า
3	ContactName	varchar(50)			ชื่อใช้ในการติดต่อ
4	Address	varchar(120)			ที่อยู่
5	City	varchar(30)			เมือง
6	Country	varchar(20)			ประเทศ
7	PostCode	varchar(5)			รหัสไปรษณีย์

ตารางที่ 3.2 Product table

Field No.	Field Name	Type(length)	Pk	Fk	Description
1	ProductID	varchar(10)	Y		รหัสผลิตภัณฑ์
2	ProductName	varchar(30)			ชื่อผลิตภัณฑ์
3	UnitPrice real	real			ราคา
4	ProdGroupID	char(2)		Y	กลุ่มผลิตภัณฑ์
5	ProdLotypeID	char(2)		Y	ผลิตภัณฑ์

ตารางที่ 3.3 ProdGroup table

Field No.	Field Name	Type(length)	Pk	Fk	Description
1	ProdGroupID	char(2)	Y		รหัสกลุ่มผลิตภัณฑ์
2	GroupName	char(25)			ชื่อกลุ่มผลิตภัณฑ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.4 ProdLotType table

Field No.	Field Name	Type(length)	Pk	Fk	Description
1	ProdLotTypeID	char(2)	Y		รหัสการบรรจุผลิตภัณฑ์
2	ProdLottypeName	char(25)			ชื่อการบรรจุผลิตภัณฑ์

ตารางที่ 3.5 Employee table

Field No.	Field Name	Type(length)	Pk	Fk	Description
1	Emp_Id	varchar(10)	Y		รหัสพนักงาน
2	Title_Thai	varchar(10)			ชื่อนำหน้า
3	Name_Thai	varchar(50)			ชื่อ(ไทย)
4	Surname_Thai	varchar(50)			นามสกุล
5	Title_Eng	varchar(10)			ชื่อนำหน้า
6	Name_Eng	varchar(50)			ชื่อ(อังกฤษ)
7	Surname_Eng	varchar(50)			นามสกุล
8	Sex_Code	varchar(2)		Y	รหัสเพศ
9	Nation	varchar(10)			สัญชาติ
10	Blood_Group	char(2)			กลุ่มเลือด
11	Birth_Date	varchar(15)			วันเกิด
12	Current_Tel	varchar(12)			เบอร์โทรศัพท์ที่ติดต่อได้
13	R_Tel	varchar(12)			เบอร์โทรศัพท์ที่บ้าน
14	Card_Id_No	varchar(13)			เลขที่บัตรประชาชน
15	Position_Desc	varchar(50)			ตำแหน่ง
16	Org_Thai	varchar(50)			สังกัด
17	RegionID	char(3)		Y	เขตในการรับผิดชอบ
18	Status_Desc	varchar(50)			สถานะ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.6 Region table

Field No.	Field Name	Type(length)	Pk	Fk	Description
1	RegionID	varchar(5)	Y		รหัสเขตการขาย
2	RegionDesc	varchar(50)			ชื่อเขตการขาย

ตารางที่ 3.7 Username table

Field No.	Field Name	Type(length)	Pk	Fk	Description
1	UserName	char(10)	Y		ชื่อ
2	Password	char(8)			รหัส
3	Position	char(25)			ระดับการเข้า
4	empID	char(10)			รหัสพนักงาน
5	Description	char(25)			รายละเอียดอื่น
6	CreateDate	datetime			วันที่เพิ่มชื่อ

ตารางที่ 3.8 Sale table

Field No.	Field Name	Type(length)	Pk	Fk	Description
1	RegionID	varchar(5)	Y		รหัสเขตการขาย
2	ProductID	varchar(10)	Y		รหัสผลิตภัณฑ์
3	CustID	varchar(10)	Y		รหัสผลิตภัณฑ์
4	SaleDate	datetime	Y		วันที่ขาย
5	Quantity	real			จำนวน
6	Discount	real			ส่วนลด
7	Totalamt	real			ผลรวมทั้งหมด

ตารางที่ 3.9 Sex table

Field No.	Field Name	Type(length)	Pk	Fk	Description
1	Sex_Code	char(1)	Y		รหัสเพศ
2	Sex	char(10)			เพศ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.10 Time table

Field No.	Field Name	Type(length)	Pk	Fk	Description
1	SaleDate	datetime	Y		วันที่ขาย
2	Smonth	varchar(10)			เดือนที่ขาย
3	Syear	varchar(10)			ปีที่ขาย
4	Squarter	varchar(2)			ไตรมาส

### 3.5 ทำการออกแบบตารางข้อมูลใน Data Warehouse

ในการออกแบบตารางของฐานข้อมูลการขาย สำหรับฐานข้อมูลในการทำ Data Warehouse จะต้องคำนึงถึงการนำไปใช้ใน OLAP เป็นสิ่งสำคัญอย่างยิ่งเพื่อให้ได้ผลลัพธ์ที่สอดคล้องกับความต้องการของผู้ใช้งาน และต้องตรงกับข้อมูลที่มีอยู่จริงในฐานข้อมูลหลัก สิ่งที่ต้องการวิเคราะห์การขายปูนซีเมนต์แบ่งเป็นการวิเคราะห์ 7 ส่วน คือ

#### 3.5.1. การวิเคราะห์ข้อมูลทางด้านลูกค้า

การวิเคราะห์ข้อมูลทางด้านลูกค้าจากการขายปูนซีเมนต์ ได้แก่ ลูกค้า ได้แก่ ลูกค้าที่เป็นบริษัท , ร้านค้าวัสดุ, ผู้รับเหมาก่อสร้าง รวมถึง การส่งออกนอกประเทศเป็นต้น เนื่องจากองค์กรมีลูกค้าหลายประเภทจำนวนมากจึงแบ่งลูกค้าตามปริมาณการซื้อปูนซีเมนต์ คือถ้าลูกค้ารายใดซื้อสินค้าปริมาณมากวัดจากน้ำหนักของสินค้าที่ใช้เป็นต้น ก็จะได้ส่วนลดสินค้ามากด้วย

จากการวิเคราะห์ข้อมูลทางด้านลูกค้าดังกล่าว เราสามารถออกแบบตารางสำหรับเก็บข้อมูลลูกค้าได้ดังตารางที่ 3.11

ตารางที่ 3.11 Customer Dimension Table

Field Name	Type	Length	Decimal	Description
CusID	Int	7	0	รหัสลูกค้า
CompName	Varchar	30	0	ชื่อบริษัท
ContactName	Varchar	30	0	ชื่อที่ใช้ติดต่อ
Address	Varchar	45	0	ที่อยู่
City	Varchar	15	0	เมือง
Country	Varchar	15	0	ประเทศ
PostCode	Varchar	15	0	ส่งจากประเทศ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 3.5.2 การวิเคราะห์ข้อมูลทางด้านสินค้า

การวิเคราะห์ข้อมูลทางด้านสินค้านั้น เนื่องจากองค์กรทำธุรกิจปูนซีเมนต์ จึงมีสินค้าที่เป็นปูนซีเมนต์เป็นหลักแต่ได้แบ่งเป็นหลายประเภทตามความต้องการ และการใช้งานของตลาด ในการแบ่งสินค้าจะแบ่งแบบกว้างตามบรรจุภัณฑ์ว่าเป็นแบบใด เช่นเป็นปูนผง หรือปูนถุง ซึ่งจะมีราคาที่แตกต่างกันออกไป สินค้าแต่ละอย่างจะถูกแยกประเภทในการใช้งานอีกกว่าใช้งานในลักษณะใด เหมาะสำหรับภูมิประเทศแบบใดอีก แต่เมื่อเราวิเคราะห์ข้อมูลจริงๆแล้ว สามารถออกแบบตารางสินค้าที่รวมกันได้เพราะ ProductID ได้บ่งบอกจำแนกประเภทของสินค้าตามรหัสสินค้าให้ จึงสามารถนำข้อมูลมาไว้ในตารางเดียวได้เลย ดังนั้นจึงออกแบบตาราง Product ที่เก็บในฐานข้อมูล Data Warehouse ได้ดังตารางที่ 3.12

ตารางที่ 3.12 Product Dimension Table

Field Name	Type	Length	Decimal	Description
ProductID	Int	6	0	รหัสสินค้า
ProductName	Varchar	30	0	ชื่อสินค้า
UniPrice	Decimal	8	0	ราคาต่อหน่วย
ProdGroupID	Int	10	0	รหัสกลุ่มสินค้า
GroupName	Text	10	0	ชื่อกลุ่มสินค้า

### 3.5.3. การวิเคราะห์ข้อมูลทางด้านเวลาขาย

ส่วนการวิเคราะห์เรื่องเวลาที่ทำการขายปูนซีเมนต์ สามารถออกแบบให้ดูยอดขายรวมทั้งปี , รายเดือน, รายไตรมาส หรือ สัปดาห์ ได้ แล้วแต่ความต้องการที่จะลงในเรื่อง ยอดขาย ดังนั้นจึงออกแบบตาราง Time ที่เก็บในฐานข้อมูล Data Warehouse ได้ดังตารางที่ 3.13

ตารางที่ 3.13 Time Dimension Table

Field Name	Type	Length	Decimal	Description
SaleDate	date/time	8	0	วันที่ขาย
Smonth	date/time	8	0	เดือนที่ขาย
Syear	date/time	4	0	ปีที่ขาย
Squarter	Char	2	0	ประจำทุกสามเดือน

### 3.5.4. การวิเคราะห์ข้อมูลทางด้านเขตการขาย

ในองค์กรมีการแบ่งความรับผิดชอบของพนักงานขาย ที่จะต้องรับผิดชอบเขตการขายแต่ละเขต เพื่อให้ความดูแลและเข้าถึงลูกค้าในท้องถิ่นนั้นมากที่สุด ดังนั้น ฝ่ายขายได้แบ่งเขตการขายในจังหวัดและเขตต่างๆ ตามภูมิภาคในประเทศไทย โดยรหัส แยกตามเขตต่างๆ ลงไปอีก เพราะฉะนั้นจึงเป็นเรื่องง่ายที่จะทำการแยกยอดขายตามเขตการขาย หรือตามภาคต่างๆ ได้เพื่อดูความแตกต่างของการขายสินค้า รวมทั้งสามารถรู้ยอดสรุป ของทีมขายต่อละทีมได้ ดังนั้นจึงออกแบบตาราง Region ที่เก็บในฐานข้อมูล Data Warehouse ได้ดังตารางที่ 3.14

ตารางที่ 3.14 Region Dimension Table

Field Name	Type	Length	Decimal	Description
RegionID	Int	2	0	รหัสเขตการขาย
RegionDesc	Varchar	30	0	รายละเอียดเขตการขาย

### 3.5.5. การวิเคราะห์ข้อมูลตามประเภทของสินค้า

การจัดหมวดหมู่ของผลิตภัณฑ์ในธุรกิจปูนซีเมนต์นั้น โดยส่วนใหญ่จะถูกแบ่งแยกจากวัตถุดิบที่ได้จากการทำเหมืองและทำการคัดเลือกตามประโยชน์ในการใช้งานก่อสร้าง โดยสินค้าหลักๆ ที่จะทำการจำแนกมีตารางดังนี้

ตารางที่ 3.15 ProdGroup Dimension Table

Field Name	Type	Length	Decimal	Description
ProdGroupID	Char	2	0	รหัสกลุ่มผลิตภัณฑ์
GroupName	Char	25	0	ชื่อกลุ่มผลิตภัณฑ์

### 3.5.6. การวิเคราะห์ข้อมูลตามตารางแบ่งการบรรจุของสินค้า

นอกจากการจัดประเภทให้กับผลิตภัณฑ์ การทำธุรกิจปูนซีเมนต์ยังสามารถแบ่งตามบรรจุภัณฑ์ เพื่อใช้เป็นการกำหนดราคา และแบ่งปริมาณการขายให้กับสินค้าแต่ละชนิดโดยมีตารางแบ่งการบรรจุสินค้านี้

ตารางที่ 3.16 ProdLotype Dimension Table

Field Name	Type	Length	Decimal	Description
ProdLotypeID	Char	2	0	รหัสการบรรจุผลิตภัณฑ์
ProdLotypeName	Char	25	0	ชื่อการบรรจุผลิตภัณฑ์

### 3.5.7. การวิเคราะห์ข้อมูลตามตารางการขายสินค้า

ตารางสุดท้ายคือตาราง Fact Table สำหรับรายงานแนวโน้มการขายสินค้า โดยมี ตาราง Sales เป็นตารางที่รวมการขายไว้ทั้งหมด เพราะข้อมูลการขายป้อนแต่ละครั้งไว้ว่าซื้อโดยลูกค้ารายใด จากทีมขายเขตอะไร และเวลาใดสามารถเรียกใช้ได้อย่างรวดเร็วเพราะสามารถหาคัดสรูปได้ เลข ตารางนี้เป็นการเก็บยอดขาย และข้อมูลต่างๆที่จะนำมาวิเคราะห์โดยรวมได้ รวมทั้งสามารถเรียกผ่านไปดูข้อมูลในตารางต่างๆที่ได้กล่าวมาแล้วได้ ดังนั้นจึงออกแบบตาราง Fact Table ที่เก็บในฐานข้อมูล Data Warehouse ได้ดังตารางที่ 3.17

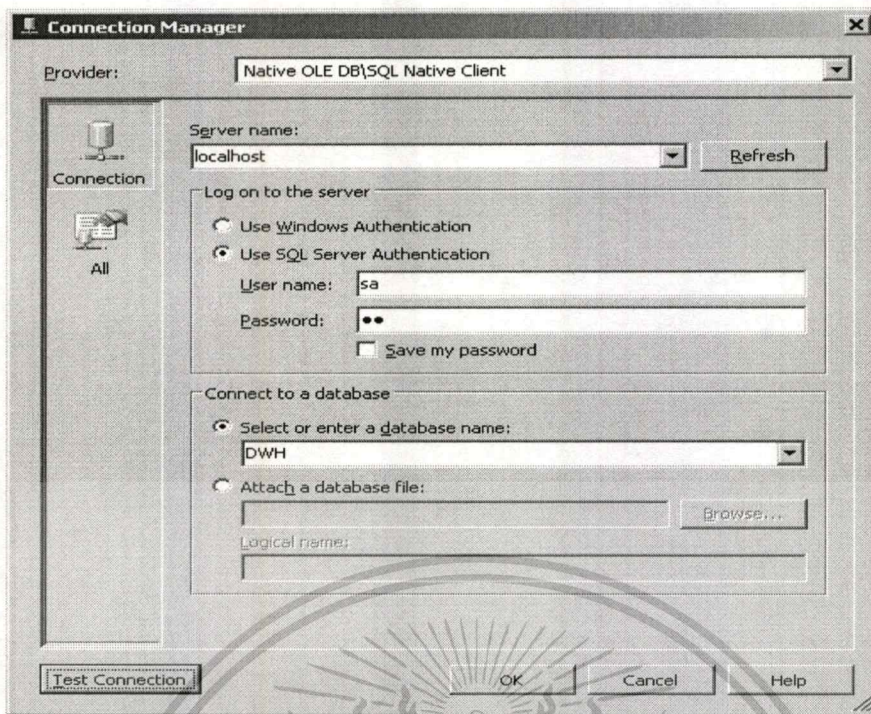
ตารางที่ 3.17 Sales Fact Table

Field Name	Type	Length	Decimal	Description
RegionID	Int	2	0	รหัสเขตการขาย
ProductID	Int	6	0	รหัสสินค้า
CustID	Int	7	0	รหัสลูกค้า
SaleDate	date/time	8	0	วันที่ขาย
Quantity	Decimal	10	2	ปริมาณการขาย
Discount	Decimal	8	0	ส่วนลด
Totalamt	Decimal	15	2	ผลรวม

ส่วนการกำหนด Measurement เพื่อที่จะเป็นตัววัดหรือค่าที่เราต้องการหาคือ Quantity (ปริมาณการขาย), Discount (ส่วนลด) และ Totalamt (ผลรวมยอดขาย)

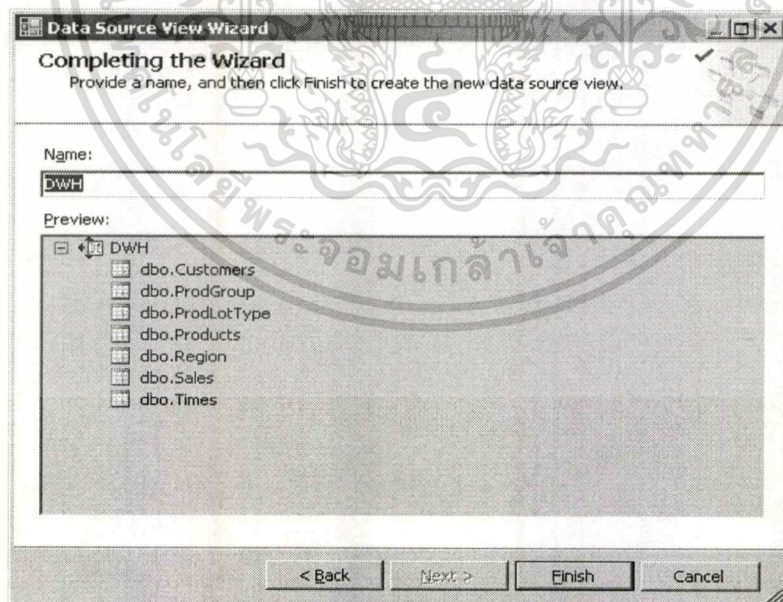
### 3.6 ความสัมพันธ์และการสร้างข้อมูลใน Sales Cube

เมื่อได้ตารางและความสัมพันธ์ของข้อมูลต่างๆแล้ว จึงนำข้อมูลมาสร้าง Sale Cube ใน Visual Studio 2005 โดยการกำหนดข้อมูล การติดต่อกฐานข้อมูลที่มีอยู่ใน MS-SQL Server 2005 โดยกำหนดค่าดังรูปที่ 3.9



รูปที่ 3.9 Connection to MS-SQL Server 2005

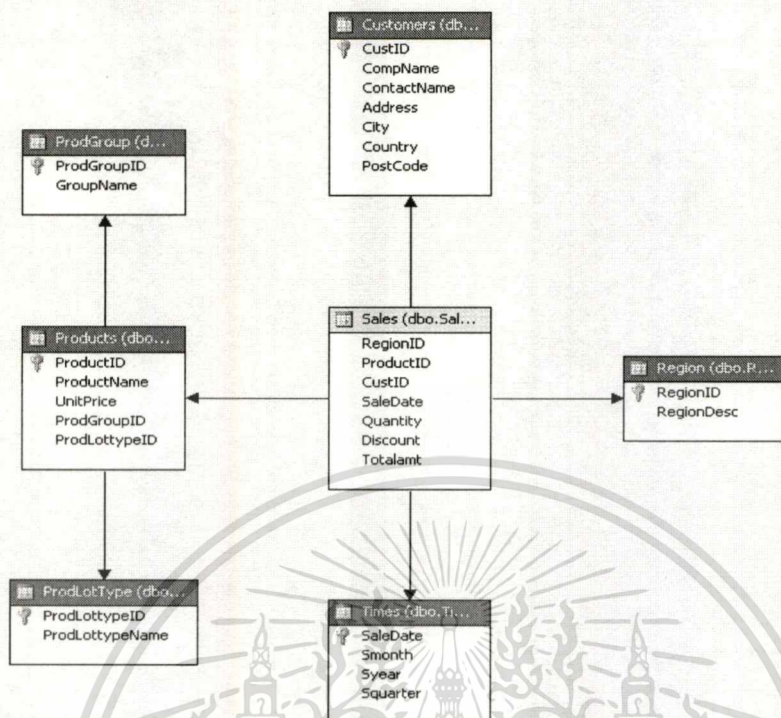
หลังจากนั้นกำหนดการสร้าง Data Source View โดย เลือก Tables ที่จัดเตรียมไว้สำหรับการสร้างความสัมพันธ์ ดังรูปที่ 3.10



รูปที่ 3.10 การกำหนด Data Source View

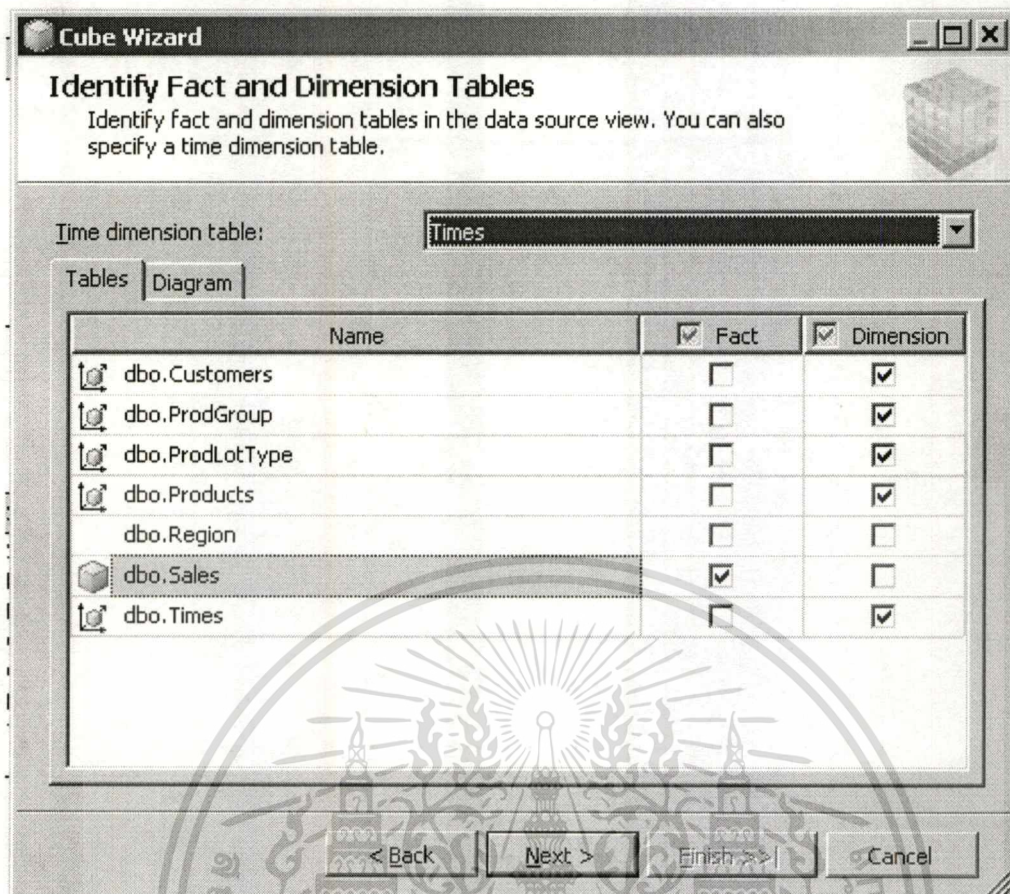
ขั้นตอนต่อไปกำหนดค่า Key ให้กับ Attribute ที่เป็น Primary Key และเชื่อมความสัมพันธ์กับ Table แต่ละTable เข้าด้วยกัน ทำการ Deploy Cube เพื่อให้ได้ความสัมพันธ์และ Dimensions ตามความต้องการ เมื่อสำเร็จจะได้ความสัมพันธ์ของข้อมูลดังรูปที่ 3.11

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



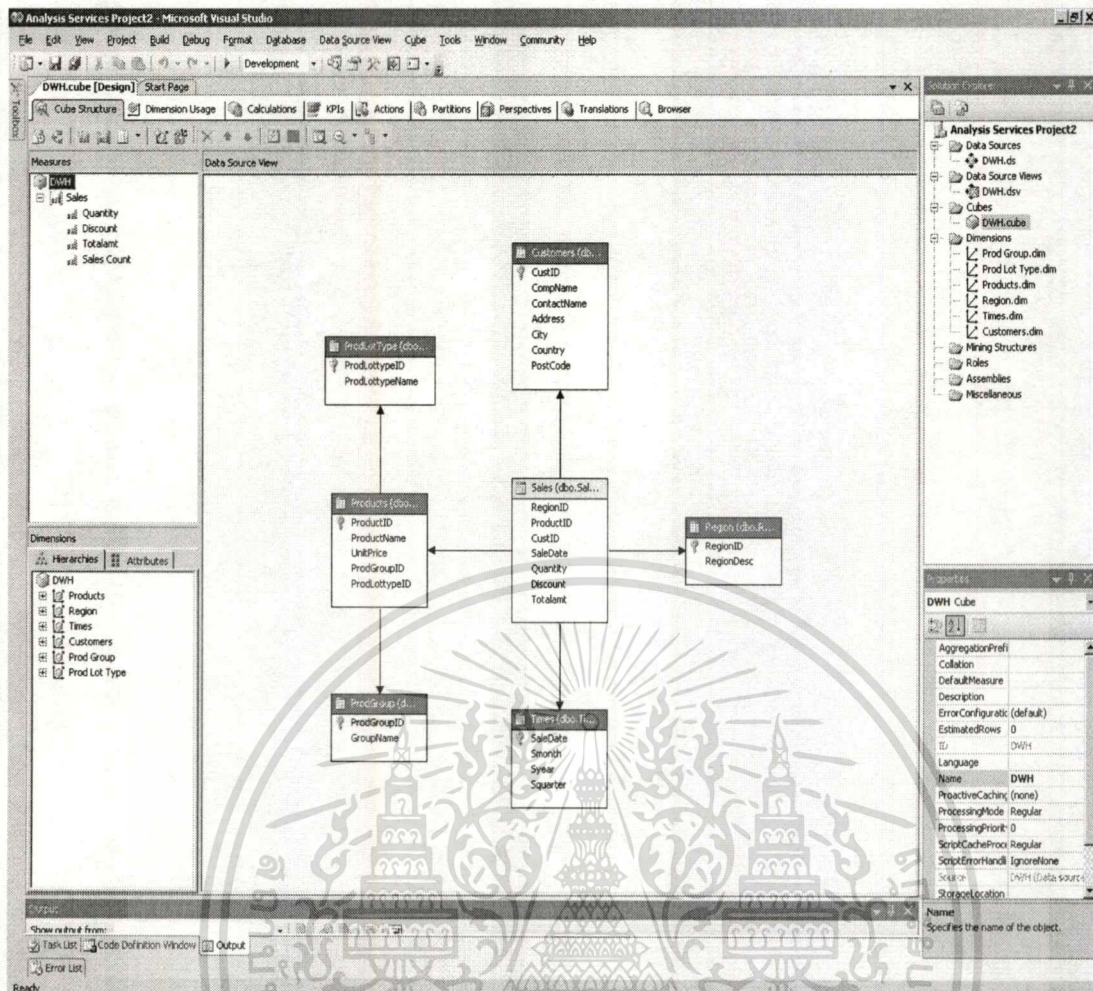
รูปที่ 3.11 ความสัมพันธ์ใน Sales Data Source View

เมื่อได้ ความสัมพันธ์ของ Sales Data Source View แล้ว ขั้นตอนต่อไปเป็นการสร้าง Cube ใหม่ขึ้นมาโดยเลือกค่า Connection และ Data Source View จากที่กล่าวมาแล้วข้างต้น เมื่อทำการ Process แล้ว Visual Studio 2005 จะให้ทำการเลือก Dimensions และ Fact Table ดังรูปที่ 3.12



รูปที่ 3.12 การกำหนดค่าในการสร้าง Cube

เมื่อเลือกค่าที่ต้องการ ได้ครบถ้วนและกำหนดค่า Measurement แล้วจึงทำการ Process ข้อมูลเพื่อให้เกิด โครงสร้างภายใน Cube ขึ้นดังรูปที่ 3.13



รูปที่ 3.13 โครงสร้าง Sale Cube

ในส่วนของโครงสร้าง Sale Cube สร้างขึ้นโดยกำหนดให้มีส่วนประกอบดังนี้

- Dimensions

- Customers (ลูกค้า)
- Products (สินค้า)
- Time (เวลาขาย)
- Region (เขตการขาย)
- ProdLotType (การบรรจุผลิตภัณฑ์)
- ProdGroup (ประเภทสินค้า)

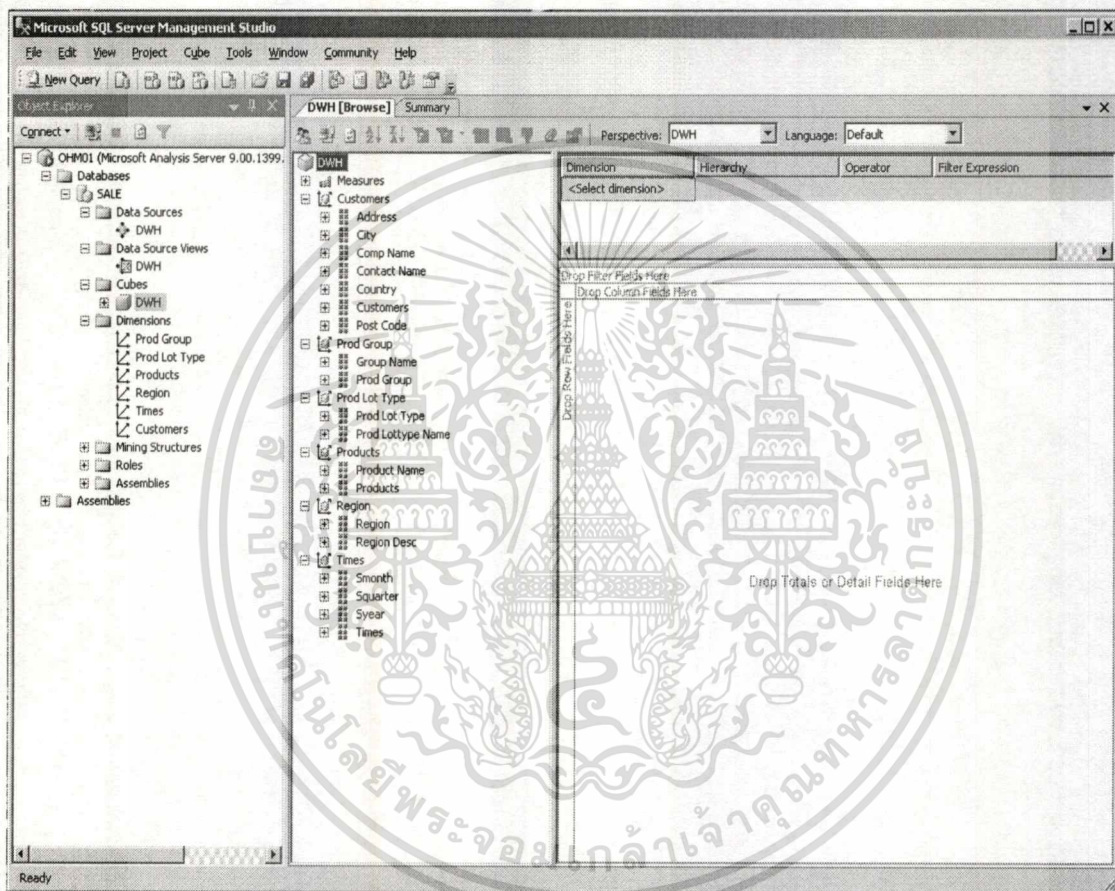
-Measurement

- Quantity (ปริมาณการขาย)
- Discount (ส่วนลด)
- Sale Count (จำนวนครั้งที่ขาย)
- Totalamt(ผลรวมการขาย)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 3.7 ทำการตรวจสอบข้อมูลของ Sales Cube ใน Analysis Services 2005

ขั้นตอนในการทำงานของ Analysis Services 2005 ต่างจาก Analysis Services 2000 โดยสิ้นเชิง โดยเฉพาะการสร้าง Cube ซึ่ง Analysis Services 2005 ใช้ Visual Studio 2005 เป็นเครื่องมือในการสร้าง Cube ขึ้นมา หลังจากทำการสร้างเสร็จแล้วจึงทดสอบโดยการเปิด Analysis Services Management Studio และทดสอบโดยการใส่ Cube Browser เรียกดูข้อมูลที่ต้องการ ดังรูปที่ 3.14



รูปที่ 3.14 Analysis Services Management Studio

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทที่ 4

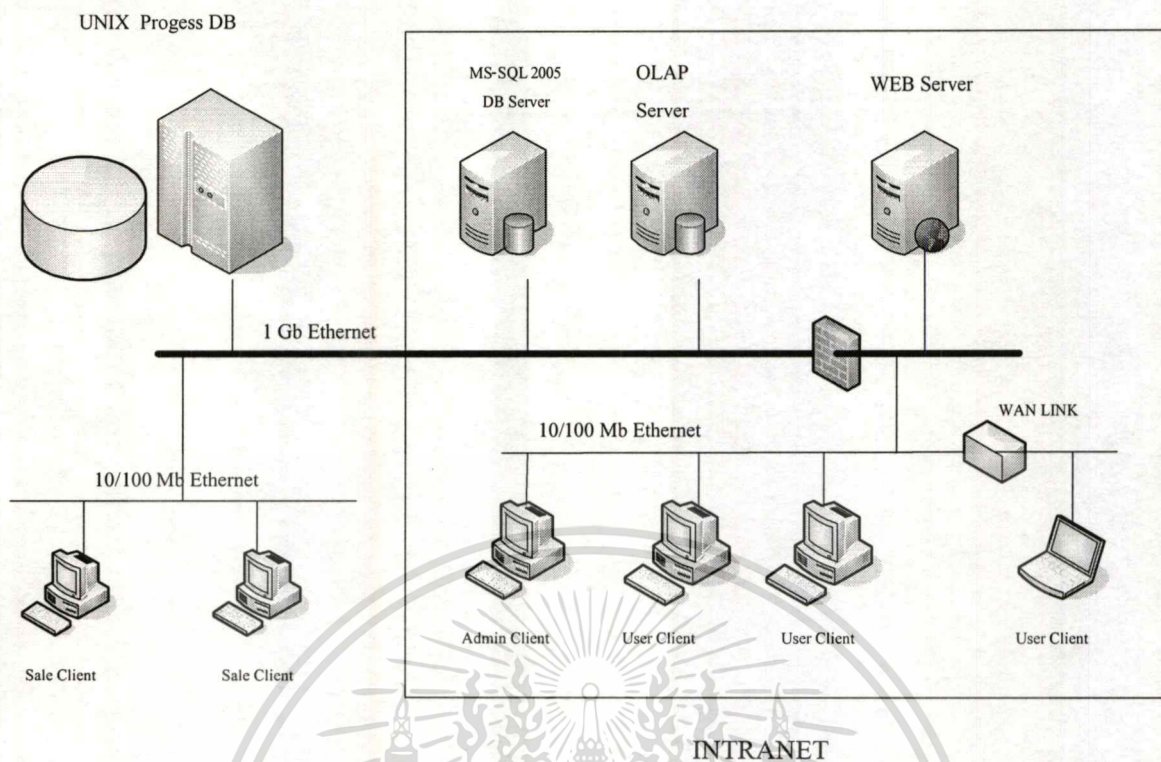
### การพัฒนาระบบวิเคราะห์การขาย

ในบทนี้จะขอก้าวในเรื่องของโครงสร้างของระบบโดยรวมทั้งหมด เช่น สถาปัตยกรรม การสื่อสาร Hardware Software โดยมีสถาปัตยกรรมอินเทอร์เน็ตเป็นพื้นฐาน ซึ่งเน้นการทำงานของผู้ใช้หลายคน โดยผู้ใช้แต่ละคนในที่นี้สามารถเข้าถึงระบบแตกต่างกันตามสิทธิที่ได้ในการเข้าถึงอินโดยผ่านหน้าเว็บ นอกจากนี้ยังกล่าวถึงการพัฒนาระบบการวิเคราะห์ข้อมูลโดยเน้นในส่วนของตัวโปรแกรม Cube Browser ที่ทำงานผ่านเว็บ เน้นความง่ายและมีประสิทธิภาพในการใช้งานสำหรับวิเคราะห์งานขายในองค์กร

#### 4.1 สถาปัตยกรรมด้านการสื่อสารและฮาร์ดแวร์

ในระบบการขายปูนซีเมนต์ขององค์กรนั้นมีระบบการขายปูนซีเมนต์เป็นหลัก โดยมุ่งเน้นข้อมูลการขายของสินค้าและการเก็บข้อมูลต่างๆขององค์กร โดยผู้ใช้งานจะอยู่ตามสาขาต่างๆ และส่วนกลาง โดยอาศัยการเชื่อมต่อผ่านระบบ LAN เข้ามาเรียกใช้บริการฐานข้อมูลกลางได้สะดวกรวดเร็ว ในส่วนของเครื่อง Server ต่างๆที่ตั้งขึ้นจะถูกสร้างเป็นกลุ่มในการจัดเก็บเรียกว่า Server Farm โดยมีการเชื่อมต่อของเครื่องข่ายวงเดียวกันเพราะฉะนั้นการสื่อสารของแต่ละ Server จะมีประสิทธิภาพและปลอดภัยมากยิ่งขึ้น ในโครงการ Data Warehouse ประกอบด้วย Server ทั้งหมด 3 เครื่องเชื่อมต่อกับเครื่อง Server ของการขาย เพื่อใช้รับส่งข้อมูลให้กันและกัน การใช้งานของ Data Warehouse ต้องผ่านทาง Web Server เท่านั้นไม่สามารถติดต่อโดยตรงกับ OLAP Server ได้ โดยมี Firewall เป็นตัวกำหนดสิทธิของเครื่อง

สถาปัตยกรรมทางด้านฮาร์ดแวร์นั้น ควรคำนึงถึงภาพรวมของของส่วนต่างๆที่เกี่ยวข้องกับระบบทั้งหมด ได้แก่ การเชื่อมต่อ , การรับส่งของ Server, Client ต่างๆที่เกี่ยวข้อง รวมถึงการนำข้อมูลมาเก็บไว้ใน Data Warehouse การสร้าง OLAP Cube จนกระทั่งการที่ผู้ใช้งานสามารถเข้ามาใช้งานฐานข้อมูลที่อยู่ใน Database Server และ OLAP Server โดยผ่านทาง Web Server ได้ ดังรูปที่ 4.1



รูปที่ 4.1 โครงสร้างและสถาปัตยกรรมของระบบ

สถาปัตยกรรมด้านฮาร์ดแวร์ที่เกี่ยวข้องกับระบบทั้งหมดประกอบด้วย

1. เครื่อง HP Server โดยใช้ OS HP-UNIX เป็นระบบปฏิบัติการ ส่วนการเก็บข้อมูลการขายใช้ฐานข้อมูล Progress V. 7 เป็น RDBMS ฐานข้อมูลนี้ทำหน้าที่ส่งข้อมูลการขายไปยัง Database Server หรือเรียกได้ว่าเป็นต้นทางของข้อมูลนั่นเอง
2. เครื่อง Data Warehouse Database Server ทำหน้าที่แปลงข้อมูลและจัดเก็บข้อมูลต่างที่ได้ผ่านกระบวนการ ETL มาแล้ว
3. เครื่อง Server ที่ทำหน้าที่เป็น OLAP Server สำหรับเก็บข้อมูลของ Cube
4. เครื่อง Web Server ทำหน้าที่ให้บริการเว็บต่างที่ใช้ในการเรียกข้อมูลในฐานข้อมูล มาแสดงโดยใช้ IIS เป็นตัวจัดการเกี่ยวกับเว็บ
5. เครื่อง User Client เป็นเครื่องที่เรียกใช้บริการระบบ โดยทำงานผ่าน Web Browser
6. อุปกรณ์ Network ต่างๆ ที่ใช้ในการเชื่อมต่อทุกระบบเข้าด้วยกัน เช่น Firewall , HUB ฯลฯ

จากอุปกรณ์เบื้องต้นที่ได้กล่าวมาแล้วเป็นอุปกรณ์ส่วนใหญ่ที่มีอยู่แล้วภายในองค์กร ซึ่งต้องการเพียงการวางแผนการพัฒนาระบบการวิเคราะห์การขายปูนซีเมนต์และ Data warehouse เข้ามาใช้เพื่อเป็นการให้บริการเพิ่มอีกรูปแบบหนึ่งของการขาย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 4.2 สถาปัตยกรรมด้านซอฟต์แวร์

การใช้งานในระบบวิเคราะห์ข้อมูลการขายบนซีเมนต์ เพื่อให้สะดวกและใช้งานได้ทุกเขตการขาย รวมถึงค่าใช้จ่าย จึงได้นำระบบเทคโนโลยีอินเทอร์เน็ตมาใช้ในการเข้าใช้งานระบบ ซึ่งได้พัฒนาตัวโปรแกรม Cube Browser ให้สามารถทำงานได้ในรูปแบบเว็บ

### 4.2.1 OS (Operating System)

ในภายองค์กรมีการใช้ระบบปฏิบัติการที่หลากหลาย จะสามารถแยกได้เป็นเครื่องคอมพิวเตอร์ที่ให้บริการและเรียกใช้บริการ

- เครื่องที่ให้บริการ ส่วนใหญ่จะเป็นเครื่อง Server ต่างๆ ใน Server Farm ซึ่งมีระบบปฏิบัติการที่ใช้เช่น Windows 2003 ใช้งานบน เครื่อง Web Server, Database Server, LAP Server ส่วนเครื่องที่ใช้ในการจัดเก็บข้อมูลการขายใช้ระบบปฏิบัติการ HP – UNIX

- เครื่อง User Client ภายในองค์กรทั้งหมดเป็นระบบปฏิบัติการ Windows ซึ่งต่างกันตรงเวอร์ชัน โดยมีทั้ง Windows 98 และ Windows XP ควบคู่กันไป ซึ่งไม่มีปัญหาการใช้งานระบบ เพราะมี Internet Explorer Browser เป็นมาตรฐานอยู่แล้ว

### 4.2.2 Database Server

เครื่องมือในการจัดการระบบฐานข้อมูลสำหรับ Relational Database หรือ Data Warehouse นั้นใช้ MS-SQL 2005 SERVER ส่วน OLAP Server ที่ทำหน้าที่ให้บริการข้อมูลจาก Cube นั้น ใช้ MS-SQL 2005 SERVER Analysis Service โดยจะใช้ Visual Studio 2005 เป็นตัวพัฒนาในส่วนต่างๆ การ Install โปรแกรมเกี่ยวกับการจัดการข้อมูลจึงเป็น โปรแกรมชุดเดียวกันทั้งหมด

### 4.2.3 Client Tool

ในส่วนของ User Client มีเครื่องมือในการจัดการกับฐานข้อมูลคือ

- เครื่องมือในการจัดการของผู้ดูแลระบบ ในที่นี้ใช้ MS-SQL 2005 Server Enterprise Manager สำหรับเป็นตัวจัดการฐานข้อมูล SQL Server 2005 และ Analysis Manager 2005 ใช้ในการจัดการ OLAP Server ซึ่งโปรแกรมดังกล่าวสามารถทำการสร้างแก้ไข Dimension และ Cube ได้เลย ซึ่งโปรแกรมทั้งสองนี้จะถูกรวมมาในชุด MS-SQL 2005 Server

- โปรแกรม Web Browser เป็นโปรแกรมเรียกใช้เว็บบน Web Server ซึ่งเป็นส่วนที่จะแสดงการวิเคราะห์ข้อมูลการขายที่อยู่ในฐานข้อมูลทั้ง Database Server และ OLAP Server มาแสดง ซึ่งรูปแบบของการใช้งานระบบที่พัฒนานั้นเน้นการปรับเปลี่ยนความสามารถให้ตรงกับความต้องการของผู้ใช้งานมากที่สุด

### 4.2.4 Middleware

การรับส่งข้อมูลระหว่างเครื่อง Server หลายๆเครื่อง จำเป็นต้องมีซอฟต์แวร์ที่ใช้จัดการรูปแบบการจับส่งข้อมูลหรือบางทีเรียกว่า Middleware หน้าหน้าของ Middleware ทำหน้าที่ให้ข้อมูลใน

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แต่ละเครื่อง Server ที่ให้บริการรับส่งข้อมูลในรูปแบบเดียวกันได้ รวมถึง User Client สามารถติดต่อกันได้โดยทาง Middleware ที่เกี่ยวข้องดังนี้

- OLE DB Provider For OLAP Services 8.0 ใช้ในการติดต่อ OLAP Server

- SQL Native Client หรือ SQLNCLI.1 For SQL Server สำหรับการติดต่อกับฐานข้อมูล SQL Server ที่เป็นฐานข้อมูล Data Warehouse

#### 4.2.5 เครื่องมือในการพัฒนาระบบ

เครื่องมือในการพัฒนาระบบวิเคราะห์ข้อมูลการขายปูนซีเมนต์นั้น เน้นในการพัฒนาให้ผู้ใช้งานผ่านเว็บ สิ่งที่สำคัญที่สุดในระบบนี้คือ โปรแกรม Cube Browser ซึ่งการพัฒนานั้นได้ใช้โปรแกรม MS Visual Studio 2005 เป็นตัวสร้าง Cube และ ตัวโปรแกรม Cube Browser

MS Visual Studio 2005 เป็นโปรแกรมที่มีความคล่องตัวและมีความยืดหยุ่นสูงเพราะมีการเรียกใช้งาน Component ต่างๆ สำหรับการเชื่อมต่อไปยังฐานข้อมูล MS-SQL 2005 Server และ OLAP Server ซึ่งสามารถนำมาพัฒนาและใช้งานได้อย่างมีประสิทธิภาพ

### 4.3 การออกแบบระบบสำหรับการวิเคราะห์ข้อมูลการขายปูนซีเมนต์

ในโครงการระบบวิเคราะห์การขายปูนซีเมนต์นี้เน้นในการพัฒนา OLAP Cube Browser ขึ้นใช้เองเพื่อลดรายจ่ายในการซื้อลิขสิทธิ์ของโปรแกรมต่างๆ ที่ใช้ในการดูข้อมูลที่เป็น OLAP Cube โดยโปรแกรมได้ถูกพัฒนาขึ้นสามารถนำไปใช้กับ OLAP Cube ที่สร้างและใช้งานใน SQL Server Analysis Service ได้ทุกระบบ ในหัวข้อของโครงการนี้ได้สร้างข้อมูลที่เกี่ยวข้องกับการขายปูนซีเมนต์ไว้แล้ว ดังนั้น โปรแกรมนี้จึงถูกเขียนผูกกับ ข้อมูล Cube ที่ได้จากกระบวนการสร้างข้อมูลการขายเท่านั้น เพื่อเป็นการให้ผู้ใช้งานได้มีทักษะในการสร้างรายงานในการดูข้อมูลต่อไป

#### 4.3.1 รูปแบบการแสดงผลของโปรแกรม

จากความเคยชินของผู้ใช้งานการสร้างเอกสารทั่วไป มักถนัดกับข้อมูลที่รูปแบบเป็นตารางคือข้อมูลที่ประกอบด้วยแถวและคอลัมน์ ดังนั้นการออกแบบของโปรแกรมจึงออกแบบเป็นตารางเพื่อที่จะนำแถวและคอลัมน์ในรูปของมิติต่างๆที่อยู่ใน Cube ออกมาแสดง การออกแบบหน้าหลักสำหรับการเข้าถึงข้อมูลนั้นสามารถดึงข้อมูลใน Cube ทั้งหมดโดยสามารถเลือก Dimension ใดๆมาแสดงก็ได้ นอกจากนี้ Dimension อื่นๆที่อยู่นอกเหนือจากการเลือกมาแสดงด้านแถวและคอลัมน์ จะสามารถนำมาใส่ไว้ที่จำนวนการแยกหน้าหรือสามารถทำเป็น Filter ของข้อมูลได้อีก เมื่อคัดเลือกรายการที่ต้องการมาแสดงในรูปแบบตารางได้แล้ว แต่ละ Dimension เราสามารถทำการเลือกดูรายละเอียดได้โดยการ Drill Down การทำลักษณะนี้เป็นการเข้าถึงข้อมูลในแต่ละ Level ได้ โดยข้อมูลที่อยู่ติดลงไปจากการ Drill Down นี้จะเป็นมาตรฐานในระดับล่างสุดของข้อมูล Cube หรือเรียกได้ว่าเป็นข้อมูลที่ลึกที่สุดของ Cube นั้น ซึ่งเหมือนกันเกือบทุกระบบ นี้เป็นสิ่งจำเป็นอีกอย่างหนึ่งซึ่งโปรแกรมที่ทำหน้าแสดงผลของ Cube ควรจะมี

ตัวโปรแกรมสามารถแสดงตารางที่ได้ทำการเลือกแถวและคอลัมน์นั้นไว้ โดยมีข้อมูลที่ต้องการใช้เพื่อชี้วัดหรือ Measurement ที่สร้างไว้ใน Cube ต่อยู่ระหว่างกลาง เป็นข้อมูลที่ผู้ใช้วัดในเชิงปริมาณและเชิงคุณภาพ เช่น ยอดขายรวม, จำนวนการขาย, ปริมาณในการขาย เป็นต้น ข้อมูล Measurement ต่างๆที่เราสามารถเลือกเพื่อให้เห็นในหน้าตารางได้ตามต้องการ ดังรูปที่ 4.2

Product Name	2004		2005	
	Quantity	Discount	Discount	Quantity
All	9154642	1.807313E+08	1.296622E+08	5684597
<MORTAR [JUST>	53.76	0	0	135.73
1 1/2-3/4"				
1 1/2-3/8"				
1 1/4-3/4"				
1/2-3/16"			0	956.46
1-0" scrap				
1-7/8"				
2 1/2-3/4"				
2-1 1/4"				
3/16-0"	885438.4	0	0	736885.9
3/4-3/16"				
3/4-3/8"	3876.2	7450		
CLINKER TYPE 1				
CLINKER TYPE 5				
M299 กงพระอาทิตย์	26254.5	0	0	180
MORTAR DUST 800				
Portland Type 5-Bag				
Portland Type 5-Bulk				
Portland Type I				
TPI เขียว-สูง 1 ตัน	1470778	2.584723E+07	2.80256E+07	781109.8
TPI เขียว-สูง 50 กก			0	422
TPI เขียว-สูงพิเศษ 50 กก	8459.711	177650	96350	4261.15
TPI เขียว-หง				
TPI เหลือง-สูง 1 ตัน				
TPI เหลือง-สูง 50 กก				
TPI เหลือง-สูงพิเศษ 50 กก				
TPI เหลือง-หง	11981.45	239350	144850	8290.79
TPI แดง-สูง 1 ตัน	68022.37	0	0	23304.72
TPI แดง-สูง 50 กก				

รูปที่ 4.2 หน้าจอของโปรแกรม Cube Browser

ในการเลือกข้อมูลที่แสดงในตารางของโปรแกรมนี้ใช้ Field List Bar ซึ่งประกอบด้วยรูป Folder แสดงแทน Dimensions ต่างๆ ที่อยู่ใน Cube และส่วนที่แตกย่อยออกมาก็คือ จำนวนของ Hierarchy ว่า มี Member เป็นอะไรขึ้นอยู่กับความลึกและข้อมูลของ Cube ส่วน Field List Bar ทำหน้าที่เป็นเครื่องมือใช้เลือกข้อมูลต่างๆ ให้แสดงรูปในแบบของตาราง Dimension โดยมีทั้งแถวและคอลัมน์เป็นส่วนประกอบ หรือ ในบางครั้งการนำข้อมูลมาแบ่งเป็นหน้าๆโดยใช้ ข้อมูลที่ต้องการเป็น Filter เป็นการแสดงข้อมูลอีกรูปแบบหนึ่งให้เห็นภาพชัดมากยิ่งขึ้น ทั้งนี้ทั้งนั้น การสร้างตารางขึ้นอยู่กับความต้องการในการวิเคราะห์ข้อมูลของผู้ใช้งาน

#### 4.3.2 รูปแบบการแสดงผลในรูปแบบรายงาน

การแสดงผลในรูปแบบรายงานเป็นอีกวิธีหนึ่งที่ใช้ในการวิเคราะห์ข้อมูลได้ชัดเจน แต่ผู้ใช้งานไม่สามารถปรับเปลี่ยนรูปแบบ Dimension หรือ แถวและคอลัมน์ ได้ตามความต้องการ ดังรูปเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่อนุญาตเห็นนำไปใช้ประโยชน์ด้านการค้า

Summary report By Product in years.rdl

1 of 2 Whole Page Find | Next

Summary report By Product in years

	2004		2005	
	Totalamt	Quantity	Totalamt	Quantity
MORTAR DUST >	37094.3984375	53.75999832153	93653.703125	135.7299987275
1-0" scrap			65646	936.4599993375
3/4-3/16"	79572456	885438.375	73783976	736985.8375
CLINKER TYPE 1	4945250	3076.200195312	5	
M259 อุตสาหกรรม	2712922	26254.5	304512	160
TPX 50 มก.	2400063744	140778.375	117563098	781109.8125
TPX 42 มก.			13926	422
TPX 35 มก.	1167304	849.7103375	562490	4261.150390525
TPX 28 มก.	20051200	11981.44921875	9963809	6290.70039062
TPX 21 มก.	2856939.75	68822.3671875	977957.375	23304.71875
TPX 14 มก.	1848013696	973667.9375	886589472	436396.65625
TPX 7 มก.	5223048	156456	3523068	96224
TPX 50 มก. 50 มก.	425328768	2439674.75	2461165568	1558202.5
TPX 42 มก. 1.5 มก.			39777.75	6651.679453125
TPX 35 มก. 50 มก.	6295856	30896.5398125	33162764	16982.25
TPX 28 มก. 50 มก.			140288	3082
TPX 21 มก. 1.5 มก.	492992320	268123.9375	291521216	167710.1125
TPX 14 มก. 1.5 มก.	1221349046875	319.8299965722	66	
TPX 7 มก. 50 มก.	8307211	4319.800397656	2923598	1430.400324484
TPX 50 มก. 50 มก.	51245	1409	28040	726
TPX 42 มก. 50 มก.	6994512	331342	3489488	225667.390525
TPX 35 มก. 50 มก.	10947097	14330.6302062	1337879	17754.40039062
TPX 28 มก. 50 มก.	3872136	47498.5398125	31740038	38898.25

รูปที่ 4.3 รายงานต่างๆ ที่ใช้ Report Service

ในการพัฒนาระบบงานนี้ได้ทำการสร้างรายงาน โดยให้รายงานต่างๆติดต่อกับฐานข้อมูล OLAP Cube รูปแบบการแสดงผลของระบบมีการนำ Report Server 2005 เป็นตัวจัดการในการออกรายงานและแสดงรายงาน ในรูปแบบเว็บ เครื่องมือที่ใช้ในการพัฒนายังคงเป็น MS Visual Studio 2005 ที่ติดตั้งมาพร้อมกับ MS-SQL 2005 Server นั้นเอง

4.3.3 ความสามารถต่างๆ และการทำงานของโปรแกรม

ความสามารถต่างๆ ของโปรแกรม ได้แก่

1. สามารถทำการเชื่อมต่อกับ ข้อมูลที่เป็นลักษณะ OLAP Cube ใน SQL 2005 Server Analysis ได้ โดยการดึงข้อมูล Dimension , Hierarchy และ Measurement มาแสดงในรูปแบบตาราง
2. สามารถ Drill Down ข้อมูลใน Dimension ในตารางเพื่อนำข้อมูลต่างๆนำมาวิเคราะห์หรือเจาะจงดูข้อมูลได้
3. ทำการเรียงลำดับข้อมูลในตารางได้ เช่น เรียงลำดับตาม Measurement หรือรายชื่อ

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ของโรงเรียนที่จัดทำขึ้นเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าของ แลัวที่ได้แสดงข้อมูลทางด้านซ้ายมือ ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4. สามารถเปลี่ยนจำนวนของตัวเลขใน Measurement ให้แสดงในรูปแบบกราฟเชิงปริมาณ อย่างง่าย เพื่อให้เห็นความแตกต่างของข้อมูล
5. สามารถนำข้อมูลออกเป็นไฟล์ EXCEL ได้เพราะในบางครั้งต้องการส่งต่อข้อมูลเพื่อไปสร้างรายงานที่ซับซ้อนต่อไป
6. สามารถพิมพ์ออกเป็นรายงานทางเครื่องพิมพ์ได้

ที่กล่าวมาทั้งหมดเป็นความสามารถหลักๆของระบบที่พัฒนาขึ้น เพื่อให้ความสะดวกและมุมมองของข้อมูลแก่ผู้ใช้งานนอกจากนี้ยังมีความสามารถอื่นๆอีกหลายๆด้าน เช่นการเลือกแสดงข้อมูลที่เจาะจงให้ตามความต้องการ การรวบรวมข้อมูลเข้าด้วยกัน ฯลฯ โดยรูปแบบของตาราง Cube Browser จะขึ้นอยู่กับความชำนาญและมุมมองของผู้ใช้งานเป็นหลักสำคัญ

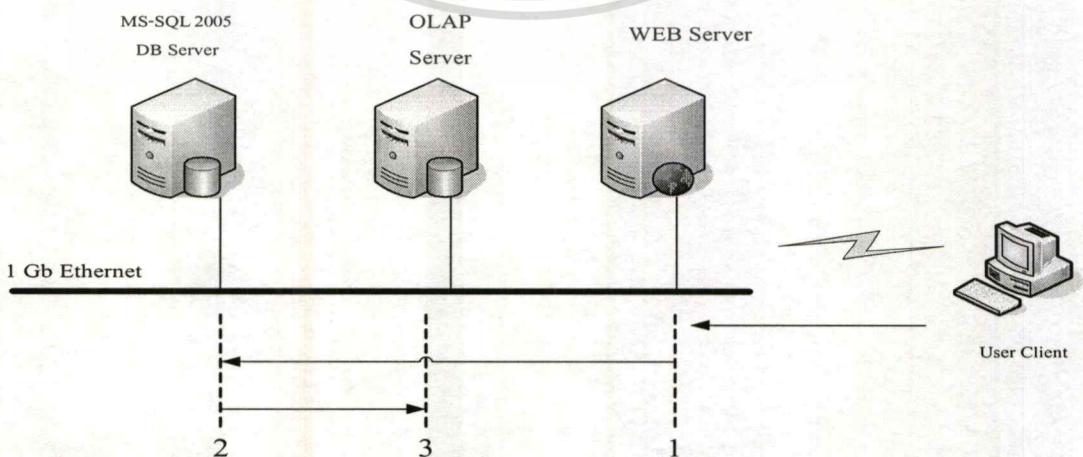


## การใช้งานระบบและบำรุงรักษาระบบ

ในบทนี้จะกล่าวถึงการใช้งานของระบบวิเคราะห์ข้อมูลการขายปูนซีเมนต์ ที่ได้พัฒนาขึ้นมาเพื่อให้ผู้ใช้งานใช้ในการวิเคราะห์การขายสินค้าขององค์กรได้อย่างแม่นยำ ซึ่งผู้ใช้งานในส่วนใหญ่จะปฏิบัติการอยู่ตามต่างจังหวัดหรือเขตการขายต่างๆ โดยตามสาขาต่างๆมีการเชื่อมต่อเครือข่ายกับสำนักงานใหญ่โดยผ่านเครือข่ายระบบอินเทอร์เน็ต ผู้ใช้งานในองค์กรสามารถเข้าใช้งานข้อมูลได้โดยผ่าน Web Browser และเชื่อมต่อกับ Web Server เข้ามาอ่านข้อมูลตามลำดับ เนื้อหาท้ายสุดจะกล่าวถึงการบริหารระบบข้อมูลให้ถูกต้อง การจัดการข้อมูลตั้งเข้าสู่ Cube เพื่อให้ผู้ใช้งานได้เข้าใช้งานและได้ข้อมูลที่ถูกต้อง

### 5.1 การเข้าสู่การใช้งานของระบบ

งานของระบบวิเคราะห์ข้อมูลการขายปูนซีเมนต์ถูกพัฒนาโดยใช้ผลิตภัณฑ์ของ Microsoft ทั้งหมด ในการพัฒนาเว็บได้ใช้ Visual Studio 2005 , ASP.NET และ VB.NET ซึ่งเครื่องมือเหล่านี้สามารถใช้ได้ดีกับฐานข้อมูล MS-SQL SERVER 2005 ดังนั้นจึงง่ายต่อการทำงานในการพัฒนาในส่วนต่างๆ รวมถึงความสะดวกต่อผู้ใช้งาน เพราะระบบทำงานบน Web server ซึ่งเครื่องของ User Client ทั้งหมดเป็นเครื่องที่ใช้ระบบปฏิบัติการ Windows XP เหมือนกันหมด รวมถึงการเชื่อมต่อกับเครือข่ายอินเทอร์เน็ตหรือต่อกับเน็ตเวิร์คขององค์กรสามารถติดต่อใช้งานได้ตลอดเวลา การเข้าใช้งานเพียงเรียก โปรแกรม Browser Internet Explorer ขึ้นมาและเรียกใช้บริการโดยต่อเข้ากับหมายเลข IP Address ของ Web Server ซึ่งขั้นตอนการทำงานจะทำงานตาม รูป 5.1



รูปที่ 5.1 ขั้นตอนการใช้งานของระบบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

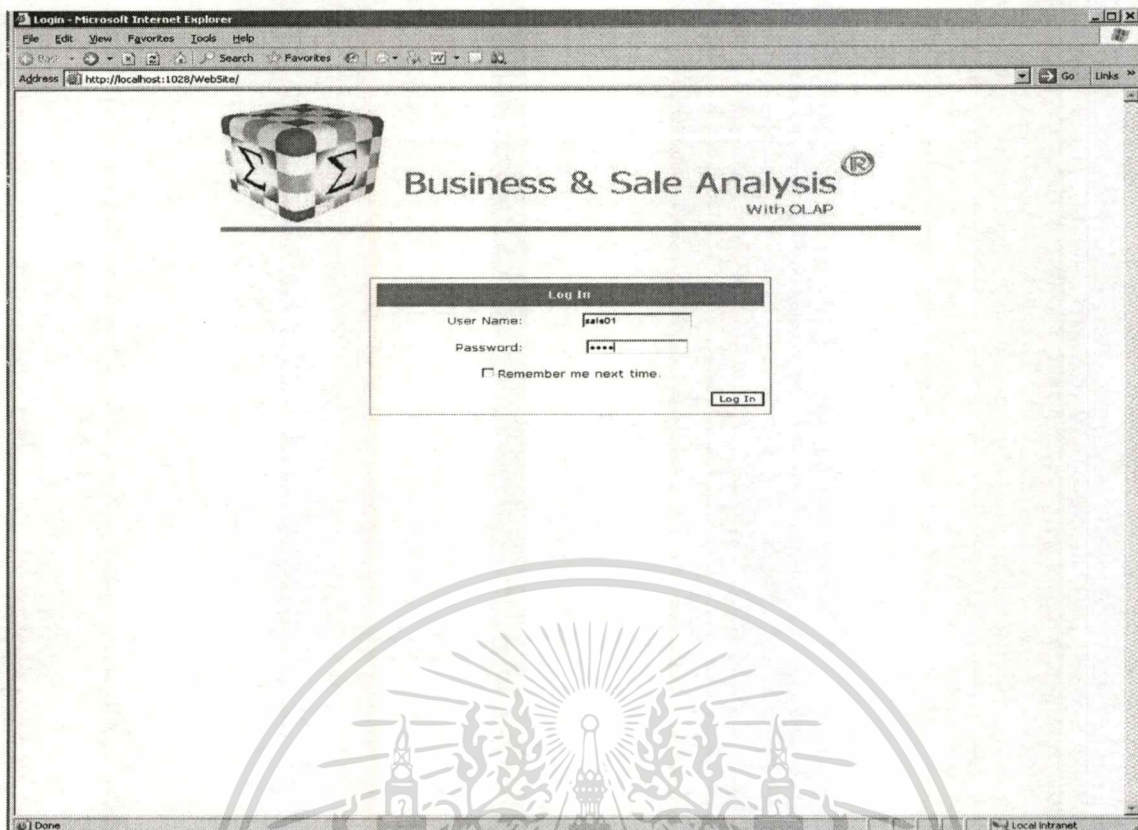
การเข้าใช้งานของระบบมีขั้นตอนดังนี้จาก

1. เมื่อ User Client ทำการเปิดเข้าเว็บโดยทำการเรียกใช้บริการมายัง Web Server ซึ่งหน้าเว็บแรกให้ผู้ใช้งานใส่ User และ Password หลังจากนั้นข้อมูลจะถูกส่งไปยังขั้นตอนต่อไป
2. ขั้นตอนนี้จะทำการตรวจสอบ User และ Password ให้กับการ login ของ User นั้นๆ โดยมีการค้นหาข้อมูลที่ได้ถูกจัดเก็บในฐานข้อมูล MS-SQL 2005 ถ้าผู้ใช้งานมีสิทธิก็จะสามารถเข้าถึงระบบฐานข้อมูลในเครื่อง MS-SQL 2005 ได้เลย
3. ขั้นตอนนี้เป็นขั้นตอนแสดงถึงผู้ที่มีสิทธิในการเข้าใช้ระบบเท่านั้นที่สามารถ ติดต่อกับ OLAP Server โดยสามารถทำการเปิดข้อมูลในลักษณะหลายมิติจาก Cube ได้ตามความต้องการ

ในการเข้าถึงข้อมูลและการจัดการข้อมูลของแต่ละระดับของผู้ใช้ขึ้นอยู่กับสิทธิ ที่ได้รับว่าอยู่ในระดับใด แต่ในทุกระดับสามารถเข้าใช้ตัวโปรแกรม Cube Browser ได้ทั้งหมด ในบางครั้งมักเกิดปัญหากับผู้ใช้งานเนื่องจากผู้มีส่วนใหญ่ขาดทักษะ และความเคยชินกับการนำข้อมูลหลายมิติมาใช้ รวมทั้งการดูข้อมูลหลายมิติด้วย เพื่อให้ระบบสามารถใช้งานได้ประโยชน์สูงสุด ทางองค์กรควรมีการอบรมให้กับพนักงานที่ใช้ระบบงานนี้ โดยเน้นการใช้งาน และการนำข้อมูลที่แสดงมาสร้างรายงานเพื่อใช้ในการวิเคราะห์การขยายต่อไป

## 5.2 สิทธิในการเข้าใช้ข้อมูลของระบบ

ในระบบวิเคราะห์ข้อมูลการขยายมุมมองข้อมูลจะถูกเก็บไว้ในฐานข้อมูล Relational และ OLAP ซึ่งแน่นอนว่าภายในองค์กรผู้ที่มีสิทธิใช้งานต้องมีสิทธิในการเข้าถึงฐานข้อมูลเท่านั้น จะมีเพียงพนักงานบางคนเท่านั้นที่สามารถเข้าดูและจัดการระบบฐานข้อมูลภายในองค์กรได้ และมีพนักงานบางคนที่สามารถจะเข้าดูข้อมูลเพียงอย่างเดียวเท่านั้น จากรูป 5.2



รูปที่ 5.2 การ Log in เข้าสู่ระบบ

ในการเพิ่มหรือแก้ไขสิทธิของผู้ใช้จะเป็นหน้าที่ของผู้จัดการระบบ เป็นผู้ดำเนินการรับผิดชอบ โดยอาศัยการทำงานของ หน้าการจัดการสิทธิให้กับผู้ใช้งาน จากรูป 5.3 หน้าเว็บนี้จะเป็นการจัดการสิทธิให้กับผู้ใช้งานระดับต่างๆได้ ในการเข้าใช้งานฐานข้อมูลของพนักงาน แต่ละคน โดยแบ่งเป็นสิทธิตามลำดับดังนี้

UserName	Password	Position	empID	Description	Create Date	Edit	Delete
MSale	*****	Sale Manager	001	Sale_Manager	2/20/2007 9:32:06 AM	<a href="#">Edit</a>	<a href="#">Delete</a>
Sale01	*****	Sale	01	sale01	2/20/2007 9:31:14 AM	<a href="#">Edit</a>	<a href="#">Delete</a>
Admin	*****	Administrator	1	Admin	2/20/2007 9:32:36 AM	<a href="#">Edit</a>	<a href="#">Delete</a>

### รูปที่ 5.3 การจัดการสิทธิให้กับผู้ใช้ระบบ

#### 1. ผู้ดูแลระบบ

ทำหน้าที่จัดการเพิ่ม , ลบ , แก้ไข รายชื่อพนักงานระดับต่างๆในการเข้าใช้งานระบบและทำการสร้าง USER กับ PASSWORD ให้กับพนักงาน สามารถแก้ไขข้อมูลในบางส่วนได้เพื่อให้สอดคล้องกับความเป็นจริง

#### 2. พนักงานระดับผู้จัดการ

มีสิทธิในการดูข้อมูล , แก้ไข, เพิ่มข้อมูลในบางส่วนเพื่อให้ได้ข้อมูลที่ถูกต้องตามความเป็นจริงและการวิเคราะห์รายละเอียดต่างๆได้เพื่อให้ได้ข้อมูลที่ถูกต้อง แต่ไม่สามารถจัดการกับสิทธิของผู้ใช้งาน

#### 3. พนักงานทั่วไป

ในส่วนนี้เป็นพนักงานที่ได้รับสิทธิทำหน้าที่วิเคราะห์ข้อมูลและสืบค้นข้อมูลได้บางส่วนเท่านั้น

### 5.3 การดูแลรักษาระบบฐานข้อมูล

เพื่อให้ได้ข้อมูลที่ต้องการอยู่เสมอ ดังนั้นจึงต้องเอาใจใส่ในการดูแลฐานข้อมูลอยู่เสมอ การนำข้อมูลเพื่อเตรียมการไปสร้าง Cube นั้นเป็นข้อมูลที่อยู่ใน Data Warehouse โดยผ่านกระบวนการ Transform มาแล้ว เพราะฉะนั้น ข้อมูลส่วนใหญ่ก็มีข้อผิดพลาดไม่มากนักเป็นข้อมูลที่เชื่อถือได้ แต่อย่างไรก็ตามต้องขึ้นอยู่กับกระบวนการและวิธีการในการ Transform ว่ามีขั้นตอนในการนำข้อมูลเข้าสู่ Data Warehouse ได้เป็นอย่างดีหรือไม่ สิ่งสำคัญที่ต้องคำนึงในข้อมูลการขายสินค้า นั่นก็คือเวลาในการจบการทำงานเสร็จสิ้นเมื่อใด ซึ่งการเก็บข้อมูลเวลาจะถูกเปลี่ยนแปลงได้เสมอ ไม่แน่นอน การที่จะเริ่มกระบวนการนำข้อมูลเข้าสู่ Data Warehouse นั้นจะต้องมั่นใจว่าการขายปูนซีเมนต์ประจำวันได้ทำการหยุดระบบทำการขายประจำวันแล้วและข้อมูลการขายสินค้าได้ค่าคงที่ไม่มีเปลี่ยนแปลง

สิ่งที่ต้องคำนึงถึงอย่างมากคือขั้นตอนเวลาในการนำข้อมูลที่อยู่ใน Data Warehouse มาสร้างเป็น Cube นั้นใช้เวลาค่อนข้างมากเนื่องจากการนำค่ามาคำนวณหลายๆค่าและเก็บเอาไว้ ทำให้มีความจำเป็นต้องเริ่มต้นกระบวนการนำข้อมูลสู่ Data Warehouse จนกระทั่งได้ทำการสร้าง Cube เสร็จสมบูรณ์ซึ่งถูกจำกัดโดยช่วงเวลาเปิดทำการถึงช่วงเวลาเปิดทำการในวันใหม่ ซึ่งข้อมูลต่างๆต้องได้ถูกเตรียมพร้อมใช้งานทันทีเมื่อวันใหม่เริ่มต้น

ถึงแม้ว่าจะทำการวางแผนที่ดีในเรื่องการจัดการเรื่องของเวลา เพื่อนำข้อมูลเข้าสู่ระบบแล้วก็ตาม แต่ในบางครั้งอาจเกิดเหตุการณ์ที่ไม่คาดคิดขึ้นได้ ทำให้ต้องมีการดึงข้อมูลเข้าไปในระบบใหม่โดยเริ่มกระบวนการตั้งแต่ต้น ถ้าระบบได้รับการออกแบบไม่ดี อาจทำให้เกิดปัญหาข้อมูลเกิดการซ้ำซ้อนขึ้นจึงควรออกแบบ Package ที่ใช้ในการ Transform ให้มีความสามารถทำงานซ้ำซ้อนได้ หรือ อาจมีการลบข้อมูลเก่าก่อนแล้วจึงเพิ่มข้อมูลใหม่ลงไป โดยมีการตรวจสอบจากข้อมูลในส่วนต่างๆ ที่ต้องการอย่างรอบคอบเพื่อไม่ให้เสียเวลาในการให้บริการของระบบ

## บทที่ 6

# สรุปผลการศึกษา และข้อเสนอแนะ

โครงการพัฒนาระบบงานสำหรับการวิเคราะห์ข้อมูลการขายปูนซีเมนต์ ได้จัดทำขึ้น เพื่อให้พนักงานขายหรือผู้ที่เกี่ยวข้องสามารถนำข้อมูลที่ถูกรวบรวมในรูปแบบหลายมิติ มาใช้งานในการวิเคราะห์ข้อมูล เพื่อนำข้อมูลเหล่านั้นมาใช้ประโยชน์ในการเพิ่มความสามารถและกลยุทธ์การขายต่อไป

### 6.1 ผลที่ได้จากการพัฒนาระบบ

เมื่อทำการพัฒนาโครงการพัฒนาระบบงานสำหรับการวิเคราะห์ข้อมูลการขายปูนซีเมนต์เป็นที่เรียบร้อยและนำไปทดลองใช้งานจริงผลที่ได้มีดังนี้

1. สามารถเรียนรู้ถึงหลักการวิเคราะห์และการออกแบบฐานข้อมูล รวมถึงการนำระบบฐานข้อมูลที่ได้เข้าใช้งาน ได้อย่างมีประสิทธิภาพ
2. สามารถเรียนรู้ถึงการใช้เครื่องมือในการออกแบบ Cube
3. เข้าใจการออกแบบและการทำงานในระบบคลังข้อมูล
4. สามารถเลือกเครื่องมือและซอฟต์แวร์ที่เหมาะสมกับระบบงาน และยังรู้ถึงข้อดีข้อเสียของเครื่องมือต่างๆ
5. สามารถแก้ไขปัญหาและในเรื่องของการใช้โปรแกรมและพัฒนางานต่อไป
6. ระบบนี้สามารถรองรับและตอบสนองกับผู้ใช้งาน ได้ระดับหนึ่ง

### 6.2 ข้อเสนอแนะ

สำหรับการพัฒนาระบบงานสำหรับการวิเคราะห์ข้อมูลการขายปูนซีเมนต์นั้น เมื่อได้ลงมือปฏิบัติตามขั้นตอนต่างในการพัฒนาแล้วได้ประสบกับปัญหาหลายอย่าง มีทั้งแก้ไขได้และแก้ไขไม่ได้จึงยกตัวอย่างบางข้อ มีดังนี้

1. การใช้เครื่องมือในการสร้าง ไม่มีเครื่องมือใดที่ให้ผลและตรงกับงานได้ 100 %
2. การใช้งานและการจัดฐานข้อมูลของเวอร์ชันนี้ใช้ทรัพยากรสูง
3. เครื่องมือที่ใช้ในการพัฒนาหลากหลายต้องเสียเวลาในการศึกษา
4. หาข้อมูลในการศึกษายากเพราะเป็นผลิตภัณฑ์ใหม่ ต้องใช้ระยะเวลามาก
5. เครื่องมือที่ใช้พัฒนามีการเก็บ Backup File ของในแต่ละงานที่ใช้ในการพัฒนาไว้หลายแห่งทำให้ต้องตามไปลบ เพื่อไม่ให้เกิดความสับสน

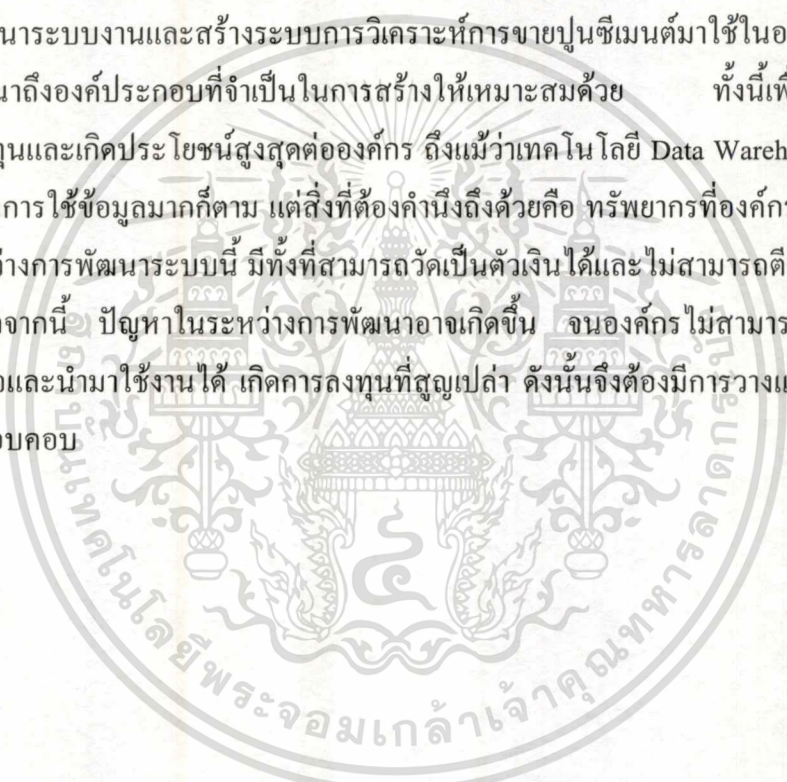
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 6.3 สรุปผลดำเนินการพัฒนา

โครงการนี้ได้นำฐานข้อมูลแบบหลายมิติ มาใช้ในการพัฒนาทั้งหมด เพราะเป็นฐานข้อมูลที่สามารถเข้าใจและใช้งานง่าย การดำเนินงานสร้างโดยใช้เครื่องมือที่เรียกว่า MS-SQL 2005 Analysis Service เป็นตัวจัดการข้อมูลในฐานข้อมูลนี้ การพัฒนา Cube ของข้อมูลแบบหลายมิตินี้สามารถทำได้โดย MS-Visual Studio 2005 เป็นตัวพัฒนา โปรแกรมดังกล่าวติดมากับการลงโปรแกรม MS-SQL 2005 Server เพื่อให้ผู้ใช้ได้ใช้งานง่ายยิ่งขึ้น

ส่วนเรื่องการพัฒนา Data Warehouse เป็นการรวบรวมข้อมูลจากฐานข้อมูลของระบบการขายประจำวันขององค์กร แล้วนำมาแปลงข้อมูลให้อยู่ในรูปแบบที่เหมาะสมในการเก็บและสะดวกในการใช้งาน แล้วจึงนำข้อมูลนั้นเข้าไปเก็บในคลังข้อมูล

การพัฒนาระบบงานและสร้างระบบการวิเคราะห์การขายปูนซีเมนต์มาใช้ในองค์กร จะต้องมีการพิจารณาถึงองค์ประกอบที่จำเป็นในการสร้างให้เหมาะสมด้วย ทั้งนี้เพื่อเกิดความคุ้มค่าในการลงทุนและเกิดประโยชน์สูงสุดต่อองค์กร ถึงแม้ว่าเทคโนโลยี Data Warehouse จะให้ประสิทธิภาพในการใช้ข้อมูลมากก็ตาม แต่สิ่งที่ต้องคำนึงถึงด้วยคือ ทรัพยากรที่องค์กรจะต้องทุ่มเทลงไปในระหว่างการพัฒนาแบบนี้ มีทั้งที่สามารถวัดเป็นตัวเงินได้และไม่สามารถตีค่าออกเป็นตัวเงินได้ นอกจากนี้ ปัญหาในระหว่างการพัฒนาอาจเกิดขึ้น จนองค์กรไม่สามารถจะพัฒนาระบบนี้จนสำเร็จและนำมาใช้งานได้ เกิดการลงทุนที่สูงเกินไป ดังนั้นจึงต้องมีการวางแผนควบคุมและจัดการให้รอบคอบ



## บรรณานุกรม

กิตติพงษ์ กลมกล่อม. 2546. การออกแบบและพัฒนาค้างข้อมูล. กรุงเทพฯ: เคทีพี คอมพ์ แอนด์ คอนซัลท์.

พงษ์พันธ์ ศิวิลัย. 2549. Server 2005 ฉบับสมบูรณ์. กรุงเทพฯ: ซีเอ็ดดูเคชั่น.

พร้อมเลิศ หล่อวิจิตร. 2549. Visual Basic 2005. กรุงเทพฯ: โปรวิชั่น.

ศุภชัย สมพานิช. 2549. คู่มือ ASP.NET 2.0 ฉบับสมบูรณ์. นนทบุรี : ไอดีซีฯ.

Robert M.Elliott. 1998. The Data warehouse lifecycle toolkit. New York: Ralph Kimball AssociatesInc.



## ภาคผนวก

การเข้าสู่โปรแกรมนี้สามารถใช้ภายใต้ Internet Explorer , Netscape และ Mozilla Firefox โดยโปรแกรมจะแบ่งการทำงาน โดยปรากฏหน้าจอเริ่มต้นเพื่อเลือกการใช้งานตามคุณสมบัติหรือสิทธิของผู้ใช้งาน โดยการเข้าใช้งานแบ่งเป็น 3 สิทธิดังนี้

### 1. การเข้าใช้งานของผู้ดูแลระบบ (Administrator)

- 1.1 การสืบค้นและจัดการข้อมูลผู้ใช้งานในระบบ
- 1.2 การจัดการข้อมูลภายในฐานข้อมูล
- 1.3 Cube Browsers

### 2. การเข้าใช้งานของระดับผู้จัดการ (Sale Manager)

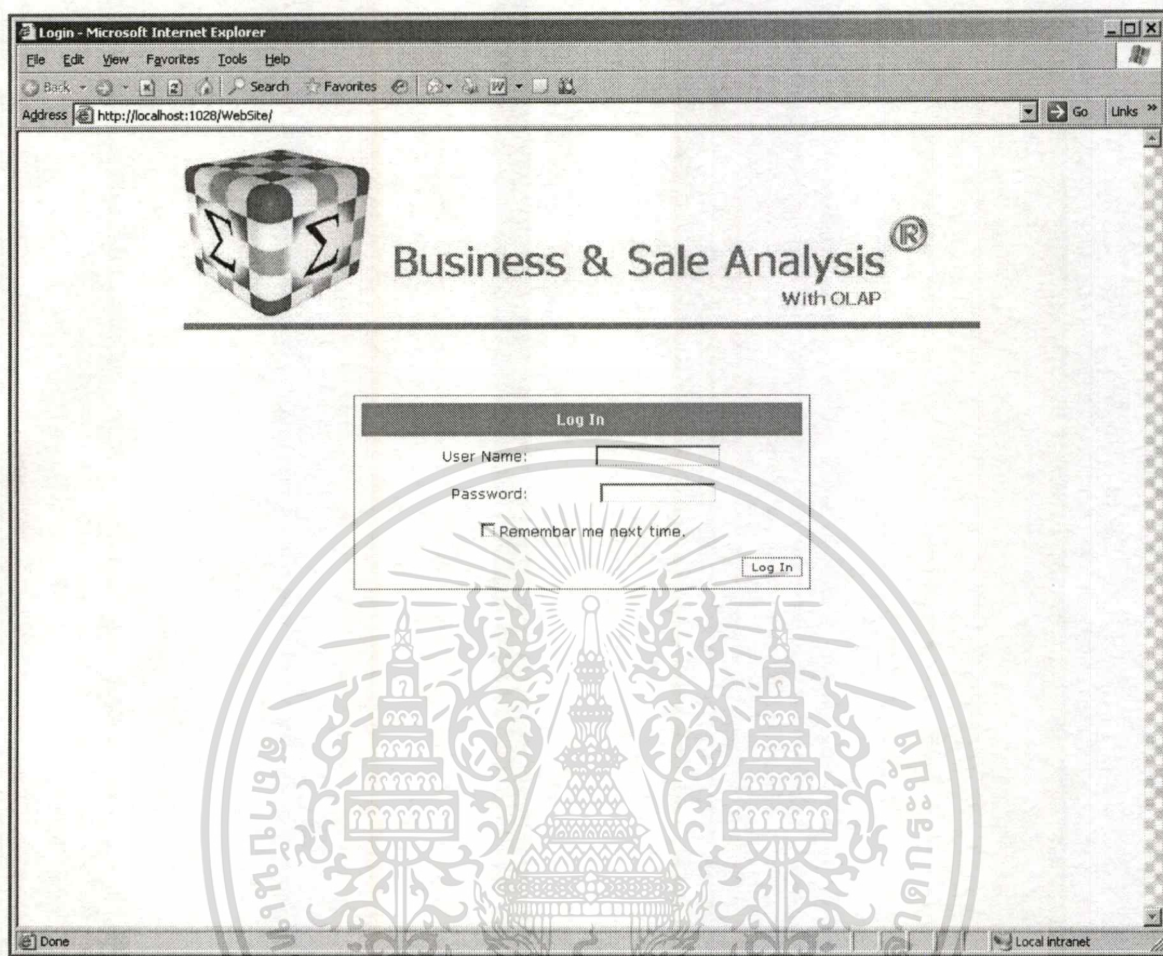
- 2.1 การสืบค้นและจัดการข้อมูลภายในฐานข้อมูล
- 2.2 การเปิดดูรายงาน
- 2.3 Cube Browsers

### 3. การเข้าใช้งานของระดับของพนักงานขาย (Sale)

- 3.1 การเปิดดูรายงาน
- 3.2 Cube Browsers



## หน้าจอเมื่อเข้าสู่ระบบ

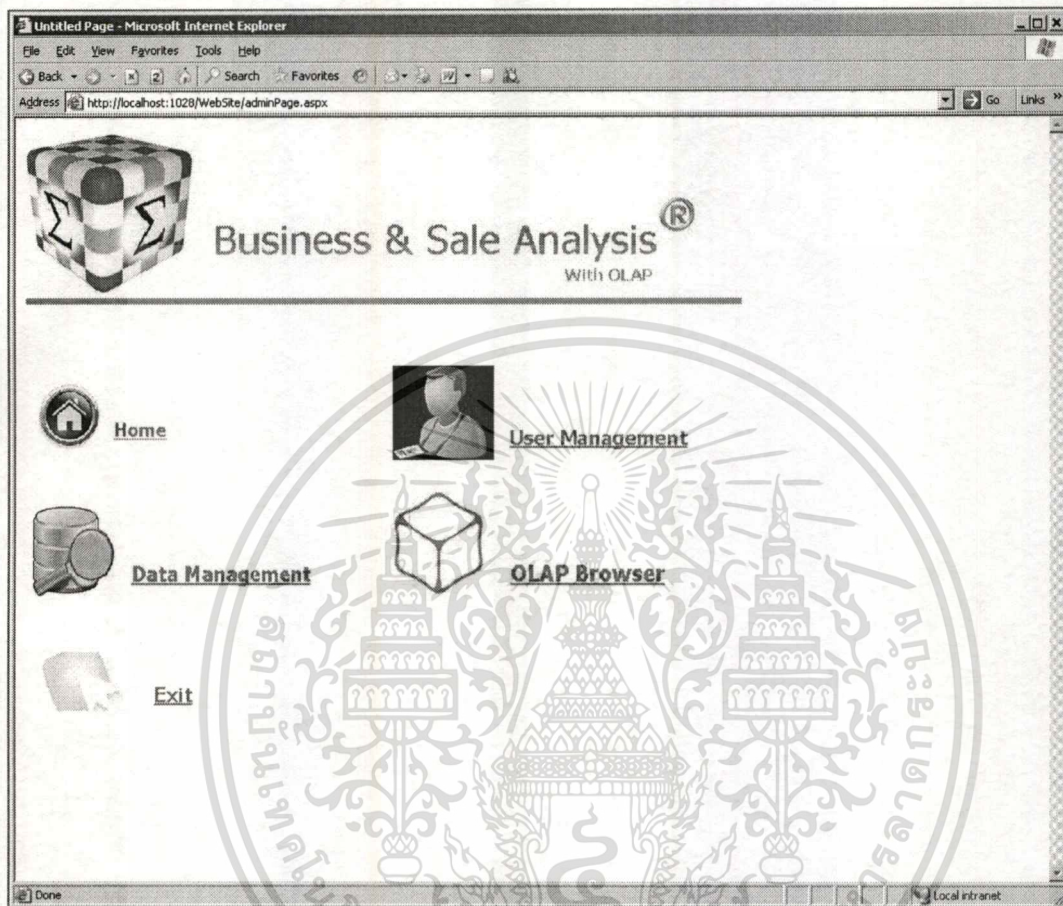


หน้าแรกของการเข้าสู่ระบบ โปรแกรมจะแบ่งการใช้งานตามสิทธิในการเข้าใช้ระบบ โดยผู้มีสิทธิในระบบเท่านั้นจึงจะมีสิทธิเข้าถึงข้อมูลต่างๆ ตามระดับที่ได้รับสิทธิ ตามที่ผู้ดูแลระบบได้อนุญาตให้สิทธิดังกล่าว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 1. การใช้งานของผู้ดูแลระบบ (Administrator)

การใช้งานในระดับนี้จะต้องมีสิทธิได้รับอนุญาตเป็นผู้ดูแลระบบเท่านั้นเมื่อทำการกรอก Username และ Password แล้วจะเข้าถึงหน้าจอของเมนู



ซึ่งหน้าจอนี้จะประกอบด้วย

- 1.1 การสืบค้นและจัดการข้อมูลผู้ใช้งานในระบบ
- 1.2 การจัดการข้อมูลภายในฐานข้อมูล
- 1.3 Cube Browsers

## 1.1 การสืบค้นและจัดการข้อมูลผู้ใช้งานในระบบ

เมื่อเลือก ปุ่ม User Management จากเมนู เว็บจะเปลี่ยนหน้าไปยัง หน้า User Management มีหน้าที่ในการจัดการ ผู้ใช้งานของระบบ ซึ่งถ้าต้องการเพิ่มรายชื่อผู้ใช้งานให้กรอกข้อมูลครบถ้วน โดยมีข้อมูล Username , Password , Position , EmployeeID , Description โดยการกำหนดสิทธิ์ในการใช้งานนั้นเป็นข้อมูลในส่วนของ Position ซึ่งแบ่งได้เป็น 3 ระดับ คือ Administrator , Sale Manager และ Sale หลังจากนั้น กดปุ่ม Add เป็นอันว่าได้เพิ่มข้อมูลลงในฐานข้อมูลเรียบร้อยแล้ว

ในส่วนของการแก้ไขข้อมูลของผู้มีสิทธิ์ใช้งานนั้นสามารถแก้ไขได้ ในตารางช่อง Edit และสามารถลบข้อมูลได้ในช่องตาราง Delete ได้ทันที

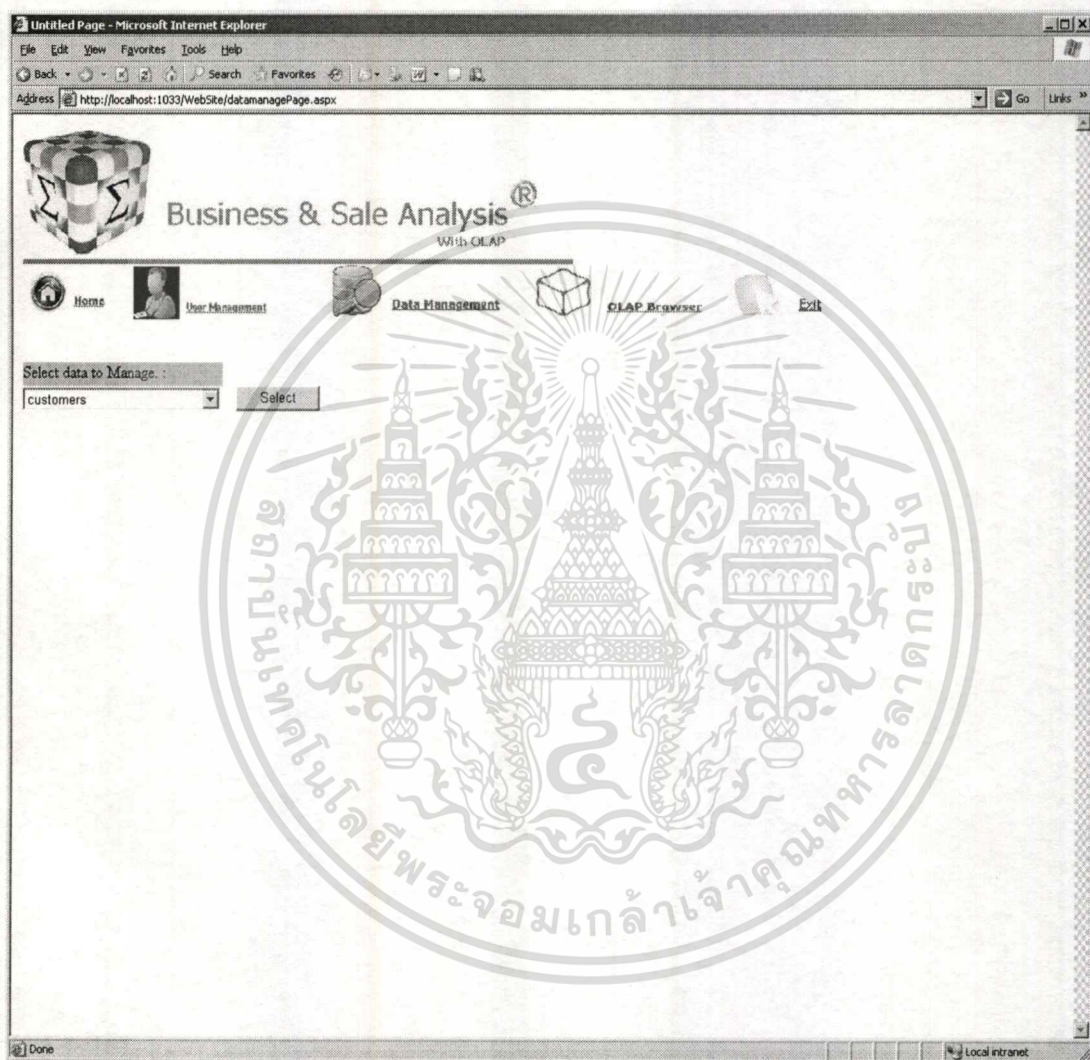
The screenshot shows a web browser window displaying a user management interface. The page title is 'Business & Sale Analysis With OLAP'. The navigation menu includes Home, User Management, Data Management, OLAP Browser, and Exit. The form contains fields for Username, Password, Position (a dropdown menu currently set to 'Sale'), Employee ID, and Description. Below the form are 'Add' and 'Clear' buttons. At the bottom, there is a table listing existing users with columns for Username, Password, Position, empID, Description, Create Date, Edit, and Delete.

UserName	Password	Position	empID	Description	Create Date	Edit	Delete
MSale	*****	Sale Manager	001	Sale_Manager	2/20/2007 9:32:06 AM	<a href="#">Edit</a>	<a href="#">Delete</a>
Sale01	*****	Sale	01	sale01	2/20/2007 9:31:14 AM	<a href="#">Edit</a>	<a href="#">Delete</a>
Admin	*****	Administrator	1	Admin	2/20/2007 9:32:36 AM	<a href="#">Edit</a>	<a href="#">Delete</a>

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 1.2 การจัดการข้อมูลภายในฐานข้อมูล

เมื่อเลือก ปุ่ม Data Management จากเมนูเว็บจะเปลี่ยนหน้าไปยัง หน้า Data Management ซึ่งมีหน้าที่ในการจัดการข้อมูลต่างๆในฐานข้อมูล โดยมีช่องให้เลือกในการจัดการกับรายละเอียดของข้อมูลต่างๆ



เมื่อทำการเลือกข้อมูลในช่องและกดปุ่ม จะแสดงข้อมูลในการจัดการ เช่น ถ้าเลือก Customers และกด Select เว็บจะเพิ่มข้อมูลการจัดการข้อมูลของ Customers เพิ่มขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับกรใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ซึ่งถ้าต้องการเพิ่มข้อมูล Customers ให้กรอกข้อมูลครบถ้วน โดยมีข้อมูล Customer ID, Company Name, Contact Name, Address, City, Country และ Post Code กดปุ่ม Add ข้อมูลจะถูกบันทึกลงในฐานข้อมูล Search, Edit, Delete ทำได้จากการใส่ Customer ID ลงในช่องค้นหา และกดปุ่ม Search ข้อมูลจะถูกแสดงออกมาในรูปแบบของตาราง สามารถการแก้ไขข้อมูลในตารางช่อง Edit และสามารถลบข้อมูลได้ในช่องตาราง Delete

Input Customer ID :

0100015

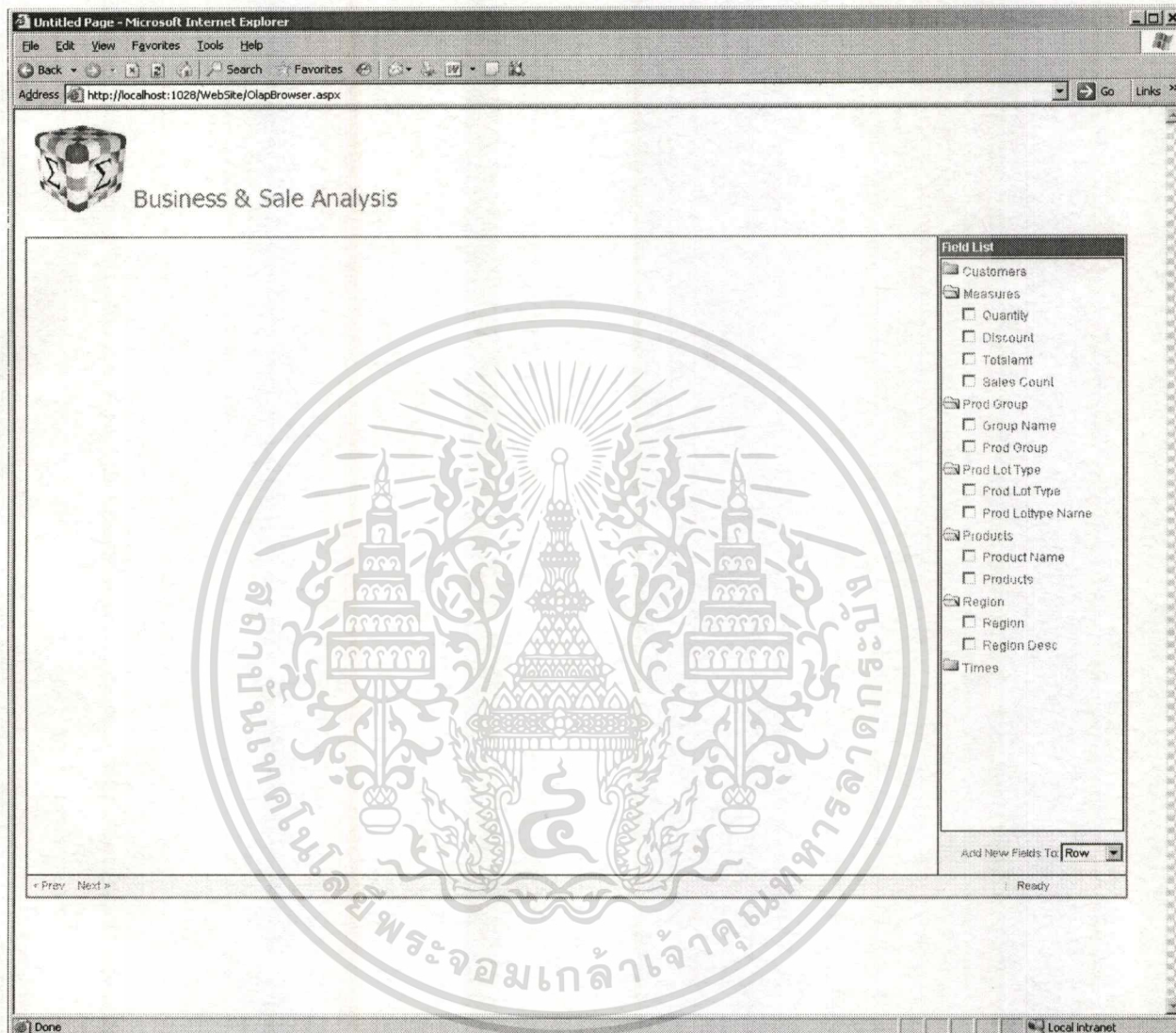
Search

CustID	CompName	ContactName	Address	City	Country	PostCode	Edit	Delete
0100015	พันแสงขนชัย จำกัด	คุณพรทิพย์ อิศาสตร์	191 ซ.สุขุมวิท 103/1	191 ซ.สุขุมวิท 103/1	กรุงเทพฯ	THAIL	Edit	Delete

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 1.3 Cube Browsers

เมื่อเลือก ปุ่ม OLAP Browser จากเมนูเว็บจะเปลี่ยนหน้าไปยัง หน้า OLAP Browser ซึ่ง ตัว Browser Cube จะทำการ ติดต่อกับฐานข้อมูลและโหลดตัวเองขึ้นมา



โดยตัว Browser จะถูกแบ่งเป็น 2 ส่วน คือ Display (ขวามือ) และ Field List (ซ้ายมือ) ซึ่ง Display จะแสดงข้อมูลที่ถูกเลือก ส่วน Field List คือข้อมูล ของ Cube ซึ่งประกอบด้วย Dimension และ Measures

การใช้ Browser ในการสร้างความสัมพันธ์ของข้อมูลนั้นเริ่มจากการเลือกข้อมูลโดยคลิกที่ช่องว่างของแต่ละฟิลด์ การกำหนดตำแหน่งของฟิลด์ต่างๆใช้ปุ่ม List Box ข้างล่างเป็นตัวกำหนด โดยมี 3 ตำแหน่งคือ Row , Column และ Filter โดยมีตัวอย่างดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Product Name	2004		2005	
	Quantity	Discount	Discount	Quantity
All	9154642	1.807313E+08	1.296822E+09	5684587
MORTAR DUST	53.76	0	0	135.73
1 1/2-3/4"				
1 1/2-3/8"				
1 1/4-3/4"				
1/2-3/8"			0	956.46
1-0" wrap				
1-7/8"				
2 1/2-3/4"				
2-1 1/4"				
3/16-0"	895438.4	0	0	736985.9
3/4-3/8"	3676.2	7450		
CLINKER TYPE 1				
CLINKER TYPE 5	28254.5	0	0	160
M299 อุตสาหกรรม				
MORTAR DUST 900				
Portland Type 5-Bag				
Portland Type 5-Bulk				
Portland Type I				
TPI เขียว-สูง 1 ตัน	1470779	2.584733E+07	2.00250E+07	781109.9
TPI เขียว-สูง 50 กก.			0	422
TPI เขียว-สูงพิเศษ 50 กก.	8459.711	177650	96350	4261.15
TPI เขียว-สูง				
TPI เหลือง-สูง 1 ตัน				
TPI เหลือง-สูง 50 กก.				
TPI เหลือง-สูงพิเศษ 50...				
TPI เหลือง-สูง	11981.45	233350	144950	6290.78
TPI แดง-สูง 1 ตัน	69022.37	0	0	23304.72
TPI แดง-สูง 1.5 ตัน				
TPI แดง-สูง 2 ตัน				

**Field List**

- Customers
- Measures
  - Quantity
  - Discount
  - Totalamt
  - Sales Count
- Prod Group
  - Group Name
  - Prod Group
- Prod Lol Type
- Products
  - Product Name
  - Products
- Region
  - Region
  - Region Desc
- Times
  - Smonth
  - Squarter
  - Syear
  - Times

Add New Fields To: **Row**

การกำหนดค่าต่างๆแสดงผล คลิกขวาในส่วนของลูกศร ตารางจะแสดงเมนูให้เลือกในการจัดการกับการแสดงค่า เช่น ถ้าเลือก Show Totals ในตารางจะมีการแสดงข้อมูลรวมของค่าทั้งหมด

รูปแบบของการแสดงในรูปตัวเลขและกราฟ รวมถึงการเรียงค่าในการแสดงผล สามารถทำได้โดยการคลิกขวาที่ หัว Column ตัว Browser จะแสดงเมนูลงมาดังรูป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Year		2004		2005	
Product Name	Quantity	Discount	Quantity	Discount	
All	53.76		08	5684587	1.296622E+08
<MORTAR DUST>			0	135.73	0
1 1/2-3/4"					
1 1/2-3/8"					
1 1/4-3/4"					
1/2-3/16"			956.46		0

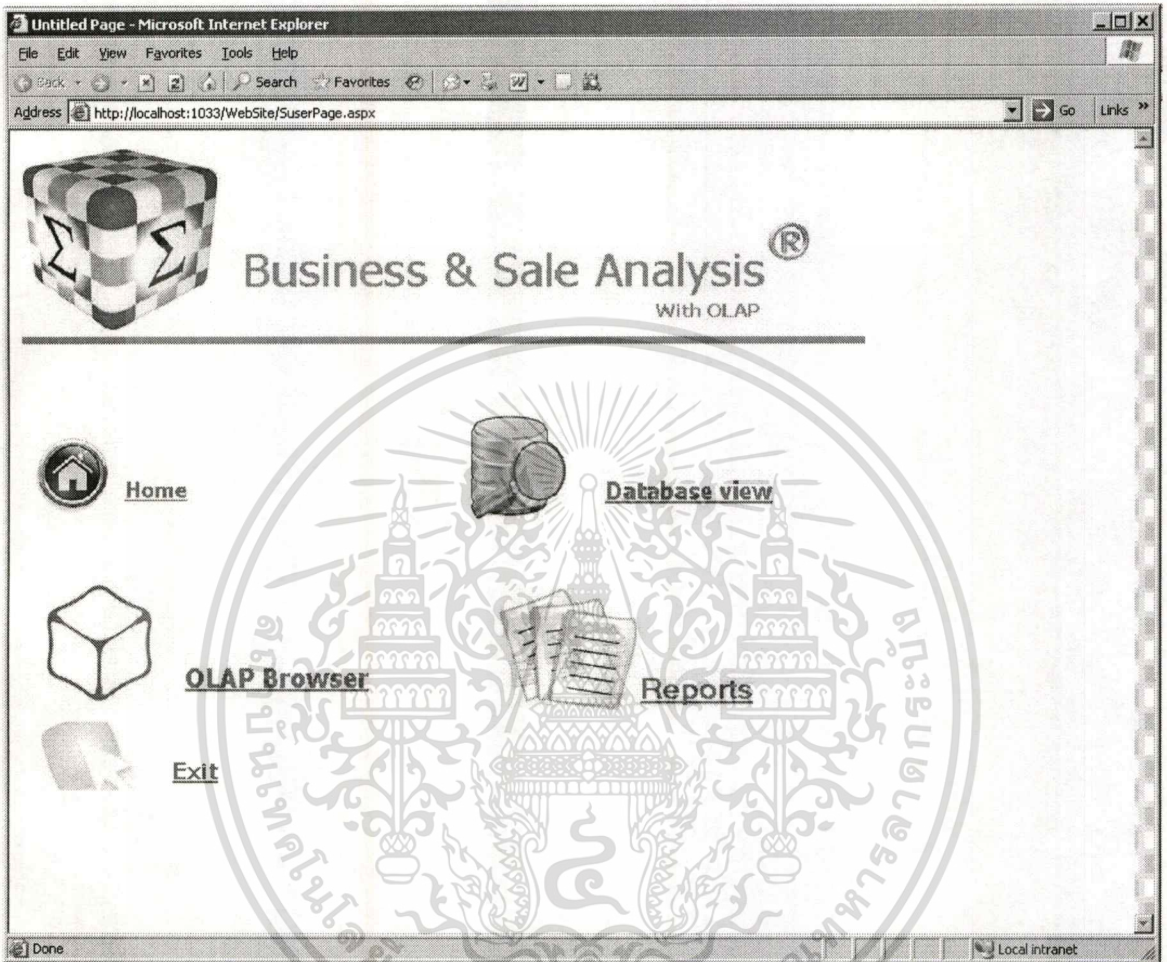
การพิมพ์ค่าออกทางเครื่องพิมพ์ ทำโดยคลิกขวาที่ บาร์ข้างล่าง จะแสดงเมนูให้เลือก Printer และเลือกเครื่องพิมพ์ ซึ่งเราสามารถนำค่าออกไปยัง โปรแกรม Excel ได้ด้วย

3/4-3/8"	Back
CLINKER TYPE 1	Forward
CLINKER TYPE 5	Save Background As...
M299 ดุงพระอาทิตย์	Set as Background
MORTAR DUST 800	Copy Background
Portland Type 5-Bag	Set as Desktop Item...
Portland Type 5-Bul	Select All
Portland Type I	Paste
TPI เขียว-ดุง 1 ตัน	Create Shortcut
TPI เขียว-ดุง 50 กก.	Add to Favorites...
TPI เขียว-ดุงพิเศษ 50	View Source
TPI เขียว-ดุง	Encoding
TPI เหลือง-ดุง 1 ตัน	Print
TPI เหลือง-ดุง 50 กก.	Refresh
TPI เหลือง-ดุงพิเศษ 5	Export to Microsoft Excel...
TPI เหลือง-ดุง	Properties
TPI แดง-ดุง 1 ตัน	
TPI แดง-ดุง 1.5 ตัน	
TPI แดง-ดุง 2 ตัน	
< Prev Next >	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 2. การใช้งานของระดับผู้จัดการ (Sale Manager)

การใช้งานในระดับนี้จะต้องมีสิทธิได้รับอนุญาตเป็นผู้ดูแลระบบเท่านั้น เมื่อทำการกรอก Username และ Password แล้วจะเข้าถึงหน้าจอของเมนู



2.1 การสืบค้นและจัดการข้อมูลภายในฐานข้อมูล

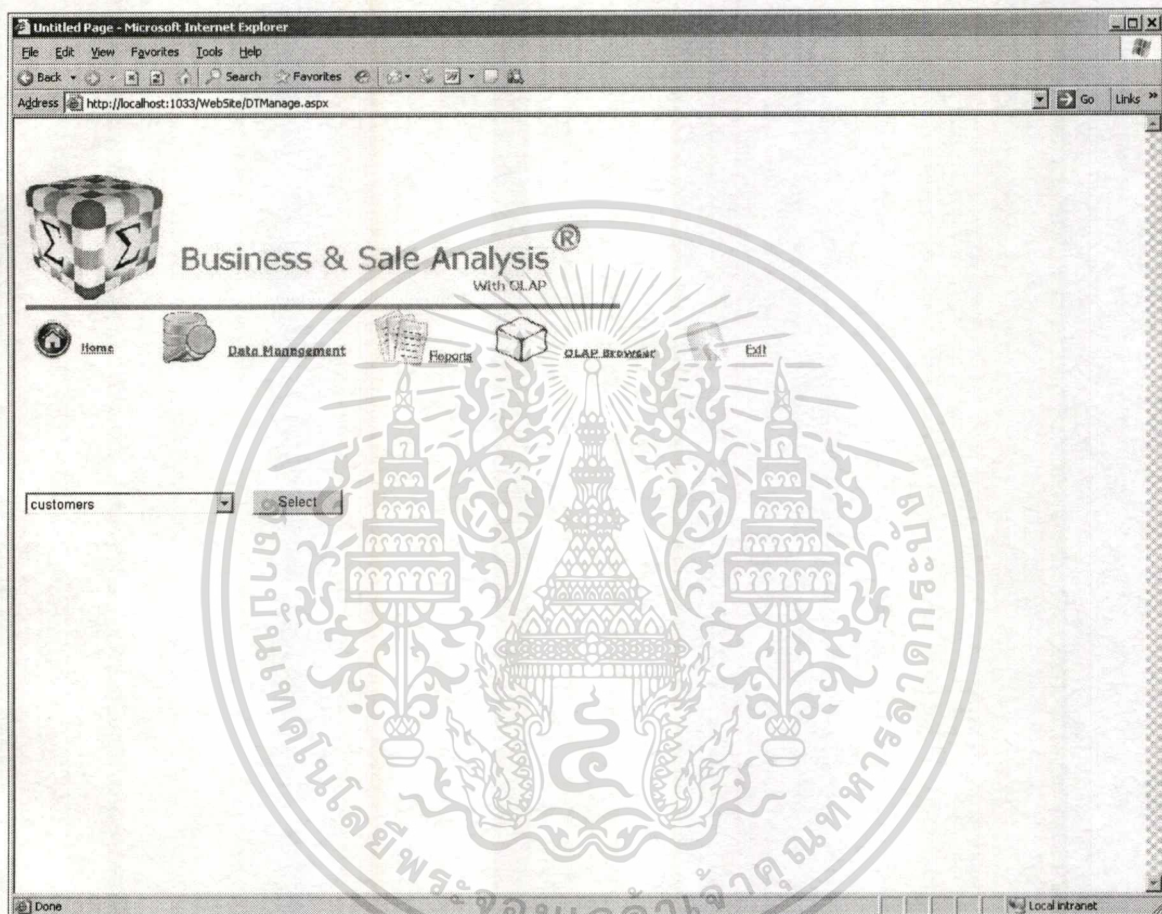
2.2 การเปิดดูรายงาน

2.3 Cube Browsers

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 2.1 การสืบค้นและจัดการข้อมูลภายในฐานข้อมูล

เมื่อเลือก ปุ่ม Data Management จากเมนูเว็บจะเปลี่ยนหน้าไปยัง หน้า Data Management ซึ่งมีหน้าที่ในการจัดการข้อมูลต่างๆในฐานข้อมูล โดยมีช่องให้เลือกในการจัดการกับรายละเอียดของข้อมูลต่างๆ



เมื่อทำการเลือกข้อมูลในช่องและกดปุ่ม จะแสดงข้อมูลในการจัดการ เช่น ถ้าเลือก Customers และกด Select เว็บ จะเพิ่มข้อมูลการจัดการข้อมูลของ Customers เพิ่มขึ้น

Search ทำได้จากการใส่ Customer ID ลงในช่องค้นหา และกดปุ่ม Search ข้อมูลจะถูกแสดงออกมาในรูปแบบของตาราง ซึ่งในที่นี้ต่างจากสิทธิของ Administrator เพราะในการแสดงผลไม่สามารถแก้ไขข้อมูล ในตารางได้

Input Customer ID :

0100015

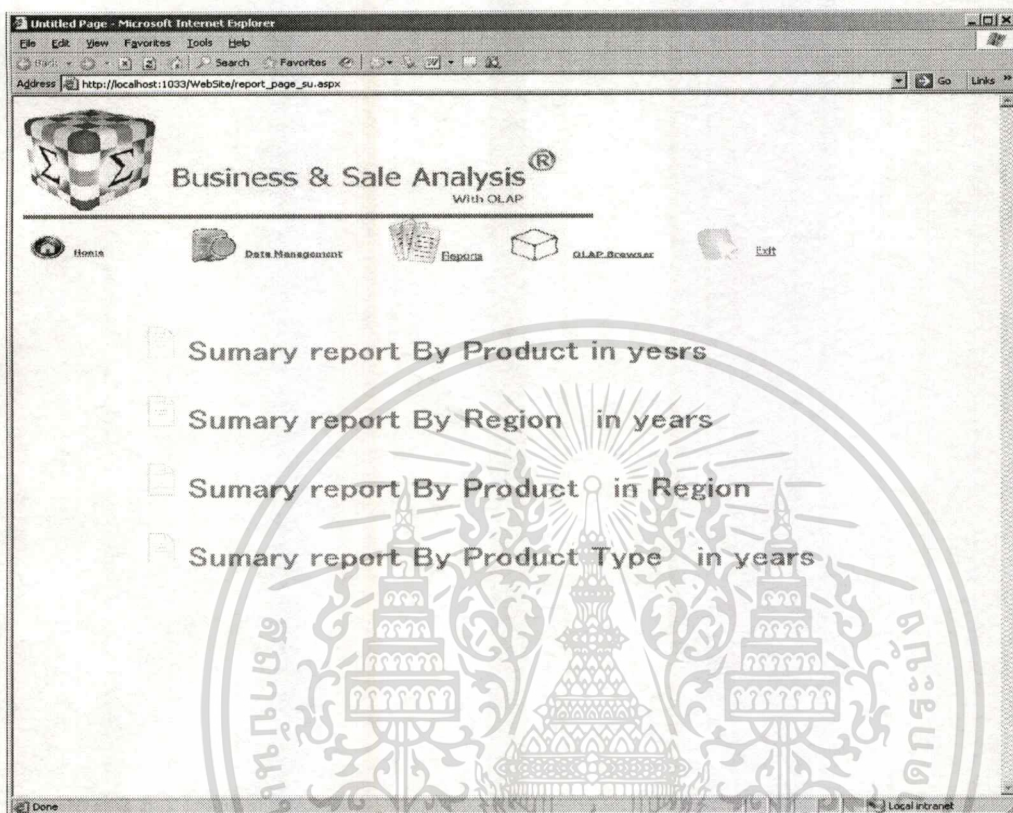
Search

CustID	CompName	ContactName	Address	City	Country	PostCode
0100015	พันแสงขนชัย จำกัด	คุณพรทิพย์ อิศาสตร์	191 ซ.สุขุมวิท 103/1	191 ซ.สุขุมวิท 103/1	กรุงเทพฯ	THAIL

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 2.2 การเปิดดูรายงาน

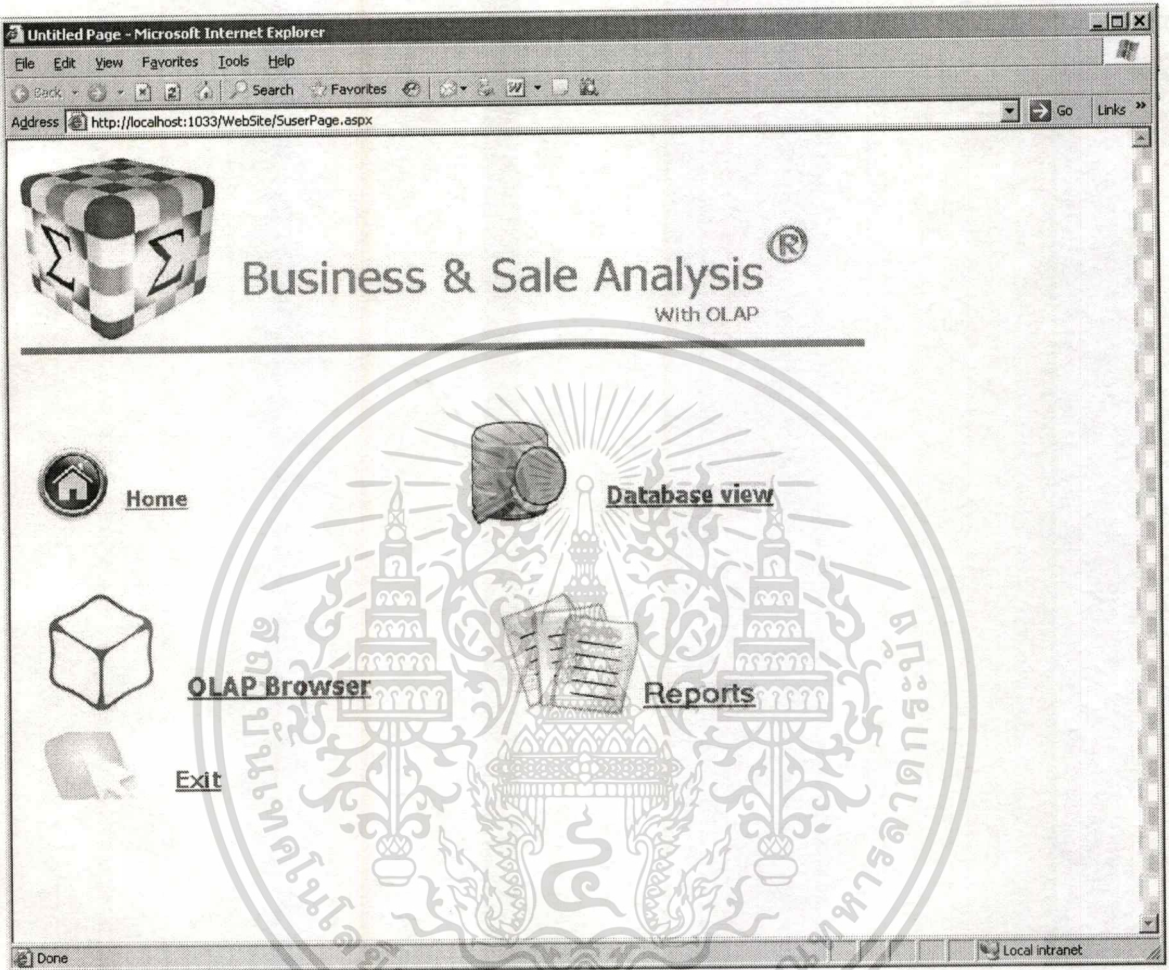
เมื่อเลือก ปุ่ม Reports จากเมนูจะเปลี่ยนหน้าไปยังหน้า Reports ซึ่งมีหน้าที่ในการแสดง Reports ต่างๆ ที่ได้จัดทำไว้ในรูปของ Reports Service



เมื่อทำการเลือกคลิกที่ Reports จะแสดง Reports ดังตัวอย่างต่อไปนี้

### 3. การใช้งานของระดับของพนักงานขาย (Sale)

การใช้งานในระดับนี้จะต้องมีสิทธิได้รับอนุญาตเป็นผู้ดูแลระบบเท่านั้นเมื่อทำการกรอก Username และ Password แล้วจะเข้าถึงหน้าจอของเมนู ซึ่งคล้ายกับของ Sale Manager



#### 3.1 การสืบค้นและจัดการข้อมูลภายในฐานข้อมูล

การใช้ คล้ายกับเว็บ ต่างกันเพียงไม่สามารถ ค้นหาและเปิดข้อมูลบางอย่างได้เช่น ข้อมูลพนักงาน

#### 3.2 การเปิดดูรายงาน

วิธีการใช้งานได้กล่าวมาแล้ว

#### 3.3 Cube Browsers

วิธีการใช้งานได้กล่าวมาแล้ว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้