การพัฒนาเว็บแอปพลิเคชันเพื่อการวิเคราะห์ข้อมูลปริมาณขยะมูลฝอยในประเทศไทย

นางสาวกมลทิพย์ วงค์อ้าย รหัสนักศึกษา 61521207084-6
นางสาวนาฏติยา สุริยะมณี รหัสนักศึกษา 61521207094-5

หลักสูตรระบบสารสนเทศทางธุรกิจ
สาขาบริหารธุรกิจ คณะบริหารธุรกิจและศิลปศาสตร์
มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา
ปีการศึกษา 2563

การพัฒนาเว็บแอปพลิเคชันเพื่อการวิเคราะห์ข้อมูล ปริมาณขยะมูลฝอยในประเทศไทย

นางสาวกมลทิพย์ วงค์อ้าย รหัสนักศึกษา 61521207084-6
นางสาวนาฏติยา สุริยะมณี รหัสนักศึกษา 61521207094-5

หลักสูตรระบบสารสนเทศทางธุรกิจ
สาขาบริหารธุรกิจ คณะบริหารธุรกิจและศิลปศาสตร์
มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา
ปีการศึกษา 2563

ชื่อโครงงาน การพัฒนาเว็บแอปพลิเคชันเพื่อการวิเคราะห์ข้อมูลปริมาณขยะมูลฝอยในประเทศไทย

โดย นางสาวกมลทิพย์ วงค์อ้าย รหัสนักศึกษา 61521207084-6
นางสาวนาฏติยา สุริยะมณี รหัสนักศึกษา 61521207094-5

อาจารย์ที่ปรึกษา ผู้ช่วยศาสตราจารย์ ดร. ชัชฎาพร ปุกแก้ว

หลักสูตร ระบบสารสนเทศทางธุรกิจ สาขาบริหารธุรกิจ
คณะบริหารธุรกิจและศิลปศาสตร์

ปีการศึกษา 2563

บทคัดย่อ

การจัดทำโครงงานในครั้งนี้มีวัตถุประสงค์ เพื่อการวิเคราะห์ข้อมูลปริมาณขยะมูลฝอยในประเทศไทยโดยนำข้อมูลมาจากกรมควบคุมมลพิษเพื่อาเพื่อเผยแพร่ข้อมูลบนเว็บไซต์มาวิเคราะห์และพัฒนาเป็นข้อมูลสารสนเทศการรายงานผลการแนะแนวแนวทางในการแก้ไขปัญหาในพื้นที่ได้อย่างถูกต้อง แม่นยำและรวดเร็วเพื่อให้เกิดการแก้ไขปัญหาอย่างทันเหตุการณ์ โดยการนำเอากระบวนการหลักในการจัดทำ เหมืองข้อมูลเพื่อวิเคราะห์ CRISP-DM หรือ (Cross Industry Standard Process for Data Mining)) ดำเนินการตั้งแต่การวิเคราะห์ปัญหาของข้อมูล จัดการกับรวบรวมข้อมูล การแปลงข้อมูลให้อยู่ในรูปที่เหมาะแก่การวิเคราะห์ข้อมูล การวัดประสิทธิภาพของผลลัพธ์ที่ได้ และนำเทคโนโลยีสารสนเทศเข้ามาช่วยสนับสนุนกระบวนการในการนำผลการวิเคราะห์ข้อมูลที่ได้สร้างเป็นสารสเทศนำไปสู่การรายงานผลทางเว็บไซต์ทำให้สามารถเข้าถึงข้อมูลได้อย่างเหมาะสมและรวดเร็ว

โดยข้อมูลที่ทำการวิเคราะห์นั้นได้ใช้เครื่องมือและโปรแกรมที่เหมาะสมกับการจัดการข้อมูลและวิเคราะห์ข้อมูลด้วยเทคนิคทางเหมืองข้อมูลการวิเคราะห์อนุกรมเวลา (Time Series Analysis) นำผลที่ได้แสดงโดยโปรแกรม Tableau public นำเสนอผ่านเว็บไซต์ซึ่งมีภาษาในการพัฒนา และจัดรูปแบบเว็บไซต์ที่นำมาใช้ได้แก่ ภาษาเอชทีเอ็มแอล (HTML) พีเอชพี (PHP) จาวาสคริปต์ (JaveScript) และซีเอสเอส (CSS3) ในการแสดงสารสนเทศข้อมูลปริมาณขยะมูลฝอยบนเว็บไซต์ที่จัดทำขึ้น

กิตติกรรมประกาศ

โครงงานครั้งนี้ สำเร็จลุล่วงไปได้ด้วยดีเนื่องจากได้รับความกรุณาเป็นอย่างดียิ่งจาก ผู้ช่วยศาสตราจารย์ดร. ชัชฎาพร ปุกแก้ว อาจารย์ที่ปรึกษาโครงงาน อาจารย์สุพงศ์ แดงสุริยศรี หัวหน้าสาขาวิชาระบบสารสนเทศทางธุรกิจ และคณาจารย์ในสาขาระบบสารสนเทศทางธุรกิจ คณะบริหารธุรกิจและศิลปศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา ทุกท่านที่ให้คำปรึกษา และคำแนะนำในการจัดทำโครงงานเพื่อวัตถุประสงค์ของโครงงาน ให้เสร็จสมบูรณ์รวมถึงช่วยแก้ไขปัญหาต่าง ๆ ที่เกิดขึ้นระหว่างการพัฒนาโครงงานจนบรรลุวัตถุประสงค์ของโครงงานด้วยความสมบูรณ์มากยิ่งขึ้น

ขอขอบคุณกรมควบคุมมวลพิษ หรือกระทรวงทรัพยากรธรรมชาติและสิ่งแวดล้อม ที่ได้แสดงรายละเอียดข้อมูลของปริมาณขยะในแต่ละปีให้รู้ถึงปริมาณที่เพิ่มขึ้นและลดลงในแต่ละปี จึงนำข้อมูลทั้งหมดมาวิเคราะห์ข้อมูล ที่จำเป็นต่อการพัฒนาโครงงานในครั้งนี้ ซึ่งเป็นส่วนที่สำคัญในการพัฒนาโครงงาน จนสำเร็จลุล่วงได้ด้วยดี

ขอขอบพระคุณคณาจารย์ประจำวิชาสาขาระบบสารสนเทศทางธุรกิจ คณะบริหารธุรกิจและศิลปศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา ทุกท่านที่กรุณาประสิทธิ์ประสาทวิชาความรู้และประสบการณ์อันมีค่าแก่ศิษย์ และค่อยชี้แนะแนวทางการจัดทำโครงงานพร้อมให้คำปรึกษา คำแนะนำ มอบข้อคิดเห็นที่เป็นประโยชน์ในการดำเนินโครงงานนี้ อันให้ส่งผลลุล่วงไปด้วยดี ขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ ที่นี่

ขอกราบพระคุณบิดา มารดาเป็นอย่างสูง ที่เป็นผู้อุปการะเลี้ยงดู ให้โอกาสในทุก ๆ ด้านขอบคุณรุ่นพี่ และเพื่อน ๆ ที่คอยให้การช่วยเหลือให้คำแนะนำ และผลักดันให้กำลังใจที่ดีเสมอมาและทุก ๆ ท่านที่ไม่ได้เอ่ยนามที่ช่วยเหลือจนกระทั่งโครงงานฉบับนี้สำเร็จไปอย่างมีประสิทธิภาพ

คณะผู้จัดทำ

สารบัญ

เรื่อง	หน้า
บทคัดย่อ	ก
กิตติกรรมประกาศ	ข
สารบัญ	ค
สารบัญตาราง	จ
สารบัญภาพ	ช
บทที่ 1 บทนำ	1
1.1 หลักการและเหตุผล	1
1.2 วัตถุประสงค์	2
1.3 ประโยชน์ที่คาดว่าจะได้รับ	2
1.4 ขอบเขต	2
1.5 เครื่องมือในการพัฒนาโปรแกรม	3
1.6 สถานที่ใช้ในการดำเนินการ	4
1.7 ระยะเวลาในการดำเนิน	4
1.8 นิยามศัพท์เฉพาะ	5
1.9 บทสรุป	6
บทที่ 2 แนวคิดทฤษฎีเอกสารที่เกี่ยวข้อง	7
2.1 แนวคิด	8
2.2 ทฤษฎี	17
2.3 เครื่องมือในการออกแบบและวิเคราะห์	40
2.4 วรรณกรรมที่เกี่ยวข้อง	44
2.5 บทสรุป	47
บทที่ 3 ผลการดำเนินงาน	48
3.1 แผนภาพกระแสข้อมูล	48
3.2 กระบวนการทำนอร์มัลไลเซชัน	59
3.3 ความสัมพันธ์ของฐานข้อมูล	66

สารบัญ

เรื่อง	หน้า
3.4 พจนานุกรมข้อมูล	79
3.5 โครงสร้างระบบ	83
3.5 กระบวนการ CRISP–DM	84
3.7 บทสรุป	114
บทที่ 4 ผลการดำเนินงาน	115
4.1 ผลการดำเนินงาน	115
4.2 การอภิปรายผล	122
4.3 บทสรุป	129
บทที่ 5 สรุปและข้อเสนอแนะ	130
5.1 บทสรุปผลการโครงงาน	130
5.2 ข้อจำกัดของโครงงาน	131
5.3 ปัญหาและอุปสรรคของโครงงาน	131
5.4 ข้อเสนอแนะ	131
บรรณานุกรม	133
ภาคผนวก ก คู่มือการใช้งานเว็บไซต์	134
ภาคผนวก ข แบบสอบถามและเอกสารที่ใช้ในโครงงาน	148
ภาคผนวก ค แบบบันทึกรายละเอียดการเข้าพบอาจารย์ที่ปรึกษา	152
ประวัติผู้จัดทำ	154

สารบัญตาราง

ตารางที่ **หน้า**

ตารางที่ 3.1 สัญลักษณ์การไหลของข้อมูล แบบ Gane and Sarson 49
ตารางที่ 3.2 สัญลักษณ์การไหลของข้อมูล แบบ Yourdon 49
ตารางที่ 3.3 แสดงคำอธิบายกระบวนการ ตรวจสอบระบบ 52
ตารางที่ 3.4 แสดงคำอธิบายกระบวนการจัดการข้อมูลพยากรณ์ปริมาณขยะมูลฝอย 52
ตารางที่ 3.5 แสดงคำอธิบายกระบวนการจัดการข้อมูลสถานที่กำจัดขยะมูลฝอย 53
ตารางที่ 3.6 แสดงคำอธิบายกระบวนการจัดการดาวน์โหลดเอกสารข้อมูลต่าง ๆ 53
ตารางที่ 3.7 แสดงคำอธิบายกระบวนการรายงานสารสนเทศ 54
ตารางที่ 3.8 แสดงคำอธิบายกระบวนการที่ 2.1 การเพิ่มข้อมูลปริมาณขยะมูลฝอย 55
ตารางที่ 3.9 แสดงคำอธิบายกระบวนการที่ 2.2 การแก้ไขข้อมูลปริมาณขยะมูลฝอย 56
ตารางที่ 3.10 แสดงคำอธิบายกระบวนการที่ 2.3 ลบข้อมูลปริมาณขยะมูลฝอย 56
ตารางที่ 3.11 แสดงคำอธิบายกระบวนการที่ 3.1 เพิ่มข้อมูลสถานที่กำจัดขยะมูลฝอย 57
ตารางที่ 3.12 แสดงคำอธิบายกระบวนการที่ 3.2 แก้ไขสถานที่กำจัดขยะมูลฝอย 58
ตารางที่ 3.13 แสดงคำอธิบายกระบวนการที่ 3.3 ลบข้อมูลสถานที่กำจัดขยะมูลฝอย 58
ตารางที่ 3.14 ตารางข้อมูลก่อนที่จะทำการนอร์มัลฟอร์มที่ 1 ในประเด็นที่ 1 59
ตารางที่ 3.15 ตารางข้อมูลที่อยู่ในรูปของนอร์มัลฟอร์มที่ 1 ในประเด็นที่ 1 60
ตารางที่ 3.16 ตารางข้อมูลที่อยู่ในรูปของนอร์มัลฟอร์มที่ 2 ในประเด็นที่ 1 61
ตารางที่ 3.17 ตาราง region ที่อยู่ในรูปของนอร์มัลฟอร์มที่ 2 (Normal Form : 2NF) 61
ตารางที่ 3.18 ตาราง Category ที่อยู่ในรูปของนอร์มัลฟอร์มที่ 2 (Normal Form : 2NF) 62
ตารางที่ 3.19 ตาราง Garbage_dumping ที่สมบูรณ์ในประเด็นที่ 1 62
ตารางที่ 3.20 ตารางข้อมูลก่อนที่จะทำการนอร์มัลฟอร์มที่ 1 ในประเด็นที่ 2 63
ตารางที่ 3.21 ตารางที่อยู่ในรูปของนอร์มัลฟอร์มที่ 1 ในประเด็นที่ 2 63
ตารางที่ 3.22 ตาราง Province ที่อยู่ในรูปของนอร์มัลฟอร์มที่ 2 (Normal Form : 2NF) 64
ตารางที่ 3.23 ตาราง District ที่อยู่ในรูปของนอร์มัลฟอร์มที่ 2 (Normal Form : 2NF) 64
ตารางที่ 3.24 ตาราง Sub-district ที่อยู่ในรูปของนอร์มัลฟอร์มที่ 2 65
(Normal Form: 2NF)

สารบัญตาราง

ตารางที่ **หน้า**

ตารางที่ 3.25 ตารางข้อมูล Category ที่อยู่ในรูปของนอร์มัลฟอร์มที่ 2 (Normal Form: 2NF) 65
ตารางที่ 3.26 ตารางข้อมูล Garbage_location ที่สมบูรณ์ในประเด็นที่ 2 65
ตารางที่ 3.27 แสดงชื่อตารางทั้งหมดของระบบฐานข้อมูล 79
ตารางที่ 3.28 แสดงรายละเอียดของตาราง Garbage_dumping 79
ตารางที่ 3.29 แสดงรายละเอียดของตาราง Garbage_location 80
ตารางที่ 3.30 แสดงรายละเอียดของตาราง Province 80
ตารางที่ 3.31 แสดงรายละเอียดของตาราง District 81
ตารางที่ 3.32 แสดงรายละเอียดของตาราง Sub-district 81
ตารางที่ 3.33 แสดงรายละเอียดของตาราง Category 82
ตารางที่ 3.34 ตารางเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE) 95

สารบัญภาพ

ภาพที่	หน้า
ภาพที่ 2.1 เว็บที่มีโครงสร้างแบบเรียงลำดับ	20
ภาพที่ 2.2 เว็บที่มีโครงสร้างแบบลำดับขั้น	21
ภาพที่ 2.3 เว็บที่มีโครงสร้างแบบตาราง	21
ภาพที่ 2.4 เว็บที่มีโครงสร้างแบบใยแมงมุม	22
ภาพที่ 2.5 แสดงการเกิดรีพีทติ้งกรุ๊ป	32
ภาพที่ 2.6 แสดงรีเลชันที่อยู่ในรูปนอร์มัลฟอร์มที่ 1	33
ภาพที่ 2.7 แสดงรีเลชันที่มีพาร์เชียลดีเพนเดนซี	34
ภาพที่ 2.8 แสดงรีเลชันที่อยู่ในรูปนอร์มัลฟอร์มที่ 2	34
ภาพที่ 2.9 แสดงตารางที่มีทรานซิทีฟดีเพนเดนซี	35
ภาพที่ 2.10 แสดงตารางที่มีรีเลชันที่อยู่ในรูปนอร์มัลฟอร์มที่ 3	36
ภาพที่ 2.11 แสดงตารางที่มีรีเลชันที่อยู่ในรูปนอร์มัลฟอร์มที่ 3 แต่ไม่อยู่ในรูปของบอยซ์	37
ภาพที่ 2.12 แสดงตารางที่มีรีเลชันที่อยู่ในรูปนอร์มัลฟอร์มที่ 3 ที่อยู่ในรูปของบอยซ์	37
ภาพที่ 2.13 แสดงตารางที่มีรีเลชันที่อยู่ในรูปนอร์มัลฟอร์มที่ 4	38
ภาพที่ 2.14 แสดงตารางที่มีรีเลชันที่อยู่ในรูปนอร์มัลฟอร์มที่ 4	39
ภาพที่ 2.15 แสดงตารางคุณสมบัติที่มีรีเลชันที่อยู่ในรูปนอร์มัลฟอร์มที่ 5	40
ภาพที่ 2.16 กราฟในลักษณะต่าง ๆ	43
ภาพที่ 3.1 แผนภาพบริบท Context Diagram	50
ภาพที่ 3.2 แสดงแผนภาพกระแสข้อมูลระดับที่ 0	51
ภาพที่ 3.3 แผนภาพกระแสข้อมูล ระดับที่ 1 กระบวนการที่ 2	55
ภาพที่ 3.4 แผนภาพกระแสข้อมูล ระดับที่ 1 กระบวนการที่ 3	57
ภาพที่ 3.5 การกำหนดเอนทิตี้ในประเด็นที่ 1 ปริมาณขยะมูลฝอยในประเทศไทย	67
ภาพที่ 3.6 การกำหนดเอนทิตี้ในประเด็นที่ 2 สถานที่กำจัดขยะมูลฝอยในประเทศไทย	67
ภาพที่ 3.7 การกำหนดแอตทริบิวต์ Garbage_dumping ในประเด็นที่ 1 ปริมาณขยะมูลฝอย	68

สารบัญภาพ

ภาพที่	หน้า
ภาพที่ 3.8 การกำหนดแอตทริบิวต์ Province ในประเด็นที่ 1 ปริมาณขยะมูลฝอย	68
ภาพที่ 3.9 การกำหนดแอตทริบิวต์ region ในประเด็นที่ 1 ปริมาณขยะมูลฝอย	69
ภาพที่ 3.10 การกำหนดแอตทริบิวต์ Category ในประเด็นที่ 1 ปริมาณขยะมูลฝอย	69
ภาพที่ 3.11 การกำหนดแอตทริบิวต์ Garbage_location ในประเด็นที่ 2	70
ภาพที่ 3.12 การกำหนดแอตทริบิวต์ Province ในประเด็นที่ 2 สถานที่กำจัดขยะมูลฝอย	70
ภาพที่ 3.13 การกำหนดแอตทริบิวต์ District ในประเด็นที่ 2 สถานที่กำจัดขยะมูลฝอย	71
ภาพที่ 3.14 การกำหนดแอตทริบิวต์ Sub-District ในประเด็นที่ 2 สถานที่กำจัดขยะมูลฝอย	71
ภาพที่ 3.15 การกำหนดแอตทริบิวต์ Category ในประเด็นที่ 2 สถานที่กำจัดขยะมูลฝอย	72
ภาพที่ 3.16 แสดงความสัมพันธ์ระหว่างตาราง Garbage_dumping กับตาราง Province	73
ภาพที่ 3.17 แสดงความสัมพันธ์ระหว่างตาราง Garbage_dumping กับตาราง region	73
ภาพที่ 3.18 แสดงความสัมพันธ์ระหว่างตาราง Garbage_dumping กับตาราง Category	74
ภาพที่ 3.19 แสดงความสัมพันธ์ระหว่างตาราง Garbage_location กับตาราง Province	74
ภาพที่ 3.20 แสดงความสัมพันธ์ระหว่างตาราง Garbage_location กับตาราง District	75
ภาพที่ 3.21 แสดงความสัมพันธ์ระหว่างตาราง Garbage_location กับตาราง Sub-District	75
ภาพที่ 3.22 แสดงความสัมพันธ์ระหว่างตาราง Garbage location กับตาราง Category	76

สารบัญภาพ

ภาพที่	หน้า
ภาพที่ 3.23 แสดงความสัมพันธ์ของฐานข้อมูลแบบ Crow's Foot Model	77
ภาพที่ 3.24 แสดงความสัมพันธ์ของฐานข้อมูลแบบ Crow's Foot Model	78
ภาพที่ 3.25 แผนผังโครงสร้างการใช้งานในระบบฐานข้อมูลปริมาณขยะ	83
ภาพที่ 3.26 แสดงกระบวนการ CRISP-DM	84
ภาพที่ 3.27 แสดงขั้นตอนการหาข้อมูลจากเว็บไซต์	86
ภาพที่ 3.28 แสดงขั้นตอนการดาวน์โหลดข้อมูล	86
ภาพที่ 3.29 ไฟล์ข้อมูลการบันทึกปริมาณขยะมูลฝอยรายปี	87
ภาพที่ 3.30 ไฟล์ข้อมูลปริมาณขยะมูลฝอยทั้งหมดที่ได้จากเว็บไซต์	88
ภาพที่ 3.31 ข้อมูลที่ปรับมาตรฐานรวมให้เป็นรายปีในแต่ละจังหวัดในประเทศไทย	88
ภาพที่ 3.32 ไฟล์ข้อมูลที่ต้องการจัดกลุ่ม	89
ภาพที่ 3.33 แสดงการจัดกลุ่มรายประเภทการกำจัดขยะมูลฝอย	89
ภาพที่ 3.34 หน้าต่างพื้นที่การทำงานของโปรแกรม Minitab	90
ภาพที่ 3.35 ข้อมูลที่นำมาใช้ในการพยากรณ์แต่ละประเภท	90
ภาพที่ 3.36 เลือกไฟล์ที่ต้องการนำข้อมูลมาพยากรณ์	91
ภาพที่ 3.37 Import ข้อมูลเพื่อหาค่าการพยากรณ์	91
ภาพที่ 3.38 แอตทริบิวท์ที่นำมาคำนวณหาค่าพยากรณ์	92
ภาพที่ 3.39 ตารางข้อมูลปริมาณขยะมูลฝอยโดยแบ่งตามจังหวัดและประเภท	92
ภาพที่ 3.40 แสดงขั้นตอนการนำข้อมูลเข้าโปรแกรม Minitab	93
ภาพที่ 3.41 แสดงขั้นตอนการกำหนดรูปแบบการพยากรณ์วิธี Single EMA	94
ภาพที่ 3.42 แสดงผลลัพธ์กราฟวิธี Single EMA	94
ภาพที่ 3.43 หน้าต่างที่ทำการเชื่อมต่อฐานข้อมูลไฟล์ Excel เพื่อนำมาวิเคราะห์	95
ภาพที่ 3.44 Import แสดงผลการโหลดฐานข้อมูลเข้าสู่โปรแกรม Tableau Public	96
ภาพที่ 3.45 หน้าต่างพื้นที่ทำงานของโปรแกรม tableau	96
ภาพที่ 3.46 หน้าต่างที่แสดงข้อมูลที่ Import ข้อมูล	97
ภาพที่ 3.47 แสดงการเตรียมกราฟข้อมูลก่อนการวิเคราะห์ข้อมูล	97
ภาพที่ 3.48 การแสดงเครื่องมือในการวิเคราะห์	98

สารบัญภาพ

ภาพที่	หน้า
ภาพที่ 3.49 การแสดงเครื่องมือในการพยากรณ์	98
ภาพที่ 3.50 การแสดงผลลัพธ์ในการพยากรณ์ช่วงเวลา 2 ปีถัดไป	99
ภาพที่ 3.51 ผลการคำนวณค่าความคลาดเคลื่อน (MAPE)	99
ภาพที่ 3.52 ผลลัพธ์การวัดประสิทธิภาพของผลลัพธ์โปรแกรม Minitab	100
ภาพที่ 3.53 ผลลัพธ์การวัดประสิทธิภาพของผลลัพธ์ด้วยโปรแกรม Tableau Public	101
ภาพที่ 3.54 Import ข้อมูลปริมาณขยะมูลฝอยในประเทศไทยเพื่อหาผลลัพธ์	102
ภาพที่ 3.55 แสดงขั้นตอนการปรับ Mark ให้เป็น Bar	102
ภาพที่ 3.56 แสดงขั้นตอนการแสดง category บน Map	103
ภาพที่ 3.57 แสดงขั้นตอนการกำหนดค่าการพยากรณ์ในรูปแบบกราฟแท่ง	103
ภาพที่ 3.58 แสดงผลลัพธ์การพยากรณ์เป็นสีและรายละเอียดต่างๆ	104
ภาพที่ 3.59 Import ข้อมูลสถานที่กำจัดขยะมูลฝอยในประเทศไทยเพื่อหาผลลัพธ์	104
ภาพที่ 3.60 แสดงขั้นตอนการปรับ Mark ให้เป็น Map	105
ภาพที่ 3.61 แสดงขั้นตอนการแสดงสถานที่กำจัดขยะในประเทศไทย บน Map	105
ภาพที่ 3.62 แสดงผลลัพธ์สถานที่กำจัดขยะมูลฝอยในประเทศไทย	106
ภาพที่ 3.63 หน้าแรกของเว็บไซต์ แสดงเมนูต่าง ๆ ของหน้าเว็บไซต์	106
ภาพที่ 3.64 หน้าเข้าสู่ระบบของผู้ดูแลระบบ	107
ภาพที่ 3.65 หน้าแหล่งความรู้ขยะมูลฝอย	107
ภาพที่ 3.66 หน้าความรู้เกี่ยวกับปัญหาขยะมูลฝอย	108
ภาพที่ 3.67 หน้าการจัดการขยะมูลฝอย	108
ภาพที่ 3.68 หน้าการนำขยะมาใช้ประโยชน์	109
ภาพที่ 3.69 หน้าการแสดงแดชบอร์ด	109
ภาพที่ 3.70 หน้าการแสดงพยากรณ์ปริมาณขยะมูลฝอย	110
ภาพที่ 3.71 หน้าดาวน์โหลดเอกสารต่างๆ	110
ภาพที่ 3.72 หน้าเพิ่มไฟล์เอกสาร	111
ภาพที่ 3.73 หน้าเพิ่มข้อมูลการพยากรณ์ขยะมูลฝอย	111
ภาพที่ 3.74 หน้าการเพิ่มสถานที่กำจัดขยะมูลฝอย	112

สารบัญภาพ

ภาพที่	หน้า
ภาพที่ 3.75 หน้าการแสดงแบบสอบถาม	112
ภาพที่ 3.76 หน้าการแสดงเกี่ยวกับเรา	113
ภาพที่ 3.77 หน้าการแสดงหน่วยงานที่เกี่ยวข้อง	113
ภาพที่ 4.1 หน้าแรกของเว็บไซต์	116
ภาพที่ 4.2 หน้าแหล่งความรขู้ยะมูลฝอย	117
ภาพที่ 4.3 หน้าปัญหาขยะมูล	117
ภาพที่ 4.4 หน้าการจัดการขยะมูลฝอย	118
ภาพที่ 4.5 หน้าการนำมาใช้ประโยชน์ใหม่	118
ภาพที่ 4.6 หน้าขอ้มูลค่าเฉลี่ยขยะมูลฝอย	119
ภาพที่ 4.7 หน้าพยากรณ์ข้อมูลขยะมูลฝอย	119
ภาพที่ 4.8 หน้าสถานที่กำจัดขยะมูลฝอย	120
ภาพที่ 4.9 หน้าดาวน์โหลดเอกสาร	120
ภาพที่ 4.10 หน้าเข้าสู่ระบบ	121
ภาพที่ 4.11 หน้าเว็บไซต์แบบสอบถามความพงึพอใจของเว็บไซต์	121
ภาพที่ 4.12 หน้าเว็บไซต์ข้อมูลเกี่ยวกับเว็บไซต์	122
ภาพที่ ก.1 หน้าเข้าสู่ระบบเว็บไซต์	135
ภาพที่ ก.2 หน้าแรกของเว็บไซต์	136
ภาพที่ ก.3 หน้าแหล่งความรู้เกี่ยวกับขยะมูลฝอย	136
ภาพที่ ก.4 หน้าข้อมูลค่าเฉลี่ยขยะมูลฝอย	137
ภาพที่ ก.4.1 หน้าการเลือกแสดงชนิดขยะมูลฝอยที่จะแสดงข้อมูล	138
ภาพที่ ก.4.2 หน้าการเลือกขอบเขตของภาค	138
ภาพที่ ก.4.3 หน้าการเลือกขอบเขตของจังหวัด	139
ภาพที่ ก.4.4 หน้าแสดงปีเริ่มต้น และปีสิ้นสุดที่จะให้แสดงข้อมูลขยะมูลฝอย	139
ภาพที่ ก.5.1 หน้าการพยากรณ์ข้อมูลขยะมูลฝอย	140
ภาพที่ ก.6 หน้าแสดงแผนที่จำกัดขยะมูลฝอย	140
ภาพที่ ก.6.1 หน้าแสดงปริมาณขยะมูลฝอยบนแผนที่	141

สารบัญภาพ

ภาพที่	หน้า
ภาพที่ ก.6.2 หน้าแสดงขนาดพื้นที่บนแผนที่	142
ภาพที่ ก.6.2 หน้าแสดงตารางข้อมูลสถานที่กำจัดขยะมูลฝอย	142
ภาพที่ ก.7 หน้าแสดงชุดข้อมูลปริมาณขยะมูลฝอย	143
ภาพที่ ก.7.1 หน้าดาวน์โหลดชุดข้อมูลปริมาณขยะมูลฝอย	143
ภาพที่ ก.7.2 หน้าเพิ่มไฟล์เอกสารขยะมูลฝอย	144
ภาพที่ ก.7.3 หน้าเพิ่มไฟล์ชุดข้อมูลการพยากรณ์ปริมาณขยะมูลฝอย	144
ภาพที่ ก.7.4 หน้าเพิ่มไฟล์ชุดข้อมูลสถานที่จำกัดขยะมูลฝอย	145
ภาพที่ ก.8 หน้าแบบสอบถามความพึงพอใจของผู้เข้าชมเว็บไซต์	146
ภาพที่ ก.9 หน้าแสดงรายละเอียดเกี่ยวกับเว็บไซต์	146

บทที่ 1
บทนำ

ในบทนี้เป็นการนำเสนอเกี่ยวกับ หลักการ เหตุผลความเป็นมา และความสำคัญของปัญหาโครการที่เกี่ยวข้องของการวิเคราะห์ข้อมูลปริมาณขยะของกระทรวงทรัพยากรธรรมชาติและสิ่งแวดล้อม เพื่อเผยแพร่ข้อมูลบนเว็บไซต์ ซึ่งได้อธิบายถึงวัตถุประสงค์และขอบเขตของการศึกษาโครงงานประกอบด้วยรายละเอียด ดังนี้

1.1 ความสำคัญและที่มาของปัญหา

ปัจจุบันวิทยาการเทคโนโลยีมีความก้าวหน้าอย่างรวดเร็ว ควบคู่ไปกับการเพิ่มของจำนวนประชากรมีแนวโน้มสูงขึ้นส่งผลให้มีการขยายตัวของชุมชนเมือง เกิดการพัฒนาประเทศหลาย ด้านทั้งด้านเทคโนโลยีอุตสาหกรรมและการท่องเที่ยว การเปลี่ยนแปลงและการพัฒนานี้จะนำมาซึ่งความเจริญทางเศรษฐกิจและสังคม จากจำนวนประชากรที่เพิ่มมากขึ้น อัตราความต้องการอุปโภคบริโภคก็เพิ่มมากขึ้น กระบวนการผลิตก็มากขึ้น เมื่อมีความต้องการเพิ่มมากขึ้นทรัพยากรที่มีอยู่ก็ถูกนำเอามาใช้อย่างสิ้นเปลือง เกิดสิ่งเหลือจากการอุปโภค และบริโภคซึ่งของเสียหรือขยะมูลฝอยเหล่านี้ ส่วนหนึ่งจะย่อยสลายเองได้ แต่ส่วนใหญ่เป็นขยะมูลฝอยที่ไม่สามารถย่อยสลายเองได้ตามธรรมชาติและยากต่อการทำลาย หากทำลายไม่ถูกวิธีหรือปล่อยทิ้งไว้ ก็จะก่อเกิดผลกระทบกับประชาชนและสิ่งแวดล้อมในหลายด้าน เช่น ปัญหาขยะล้นเมือง น้ำเน่าเสีย และก่อเกิดแหล่งเพาะพันธุ์สะสมเชื้อโรค

การดำเนินกิจกรรมในชีวิตประจำวันของประชากร ส่งผลให้มีผลิตภัณฑ์ใหม่เกิดขึ้นอย่างสม่ำเสมอให้เลือกอุปโภค และบริโภคได้สะดวกซึ่งเป็นสาเหตุให้มีสิ่งเหลือใช้ที่ไม่เป็นที่ต้องการมากขึ้น มีผลต่อปริมาณขยะเพิ่มขึ้นอย่างรวดเร็วตามปริมาณการอุปโภคบริโภคและจำนวนประชากรที่เพิ่มสูงขึ้น ปัญหาขยะมูลฝอยที่มีปริมาณมากขึ้นเหล่านี้ ทำให้เกิดปัญหา “ขยะมูลฝอยล้นเมือง” และหากขยะที่ไม่ได้รับการจัดการที่ถูกต้อง ไม่มีการแก้ไขจะส่งผลให้เกิดผลกระทบต่อสิ่งแวดล้อม และเป็นอันตรายคุกคามต่อสุขภาพอนามัยของประชาชนได้ จากการศึกษาปริมาณขยะมูลฝอยทั้งในเขตเทศบาล และนอกเทศบาลของแต่ละพื้นที่ในจังหวัดต่าง ๆ ซึ่งในแต่ละเขตพื้นที่ของประเทศไทยขึ้นอยู่กับความเจริญทางเศรษฐกิจและสังคม โดยขยะมูลฝอยสามารถจำแนกเหมือนกัน 4 ประเภท คือ 1.) ขยะอินทรีย์ 2.) ขยะรีไซเคิล

3.) ขยะทั่วไป 4.) ขยะอันตราย ซึ่งขยะมูลฝอยส่วนใหญ่นั้นเกิดจากบ้านเรือนของประชาชน ถ้ามีการจัดการมูลฝอยที่ไม่ถูกต้องในแต่ละชุมชน ก็จะเกิดปัญหาด้านขยะมูลฝอยได้ ถ้าหากมีการจัดการขยะมูลฝอยอย่างถูกวิธี และมีคัดแยกมูลฝอยที่สามารถใช้ประโยชน์ได้นำกลับมาใช้ใหม่ ก็จะสามารถช่วยลดปริมาณมูลฝอยที่จะต้องนำไปกำจัดลงได้

จากปัญหาดังกล่าวคณะผู้วิเคราะห์จึงได้มองเห็นแนวทางในการแก้ปัญหา โดยการนำข้อมูลปริมาณขยะมูลฝอยในประเทศไทยมาเข้าสู่กระบวนการวิเคราะห์ข้อมูล การรายงานผลการพยากรณ์ผ่านเว็บไซต์ ซึ่งเป็นแนวทางในการรับมือขยะที่จะเพิ่มปริมาณมากขึ้น โดยผู้วิเคราะห์ข้อมูลได้จำแนกข้อมูลออกเป็นประเภทต่าง ๆ โดยนำโมเดล Time series analysis มาช่วยในการหาค่าความคลาดเคลื่อนของการพยากรณ์ มาเปรียบเทียบตัวแบบการพยากรณ์ด้วยวิธีการคำนวณหาค่าเฉลี่ย เปอร์เซ็นต์ความคลาดเคลื่อนสมบูรณ์ (MAPE) ที่ต่ำที่สุดเพื่อพิจารณาว่าวิธีไหนให้ความถูกต้องแม่นยำมากที่สุด นำมาใช้ในการพยากรณ์ปริมาณขยะมูลฝอย เพื่อช่วยสนับสนุนการให้ข่าวสาร เกี่ยวกับปริมาณขยะมูลฝอย โดยผ่านการจัดรูปแบบการรายงานข้อมูล การพยากรณ์ผ่านทาง Visualization และนำเสนอผ่านทางเว็บไซต์

1.2 วัตถุประสงค์

1.2.1 เพื่อวิเคราะห์ข้อมูลและการพยากรณ์ปริมาณขยะมูลฝอยจากกระทรวงทรัพยากรธรรมชาติและสิ่งแวดล้อม

1.2.2 เพื่อเผยแพร่ข้อมูลบนเว็บไซต์

1.3 ประโยชน์ที่จะได้รับ

1.3.1 ได้ข้อมูลสารสนเทศจากการวิเคราะห์ข้อมูลปริมาณขยะมูลฝอยจากกระทรวงทรัพยากรธรรมชาติและสิ่งแวดล้อม

1.3.2 ได้เผยแพร่ข้อมูลบนเว็บไซต์

1.4 ขอบเขต

1.4.1 ขอบเขตผู้วิเคราะห์ข้อมูล

1.4.1.1 ผู้วิเคราะห์ได้ทำการรวบรวมข้อมูลปริมาณขยะมูลฝอยในประเทศไทยจากเว็บไซต์ https://www.pcd.go.th โดยนำมาวิเคราะห์ข้อมูล ได้แก่ ปริมาณขยะมูลฝอยแต่ละจังหวัดในปี 2556-2562

1.4.1.2 ผู้วิเคราะห์ข้อมูลได้การพยากรณ์ข้อมูลแบบอนุกรมเวลา เพื่อใช้พยากรณ์แนวโน้มปริมาณขยะมูลฝอยที่เพิ่มมากขึ้นในแต่ละปีและในแต่ละจังหวัดโดยใช้โปรแกรม Minitab 19 และโปรแกรม Tableau public 2019.3 เพื่อทดสอบประสิทธิภาพของการวิเคราะห์ข้อมูล ด้วยการหาค่าเฉลี่ยเปอร์เซ็นต์ความคลาดเคลื่อนสมบูรณ์ (MAPE) ของข้อมูลที่ใช้ในการสร้างตัวแบบหรือโมเดล

1.4.1.3 การวิเคราะห์และนำเสนอข้อมูลภาพจะถูกดำเนินการบนโปรแกรม XAMPP เพื่อใช้เป็นเว็บเซิร์ฟเวอร์จำลองสำหรับทดสอบและทดลองเว็บไซต์

1.4.1.4 ด้านการจัดการข้อมูลและสารสนเทศ โครงงานนี้จะดำเนินการโดยใช้ phpMyAdmin ในการจัดการฐานข้อมูลบนระบบจัดการฐานข้อมูล MySQL

1.4.1.5 ด้านการนำเสนอข้อมูลภาพ (Data Visualization) โครงงานนี้จะนำเสนอและเผยแพร่ข้อมูลสารสนเทศผ่านเว็บแอปพลิเคชัน ที่สามารถอัปโหลดข้อมูลใหม่เข้าสู่ระบบได้ โดยเว็บแอปพลิเคชันนี้จะพัฒนาโดยใช้ภาษา HTML, PHP, javascript และชุดคำสั่ง CSS3 เพื่อนำเข้า วิเคราะห์ข้อมูลด้วยโมเดลที่เลือกใช้และผ่านการทดสอบประสิทธิภาพของการวิเคราะห์ข้อมูลแล้ว และนำเสนอข้อมูลภาพด้วยเฟรมเวิร์คการแสดงข้อมูลภาพชื่อ highcharts.js เพื่อสร้างกราฟและ dashboard สำหรับแสดงผลข้อมูลภาพจากข้อมูลที่ผ่านการวิเคราะห์ แล้วภายในเว็บแอปพลิเคชันที่พัฒนาขึ้น

1.4.2 ผู้เยี่ยมชมเว็บไซต์ (Visitor)

1.4.2.1 สามารถเรียกดูรายงานย้อนหลังปริมาณขยะมูลฝอยในประเทศไทยได้

1.4.2.2 สามารถดูข้อมูลการพยากรณ์ปริมาณขยะมูลฝอยได้

1.4.2.3 สามารถดูข้อมูลข่าวสารเกี่ยวกับปริมาณขยะได้

1.5 เครื่องมือที่ใช้ในการพัฒนาโปรแกรม

1.5.1 Hardware

1.5.1.1 Notebook: AMD Ryzen5 2600 Six-Core RAM 16 GB

1.5.1.2 Notebook: Intel(R) Core (TM) i5-8250U RAM 4 GB

1.5.2 Software

1.5.2.1 โปรแกรม Tableau Public ใช้ในการแสดงผลข้อมูลในรูปแบบของภาพ

1.5.2.2 โปรแกรม Adobe XD ใช้สำหรับการออกแบบหน้าเว็บไซต์

1.5.2.3 โปรแกรม Minitab 19 ใช้สำหรับการพยากรณ์ข้อมูล

1.5.2.4 ระบบปฏิบัติการ Window 10 ใช้ในการเปิดซอฟต์แวร์ต่าง ๆ

1.5.2.5 Microsoft office Word 2016 ใช้ในการทำเอกสารต่าง ๆ

1.5.2.6 Microsoft office Excel 2016 ใช้ในการ Clean Data กับชุดข้อมูล

1.5.2.7 ชุดคำสั่ง HTML, JavaScript, CSS สำหรับพัฒนาส่วนหน้าจอแสดงผล

1.6 สถานที่ใช้ในการดำเนินการศึกษาและรวบรวมข้อมูล

1.6.1 สถานที่ใช้ในการดำเนินการศึกษา

มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา 128 ถนนห้วยแก้ว ตำบลช้างเผือก อำเภอเมือง จังหวัดเชียงใหม่ 50300

1.6.2 แหล่งรวบรวมข้อมูล

เว็บไซต์ http://www.pcd.go.th หรือ จากกระทรวงทรัพยากรธรรมชาติและสิ่งแวดล้อม

1.7 ระยะเวลาในการดำเนินการ

แผนการดำเนินงาน	**พ.ศ 2563**				
	พ.ย	**ธ.ค**	**ม.ค**	**ก.พ**	**มี.ค**
1. ทำความเข้าใจในปัญหา	←	→			
2. การเก็บรวบรวมข้อมูล	←	→			
3. การเตรียมข้อมูล data cleaning		←	→		
4. การสร้างแบบจำลอง			←	→	
5. การประเมินประสิทธิภาพ			←	→	
6. นำข้อมูลไปแสดงผลทางเว็บไซต์				←	→
7. จัดทำเอกสารประกอบโครงการ	←				→

1.8 นิยามศัพท์เฉพาะ

1.8.1 ข้อมูลอนุกรมเวลา (Time Series data) คือ ชุดของข้อมูลที่เก็บรวบรวมตามระยะเวลาเป็นช่วง ๆ อย่างต่อเนื่องกัน เช่น ข้อมูลยอดขายสินค้าที่เก็บรวบรวมต่อเนื่องกันไปเป็นระยะเวลาหลาย ๆ เดือน ข้อมูลรายได้ประชาชาติปีต่าง ๆ ที่เก็บรวบรวมต่อเนื่องกันไปเป็นระยะเวลาหลาย ๆ ปี เป็นต้น ข้อมูลอนุกรมเวลาอาจอยู่ในลักษณะที่เป็นข้อมูลรายปี รายไตรมาส หรือรายเดือนก็ได้ ทั้งนี้ขึ้นอยู่กับความเหมาะสมในการนำไปใช้ประโยชน์

1.8.2 Data Visualization คือ การแสดงข้อมูลเป็นภาพ ทำให้ข้อมูลดูง่ายและน่าสนใจมากขึ้นเป็นข้อมูลแบบตาราง เป็นตัวเลข อาจทำให้ข้อมูลดูน่าเบื่อ แต่ถ้าแสดงข้อมูลเป็น Data Visualization จะทำให้เห็นรูปแบบอย่างว่าถ้าค่าสูงขึ้น-น้อยลง เราก็จะพยายามหาสาเหตุว่ามันสูงขึ้นเพราะอะไร น้อยลงเพราะอะไร

1.8.3 High charts คือ แผนภูมิที่แสดงข้อมูลบนเว็บไซต์ โดยออกแบบและพัฒนาด้วย JavaScript มีหลายหลายรูปแบบในการใช้งาน

1.8.4 ขยะหรือขยะมูลฝอย (Refuse or Solid Waste) คือ ของเสียที่อยู่ในรูปของแข็ง ซึ่งอาจจะมีความชื้นปะปนมาด้วยจำนวนหนึ่ง ขยะที่เกิดขึ้นจากอาคารที่พักอาศัย สถานที่ทำการโรงงานอุตสาหกรรม หรือตลาดสดก็ตามจะมีปริมาณและลักษณะแตกต่างกันออกไป โดยปกติแล้ววัตถุต่าง ๆ ที่ถูกทิ้งมาในรูปของขยะนั้น จะมีทั้งอินทรีย์สารและอนินทรีย์สาร สารวัตถุต่าง ๆ เหล่านี้บางชนิดก็สามารถย่อยสลายได้ด้วยจุลินทรีย์ในเวลาอันรวดเร็ว โดยเฉพาะพวกเศษอาหารเศษพืชผัก แต่บางชนิดก็ไม่อาจจะย่อยสลายได้เลย เช่น พลาสติก เศษแก้ว เป็นต้น

1.8.5 การพยากรณ์ คือ การคาดการณ์สิ่งที่จะเกิดขึ้นในอนาคต โดยอาศัยการวิเคราะห์ข้อมูลในอดีต ข้อมูลปัจจุบันและจากประสบการณ์ สามารถนำไปใช้เพื่อให้ทราบถึงแนวโน้มการเปลี่ยนแปลงของสถานการณ์หรือสภาพแวดล้อมที่จะมีผลในอนาคตและทำให้สามารถที่จะวางแผนหรือกำหนดนโยบายเพื่อให้บรรลุวัตถุประสงค์ได้

1.8.6 อุตสาหกรรม คือ กิจกรรมที่ใช้ทุนและแรงงาน เพื่อที่จะผลิตสิ่งของ หรือ จัดให้มีบริการ เช่น อุตสาหกรรมสิ่งทอ ในยุควิกตอเรีย นักประวัติศาสตร์เรียกช่วงเวลานั้นว่าการปฏิวัติอุตสาหกรรม โดยมีการผลิตเครื่องทุ่นแรงต่าง ๆ มากมาย และทำให้อุตสาหกรรมเจริญรุดหน้าอย่างรวดเร็วและมีระเบียบ เป็นมาตรฐานเดียวกันทั้งหมด

ในปัจจุบันอุตสาหกรรมถือเป็นสิ่งหนึ่งที่จำเป็นต่อมนุษย์อย่างยิ่ง ด้วยว่ามนุษย์ต้องพึ่งพาการผลิตสิ่งที่จำเป็นต่อชีวิตประจำวันหรือเรียกรวมว่าปัจจัยสี่ โดยสิ่งที่สามารถผลิตปัจจัยสี่ให้ดี มีคุณภาพและไม่ก่ออันตรายหรือก่ออันตรายให้กับร่างกายและทรัพย์สินน้อยที่สุดคือการผลิตจากอุตสาหกรรม

1.9 บทสรุป

จากบทนำที่ได้กล่าวมาในข้างต้นทั้งหมดนั้น ผู้วิเคราะห์ได้เล็งเห็นความสำคัญของการนำเทคโนโลยีมาประยุกต์ใช้ในการวิเคราะห์ข้อมูลปริมาณขยะมูลฝอยจากกรมควบคุมมลพิษ ด้วยกระบวนการวิเคราะห์ข้อมูลด้วย CRISP-DM จากเทคนิค Neural network ในรูปแบบอนุกรมเวลา หรือที่เรียกว่า Time series analysis เข้ามาช่วยในการหาค่าความคลาดเคลื่อนของการพยากรณในวิธีต่าง ๆ ได้แก่ วิธีการปรับค่าเฉลี่ยเคลื่อนที่แบบมูฟวิ่ง, วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล โดยผู้วิเคราะห์จะนำการพยากรณ์ทั้ง 2 วิธีมาเปรียบเทียบเพื่อหาวิธีที่แม่นยำมากที่สุดมาใช้ในการพยากรณ์ปริมาณขยะมูลฝอยในประเทศไทย โดยผ่านเสนอนำข้อมูลต่าง ๆ ผ่านการ Visualization จัดนำเสนอผ่านทางเว็บไซต์

บทที่ 2
แนวคิดทฤษฎีเอกสารที่เกี่ยวข้อง

ในบทนี้เป็นการนำเสนอเกี่ยวกับ แนวคิด ทฤษฎี เครื่องมือและวรรณกรรมที่เกี่ยวข้องของการพัฒนาเว็บแอปพลิเคชันเพื่อการวิเคราะห์ข้อมูลปริมาณขยะมูลฝอยในประเทศไทย ซึ่งได้รวบรวมการศึกษาเอกสารงานวิจัยที่เกี่ยวข้องกับระบบ เพื่อใช้เป็นแนวทางการศึกษา ประกอบด้วยรายละเอียดตามลำดับ ดังนี้

2.1 แนวคิด

2.1.1 แนวคิดเกี่ยวกับการวิเคราะห์ข้อมูล (Data analytic)

2.1.2 แนวคิดเกี่ยบกับการพยากรณ์ข้อมูล (Forecasting data)

2.1.3 แนวคิดเกี่ยวกับแผงหน้าปัด (Dashboard)

2.1.4 แนวคิดเกี่ยวกับการแสดงผลข้อมูล (Data visualization)

2.2 ทฤษฎี

2.2.1 ทฤษฎีเกี่ยวกับการออกแบบเว็บไซต์

2.2.2 ทฤษฎีเกี่ยวกับอนุกรมเวลา (Time Series)

2.2.3 ทฤษฎีเกี่ยวกับการออกแบบ (visualization)

2.2.4 ทฤษฎีเกี่ยวออกแบบ (Dashboard)

2.2.5 ทฤษฎีการนอร์มัลไลเซชั่น (Normalization)

2.3 เครื่องมือในการออกแบบ และวิเคราะห์ข้อมูล

2.3.1 แบบจำลองอนุกรมเวลา (Time series analysis)

2.3.2 กระบวนการวิเคราะห์ข้อมูลด้วย (CRISP-DM)

2.3.3 การแสดงผลแบบ (visualization)

2.4 วรรณกรรมที่เกี่ยวข้อง

2.5 บทสรุป

2.1 แนวคิด

2.1.1 การวิเคราะห์ข้อมูลเชิงคุณภาพ

การวิเคราะห์ข้อมูลเชิงคุณภาพ นับเป็นขั้นตอนที่สำคัญในกระบวนการวิจัย วิธีการหลักที่ใช้ในการวิเคราะห์ข้อมูลเชิงคุณภาพ เป็นวิธีการสร้างข้อสรุปจากการศึกษาจากข้อมูล จำนวนหนึ่ง ซึ่งมักไม่ใช้สถิติในการวิเคราะห์หรือถ้าใช้สถิติก็ไม่ได้ถือว่าสถิติเป็นวิธีการวิเคราะห์หลักแต่จะถือเป็นข้อมูลเสริมในการวิเคราะห์ข้อมูลเชิงคุณภาพนั้น ผู้วิเคราะห์ข้อมูลนับเป็นผู้ที่มีบทบาทสำคัญยิ่ง และควรมีความรอบรู้ในเรื่องแนวคิดทฤษฎีอย่างกว้างขวาง มีความเป็นสหวิทยากรอยู่ในตัวเอง มีความสามารถทางภาษาสามารถเชื่อมโยงข้อความ และสร้างข้อสรุปเป็นกรอบแนวคิดและตีความหมายของข้อมูลได้หลาย ๆ แบบทั้งนี้การวิเคราะห์ข้อมูลเชิงคุณภาพนั้นอาจไม่จำ เป็นต้องใช้เฉพาะการวิจัยเชิงคุณภาพเท่านั้น แต่อาจใช้กับการวิจัยเชิงปริมาณที่ผู้วิจัยมีการเก็บรวบรวมข้อมูลเชิงคุณภาพ เช่น แบบสอบถามปลายเปิดการสัมภาษณ์ การสังเกต จดบันทึก มาทำการวิเคราะห์เพื่อให้ข้อมูลที่รวบรวมมามีความหมายและตอบคำถามหรือจุดมุ่งหมายของการวิจัย สำหรับสาระในการวิเคราะห์ข้อมูลเชิงคุณภาพในบทความนี้ ผู้เขียนมุ่งนำ เสนอสาระเกี่ยวกับเทคนิควิเคราะห์ข้อมูลเพื่อนำไปสู่การได้ข้อสรุปที่เป็นคำตอบของการวิจัย โดยเทคนิคการวิเคราะห์ข้อมูลเชิงคุณภาพที่สำคัญ ๆ และนักวิจัยนิยมใช้ ได้แก่ การจำแนกหรือการจัดกลุ่มข้อมูล การเปรียบเทียบเหตุการณ์การวิเคราะห์ส่วนประกอบการวิเคราะห์แบบอุปนัย และการวิเคราะห์ข้อมูลที่เป็นเอกสาร ซึ่งเทคนิคการวิเคราะห์ข้อมูลเชิงคุณภาพแต่ละเทคนิคนั้นผู้เขียนได้นำ เสนอรายละเอียดพร้อมตัวอย่างประกอบตามลำดับ ดังนี้

2.1.1.1 เทคนิควิเคราะห์ข้อมูลเชิงคุณภาพ

การวิเคราะห์โดยการจำแนกชนิดข้อมูล(Typological Analysis) การวิเคราะห์โดยการจำแนกชนิดข้อมูล คือ การจำแนกข้อมูลเป็นชนิด (Typologles) คำว่า “Typologles” หมายถึง ขั้นตอนของเหตุการณ์ที่เกิดขึ้นอย่างต่อเนื่องซึ่งในกรณีที่นักวิจัยทำการเก็บรวบรวมข้อมูลแล้วนักวิจัยจำเป็นที่จะต้องจัดระบบข้อมูลโดยอาศัยหลักเกณฑ์ ที่ผู้วิจัยกำหนดขึ้น ซึ่งการจำแนก หรือการจัดกลุ่มข้อมูลการจำแนกข้อมูลในระดับจุลภาค การจำแนกข้อมูลระดับนี้แบ่งเป็น 2 ประเภทย่อยได้แก่ การวิเคราะห์คำหลัก (Domain Analysis) กับการวิเคราะห์สารระบบ (Taxonomy Analysis) การวิเคราะห์คำหลัก หมายถึงการจำแนก จัดกลุ่มคำชุดหนึ่งให้อยู่ภายใต้คำอีกชุดหนึ่ง ซึ่งคำดังกล่าวนี้มีความหมายครอบคลุมคำชุดนั้น ๆ หรือ อาจกล่าวได้ว่าการวิเคราะห์คำหลัก เป็นการจัดกลุ่มคำชุดหนึ่งให้อยู่ร่วมกันโดยอาศัยลักษณะความสัมพันธ์บางอย่างของคำแต่ละคำที่นำมาใช้จัดกลุ่ม ทั้งนี้ลักษณะความสัมพันธ์บางอย่างที่ว่านี้เป็น

ความสัมพันธ์ เชิงวัฒนธรรมที่บุคคล แต่ละสังคมเป็นผู้จัดจำแนก เช่นคำว่า “คุณลักษณะครูที่ดี”นั้นมิได้ หมายถึงครูที่สอนเท่านั้นแต่คำนี้หมายรวมถึงการเป็นผู้ที่มีความสามารถด้านการสอนมีคุณธรรมจริยธรรมและบุคลิกที่เป็นแบบอย่างแก่นักเรียน มีความเข้าใจ สามารถทำงานร่วมกับผู้อื่นได้เป็นต้น

ซึ่งถ้าเปรียบเทียบกับการวิเคราะห์ข้อมูล เชิงปริมาณก็คือการวิเคราะห์องค์ประกอบ (Factor Analysis) เพื่อจัดกลุ่มตัวแปรย่อยให้อยู่ภายใต้องค์ประกอบ เดียวกันนั่นเอง (รัตนะ บัวสนธ์, 2551. หน้า 198)

การจำแนกข้อมูลในระดับมหาภาค เป็นการจำแนกข้อมูลตามเหตุการณ์ (Event) หรือการวิเคราะห์เหตุการณ์ ตามเรื่องราว (Event Analysis) ที่ปรากฏ ซึ่งการวิเคราะห์ข้อมูล ในระดับมหาภาคแบ่งได้เป็น 2 วิธี คือ การวิเคราะห์เหตุการณ์แบบอิงทฤษฎีและแบบไม่อิงทฤษฎีการวิเคราะห์เหตุการณ์แบบอิงทฤษฎี

1) การวิเคราะห์เหตุการณ์แบบอิงทฤษฎี คือ การแยกชนิดในเหตุการณ์นั้น ๆ โดยการยึดแนวคิดหรือกรอบการจำแนกเหตุการณ์โดยอาศัยทฤษฎีใดทฤษฎีหนึ่งเป็นกรอบการจำแนก ซึ่งกรอบการจำแนกเหตุการณ์ที่นิยมใช้หรือมีลักษณะกลาง ๆ ที่มักนำมาใช้ร่วมกัน คือ การวิเคราะห์เหตุการณ์ออกเป็น 6 ประเภท คือ ใคร ทำอะไร ที่ไหน เมื่อไร อย่างไร และทำไม

2) การวิเคราะห์เหตุการณ์แบบไม่อิงทฤษฎี คือ การจำแนกข้อมูลในเหตุการณ์หนึ่ง ๆ ที่จะวิเคราะห์ตามความเหมาะสมกับข้อมูล ซึ่งอาจใช้สามัญสำนึกหรือประสบการณ์ ของผู้วิจัย ซึ่งผู้วิจัยจะจำแนกข้อมูลเป็นชนิดง่าย ๆ ตามประเภทที่ผู้วิจัยสังเกต เมื่อจำแนกข้อมูลเป็นชนิดแล้ว ผู้วิจัยจะพิจารณาดูความสม่ำเสมอของการเกิดของข้อมูลต่าง ๆ ซึ่งจะเป็นพื้นฐานในการอธิบายสาเหตุของปรากฏการณ์ นอกจากนี้แม้ว่าผู้วิจัยจะจำแนกเองแต่ก็ต้องอาศัยมุมมองของบุคคลในสนามวิจัยเป็นพื้นฐานในการจำแนก ดังตัวอย่างจากงานวิจัยของรัตนะ บัวสนธ์ (2535) ที่คณะกรรมการศึกษาซึ่งเป็นบุคคลในสนามวิจัยได้จำแนกครูออกเป็น 3 กลุ่ม

2.1 การวิเคราะห์โดยการเปรียบเทียบเหตุการณ์ (Constant Comparison) การวิเคราะห์โดยการเปรียบเทียบเหตุการณ์ คือ การใช้วิธีการเปรียบเทียบ โดยการนำข้อมูลมาเทียบ เป็นปรากฏการณ์ วิธีการนี้ สามารถทำ ได้โดยการที่ผู้วิจัยสังเกต หรือรวบรวมข้อมูลได้หลาย ๆ อย่างแล้วนำ มาแยกตามชนิด นำมาเปรียบเทียบกันโดยทำตารางหาความสัมพันธ์จากสิ่งต่าง ๆ เหล่านั้นและสรุปผลออกมาผลที่ได้จากการวิเคราะห์ด้วยวิธีการนี้จะทำให้ได้ข้อสรุปที่มีความเป็นนามธรรมมากขึ้นและครอบคลุมหรือสามารถใช้อ้างอิงเหตุการณ์

ที่เหมาะสม โดยมีวิธีการวิเคราะห์โดยการเปรียบเทียบเหตุการณ์ ขั้นตอนในการวิเคราะห์ย่อย ๆ 4 ขั้นตอนดังนี้

- จัดชุดเหตุการณ์ใส่ตาราง ในขั้นตอนนี้เป็นการนำข้อมูลเหตุการณ์แต่ละเหตุการณ์ซึ่งผู้วิจัยได้ทำการวิเคราะห์จำแนกไว้มาใส่ในตารางเพื่อแยกประเด็นต่าง ๆ ของแต่ละเหตุการณ์โดยแยกประเด็นในแต่ละเหตุการณ์เป็น 6 ประเด็นดังตัวอย่างการวิเคราะห์โดยการเปรียบเทียบเหตุการณ์
- การเปรียบเทียบเหตุการณ์ หลังจากนำเหตุการณ์แต่ละเหตุการณ์ใส่ตารางแล้วดำเนินการเปรียบเทียบระหว่างเหตุการณ์ใหม่ ๆ กับเหตุการณ์เดิมซึ่งเหตุการณ์เหล่านี้ อาจเกิดขึ้นจากการแสดงพฤติกรรมของกลุ่มบุคคลเดิมหรือกลุ่มบุคคลอื่น ๆ แต่เหตุการณ์ ดังกล่าวนี้ได้รับการบันทึกลงตารางโดยการจำแนกประเด็นต่าง ๆ
- ประมวลข้อมูลเหตุการณ์แต่ละชุด แต่ละประเด็นของข้อมูลเหตุการณ์เข้าด้วยกันในขั้นตอนนี้เป็นการนำเหตุการณ์แต่ละชุดที่จัดลงตารางตามประเด็นต่าง ๆ มาเปรียบเทียบกันเพื่อหาลักษณะร่วม และลักษณะที่แตกต่างกัน หลังจากนั้นก็เขียนสรุปบรรยายเชื่อมโยงข้อมูลเหตุการณ์แต่ละชุดเข้าด้วยกัน เป็นการสั่งสมข้อค้นพบหรือข้อสรุปย่อย ๆ จากการเปรียบเทียบประเภทของข้อมูลและคุณลักษณะของประเภทซึ่งกันและกัน เมื่อเปรียบเทียบแล้ว ผู้วิจัยก็จะเริ่มเชื่อมโยงข้อมูลเหล่านั้นทำให้เกิดความสัมพันธ์ และเกิดเป็นแนวคิดย่อย ๆ
- ขยายวงของการเปรียบเทียบแล้วเลือกเป็นเหตุการณ์ที่เป็นกุญแจสำคัญ ผู้วิจัยจะใช้กรอบแนวคิดที่ได้จากการสรุปลักษณะความสัมพันธ์ของข้อมูลในเหตุการณ์ต่าง ๆ ตามขั้นตอนที่ 3 มาพิจารณาเหตุการณ์อื่น ๆ ที่มีอยู่เมื่อขยายวงของการเปรียบเทียบออกไป คุณสมบัติที่ได้คล้ายคลึงและที่แตกต่างกันของข้อมูลก็ยิ่งมีความชัดเจนมากขึ้น ซึ่งทำให้ผู้วิจัยพบข้อสรุปได้ ซึ่งข้อสรุปดังกล่าวจะมีลักษณะเป็นกรอบแนวคิดเชิงนามธรรมเป็นฐานขั้นแรกของการนำไปสู่ทฤษฎีหรือข้อสรุปเชิงนามธรรมที่ใหญ่ขึ้น

การวิเคราะห์ส่วนประกอบ (Component Analysis) ข้อมูลเป็นการวิเคราะห์คุณสมบัติของส่วนประกอบของข้อมูลแต่ละชุด แล้วนำคุณสมบัติของส่วนประกอบของข้อมูลมาเปรียบเทียบเพื่อหาลักษณะร่วมที่เหมือนกันและแตกต่างกัน หลังจากนั้นจึงทำการสรุปบรรยายให้เห็นถึงความหมายของข้อมูลเหล่านั้น โดยการวิเคราะห์ส่วนประกอบจะกระทำได้ก็ต่อเมื่อมีข้อมูลตั้งแต่สองชุดขึ้นไป แต่ไม่ควรมากเกินสิบชุด (สุภางค์ จันทวนิช, 2540 : 100) เพราะถ้าหากข้อมูลมากเกินไปจะทำให้ยากแก่การลงสรุปเกี่ยวกับคุณสมบัติของส่วนประกอบข้อมูลนั้น นอกจากนั้นแล้วการลงสรุปข้อมูลที่เหมาะสมสำหรับนำมาวิเคราะห์ส่วนประกอบ

ควรเป็นข้อมูลที่มีความละเอียดและได้จากการเก็บรวบรวมด้วยการวิเคราะห์ที่เจาะลึกหรือเน้นจุดสนใจ ทั้งนี้เพราะว่าข้อมูลดังกล่าวสามารถนำมาแยกส่วนประกอบได้หลายส่วน สำหรับการวิเคราะห์ส่วนประกอบของข้อมูลอาจสรุปได้เป็น 4 ขั้นตอนตามลำดับดังนี้

1) เลือกข้อมูล (ที่ทำการวิเคราะห์จัดกลุ่มหรือกำหนดชื่อข้อมูลแล้ว) ที่จะนำมาแยกส่วนประกอบเพื่อหาคุณสมบัติที่ต้องการเปรียบเทียบ ซึ่งควรมีข้อมูลอย่างน้อยตั้งแต่สองชุดขึ้นไป แต่ก็ไม่ควรมากเกินไปควรพิจารณาให้เหมาะสมสะดวกต่อการเปรียบเทียบ

2) จดรายชื่อข้อมูล และส่วนประกอบที่จะแยกข้อมูลชุดนั้น ๆ ไว้ในกระดาษบันทึก เพื่อป้องกันการลืม

3) จัดทำตารางเปรียบเทียบคุณสมบัติข้อมูลแต่ละชุดแยกตามส่วนประกอบและใส่คุณสมบัติข้อมูลแต่ละชุดแยกตามส่วนประกอบลงในตารางถ้าคุณสมบัติของข้อมูลและส่วนประกอบใดขาดหายไปอาจต้องเก็บข้อมูลเพิ่มเติม

4) เปรียบเทียบคุณสมบัติของข้อมูลทั้งหมดตามส่วนประกอบ โดยพิจารณาความเหมือนและความแตกต่างและสร้างข้อสรุปที่ได้จากการเปรียบเทียบโดยบรรยายโยงให้เห็นคุณสมบัติของข้อมูลตามส่วนประกอบนั้น ๆ

การวิเคราะห์สรุปอุปนัย (Analytic Induction) คือ การตีความสร้างข้อสรุปข้อมูลจากสิ่งที่เป็นรูปธรรมหรือปรากฏการณ์ที่มองเห็นที่เก็บรวบรวมมาได้จากข้อมูลตั้งแต่ 2 ชุดขึ้นไป เช่นการปฏิบัติงาน พฤติกรรมการสอน ตลอดจนการดำเนินชีวิต ความเป็นอยู่ ฯลฯ เมื่อผู้วิจัยได้เห็นหรือสังเกตหลาย ๆ เหตุการณ์ต่าง ๆ แล้วจึงลงมือสรุปแต่หากข้อสรุปนั้นยังไม่ได้รับการตรวจสอบอื่น ๆ ก็ถือว่าผลที่ได้เป็นสมมติฐาน หากได้รับการยืนยันก็ถือว่าเป็นข้อสรุปซึ่งมีความเป็นนามธรรมในระดับต้น ๆ ซึ่งการวิเคราะห์สรุปอุปนัย จัดได้ว่าเป็นวิธีการวิเคราะห์ข้อมูลเชิงคุณภาพที่ต้องนำ มาใช้สำ หรับวิเคราะห์ข้อมูลในการวิจัยเชิงคุณภาพทุกเรื่อง ทั้งนี้เพราะการวิเคราะห์สรุปอุปนัย เป็นการพิจารณาลักษณะร่วมกันของข้อมูลรูปธรรมเพื่อสรุปร่วมลักษณะดังกล่าว ซึ่งเป็นไปตามหลักของคำว่า “อุปนัย” (induction) ซึ่งหมายรวมถึงการหาความจริงจากข้อเท็จจริง (fact) ส่วนย่อยหลาย ๆ ส่วนที่มีลักษณะเป็นรูปธรรมแล้วสรุปความจริงชุดใหญ่ที่มีลักษณะเป็นนามธรรมครอบคลุมข้อเท็จจริงส่วนย่อยเหล่านั้น ยกตัวอย่าง เช่น จากการสังเกตและสัมภาษณ์ครูต้นแบบที่มีลักษณะของครูที่ดีจำนวน 3–5 คนเราได้ข้อเท็จจริงว่า ครูต้นแบบที่มีลักษณะครูที่ดี จะอุทิศเวลาให้กับการสอนทั้งในเวลาราชการ และนอกเวลาราชการ เตรียมการสอนและสื่อการเรียนการสอนล่วงหน้า หมั่นศึกษาค้นคว้าเพื่อพัฒนาการสอนของตนอย่างต่อเนื่องมีมนุษย์สัมพันธ์ สำหรับวิธีการวิเคราะห์สรุปอุปนัยในการวิเคราะห์ข้อมูลเชิงคุณภาพนั้น สามารถนำมาใช้ได้ 3 ลักษณะ

ได้แก่ การวิเคราะห์ สรุปข้อมูลจากบันทึกภาคสนามที่เป็นส่วนบันทึกละเอียดหรือบันทึกพรรณนา การวิเคราะห์ สรุปข้อมูลที่ได้จากการวิเคราะห์จำแนกหรือจัดกลุ่ม และการวิเคราะห์ข้อมูลที่ทำการวิเคราะห์ส่วนประกอบแล้ว

การวิเคราะห์ข้อมูลเอกสาร สามารถทำได้โดยวิธีการเชิงปริมาณและเชิงคุณภาพ วิธีเชิงปริมาณ คือ การทำให้ข้อมูลของเอกสารนั้น ได้แก่ ถ้อยคำประโยคหรือใจความที่ปรากฏในเอกสารเป็นจำนวนที่วัดได้แล้วแจงนับจำนวนของถ้อยคำประโยคหรือใจความเหล่านั้น วิธีการวิเคราะห์ข้อมูลแบบนี้ที่รู้จักกันดี คือ การวิเคราะห์เนื้อหา (Content analysis) ซึ่งโดยปกติการวิเคราะห์เนื้อหาจะทำตามเนื้อหาที่ปรากฏ

2.1.2 แนวคิดเกี่ยบกับการพยากรณ์ข้อมูล (Forecasting data)

การพยากรณ์ หมายถึงการคาดการณ์ (Predict) เกี่ยวกับลักษณะ หรือแนวโน้มของสิ่งที่สนใจที่จะเกิดขึ้นในอนาคต เพื่อใช้เป็นสารสนเทศ (Information) ประกอบการตัดสินใจซึ่งการพยากรณ์จะต้องดำเนินการเป็นส่วนแรกสุดก่อนการวางแผน หรือการเตรียมการที่จะเริ่มกระทำการใด ๆ เพื่อความถูกต้องและแม่นยำในการตัดสินใจ ดังนั้นในการดำเนินธุรกิจภายใต้ความไม่แน่นอนและการเปลี่ยนแปลงที่รวดเร็วเช่นปัจจุบัน จึงจำเป็นที่จะต้องทราบถึงความเป็นไปในอนาคต โดยอาศัยเทคนิคการพยากรณ์ต่าง ๆ เพื่อเพิ่มประสิทธิภาพในการดำเนินงานทางธุรกิจ สามารถตอบสนองต่อความต้องการของลูกค้า เพิ่มผลตอบแทนทางธุรกิจ ลดต้นทุนและความสูญเสียต่าง ๆ ที่อาจเกิดขึ้นเป็นต้น
การพยากรณ์เป็นทั้งศาสตร์และศิลป์ในการทำนายเหตุการณ์ในอนาคต ซึ่งอาจนำหลาย ๆ วิธีมาใช้แล้วแต่สถาณการณ์ เช่น นำข้อมูลในอดีตมาพยากรณ์อนาคต โดยอาศัยหลักการทางคณิตศาสตร์เข้าช่วย หรือใช้ดุลยพินิจของผู้พยากรณ์เพียงอย่างเดียว หรืออาจใช้หลาย ๆ วิธีร่วมกันเพื่อให้การพยากรณ์มีความแม่นยำมากขึ้น

2.1.2.1 ช่วงเวลาของการพยากรณ์ สามารถแบ่งตามระยะเวลาของการพยากรณ์ได้ 3 ประเภทคือ

1) การพยากรณ์ในระยะสั้น เป็นการพยากรณ์เหตุการณ์ที่ไม่เกิน 1 ปี โดยทั่วไปมักจะอยู่ในช่วงไม่เกิน 3 เดือน เช่น การพยากรณ์การวางแผนจัดซื้อ การจัดตารางการทำงาน การมอบหมายงาน การพยากรณ์ยอดขาย และการพยากรณ์ระดับการผลิต

2) การพยากรณ์ระยะกลาง เป็นการพยากรณ์เหตุการณ์ที่อยู่ในช่วง 3 เดือน ถึง 3 ปี จะใช้มากในการพยากรณ์การวางแผนการขาย การวางแผนการผลิต การวางแผน ด้านงบประมาณเงินสด และการวิเคราะห์การวางแผนการดำเนินงานต่าง ๆ

3) การพยากรณ์ระยะยาว เป็นการพยากรณ์เหตุการณ์ที่มากกว่า 3 ปีขึ้นไปมักใช้สำหรับการวางแผนผลิตภัณฑ์ใหม่ ค่าใช้จ่ายในการลงทุน การขยายทำเลที่ตั้ง และการวิจัยพัฒนา

การวิเคราะห์อนุกรมเวลาเป็นเทคนิคทางสถิติที่ภาครัฐบาลและภาคธุรกิจสามารถนำไปใช้พยากรณ์ค่าของตัวแปรที่สนใจได้ เช่น ภาคธุรกิจใช้เทคนิคการวิเคราะห์อนุกรมเวลาในการพยากรณ์ยอดขาย ภาครัฐบาลใช้เทคนิคการวิเคราะห์อนุกรมเวลา ในการพยากรณ์อัตราเงินเฟ้อช่วงหลายสิบปี ที่ผ่านมานักเศรษฐมิติก็ได้มีการวิจัยและพัฒนาเทคนิคการวิเคราะห์อนุกรมเวลาให้ลึกซึ้งมากขึ้นเรื่อย ๆ เพื่อจะนำไปประยุกต์ใช้กับทฤษฎี ทางเศรษฐศาสตร์หรือทฤษฎีทางธุรกิจ และการเงินได้อย่างถูกต้องมากขึ้น เราจะเห็นว่ามีงานวิจัยทางเศรษฐศาสตร์ธุรกิจ และการเงินนำเทคนิคการวิเคราะห์อนุกรมเวลามาประยุกต์ใช้อย่างแพร่หลาย ดังนั้นควรทำความเข้าใจถึงเทคนิคการวิเคราะห์อนุกรมเวลา ให้ถูกต้องเพื่อที่จะนำมาใช้ประโยชน์ได้อย่างถูกต้องที่สุด

การวิเคราะห์อนุกรมเวลา เป็นการศึกษาถึงรูปแบบ และสาเหตุในการเคลื่อนไหวของข้อมูลประกอบด้วยส่วนสำคัญ 4 ส่วน คือ

1) ค่าแนวโน้ม (Secular Trend) ใช้สัญลักษณ์ T ค่าผันแปรตามฤดูกาล(Seasonal Variation) ใช้สัญลักษณ์ S การเคลื่อนไหวตามวัฏจักร (Cyclical Variation) ใช้สัญลักษณ์ C ส่วนสุดท้ายได้แก่ความผันแปรผิดปกติ (Irregular Variation) ใช้สัญลักษณ์ I โดยวีนัส ฤาชัย (2547) กล่าวถึงการวิเคราะห์ อนุกรมเวลาว่าเป็นการศึกษาเกี่ยวกับตัวแปรต่าง ๆ ที่เปลี่ยนแปลงไปตามเวลาหรือเป็นฟังก์ชั่นกับเวลาได้แก่ Y = f (t) โดยที่ Y คือตัวแปรตาม และ t คือตัวแปรอิสระ รูปแบบจำลองอนุกรมเวลา (Time series Model) ในทางเศรษฐศาสตร์นิยมใช้แบบจำลองในรูปผลคูณโดยที่ส่วนประกอบแต่ละส่วนมีผลกระทบต่อกันโดยวิธีแยกส่วนประกอบ (Decomposition Method) ทำได้โดยการแทนค่าพยากรณ์ของส่วนประกอบแต่ละส่วนลงไปในรูปแบบจำลองของส่วนประกอบอนุกรมเวลาโดยไม่มีค่าพยากรณ์ของ I ซึ่งเกิดขึ้นไม่แน่นอน Y = T*S*C ค่าพยากรณ์ของ T คำนวณได้โดยการแทนค่า Xt ของช่วงเวลาที่ต้องการพยากรณ์ไปในสมการแนวโน้มที่คำนวณขึ้นมาได้ค่าพยากรณ์ S ของแต่ละฤดูกาลจะมีรูปแบบซ้ำเดิมทุก ๆ รอบค่าพยากรณ์ C มักใช้วิจารณญาณของผู้วิเคราะห์ว่าควรจะมีค่าเพิ่มขึ้นหรือลดลงเท่าใด หรืออาจใช้วิธีคำนวณค่าเฉลี่ยของ C ของช่วงเวลานั้น ๆ ในอดีต

2) ข้อมูลอนุกรมเวลา คือ ข้อมูลที่เกิดขึ้นที่มีระยะห่างเท่า ๆ กัน และต่อเนื่องกัน

3) การวิเคราะห์อนุกรมเวลา คือ การศึกษาหารูปแบบการเปลี่ยนแปลงของตัวแปร ที่เปลี่ยนไปตามเวลาในอดีตปัจจุบัน แล้วนำรูปแบบนั้นมาวิเคราะห์เพื่อพยากรณ์ค่าของตัวแปรนั้นในอนาคตส่วนประกอบของอนุกรมเวลา แบ่งออกเป็น 4 ส่วน ได้แก่

3.1 แนวโน้ม (Trend) เป็นข้อมูลที่มีการเปลี่ยนแปลงในระยะเวลาที่นานพอที่จะเห็นแนวโน้มของข้อมูลว่าในอนาคตจะเพิ่มขึ้น หรือลดลง ระยะเวลาที่จะทำให้สามารถเห็นแนวโน้มจะต้องนานกว่า 1 ปี แนวโน้มอาจจะอยู่ในรูปเส้นตรง หรือเส้นโค้ง

3.2 ความผันแปรตามฤดูกาล (Seasonal Variation) หมายถึง การเปลี่ยนแปลงของข้อมูลที่เกิดขึ้นเนื่องจากอิทธิพลของฤดูกาล ซึ่งจะเกิดขึ้นซ้ำ ๆ กันช่วงเวลาเดียวกันของแต่ละปี

3.3 ความผันแปรตามวัฏจักร (Cyclical Variation) หมายถึง การเคลื่อนไหวของข้อมูลที่เกิดขึ้นซ้ำ ๆ กันในระยะเวลายาวมากกว่า 1 ปี

3.4 ความผันแปรที่ไม่แน่นอน (Irregular Variation) หรือความผันแปรอย่างสุ่ม (Random Variable) หมายถึง การเปลี่ยนแปลงของข้อมูลที่ไม่มีรูปแบบที่แน่นอน ดังนั้นจึงไม่สามารถคาดคะเน หรือพยากรณ์ความผันแปรที่ไม่แน่นอนโดยใช้ข้อมูลจากอดีตได้

2.1.3 แนวคิดเกี่ยวกับแผงหน้าปัด (Dashboard)

แผงหน้าปัด (Dashboard) คือการแสดงข้อมูลที่จะต้องใช้ในการบรรลุวัตถุประสงค์หนึ่งหรือมากกว่าหนึ่งฝากรวมและการจัดการภายในหน้าจอเดียวสามารถดูข้อมูลได้เพียงการมองครั้งเดียว (Pauwels et al., 2009) หรือการแสดงภาพผ่านหน้าจอให้ผู้ใช้งานโดยไม่ต้องมีตัวชี้วัดแผงหน้าปัดคือการแสดงภาพและการถ่ายภาพด้วยเครื่องมือการจัดการแสดงให้เห็นผ่านหน้าจอเดียวด้วยสารสนเทศเพื่อให้บรรลุวัตถุประสงค์ของผู้ใช้งานหรือวัตถุประสงค์ดังต่อไปนี้

1) ความสอดคล้อง (Consistency) เพื่อให้เกิดความสอดคล้องระหว่างการวัดผลและกระบวนการที่ใช้ในการวัดผล

2) การตรวจสอบ (Monitoring) เพื่อเป็นการตรวจสอบการดำเนินงาน หรืออาจเป็นทั้งการตรวจสอบและการประเมินว่าใครหรืออะไรที่ทำงานได้ดี

3) การวางแผน (Planning) ใช้ในการวางแผนถึงเป้าหมายและกลยุทธ์ในอนาคตที่ควรจะเป็นจากสถานะปัจจุบันขององค์กร

4) การสื่อสาร (Communication) เพื่อใช้เป็นเครื่องมือสื่อสารการดำเนินธุรกิจต่อหุ้นส่วนสำคัญ และไม่เฉพาะการดำเนินงานแต่รวมไปถึงการวัดเปรียบเทียบมูลค่าขององค์กร ในการดำเนินการแผงหน้าปัดที่มีประสิทธิภาพนั้นไม่จำเป็นต้องให้ผู้ใช้ดูทุกรายละเอียดว่าเกิด

อะไรขึ้นบ้าง แต่จะแสดงตัวชี้วัดผลการดำเนินงานหลักที่เป็นตัวเลข ซึ่งเป็นการแสดงยอดสรุป มีส่วนที่ให้ผู้ใช้สามารถดูข้อมูลเพิ่มเติมได้เองแผงหน้าปัดที่มีคุณสมบัติพิเศษไม่เหมือนใครเกิดอะไรขึ้นบ้าง แต่จะแสดงตัวชี้วัดผลการดำเนินงานมีส่วนที่ให้ผู้ใช้สามารถดูข้อมูลเพิ่มเติมได้เอง (Domo Inc, 2012) กล่าวคือแผงหน้าปัดคือการแสดงหน่วยงานที่เป็นตัวชี้วัดผลการปฏิบัติงานหลักผู้ใช้งานสามารถมองและเข้าใจข้อมูลได้อย่างรวดเร็วและถูกต้องของผู้ใช้ส่งผลให้การดำเนินงานภายในองค์กรมีประสิทธิภาพมากขึ้น

2.1.4 แนวคิดเกี่ยวกับการแสดงผลข้อมูล (Data visualization)

Data Visualization หรือ Information Visualization คือ การถ่ายทอดข้อมูลในเชิงปริมาณที่มีความซับซ้อนได้อย่างมีประสิทธิผล ซึ่งคำว่า “ประสิทธิผล” ในที่นี้หมายถึงมีความชัดเจน (Clarity), มีความแม่นยำ (Precision), และมีประสิทธิภาพ (Efficiency) หากไม่มีการทำData Visualization อาจทำให้เราไม่สามารถค้นพบนัยยะของข้อมูลในแง่ของแนวโน้ม, รูปแบบพฤติกรรม และความสัมพันธ์เชื่อมโยงได้ ส่วนผสมที่สำคัญในการทำให้เกิดประสิทธิผลในการทำ Visualization ได้แก่

1) Visual cues / Visual encoding: การแปลงข้อมูลให้แสดงในลักษณะของรูปร่าง, สี หรือขนาด เช่น ถ้าลองดูกราฟด้านบน จะสังเกตได้ว่าจุดแต่ละจุดแสดงข้อมูลจำนวนโฮมรันในแต่ละปี หรือเส้นที่ลากระหว่างแต่ละจุดแสดงถึงอัตราการเปลี่ยนแปลงของจำนวนโฮมรันจากปีหนึ่งไปอีกปีหนึ่ง

2) Coordinate systems: หมายถึงระบบพิกัดที่ช่วยทำให้เกิดพื้นที่ที่มีโครงสร้างสำหรับ ใส่ข้อมูลที่ต้องการแสดง รูปแบบพื้นฐานที่เราเห็นได้ทั่วไป

3) Scale / Data types: จะต้องสอดคล้องกับระบบพิกัดข้างต้น โดย Data types พื้นฐานนั้นมีอยู่ 3 ประเภทได้แก่ ข้อมูลเชิงตัวเลข (Numeric), ข้อมูลที่เป็นหมวดหมู่ (Categorical) และข้อมูลตามลำดับเวลา (Time Series)

4) Context: ในกรณีที่ผู้รับสารไม่คุ้นเคยกับข้อมูล เป็นหน้าที่ของ Data Scientist ที่จะต้องระบุบริบทของข้อมูลให้ชัดเจน และอธิบายว่าควรจะอ่านข้อมูลจากกราฟอย่างไร อย่างเช่น การตั้งชื่อกราฟ, ระบุแกน X แกน Y ว่าคืออะไร, หรือการระบุเหตุการณ์บางอย่างที่สำคัญลงบนแกนลำดับเวลา เพื่อให้ผู้อ่านสามารถตีความและเข้าใจได้มากขึ้น

การเลือกรูปแบบ Visualization ให้เหมาะสมกับข้อมูล ในปัจจุบันเป็นยุคที่เทคโนโลยีเข้าถึงทุกคน ทำให้การรับรู้ข่าวสาร ข้อมูลต่าง ๆ เป็นไปได้ง่าย และรวดเร็วมากขึ้นคนที่นำเสนอข้อมูลจึงต้องนำเสนอข้อมูลให้น่าสนใจ เข้าใจง่าย และรวดเร็ว จึงเกิดการสร้าง Data Visualization ขึ้นมา Data Visualization เป็นการใช้ภาพเพื่อแสดงข้อมูลในเชิงปริมาณที่วัด

ได้ซึ่งอาจนำเสนอออกมาในรูปแบบ แผนภูมิ กราฟ กราฟิก และอื่น ๆ อีกมากมาย เพื่อให้เข้าใจได้ง่าย และรวดเร็ว การเลือกรูปแบบ Visualization ให้เหมาะสมกับข้อมูลที่มีอยู่ เพื่อให้การนำเสนอข้อมูลน่าสนใจและมีประสิทธิภาพสูงสุด ซึ่งเลือกแผนภูมิยอดนิยมที่ใช้กันบ่อย ๆ ดังนี้

1) แผนภูมิแท่ง (Bar Charts) เป็นแผนภูมิที่ประกอบด้วยแกนนอน แกนตั้ง ที่นิยมแสดงออกมาในรูปแท่งสี่เหลี่ยมที่สามารถบอกความสูงได้ เหมาะสำหรับใช้การเปรียบเทียบจำนวนของข้อมูลในแต่ละชุด เช่น รายรับในแต่ละเดือน, ยอดขายที่ขายได้จริงเปรียบเทียบกับเป้าหมายยอดขายที่ตั้งไว้ เป็นต้น ซึ่งแผนภูมิแท่งยังสามารถแบ่งออกเป็นหลายประเภท

1.1 แผนภูมิแท่งแบบจัดกลุ่มแผนภูมิแท่งแบบจัดกลุ่ม เหมาะสมกับการนำเสนอข้อมูลที่มีข้อมูลย่อย ๆ อยู่ภายใต้ข้อมูลใหญ่เป็นการเน้นให้เห็นข้อมูลย่อยนั้น ๆ

1.2 แผนภูมิแท่งแบบวางซ้อนกัน แผนภูมิแท่งแบบวางซ้อนกัน เหมาะสมกับการนำเสนอข้อมูลที่ต้องการให้เห็นข้อมูลย่อยในแต่ละข้อมูลใหญ่และยังแสดงให้เห็นสัดส่วนของข้อมูลย่อยต่าง ๆ เหล่านั้นได้ด้วย ใช้แผนภูมินี้เมื่อมีชุดข้อมูล หลายชุดและต้องการเน้นผลรวมทั้งหมด

2) แผนภูมิเส้น (Line Charts) แผนภูมิเส้น มีลักษณะคล้ายแผนภูมิแท่ง ซึ่งประกอบด้วยแกนตั้งและนอน เพียงแต่เปลี่ยนจากแท่งข้อมูลเป็นจุดบนแผนภูมินั่นเอง แผนภูมิประเภทนี้เหมาะกับการนำเสนอข้อมูลตัวเลขที่มีความสัมพันธ์กับข้อมูลที่มีลักษณะเป็นช่วงใช้แสดงให้เห็นการเปลี่ยนแปลงที่เกิดขึ้นตามเวลาเพื่อดูแนวโน้ม

3) แผนภูมิวงกลม (Pie Charts) แผนภูมิวงกลมเหมาะกับการนำเสนอข้อมูลที่มีส่วนประกอบย่อยที่รวมกันเป็นส่วนใหญ่ มีการแบ่งส่วนให้ดูง่าย และสวยงามแต่ในทางกลับกันอาจจะดูยากในเรื่องของการประมาณขนาดของแต่ละชิ้น ยิ่งถ้ามีจำนวนชิ้นมาก จะยิ่งแยกยากเพราะต้องใช้หลายสี ในการนำเสนอข้อมูล

4) แผนภูมิโดนัท (Doughnut Charts) แผนภูมิโดนัทมีหลักการออกแบบเช่นเดียวกับแผนภูมิวงกลมแต่สามารถแสดงชุดข้อมูลได้มากกว่า 1 ชุด โดยนำเสนอข้อมูลเป็นวงกลมซ้อนกันหลาย ๆ

5) แผนภูมิพื้นที่ (Area Charts) มีหน้าตาคล้ายแผนภูมิเส้น แต่มีการแรเงาพื้นที่ใต้เส้นข้อมูล หรือระหว่าง 2 เส้นเพื่อแสดงให้เห็นปริมาณความแตกต่างระหว่างเส้นเหมาะสำหรับเน้นความสำคัญของการเปลี่ยนแปลงในช่วงเวลาแสดงให้เห็นผลรวมของความแตกต่างระหว่างข้อมูล

6) แผนภูมิเรดาร์ (Radar Charts) มีลักษณะคล้ายแผนภูมิเส้นที่มีการแสดงผลแบบวงกลมจำนวนเหลี่ยมของเรดาร์เท่ากับจำนวนหัวข้อของข้อมูล แผนภูมินี้ไม่ได้บอกถึงความต่อเนื่องของข้อมูลแต่เหมาะสำหรับการนำเสนอข้อมูลเป็นหัวข้อ แล้วนำมาวิเคราะห์หาจุดอ่อน จุดแข็งของข้อมูล

7) แผนภูมิต้นไม้ (Tree Maps) คือการนำเสนอข้อมูลแบบแสดงให้เห็นพื้นที่แสดงผลได้ในแบบลำดับชั้น เหมือนแบบโครงสร้างต้นไม้ อาจจะนำเสนอข้อมูลที่ต้องการให้เห็นถึงเขตพื้นที่ แสดงพื้นที่สีที่แตกต่างกันได้

8) แผนภูมิรูปภาพ (Picture Graph) เป็นแผนภูมิที่ประกอบไปด้วยแกนนอนและแกนตั้ง แต่เลือกใช้รูปภาพ หรือไอคอนแทนจำนวนของสิ่งของนั้น ๆ เช่น การแสดงผลจำนวนของนมที่ขายได้ในแต่ละเดือน โดยนำเสนอทั้งนมรสจืด รสช็อกโกแลต เปรียบเทียบในแต่ละเดือนซึ่งมีการนำเสนอลักษณะคล้ายกับกราฟแท่ง แต่เปลี่ยนจากแท่งเป็นรูปกราฟิกของนม 2 รสชาติแทน ก็ทำให้การนำเสนอข้อมูลน่าสนใจมากยิ่งขึ้น

2.2 ทฤษฎี

2.2.1 ทฤษฎีเกี่ยวกับการออกแบบเว็บไซต์

เว็บไซต์เป็นสื่อที่ได้รับความนิยมอย่ามากบนอินเตอร์เน็ต ซึ่งเว็บไซต์เป็นสื่อที่อยู่ในความควบคุมของผู้ใช้โดยสมบูรณ์ กล่าวคือ ผู้ใช้สามารถตัดสินใจเลือกได้ว่าจะดูเว็บไซต์ใดและจะไม่เลือกดูเว็บไซต์ใด ได้ตามต้องการ จึงทำให้ผู้ใช้ไม่มีความอดทนต่ออุปสรรคและปัญหาที่เกิดจากการออกแบบเว็บไซต์ผิดพลาดถ้าผู้ใช้เห็นว่าเว็บที่กำลังดูอยู่นั้นไม่มีประโยชน์ต่อตัวเขาหรือไม่เข้าใจว่าเว็บไซต์นี้จะใช้งานอย่างไร เขาก็สามารถที่จะเปลี่ยนไปดูเว็บไซต์อื่น ๆ ได้อย่างรวดเร็ว เนื่องจากในปัจจุบันมีเว็บไซต์อยู่มากมาย และยังมีเว็บไซต์ที่เกิดขึ้นใหม่ ๆ ทุกวัน ผู้ใช้จึงมีทางเลือกมากขึ้น และสามารถเปรียบเทียบคุณภาพของเว็บไซต์ต่าง ๆ ได้เอง

2.2.1.1 องค์ประกอบของการออกแบบเว็บไซต์

1) ความเรียบง่าย (Simplicity) หมายถึง การจำกัดองค์ประกอบเสริมให้เหลือเฉพาะองค์ประกอบหลัก กล่าวคือในการสื่อสารเนื้อหากับผู้ใช้นั้น เราต้องเลือกเสนอสิ่งที่เราต้องการนำเสนอจริง ๆ ออกมาในส่วนของกราฟิก สีสัน ตัวอักษรและภาพเคลื่อนไหวต้องเลือกให้พอเหมาะ

2) ความสม่ำเสมอ (Consistency) หมายถึง การสร้างความสม่ำเสมอให้เกิดขึ้นตลอดทั้งเว็บไซต์ โดยอาจเลือกใช้รูปแบบเดียวกันตลอดทั้งเว็บไซต์ก็ได้ เพราะถ้าหา

กว่าแต่ละหน้าในเว็บไซต์นั้นมีความแตกต่างกันมากจนเกินไป อาจทำให้ผู้ใช้เกิดความสับสนและไม่แน่ใจว่ากำลังอยู่ในเว็บไซต์เดิมหรือไม่

3) ความเป็นเอกลักษณ์ (Identity) ในการออกแบบเว็บไซต์ต้องคำนึงถึงลักษณะขององค์กรเป็นหลัก เนื่องจากเว็บไซต์จะสะท้อนถึงเอกลักษณ์และลักษณะขององค์กร การเลือกใช้ตัวอักษร ชุดสี รูปภาพหรือกราฟิก จะมีผลต่อรูปแบบของเว็บไซต์เป็นอย่างมาก ตัวอย่างเช่น ถ้าเราต้องออกแบบเว็บไซต์ของธนาคารแต่เรากลับเลือกสีสันและกราฟิกมากมาย อาจทำให้ผู้ใช้คิดว่าเป็นเว็บไซต์ของสวนสนุกซึ่งส่งผลต่อความเชื่อถือขององค์กรได้

4) เนื้อหา (Useful Content) ถือเป็นสิ่งสำคัญที่สุดในเว็บไซต์ เนื้อหาในเว็บไซต์ต้องสมบูรณ์และได้รับการปรับปรุงพัฒนาให้ทันสมัยอยู่เสมอ ผู้พัฒนาต้องเตรียมข้อมูล และเนื้อหาที่ผู้ใช้ต้องการให้ถูกต้องและสมบูรณ์ เนื้อหาที่สำคัญที่สุดคือเนื้อหาที่ทีมผู้พัฒนาสร้างสรรค์ขึ้นมาเอง และไม่ไปซ้ำกับเว็บอื่น

5) ระบบเนวิเกชัน (User–Friendly Navigation) เป็นส่วนประกอบที่มีความสำคัญต่อเว็บไซต์มาก เพราะจะช่วยไม่ให้ผู้ใช้เกิดความสับสนระหว่างดูเว็บไซต์ ระบบเนวิเกชัน จึงเปรียบเสมือนป้ายบอกทาง

6) คุณภาพของสิ่งที่ปรากฏให้เห็นในเว็บไซต์ (Visual Appeal) ลักษณะที่น่าสนใจของเว็บไซต์นั้น ขึ้นอยู่กับความชอบส่วนบุคคลเป็นสำคัญ แต่โดยรวมแล้วก็สามารถสรุปได้ว่าเว็บไซต์ที่น่าสนใจนั้นส่วนประกอบต่าง ๆ ควรมีคุณภาพ

7) ความสะดวกของการใช้ในสภาพต่าง ๆ (Compatibility) การใช้งานของเว็บไซต์นั้นไม่ควรมีขอบจำกัด กล่าวคือ ต้องสามารถใช้งานได้ดีในสภาพแวดล้อมที่หลากหลาย ไม่มีการบังคับให้ผู้ใช้ต้องติดตั้งโปรแกรมอื่นใดเพิ่มเติม นอกเหนือจากเว็บบราวเซอร์

8) ความคงที่ในการออกแบบ (Design Stability) ถ้าต้องการให้ผู้ใช้งานรู้สึกว่าเว็บไซต์มีคุณภาพ ถูกต้อง และเชื่อถือได้ ควรให้ความสำคัญกับการออกแบบเว็บไซต์เป็นอย่างมาก

9) ความคงที่ของการทำงาน (Function Stability) ระบบการทำงานต่าง ๆ ในเว็บไซต์ควรมีความถูกต้องแน่นอน ซึ่งต้องได้รับการออกแบบสร้างสรรค์และตรวจสอบอยู่เสมอ ตัวอย่างเช่น ลิงค์ต่าง ๆ ในเว็บไซต์

2.2.1.2 การออกแบบเว็บไซต์

ในการออกแบบเว็บไซต์นั้นประกอบด้วยกระบวนการต่าง ๆ มากมาย เช่น การออกแบบโครงสร้าง ลักษณะหน้าตา หรือการเขียนโปรแกรม แต่มีหลายคนที่พัฒนาเว็บไซต์ โดยขาดการวางแผนและทำงานไม่เป็นระบบ ตัวอย่างเช่น การลงมือออกแบบ

โดยการใช้โปรแกรมช่วยสร้างเว็บ เนื้อหาและรูปแบบก็เป็นไปตามที่นึกขึ้นได้ขณะนั้น และ เมื่อเห็นว่าดูดีแล้วก็เปิดตัวเลย ทำให้เว็บนั้นมีเป้าหมายและแนวทางที่ไม่แน่นอน ผลลัพธ์ที่ได้จึงเสี่ยง กับความล้มเหลวค่อนข้างมาก กระบวนการแรกของการออกแบบเว็บไซต์ คือการกำหนดเป้าหมายของเว็บไซต์กำหนดกลุ่มผู้ใช้ ซึ่งการจะให้ได้มา ซึ่งข้อมูลผู้พัฒนาต้องเรียนรู้ผู้ใช้ หรือจำลองสถานการณ์ สิ่งเหล่านี้จะช่วยให้เราสามารถออกแบบเนื้อหาและการใช้งานเว็บไซต์ได้อย่างเหมาะสม ตรงกับความต้องการของผู้ใช้อย่างแท้จริง

1) กำหนดเป้าหมายของเว็บไซต์

2) กำหนดกลุ่มผู้ใช้เป้าหมาย

2.2.1.3 ออกแบบหน้าเว็บไซต์ (Page Design)

หลักสำคัญในการออกแบบหน้าเว็บก็คือ การใช้รูปภาพและองค์ประกอบต่าง ๆ ร่วมกันเพื่อสื่อความหมาย เกี่ยวกับเนื้อหาหรือลักษณะสำคัญของเว็บไซต์ โดยมีเป้าหมายสำคัญเพื่อการสื่อความหมายที่ชัดเจนและน่าสนใจ บนพื้นฐานของความเรียบง่ายและความสะดวกของผู้ใช้ ดังนี้

1) ความเรียบง่าย ได้แก่ มีรูปแบบที่เรียบง่าย ไม่ซับซ้อน และใช้งานได้สะดวก

2) ความสม่ำเสมอ ได้แก่ ใช้รูปแบบเดียวกันตลอดทั้งเว็บไซต์

3) ความเป็นเอกลักษณ์ การออกแบบเว็บไซต์ควรคำนึงถึงลักษณะขององค์กร เพราะรูปแบบของเว็บไซต์จะสะท้อนถึงเอกลักษณ์และลักษณะขององค์กรนั้น ๆ เช่น ถ้าเป็นเว็บไซต์ของทาง ราชการ จะต้องดูน่าเชื่อถือไม่เหมือนสวนสนุก ฯลฯ

4) เนื้อหาที่มีประโยชน์ เนื้อหาเป็นสิ่งที่สำคัญที่สุดในเว็บไซต์

5) ระบบเนวิเกชันที่ใช้งานง่าย ต้องออกแบบให้ผู้ใช้เข้าใจง่ายและใช้งานสะดวก ใช้กราฟิกที่สื่อความหมายร่วมกับคำอธิบายที่ชัดเจน

6) ลักษณะที่น่าสนใจ หน้าตาของเว็บไซต์จะต้องมีความสัมพันธ์กับคุณภาพขององค์ประกอบต่าง ๆ เช่น คุณภาพของกราฟิกที่จะต้องสมบูรณ์ การใช้สี

7) คุณภาพในการออกแบบ การออกแบบและเรียบเรียงเนื้อหาอย่างรอบคอบ สร้างความรู้สึกว่าเว็บไซต์มีคุณภาพ ถูกต้อง และเชื่อถือได้

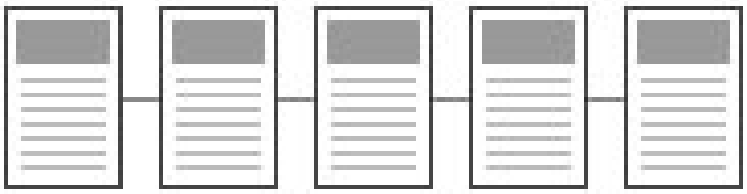
8) การใช้งานอย่างไม่จำกัด ผู้ใช้ส่วนใหญ่สามารถเข้าถึงได้มากที่สุดเลือกบราวเซอร์ชนิดใดก็ได้ในการเข้าถึงเนื้อหาสามารถแสดงผลได้ทุกระบบปฏิบัติการ

9) ลิงค์ต่าง ๆ จะต้องเชื่อมโยงไปหน้าที่มีอยู่จริงและถูกต้อง ระบบการทำงาน ต่าง ๆ ในเว็บไซต์จะต้องมีความแน่นอนและทำหน้าที่ได้อย่างถูก

2.2.1.4 การออกแบบโครงสร้างเว็บไซต์ (Site Structure Design)

โครงสร้างเว็บไซต์ (Site Structure) การออกแบบโครงสร้างของเว็บไซต์สามารถทำได้หลากหลายแบบ ซึ่งก็ขึ้นอยู่กับความชอบและความถนัดของแต่ละบุคคล นอกจากนี้ยังขึ้นอยู่กับกลุ่มเป้าหมายที่ต้องการนำเสนอ เพราะจะต้องออกแบบให้เหมาะกับการใช้งานของกลุ่มเป้าหมายมากที่สุด โดยโครงสร้างของเว็บไซต์ส่วนใหญ่ก็จะประกอบไปด้วย 4 รูปแบบดังนี้

1) โครงสร้างแบบเรียงลำดับ (Sequential Structure) จะเป็นโครงสร้างแบบธรรมดาที่นิยมใช้งานกันมากที่สุด เนื่องจากมีความง่ายต่อการจัดระบบข้อมูล และสามารถนำเสนอเรื่องราวตามลำดับได้เป็นอย่างดี เหมาะกับเว็บไซต์ที่มีขนาดเล็ก มีเนื้อหาที่ไม่ซับซ้อน ส่วนใหญ่ก็จะเป็นพวกเว็บไซต์ที่ให้ความรู้ หรือเว็บไซต์องค์กรขนาดย่อม โดยลักษณะการลิ้งค์เนื้อหา ก็จะลิ้งค์ไปทีละหน้า มีทิศทางการเข้าสู่เนื้อหาต่าง ๆ ในแบบเส้นตรง ใช้ปุ่มเดินหน้าถอยหลังในการกำหนดทิศทาง จึงทำให้การใช้งานเป็นไปอย่างง่าย แต่โครงสร้างเว็บไซต์แบบเรียงลำดับก็มีข้อเสีย คือจะทำให้ผู้ใช้งานต้องเสียเวลาในการเข้าสู่เนื้อหาเพราะไม่สามารถกำหนดทิศทางการเข้าสู่เนื้อหาด้วยตัวเองได้



ภาพที่ 2.1 เว็บที่มีโครงสร้างแบบเรียงลำดับ (Sequential Structure)
ที่มา : https://www.1belief.com/article/website-design/

2) โครงสร้างแบบลำดับขั้น (Hierarchical Structure) นิยมใช้กับเว็บที่มีความซับซ้อนของข้อมูล เพื่อให้สามารถเข้าถึงข้อมูลต่าง ๆ ได้ง่ายขึ้น โดยจะมีการแบ่งเนื้อหาออกเป็นส่วนๆ และมีการนำเสนอรายละเอียดย่อย ๆ ที่ลดหลั่นกันมา ทำให้สามารถทำความข้าใจกับโครงสร้างเนื้อหาได้ง่ายขึ้น โดยจะมีโฮมเพจเป็นจุดเริ่มต้น และจุดร่วมจุดเดียวที่จะนำไปสู่การเชื่อมโยงเนื้อหาเป็นลำดับจากบนลงล่าง

ภาพที่ 2.2 เว็บที่มีโครงสร้างแบบลำดับขั้น (Hierarchical Structure)
ที่มา : https://www.1belief.com/article/website-design/

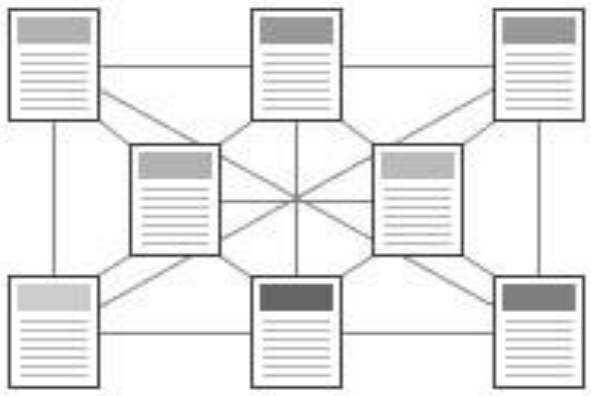
3) โครงสร้างแบบตาราง (Grid Structure) เป็นโครงสร้างการออกแบบเว็บไซต์ที่มีความซับซ้อน แต่ก็มีความยืดหยุ่นในระดับหนึ่ง เพื่อให้ผู้ใช้งานสามารถเข้าสู่เนื้อหาต่าง ๆ ได้ง่ายขึ้น การออกแบบในลักษณะนี้จะมีการเชื่อมโยงเนื้อหาในแต่ละส่วนซึ่งกันและกัน ทำให้ผู้ใช้งานสามารถเปลี่ยนทิศทาง หรือกำหนดทิศทางในการเข้าสู่เนื้อหาด้วยตัวเองได้ จึงไม่ทำให้เสียเวลา แถมยังทำให้เว็บไซต์มีความทันสมัยขึ้น

ภาพที่ 2.3 เว็บที่มีโครงสร้างแบบตาราง (Grid Structure)
ที่มา : https://www.1belief.com/article/website-design/

ในการจัดระบบโครงสร้างแบบนี้ เนื้อหาที่นำมาใช้แต่ละส่วนควรมีลักษณะที่เหมือนกัน และ สามารถใช้รูปแบบร่วมกัน หลักการออกแบบคือนำหัวข้อทั้งหมดมาบรรจุลงในที่เดียวกันซึ่งโดยทั่วไป จะเป็นหน้าแผนภาพ (Map Page) ที่แสดงในลักษณะเดียวกับโครงสร้างของเว็บ เมื่อผู้ใช้คลิกเลือก หัวข้อใด ก็จะเข้าไปสู่หน้าเนื้อหา (Topic Page) ที่แสดงรายละเอียด

ของหัวข้อนั้น ๆ และภายในหน้านั้น ก็จะมีการเชื่อมโยงไปยังหน้ารายละเอียดของหัวข้ออื่นที่เป็นเรื่องเดียวกัน นอกจากนี้ยังสามารถนำโครงสร้างแบบเรียงลำดับและแบบลำดับขั้นมาใช้ร่วมกันได้อีกด้วย ถึงแม้โครงสร้างแบบนี้ อาจจะสร้างความยุ่งยากในการเข้าใจได้และอาจเกิดปัญหาการคงค้าง ของหัวข้อ (Cognitive Overhead) ได้ แต่จะเป็นประโยชน์ที่สุดเมื่อผู้ใช้ได้เข้าใจถึงความสัมพันธ์ ระหว่างเนื้อหา ในส่วนของการออกแบบจ าเป็นจะต้องมีการวางแผนที่ดีเนื่องจากมีการเชื่อมโยงที่เกิดขึ้น ได้หลายทิศทาง นอกจากนี้การปรับปรุงแก้ไขอาจเกิดความยุ่งยากเมื่อต้องเพิ่มเนื้อหาในภายหลัง

4) เว็บที่มีโครงสร้างแบบใยแมงมุม (Web Structure) โครงสร้างประเภทนี้จะมีความยืดหยุ่นมากที่สุด ทุกหน้าในเว็บสามารถจะเชื่อมโยงไปถึงกัน ได้หมด เป็นการสร้างรูปแบบการเข้าสู่เนื้อหาที่เป็นอิสระ ผู้ใช้สามารถก าหนดวิธีการเข้าสู่เนื้อหาได้ด้วยตนเองการเชื่อมโยงเนื้อหาแต่ละหน้าอาศัยการโยงใยข้อความที่มีมโนทัศน์ (Concept) เหมือนกันของแต่ละหน้าในลักษณะของไฮเปอร์เท็กซ์หรือไฮเปอร์มีเดีย โครงสร้างลักษณะนี้จัดเป็นรูปแบบที่ไม่มีโครงสร้างที่แน่นนอนตายตัว (Unstructured) นอกจากนี้การเชื่อมโยงไม่ได้จำกัดเฉพาะเนื้อหาภายในเว็บนั้น ๆ แต่สามารถเชื่อมโยงออกไปสู่เนื้อหาจากเว็บภายนอกได้



ภาพที่ 2.4 เว็บที่มีโครงสร้างแบบใยแมงมุม (Web Structure)

ที่มา : https://www.1belief.com/article/website-design/

2.2.1.5 การใช้สีในการออกแบบเว็บไซต์

การสร้างสีสันบนหน้าเว็บเป็นสิ่งที่สื่อความหมายของเว็บไซต์ได้อย่างชัดเจน การเลือกใช้สีให้เหมาะสมกลมกลืน ไม่เพียงแต่จะสร้างความพึงพอใจให้กับผู้ใช้แต่ยังสามารถทำให้เห็นถึงความแตกต่างระหว่างเว็บไซต์ได้ สีเป็นองค์ประกอบหลักสำหรับการตกแต่งเว็บจึงจำเป็นอย่างยิ่งที่จะต้องทำความเข้าใจเกี่ยวกับการใช้สีระบบสีที่แสดงบนจอคอมพิวเตอร์ มี

ระบบการแสดงผลผ่านหลอดลำแสงที่เรียกว่า CRT (Cathode ray tube) โดยมีลักษณะระบบสีแบบบวก อาศัยการผสมของของแสงสีแดง สีเขียว และสีน้ำเงิน หรือระบบสี RGB สามารถกำหนดค่าสีจาก 0 ถึง 255 ได้ จากการรวมสีของแม่สีหลักจะทำให้เกิดแสงสีขาว มีลักษณะเป็นจุดเล็ก ๆ บนหน้าจอไม่สามารถมองเห็นด้วยตาเปล่าได้ จะมองเห็นเป็นสีที่ถูกผสมเป็นเนื้อสีเดียวกันแล้ว จุดแต่ละจุดหรือพิกเซล (Pixel) เป็นส่วนประกอบของภาพบนหน้าจอคอมพิวเตอร์ โดยจำนวนบิตที่ใช้ในการกำหนดความสามารถของการแสดงสีต่าง ๆ เพื่อสร้างภาพบนจอนั้นเรียกว่า บิตเด็ป (Bit–depth) ในภาษา HTML มีการกำหนดสีด้วยระบบเลขฐานสิบหก ซึ่งมีเครื่องหมาย (#) อยู่ด้านหน้าและตามด้วยเลขฐานสิบหกจำนวนอักษรอีก 6 หลัก โดยแต่ละไบต์ (byte) จะมีตัวอักษรสองตัว แบ่งออกเป็น 3 กลุ่ม เช่น #FF12AC การใช้ตัวอักษรแต่ละไบต์นี้เพื่อกำหนดระดับความเข้มของแม่สีแต่ละสีของชุดสี RGB โดย 2 หลักแรก แสดงถึงความเข้มของสีแดง 2 หลักต่อมา แสดงถึงความเข้มของสีเขียว 2 หลักสุดท้ายแสดงถึงความเข้มของสีน้ำเงิน

สีมีอิทธิพลในเรื่องของอารมณ์การสื่อความหมายที่เด่นชัด กระตุ้นการรับรู้ทางด้านจิตใจมนุษย์ สีแต่ละสีให้ความรู้สึก อารมณ์ที่ไม่เหมือนกัน สีบางสีให้ความรู้สึกสงบ บางสีให้ความรู้สึกตื่นเต้นรุนแรง สีจึงเป็นปัจจัยสำคัญอย่างยิ่งต่อการออกแบบเว็บไซต์ ดังนั้นการเลือกใช้โทนสีภายในเว็บไซต์เป็นการแสดงถึงความแตกต่างของสีที่แสดงออกทางอารมณ์ มีชีวิตชีวาหรือเศร้าโศก รูปแบบของสีที่สายตาของมนุษย์มองเห็น สามารถแบ่งออกเป็น 3 กลุ่ม คือ

1) สีโทนร้อน (Warm Colors) เป็นกลุ่มสีที่แสดงถึงความสุขความปลอบโยน ความอบอุ่น และดึงดูดใจ สีกลุ่มนี้เป็นกลุ่มสีที่ช่วยให้หายจากความเฉื่อยชามีชีวิตชีวามากยิ่งขึ้น

2) สีโทนเย็น (Cool Colors) แสดงถึงความที่ดูสุภาพ อ่อนโยนเรียบร้อย เป็นกลุ่มสีที่มีคนชอบมากที่สุด สามารถโน้มนาวในระยะไกลได้

3) สีโทนกลาง (Neutral Colors) สีที่เป็นกลาง ประกอบด้วย สีดำ สีขาว สีเทา และสีน้ำตาล กลุ่มสีเหล่านี้คือ สีกลางที่สามารถนำไปผสมกับสีอื่น ๆ เพื่อให้เกิดสีกลางขึ้นมา

2.2.2 ทฤษฎีเกี่ยวกับอนุกรมเวลา

การวิเคราะห์ข้อมูลอนุกรมเวลาจะขึ้นอยู่กับการเปลี่ยนแปลงของเวลาในอดีตเป็นพื้นฐานกล่าวคือ การวิเคราะห์อนุกรมเวลาเป็นวิธีที่ใช้ในการวิเคราะห์ข้อมูลหรือค่าสังเกตที่มีการเปลี่ยนแปลงไปตามลำดับเวลาที่เกิดขึ้น หรือการเปลี่ยนแปลงของตัวแปรในช่วงเวลาในอดีต ที่อนุกรมเวลามีรูปแบบการเปลี่ยนแปลงในอดีตนั้น ทำให้สามารถคาดการณ์การเปลี่ยนแปลงในอนาคตได้ว่าในอนาคตลักษณะการเปลี่ยนแปลงควรอยู่ในทิศทางการเคลื่อนไหวแบบใด หรือรูปแบบใดได้ข้อมูลอนุกรมเวลาที่มีลักษณะนิ่ง (Stationary) คือ ข้อมูลที่ค่าเฉลี่ยและความแปรปรวนของกระบวนการเชิงสุ่ม (Random process) นั้นมีค่าคงที่เมื่อเวลาได้เปลี่ยนไปและค่าความแปรปรวนระหว่างสองคาบเวลาขึ้นอยู่กับความล่า (lag) ระหว่างคาบเวลาทั้งสองนั้นโดยเขียนเป็นสมการได้ดังนี้ การวัดความถูกต้องเป็นฟังก์ชั่นของค่าคลาดเคลื่อนของการพยากรณ์โดยไม่คำนึงถึงทิศทางนั่น คือพิจารณาค่า $|e_t|, e_t^2|\frac{|et|}{yt}$ แทนการพิจารณาค่าคลาดเคลื่อนธรรมดาโดยฟังก์ชั่นของค่าคลาดเคลื่อนดังกล่าวจะใช้ช่วยวัดความถูกต้อง ได้แก่ค่า *MAPE* ตามลำดับ การที่ค่าเหล่านี้มีค่าน้อยเป็นผลเนื่องมาจากค่าคลาดเคลื่อนของการพยากรณ์ (e_t) ต่ำที่สุด ดังนี้ MAPE (Mean Absolute Percentage Error) เป็นค่าวัดความถูกต้องของการพยากรณ์ที่วัดจากขนาดของค่าความคลาดเคลื่อนของการพยากรณ์เทียบกับค่าจริง โดยไม่คิดเครื่องหมาย ค่า MAPE เป็นค่าวัดความแม่นยำที่ไม่มีหน่วย จึงเหมาะที่จะใช้กับการเปรียบเทียบวิธีการพยากรณ์เดียวกันหรือเปรียบเทียบวิธีพยากรณ์หลายวิธีเมื่อใช้อนุกรมเวลาชุดเดียวกันต่างกับการใช้ค่า MAD ตรงที่หากข้อมูลจริงเป็น 0 การวัดประสิทธิภาพการพยากรณ์โดยใช้ MAPE จะไม่มีประสิทธิภาพเท่ากับการใช้ค่า MAD ซึ่งการวัดค่า MAPE มีสูตรดังต่อไปนี้

$$\text{MAPE} = \sum \frac{\sum_{t=1}^{n}|\frac{e_t}{y_t}|}{n}\times 100$$

เมื่อ e_t คือ ความคลาดเคลื่อน y_t คือ ข้อมูลจริง และ n คือ จำนวนข้อมูลค่า MAPE เป็นค่าที่เหมาะกับการนำไปใช้ในการเปรียบเทียบกับอนุกรมเวลาต่างชุดกันเมื่อพยากรณ์ด้วยวิธีการพยากรณ์เดียวกันค่าวัดความถูกต้องที่ใช้กันอยู่มีหลายค่าและยังไม่มีผลการวิจัยใดที่ระบุว่าค่าวัดความถูกต้องใดเหมาะสมที่สุดซึ่งในการศึกษาครั้งนี้จะใช้ค่า RMSE

ในการตัดสินใจเลือกวิธีพยากรณ์เพราะเป็นค่าที่ให้ความคลาดเคลื่อนที่สม่ำเสมอในทุก ๆ รหัสเวลา

2.2.2.1 องค์ประกอบของอนุกรมเวลา

ในการวิเคราะห์อนุกรมเวลา ผู้วิเคราะห์จะแยกองค์ประกอบต่าง ๆที่ประกอบกันขึ้นเป็นอนุกรมเวลา โดยจะมีการเปลี่ยนแปลงไปตามอิทธิพลต่าง ๆ เช่น การเปลี่ยนแปลงการผลิต เทคโนโลยี สภาวะอากาศ เป็นต้น ในการหาคุณลักษณะของอนุกรมเวลาเราสามารถใช้แบบจำลองได้หลายแบบ แบบจำลองที่ใช้โดยนักเศรษฐศาสตร์แบบหนึ่ง คือ แบบจำลองแบบคลาสสิก (classical model) เป็นการอธิบายถึงองค์ประกอบของการแปรผันของอนุกรมเวลา 4 ส่วน ดังนี้

1) ค่าแนวโน้ม (Secular trend) แทนด้วย Tt เป็นการเปลี่ยนแปลงข้อมูลมีลักษณะราบเรียบ แนวโน้ม อาจมีลักษณะเป็นเส้นตรงหรือเส้นโค้งในทางเพิ่มขึ้นหรือลดลง ค่าแนวโน้มของข้อมูลเป็นการเคลื่อนไหวในช่วงระยะเวลาที่ค่อนข้างนานพอสมควร ควรเป็นข้อมูลรายปี และควรมีข้อมูลอย่างน้อย 15 ปี ซึ่งจะแสดงทิศทางของอนุกรมเวลา ตัวอย่าง กราฟแสดงค่าแนวโน้มของผลผลิตเคมีภัณฑ์ชนิดหนึ่ง

2) การเปลี่ยนแปลงหรือความแปรผันตามฤดูกาล (Seasonal Variation) แทนด้วย St เป็นการเปลี่ยนแปลงข้อมูลมีลักษณะการเพิ่มขึ้น หรือลดลงในลักษณะเดียวกันของรอบระยะเวลาหนึ่งที่แน่นอน เรียกว่า การเปลี่ยนแปลงตามฤดูกาล หน่วยของระยะเวลาสำหรับข้อมูลอาจเป็นรายชั่วโมง รายวัน รายสัปดาห์ รายเดือน รายไตรมาส สำหรับข้อมูลรายปีไม่มีการแปรผันตามฤดูกาล การเปลี่ยนแปลงตามฤดูกาลนั้นกำหนดระยะเวลาการเกิดซ้ำในรอบหนึ่ง ๆ ได้ค่อนข้างแน่นอน ตัวอย่างเช่น ยอดขายรายเดือนของห้างสรรพสินค้าแห่งหนึ่ง

3) การเปลี่ยนแปลงหรือความแปรผันตามวัฏจักร (*Cyclical Variation)* แทนด้วย *Ct* การเปลี่ยนแปลงตามวัฏจักร มีการเปลี่ยนแปลงเคลื่อนไหวในลักษณะซ้ำ ๆ กันและจะมีลักษณะคล้ายคลึงกับการเปลี่ยนแปลงตามฤดูกาล จะต่างกันก็ตรงที่การเปลี่ยนแปลงตามวัฏจักรแต่ละรอบจะใช้ระยะเวลาที่นานกว่า คือ ตั้งแต่ 5 ปีขึ้นไป ข้อมูลที่มีการเปลี่ยนแปลงตามวัฏจักรในทางธุรกิจ เรียกว่า "วัฏจักรธุรกิจ" (Business Cyclical) โดยทั่วไปประกอบด้วยระยะเจริญรุ่งเรือง (prosperity) ระยะฝืดเคือง (recession) ระยะตกต่ำ (depression) และระยะขยายตัว (recovery)

2.2.2.2 รูปแบบของอนุกรมเวลา

จากปัจจัยทั้ง 4 ข้างต้น ถ้า Y แทนข้อมูลอนุกรมเวลาชุดหนึ่ง ๆ เราสามารถกำหนดแบบจำลองได้ 2 แบบ ดังนี้

1) แบบจำลองผลบวก (Additive model) ถือว่าข้อมูลในแต่ละอนุกรมเวลาประกอบด้วยผลบวกขององค์ประกอบทั้ง 4 อย่าง Yt = Tt + St + Ct + It

2) แบบจำลองผลคูณ (*Multiplicative model)* ถือว่าข้อมูลในแต่ละอนุกรมเวลาประกอบด้วยผลคูณขององค์ประกอบทั้ง 4 อย่าง Y t = Tt * St * Ct * It โดยทั่วไปข้อมูลอนุกรมเวลา ในทางธุรกิจจะมีความสัมพันธ์ในรูปแบบจำลองผลคูณ เนื่องจากเป็นการพิจารณาการเปลี่ยนแปลงในรูปอัตราร้อยละ ซึ่งจะทำให้ผลการวิเคราะห์ใกล้เคียงความเป็นจริงมากกว่าการใช้แบบจำลองผลบวก

2.2.3 ทฤษฎีเกี่ยวกับการออกแบบ (Visualization)

รูปแบบของจินตทัศน์ (Visualization) รูปแบบของจินตทัศน์ที่ใช้ในการแสดงผลสารสนเทศนั้นมีจำนวนมาก ซึ่งการเลือกใช้ก็ขึ้นอยู่กับสารสนเทศที่ต้องการแสดง รวมถึง ผลลัพธ์ที่ต้องการให้เกิดการตอบสนองต่อ สารสนเทศนั้น ๆ ในงานวิจัยนี้มุ่งศึกษาแนวทางการ ออกแบบแผงหน้าปัดโดยใช้ซอฟต์แวร์ IBM Cognos ดังนั้นจึงมีรูปแบบแผนภูมิที่สามารถสร้างบนซอฟต์แวร์ IBM Cognos ตามที่ได้อธิบายไว้ในคู่มือการใช้งาน IBM Cognos *Business Intelligence V10.1 Handbook* ดังนี้

1) แผนภูมิแท่ง (Column and Bar Chart) แผนภูมิแท่งใช้ในการแสดงข้อมูลตามช่วงเวลาหรือข้อมูลที่ไม่ต่อเนื่องกัน โดยมีทั้งแผนภูมิแท่งแนวตั้งและแนวนอน ซึ่งสามารถเปรียบเทียบข้อมูลได้ทั้งข้อมูลเดียว หรือหลายข้อมูล เหมาะต่อการดูข้อมูลที่เรียงลำดับตามค่าต่ำสุดหรือสูงสุด นอกจากนี้แผนภูมิแท่ง สามารถแสดงข้อมูลหลาย ๆ ข้อมูลเรียงกันเป็นชั้นภายในแผนภูมิแท่งเดียวกันเป็นแผนภูมิแท่งแบบชั้น (Stack Charts) (Browne et al., 2010) ซึ่งเป็นรูปแบบที่เหมาะกับการแสดงข้อมูลที่มีความสัมพันธ์เป็นชุด เพื่อให้ค่าที่แสดงสามารถเปรียบเทียบกันได้ง่ายเมื่อมีค่าข้อมูลที่ต้องเปรียบเทียบแสดงอยู่ด้วยกัน อาจใช้ในการแสดงข้อมูลตามช่วงเวลาสั้น ๆ ข้อมูลการขายตามหมวดหมู่สินค้า (*Kerzner , 2011)* สามารถใช้การวัดหลายอย่างได้ภายในแผนภูมิเดียวกันโดยที่ไม่ทำให้ข้อมูลดูหนาแน่นเกินไป สามารถใช้ในการแจ้งเตือนผู้ใช้แผงหน้าปัดถึงค่าที่ผิดปกติไปจากเป้าหมายที่วางไว้ ข้อควรระวังคือ ต้องมั่นใจในความถูกต้องของข้อมูล เนื่องจากถ้าค่าข้อมูลผิดก็จะทำให้แผนภูมิแสดงความสัมพันธ์ และแนวโน้มที่ผิด (Smietana, 2010)

2) แผนภูมิเส้น (Line Charts) แผนภูมิเส้นมีความคล้ายคลึงกับแผนภูมิแท่งเพียงแต่จะใช้จุดเป็นตัวบอกค่าของข้อมูลแทนตัวแท่ง แล้วมีการลากเส้นเชื่อมแต่ละจุดจึงเหมาะสำหรับการแสดงข้อมูล แนวโน้มหลาย ๆ แนวโน้มเปรียบเทียบกัน แต่ก็ไม่ควรใช้ในการเปรียบเทียบข้อมูลหลาย ๆ ค่า เพราะจะแยกความแตกต่างระหว่างข้อมูลได้ยากกว่าการแสดงแบบเป็นชั้นในแผนภูมิแท่ง (*Browne et al., 2010)* เหมาะสำหรับการวิเคราะห์อนุกรมเวลาที่ต้องการที่จะเห็นการเปลี่ยนแปลงที่เกิดขึ้นของข้อมูล หนึ่งหรือมากกว่าหนึ่งข้อมูลการวัดค่าในช่วงเวลา แผนภูมิเส้นยังให้การเปรียบเทียบการวิเคราะห์แนวโน้มใช้ข้อมูลหลายชุดซ้อนกันในหนึ่งแผนภูมิ (*Kerzner, 2011)* ในแผนภูมิเส้นสามารถใช้จุดบอกตำแหน่งของข้อมูลที่ชัดเจนได้ เนื่องจากถ้าใช้แต่เส้นก็จะทราบเฉพาะแนวโน้มของข้อมูล หรือการใช้จุดอย่างเดียว ก็จะเห็นแต่ข้อมูล ณ ตำแหน่งนั้น ๆ แต่จะไม่เห็นแนวโน้ม ปัจจุบันจึงมีการใช้แผนภูมิเส้นและจุด (*Spline)* ด้วยกันเพื่อให้เห็นทั้งค่าข้อมูลและแนวโน้ม (*Smietana, 2010)* ควรทำเส้นให้จุด ของข้อมูลเด่นว่ามีข้อมูลที่ตำแหน่งใด*(Juice Inc., 2009)*

3) แผนภูมิพื้นที่ (*Area Chart)* แผนภูมิพื้นที่สามารถใช้สีในบริเวณพื้นที่ใต้กราฟได้แทนการลากเส้น เพื่อแสดงแนวโน้มของข้อมูลเพื่อใช้ในการเปรียบเทียบ *(Browne et al., 2010)* การใช้สีเพื่อแสดงพื้นที่ใต้หรือเหนือกราฟของแผนภูมิเส้นใช้เมื่อต้องการเปรียบเทียบแนวโน้มตามช่วงเวลาของข้อมูล (*Kerzner, 2011)*

4) แผนภูมิจุด (*Point Charts)* แผนภูมิจุดมีลักษณะคล้ายแผนภูมิเส้นเพียงแต่แสดง เฉพาะตำแหน่งค่าข้อมูล โดยไม่มีการลากเส้นเพื่อแสดงแนวโน้ม แผนภูมิจุดมีประโยชน์สำหรับ การแสดงข้อมูลเชิงปริมาณ โดยใช้จุดแทนข้อมูลตามแกนโดยวางตามลำดับ (*Browne et al., 2010)*

5) แผนภูมิแบบผสม (*Combination Charts)* แผนภูมิแบบผสมใช้ในการแสดงค่าข้อมูลหลายค่าโดยใช้แผนภูมิแท่ง แผนภูมิพื้นที่ หรือแผนภูมิเส้น รวมไว้ในแผนภูมิเดียวกัน มีประโยชน์ในการเน้นความสัมพันธ์ของข้อมูลแต่ละชุด (*Browne et al.,2010)*

6) แผนภูมิแบบกระจาย (Scatter Plot) แผนภูมิแบบกระจายมักใช้ในการเปรียบเทียบความสัมพันธ์ของข้อมูลระหว่างชุดเพื่อดูความแตกต่างตามตำแหน่งของข้อมูลที่ปรากฏบนแผนภูมิ (Browne et al., 13 2010) ใช้ในกรณีที่ต้องการดูความสัมพันธ์ของข้อมูล มากกว่าค่าของข้อมูลจะเป็นการดูแนวโน้มใน ภาพรวมว่าข้อมูลชุดนี้มีค่าไปในทิศทางใด (Smietana, 2010)

7) แผนภูมิฟอง (*Bubble Charts)* แผนภูมิฟองมีลักษณะคล้ายแผนภูมิแบบกระจายเพียงแต่จะมีตัววัดเพิ่มขึ้นโดยเป็นขนาดของฟองตามค่าของข้อมูล โดยมักใช้ในการแสดงข้อมูลด้านการเงิน (*Browne et al., 2010)* แสดงค่าเป็นกลุ่มของข้อมูลด้วยขนาดของฟอง

(วงกลม หรือ ทรงกลม) ที่แตกต่างกันตามค่าของข้อมูล ใช้ในการเปรียบเทียบระหว่างข้อมูลเดียวกันแต่ ต่างช่วงเวลา หรือเปรียบเทียบ ระหว่างข้อมูลที่แตกต่างกันด้วย (*Kerzner, 2011)* การเลือกใช้ ควรแน่ใจในเรื่องการอธิบายขนาดของฟองสบู่แต่ละฟอง ว่ามีการกำหนดค่าได้ถูกต้องและ ชัดเจน (*Smietana, 2010)*

8) แผนภูมิสี่ส่วน (*Quadrant Charts)* โดยหลักแล้วจะใช้แผนภูมิสี่ส่วนเป็นแกนพื้นหลังให้แผนภูมิฟอง ซึ่งจะแบ่งพื้นที่เป็นสี่ส่วนเท่า ๆ กัน ในการแบ่งส่วนของแผนภูมิอาจใช้สีในการแยกความแตกต่างของแต่ละส่วน ซึ่งขนาดของส่วนที่แบ่งอาจแบ่งย่อยลงไปได้อีกเสมือนเป็นมาตรวัดในแผนภูมิ สามารถใช้แผนภูมิสี่ส่วนในการแบ่งหมวดหมู่เพื่อวางข้อมูลตามหมวดหมู่เช่นการวิเคราะห์จุดอ่อนจุดแข็งโอกาส และอุปสรรค (*SWOT Analysis)* เป็นต้น

9) แผนภูมิวงกลม (*Pie Charts)* แผนภูมิวงกลมใช้ในการแสดงความสัมพันธ์ของข้อมูลในส่วนนั้น เปรียบเทียบกับสัดส่วนทั้งหมดของแผนภูมิ ซึ่งหากต้องการแสดงข้อมูลจริงเป็นแผนภูมิแท่งแบบเป็น ชั้นมากกว่าแผนภูมิวงกลมไม่ใช่ตัวเลือกที่ดีในการแสดงข้อมูลที่มีค่าเป็นศูนย์หรือมีค่าติดลบ (*Browne et al., 2010)* เป็นแผนภูมิที่มีองค์ประกอบมากกว่าแผนภูมิแบบอื่น โดยเป็นเรื่องยากมากที่จะแยกแยะความแตกต่างของสัดส่วนของข้อมูลกับการแบ่งพื้นที่ตามค่าข้อมูล นอกจากจะเป็นกรณีของชุดข้อมูลขนาดเล็กที่มีความแตกต่างของค่าข้อมูลอยู่มาก และแผนภูมิวงกลมที่ยังก่อให้เกิดปัญหาสำหรับการใส่คำอธิบายเพราะมีทั้งสีของ แผนภูมิและรายละเอียดต่าง ๆ ภายใน ซึ่งการใส่คำอธิบายอื่นเพิ่มอาจก่อให้เกิดการสับสนในการดูข้อมูล (*Kerzner, 2011)* แผนภูมิวงกลมเป็นแผนภูมิสำหรับการแสดงองค์ประกอบข้อมูลที่ มีส่วนประกอบค่อนข้างน้อยและมีเพียงหนึ่งตัวชี้วัดเชิงปริมาณ และการแสดงแผนภูมิวงกลม ต้องคิดค่าข้อมูลรวมเป็น *100%* และข้อมูลที่ไม่เป็นค่าลบ ใช้การแสดงค่าข้อมูลทั้งหมดด้านนอกแผนภูมิ โดยเรียงค่าข้อมูลตามลำดับที่แสดงในแผนภูมิ (*Smietana, 2010)*

10) แผนภูมิจุดนำ (*Bullet Charts)* แผนภูมิจุดนำเป็นรูปแบบหนึ่งของแผนภูมิแท่ง เป็นการแสดงตัว วัดที่สำคัญ ใช้ในการเปรียบเทียบตัววัดตั้งแต่ *1* ตัวหรือมากกว่านั้น โดยสามารถวัดโดยเปรียบเทียบ กับสีของพื้นหลังที่จะแสดงตัววัดที่มากขึ้น เช่น ระดับความพึงพอใจที่ดี จนถึงระดับที่ไม่ดี และเนื่องจากเป็นแสดงสารสนเทศที่ไม่จำเป็นต้องใช้พื้นที่มาก จึงสามารถใส่แผนภูมิจุดนำเป็นส่วนหนึ่งในการแสดงรายงานบนแผงหน้าปัด (*Browne et al., 2010)* จะแสดงการวัดมูลค่าในแกนแนวนอน ซึ่งอาจมีการเพิ่มบริบทข้องข้อมูลเข้าไปโดยการใช้สีเป็นตัวบอกระดับความพึงพอใจ ทำให้มองข้อมูลเข้าใจได้ในครั้งเดียว (*Smietana, 2010)*

11) แผนภูมิมาตรวัด (*Gauge Charts)* แผนภูมิมาตรวัดหรืออาจเรียกว่าแผนภูมิหน้าปัด (*Dial Charts)* หรือแผนภูมิมาตรอัตราเร็ว (*Speedometer Charts)* มีความคล้ายกับแผนภูมิจุดนำในแง่การเปรียบเทียบตัววัดหลายค่าเพียงแต่ใช้เข็มในการแสดงค่า ซึ่งการอ่านค่าสามารถทำได้ง่ายเหมือนการอ่านค่าจากหน้าปัดและแต่ละค่าที่แสดงจะเปรียบเทียบด้วยช่วงสีที่อยู่บนแผนภูมิ โดยแผนภูมิมาตรวัดนั้นเป็นทางเลือกที่ดีกว่าแผนภูมิจุดนำเมื่อต้องเปรียบเทียบค่ามากกว่าสองค่าขึ้นไป โดยแผนภูมินี้มักถูกใช้ในการแสดงตัวชี้วัดผลการดำเนินงานหลักบนแผงหน้าปัดสำหรับผู้บริหารระดับสูง สำหรับการออกรายงานโดยใช้ PDF และ HTML จะจำกัดอยู่ที่ แผนภูมิ และใน Microsoft Excel จะไม่สนับสนุนการแสดงผลแผนภูมิประเภทนี้ (*Browne et al., 2010)* ใช้ในการบอกสถานะปัจจุบันโดยจะใช้สีแดง ส้ม เหลือง เขียว ในการบอกถึงระดับของปัญหาตั้งแต่มาก (สีแดง) จนถึงระดับปกติ (สีเขียว) สามารถช่วยให้เห็นความสัมพันธ์ ระหว่างค่าที่วัดแต่ว่าจะค่อนข้างสิ้นเปลืองเนื้อที่ จึงควรเลือกใช้เมื่อมีพื้นที่มากเพียงพอ (*Smietana,2010)*

12) แผนภูมิพาเรโต (*Pareto Charts)* แผนภูมิพาเรโตเป็นลำดับของหมวดหมู่ที่เรียงจากความถี่มากที่สุดไปจนถึงน้อยที่สุด โดยจะมีเส้นการสั่งสม (*Cumulation Line)* ซึ่งเป็นการแสดงอัตราส่วนร้อยละยอดรวมสะสมทั้งหมดของแผนภูมิแท่ง โดยแผนภูมิชนิดนี้จะใช้ในการ ควบคุมคุณภาพ ทำให้สามารถระบุและลดสาเหตุของปัญหาที่เกิดขึ้นตามลำดับความสำคัญของปัญหาและสามารถที่จะสร้างแผนภูมิพาเรโตได้ทั้งก่อนและหลังการแก้ปัญหาเพื่อเป็นการ เปรียบเทียบให้เห็นถึงส่งที่เปลี่ยนแปลงไป *(Browne et al., 2010)*

13) แผนภูมิแท่งแบบก้าวหน้าหรือแผนภูมิน้ำตก (Progressive Column Charts or Waterfall Charts) เป็นแผนภูมิที่แสดงการเปลี่ยนแปลงที่เกิดขึ้นต่อจากแผนภูมิแท่งหรือแผนภูมิแท่งแบบเป็นชั้น โดยมักใช้ในการเปรียบเทียบค่าข้อมูล 1 ข้อมูลต่อข้อมูลทั้งหมด โดยใน Microsoft Excel ไม่สนับสนุนการแสดงผลแผนภูมิประเภทนี้ (Browne et al., 2010)

14) แผนภูมิมาริเมกโก (*Marimekko Charts)* เป็นแผนภูมิชั้นซึ่งความกว้างของชั้นของแท่งแผนภูมิจะเป็น สัดส่วนของข้อมูล 1 ข้อมูลต่อข้อมูลทั้งหมด โดยส่วนที่เป็นความสูงของแต่ละแนวแท่งแผนภูมินั้นจะเป็นอัตราร้อยละของมูลค่ารวมแนวตามลำดับ (*Browne et al., 2010)*

15) แผนภูมิเรดาร์หรือแผนภูมิแมงมุม (*Radar or Spider Charts)* แผนภูมิเรดาร์เป็นแผนภูมิที่ใช้ในการเปรียบเทียบข้อมูลหลายค่าตามแกนหลายแกนโดยจะใส่ค่าข้อมูลโดยเริ่มจากตรงกลางที่มีแกนวัดค่าเป็นรัศมี มีประโยชน์ในการเปรียบเทียบข้อมูลหลายชุดและหลายตัววัด และมีประโยชน์ในการดูค่าผิดปกติ (*Browne et al., 2010)* ข้อมูลจะถูกใส่ค่าตามระยะห่าง

จากจุดศูนย์กลาง แต่ละจุดข้อมูลที่มีส่วนประกอบในแนวตั้ง คือ ระยะห่างจากจุดศูนย์กลาง ของ แกน แผนภูมิเรดาร์มีประโยชน์สำหรับการแสดงข้อมูลวัฏจักร เช่น ระดับการรับพนักงานรายวัน หรือผลรวมรายได้รายเดือนในการออกแบบสามารถวางซ้อนเป้าหมาย และเกณฑ์การให้บริบท สำหรับตัวชี้วัดเชิงปริมาณ (Smietana, 2010)

2.2.4 ทฤษฎีเกี่ยวกับการออกแบบ (Dashboard)

แดชบอร์ด เป็นแบบหน้าเดียว ซึ่งมักเรียกว่าพื้นที่ว่างเปล่าที่ใช้การแสดภาพเพื่อบอกเล่าเรื่องราวหนึ่ง ๆ ได้เนื่องจากจำกัดอยู่เพียงหนึ่งหน้าแดชบอร์ดที่ออกแบบมาอย่างดีจะประกอบด้วยองค์ประกอบที่สำคัญที่สุดของเรื่องราวเท่านั้น แดชบอร์ดเป็นวิธีที่ยอดเยี่ยมในการตรวจดูธุรกิจของคุณ ในการค้นหาคำตอบ และดูเมตริกที่สำคัญที่สุดของคุณอย่างรวดเร็ว การแสดงภาพบนแดชบอร์ดอาจมาจากหนึ่งหรือหลายชุดข้อมูลพื้นฐาน และจากรายงานหนึ่งหรือหลายรายงานพื้นฐาน แดชบอร์ดสามารถรวมข้อมูลภายในองค์กรและข้อมูลบนระบบคลาวด์ ให้มุมมองแบบรวมโดยไม่คำนึงถึงตำแหน่งที่ข้อมูลอยู่ แดชบอร์ดไม่ได้เป็นเพียงภาพที่สวยงาม เนื่องจากแดชบอร์ดสามารถโต้ตอบได้และมีการอัปเดตไทล์เมื่อมีการเปลี่ยนแปลงข้อมูลพื้นฐาน

2.2.4.1 Dashboard ใน Pentaho

1) Dashboard by Pentaho จะมีลักษณะพิเศษคือสามารถ set template, settheme, และสามารถตั้งเวลา ในการ refresh ตัวเอง เพื่อ update ข้อมูลที่อาจมีการอัพเดทลงบน Database ได้การ Set Template นั้นทำเพื่อหากมีกรณีที่ต้องการเปรียบเทียบข้อมูลเกิดขึ้นจะทำให้สามารถเห็นข้อแตกต่างระหว่างข้อมูลได้อย่างชัดเจน

2) Dashboard by CDE *จะแตกต่างจาก* Dashboard by Pentaho คือผู้ใช้จะต้องมีความรู้พื้นฐานเกี่ยวกับ Html, Java เนื่องจาก Dashboard by CDE เราจะสามารถปรับแต่งส่วนแสดงผลได้ โดยเราสามารถแก้ไขหรือปรับแต่งได้โดยเปลี่ยนแปลงโค้ดในส่วนของฟังก์ชัน Dashboard by CDE มีข้อดีคือเราจะสามารถปรับแต่ง และเพิ่มฟังก์ชันการแสดงผลได้มากกว่า Dashboard by Pentaho

2.2.4.2 หลักในการออกแบบแดชบอร์ด

1) Users ต้องเข้าใจผู้ใช้ก่อนเป็นอันดับแรก การเข้าใจผู้ใช้นั้น ให้ยึดตามหลักการของ Design Thinking จงหา Insight ของผู้ใช้ หรือ ตอบให้ได้ว่า ทำไมเขาถึงใช้คำถามหลัก ของการออกแบบ Dashboard ก็คือ ผู้ใช้จะเอา Dashboard ไปใช้ทำอะไรเช่น ช่วยตัดสินใจติดตามผลการดำเนินงาน ใช้เตือนเมื่อมีสิ่งผิดปกติ ในอดีต เวลาส่งคนไปเก็บ Requirement กับผู้ใช้ ก็จะชอบไปถามว่า อยากดูข้อมูลอะไรบ้าง อยากดูเป็นกราฟไหน ซึ่งพอไปทำกราฟมา

เสร็จเรียบร้อย ผู้ใช้มักจะบอกว่า ไม่เห็นตรงกับที่อยากได้เลย อันนี้เกิดจากการที่เราไม่ได้เข้าใจจริง ๆ ว่าเค้าอยากจะเอาไปใช้ทำอะไร ถ้าเราเข้าใจถึงรูปแบบว่าเค้าจะเอา Dashboard ไปใช้อย่างไรเราก็จะออกแบบได้ตรงใจมากขึ้น

2) Content เมื่อเข้าใจแล้วว่า ผู้ใช้อยากเอา Dashboard ไปใช้อย่างไร ก็ต้องมาคิดต่อว่า เนื้อหา หรือ สิ่งที่เค้าจะต้องเอาไปใช้นั้น มีอะไรบ้าง ซึ่งก็ขอแบ่งเป็น measures หรือ ตัวเลขที่เราสนใจ เช่น ยอดขาย จำนวนลูกค้า dimensions หรือ มุมมองที่เราอยากวิเคราะห์ข้อมูล เช่น ตามช่วงเวลา ตามกลุ่มสินค้า ตามพื้นที่ การเลือก measures นั้น ถ้าเราสามารถช่วยคิด ช่วยออกแบบให้การวัดผลนั้น มีความน่าสนใจ หรือ ตรงประเด็นมากขึ้น ก็จะทำให้ Dashboard นั้นมีความน่าสนใจมากขึ้นไปอีก

3) Presentation หลังจากที่รู้แล้วว่า จะใช้ measures และ dimensions แบบไหนก็ถึงเวลาที่มาเลือกกราฟให้

4) Navigation เมื่อมีหลายกราฟแล้ว จะเอามาประกอบกันเป็น Dashboard การจัดวางกราฟก็เป็นส่วนสำคัญ หลักง่ายๆ คือ กราฟที่เป็นเรื่องเดียวกัน ก็ควรวางไว้ใกล้ๆกัน อย่าให้คนต้องอ่านกราฟนึงแล้วกระโดดข้ามไปอีกกราฟ แบบกระโดดไปมา

2.2.5 การนอร์มัลไลเซชั่น (Normalization)

การทำนอร์มัลไลเซชัน เป็นวิธีการในการกำหนดแอตทริบิวต์ให้กับแต่ละเอนทิตี เพื่อให้ได้โครงสร้างของตารางที่ดี สามารถควบคุมความซ้ำซ้อนของข้อมูลหลีกเลี่ยงความผิดปกติของข้อมูล โดยทั่วไปผลลัพธ์ของการนอร์มัลไลเซชัน จะได้ตารางที่มีโครงสร้างซับซ้อนน้อยลง แต่จำนวนของตารางจะมากขึ้น

การทำนอร์มัลไลเซชัน จะประกอบด้วยนอร์มัลฟอร์ม (*Normal Form)* แบบต่าง ๆ ที่มีเงื่อนไขของการทำให้อยู่ในรูปของนอร์มัลฟอร์มที่แตกต่างกันไป ขึ้นอยู่กับผู้ออกแบบฐานข้อมูลว่า ต้องการลดความซ้ำซ้อนในฐานข้อมูลให้อยู่ในระดับใด ซึ่งประกอบด้วยนอร์มัลฟอร์มแบบต่าง ๆ ดังต่อไปนี้

1) นอร์มัลฟอร์มที่ 1 (First Normal Form : 1NF)
2) นอร์มัลฟอร์มที่ 2 (Second Normal Form : 2NF)
3) นอร์มัลฟอร์มที่ 3 (Third Normal Form : 3NF)
4) บอยซ์คอดด์นอร์มัลฟอร์ม (Boyce–Codd Normal Form : BCNF)
5) นอร์มัลฟอร์มที่ 4 (Fourth Normal Form : 4NF)
6) นอร์มัลฟอร์มที่ 5 (Fifth Normal Form : 5NF)

2.2.5.1 การแปลงให้อยู่ในรูปนอร์มัลฟอร์มที่ 1 (First Normal Form: 1NF)

คุณสมบัติของรีเลชันของแบบจำลองข้อมูลเชิงสัมพันธ์ คือข้อมูลในแต่ละทัปเพิลจะต้องไม่ซ้ำกัน และค่าในแต่ละแอตทริบิวต์จะต้องไม่สามารถถูกแบ่งแยกย่อยลงไปได้อีก หรือมีความเป็นอะตอมมิค(Atomic) รวมถึงจะต้องมีค่าเพียงค่าเดียวที่อยู่ในแต่ละแอตทริบิวต์หรือมีความเป็นซิงเกิลแวลู (Single Value) ซึ่งในการทำนอร์มัลไลเซชันให้อยู่ในนอร์มัลฟอร์มที่ 1 ก็อาศัยคุณสมบัติดังที่กล่าวไว้ข้างต้น

1.1) รีพีทติ้งกรุ๊ป (Repeating Group)

การที่ข้อมูลใน 1 ทัปเพิล สามารถมีค่าในแต่ละแอตทริบิวต์ได้มากกว่าหนึ่งค่า (Multivalued) จะทำให้เกิดรีพีทติ้งกรุ๊ป ดังตารางที่แสดงในภาพข้างล่าง ซึ่งเลขที่โครงการหนึ่งหมายเลขประกอบด้วยกลุ่มข้อมูลหลายกลุ่ม ซึ่งทำให้รีเลชันดังกล่าวขาดคุณสมบัติซิงเกิลแวลู

การทำงานของพนักงานในโครงการ

เลขที่โครงการ	ชื่อโครงการ	รหัสนักงาน	ชื่อพนักงาน	ตำแหน่งงาน	ค่าแรง/ ชม.	จำนวน ชม.
11	RFID	103	สมชาย	Engineer	500	23.8
		101	วิชา	Programmer	500	19.4
		102	สุรชัย	Administrator	200	12.6

ภาพที่ 2.5 แสดงการเกิดรีพีทติ้งกรุ๊ป

ที่มา : https://msit5.wordpress.com

1.2) นิยามของนอร์มัลฟอร์มที่ 1

รีเลชันจะอยู่ในรูปของนอร์มัลฟอร์มที่ 1 ก็ต่อเมื่อมีคุณสมบัติตามเงื่อนไขดังต่อไปนี้

1) มีการกำหนดแอตทริบิวต์ที่เป็นคีย์
2) ต้องไม่มีรีพีทติ้งกรุ๊ปแต่ละแถวหรือคอลัมน์จะมีค่าได้เพียง 1
3) แอตทริบิวต์ทุกตัวต้องขึ้นอยู่กับคีย์หลัก

จากภาพข้างบน เมื่อการการนอร์มัลไลเซชันให้อยู่ในรูปนอร์มัลฟอร์มที่ 1 จะได้ตารางที่แตกย่อยออกมาเป็น 2 ตาราง ดังภาพข้างล่าง ซึ่งมีคุณสมบัติตามนอร์มัลฟอร์มที่ 1 แล้ว

การทำงานของพนักงาน

รหัสนักงาน	ชื่อพนักงาน	ตำแหน่งงาน	ค่าแรง/ ชม.	จำนวน ชม.
103	สมชาย	Engineer	500	23.8
101	วิชา	Programmer	500	19.4
102	สุรชัย	Administrator	200	12.6

ชื่อโครงการ

เลขที่โครงการ	ชื่อโครงการ
11	RFID

ภาพที่ 2.6 แสดงรีเลชันที่อยู่ในรูปนอร์มัลฟอร์มที่ 1

ที่มา : https://msit5.wordpress.com

2.2.5.2 การแปลงให้อยู่ในรูปนอร์มัลฟอร์มที่ 2 (Second Normal Form : 2NF)

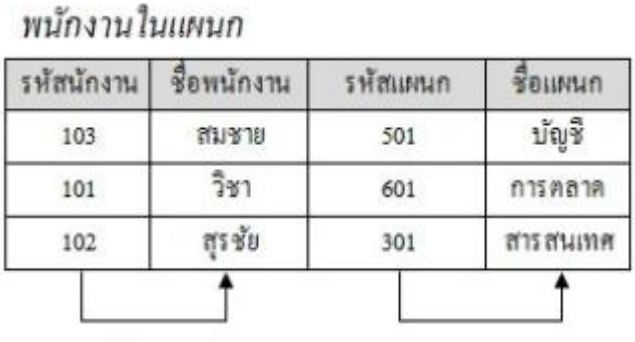
ในหนึ่งรีเลชันจะประกอบด้วยแอตทริบิวต์ต่าง ๆ ที่มีความสัมพันธ์ที่ขึ้นต่อกัน ซึ่งความสัมพันธ์ดังกล่าวจะเป็นตัวกำหนดว่าแอตทริบิวต์ใดเป็นตัวกำหนดข้อมูลหรือคีย์แอตทริบิวต์ (Key Attribute) และแอตทริบิวต์ใดเป็นข้อมูลที่ถูกกำหนดหรือนอนคีย์แอตทริบิวต์ (Nonkey Attribute)

1.1) ฟังก์ชันนัลดีเพนเดนซี (Functional Dependency: FD)

ในการทำนอร์มัลไลเซชัน จะต้องมีความเข้าใจหลักการของฟังก์ชันดีเพนเดนซี (Function Dependency: FD) เสียก่อน โดยมีคำจำกัดความคือ B ขึ้นอยู่กับ A ถ้าทราบค่าของ A ก็จะทำให้รู้ค่าของ B ได้ ฟังก์ชันนัลดีเพนเดนซี สามารถแสดงด้วยการใช้เครื่องหมายลูกศร (–>) ตัวอย่างเช่น A–>B แสดง B เป็นฟังก์ชันนัลดีเพนเดนต์กับ A กล่าวคือถ้ารู้ค่า A ก็จะทำให้ทราบค่าของ B ด้วย ทุกค่าของ A ที่มีค่าเท่ากัน จะได้ค่าเท่ากันเสมอ

1.2) พาเชียลดีเพนเดนซี (Partial Dependency)

พาร์เชียลดีเพนเดนซี หมายถึง การที่มีแอตทริบิวต์บางแอตทริบิวต์ที่ขึ้นอยู่กับเพียงบางส่วนของคีย์หลักเท่านั้น ตัวอย่างเช่น จากตารางในภาพข้างล่าง แอตทริบิวต์ชื่อพนักงานจะขึ้นอยู่กับคีย์รหัสพนักงานในขณะที่แอตทริบิวต์ชื่อแผนก จะขึ้นอยู่กับคีย์รหัสแผนก จะเห็นว่า ข้อมูลที่อยู่ในรีเลชันเดียวกันแต่ไม่ได้ขึ้นอยู่กับคีย์ใดคีย์หนึ่งทั้งหมด



ภาพที่ 2.7 แสดงรีเลชันที่มีพาร์เชียลดีเพนเดนซี

ที่มา : https://msit5.wordpress.com

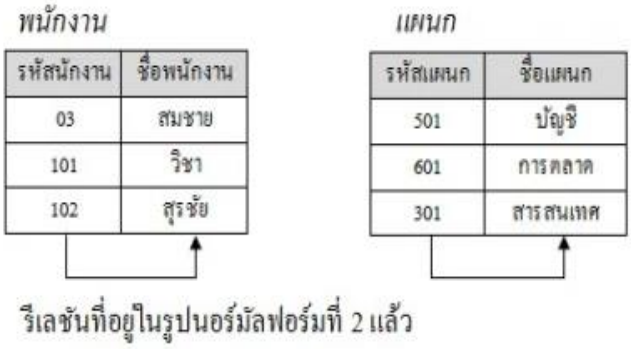
1.3) นิยามของนอร์มัลฟอร์มที่ 2

รีเลชันจะอยู่ในรูปของนอร์มัลฟอร์มที่ 2 ก็ต่อเมื่อมีคุณสมบัติตามเงื่อนไข

1) รีเลชันนั้นเป็นนอร์มัลฟอร์มที่ 1 อยู่แล้ว

2) รีเลชันนั้นไม่มีพาร์เชียลดีเพนเดนซี

ตัวอย่างรีเลชันพนักงานในแผนกในภาพข้างบน เมื่อทำการแตกออกเป็นรีเลชันย่อยที่ไม่มีพาร์เชียลดีเพนเดนซีแล้ว จะได้เป็นรีเลชันสองรีเลชัน คือรีเลชันพนักงานและรีเลชันแผนก ซึ่งอยู่ในรูปของนอร์มัลฟอร์มที่ 2 แล้ว ดังภาพข้างล่าง



ภาพที่ 2.8 แสดงรีเลชันที่อยู่ในรูปนอร์มัลฟอร์มที่ 2

ที่มา : https://msit5.wordpress.com

2.2.5.3 การแปลงให้อยู่ในรูปนอร์มัลฟอร์มที่ 3 (Third Normal Form: 3NF)

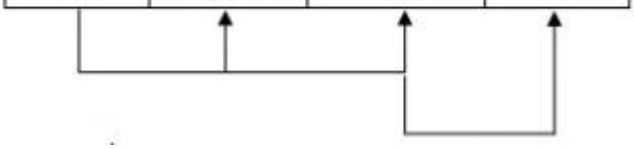
ในหนึ่งรีเลชันจะประกอบคีย์แอตทริบิวต์และนอนคีย์แอตทริบิวต์ คีย์แอตทริบิวต์จะต้องเป็นตัวกำหนดความหมายหรือการมีอยู่ของแอตทริบิวต์อื่น ๆ ที่อยู่ในรีเลชันเสมอ

1) ทรานซิทีฟดีเพนเดนซี (Transitive Dependency)

ทรานซิทีฟดีเพนเดนซี หมายถึง การที่มีฟังก์ชันนัลดีเพนเดนซีระหว่างแอตทริบิวต์ที่ไม่ได้เป็นส่วนของคีย์ใด ๆ แต่มีแอตทริบิวต์อื่น ๆ มาขึ้นกับแอตทริบิวต์นั้น ตัวอย่างเช่น จากตารางในภาพข้างล่าง แอตทริบิวต์ชื่อพนักงาน และรหัสตำแหน่งงานจะขึ้นอยู่กับคีย์รหัสพนักงาน ในขณะที่แอตทริบิวต์ค่าแรงต่อชั่วโมงของพนักงาน จะขึ้นอยู่กับแอตทริบิวต์รหัสตำแหน่งงานซึ่งไม่ใช่คีย์อีกต่อหนึ่งทำให้มีทรานซิทีฟดีเพนเดนซีเกิดขึ้นในรีเลชันนี้

การทำงานของพนักงาน

รหัสนักงาน	ชื่อพนักงาน	รหัสตำแหน่ง	ค่าแรง/ ชม.
103	สมชาย	702	500
101	วิชา	704	500
102	สุรชัย	705	200



ภาพที่ 2.9 แสดงตารางที่มีทรานซิทีฟดีเพนเดนซี

ที่มา : https://msit5.wordpress.com

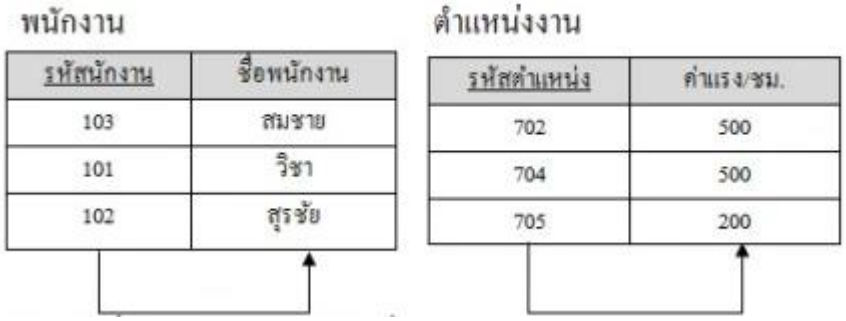
2) นิยามของนอร์มัลฟอร์มที่ 3

รีเลชันจะอยู่ในรูปของนอร์มัลฟอร์มที่ 3 ก็ต่อเมื่อมีคุณสมบัติตามเงื่อนไข

1.1) รีเลชันนั้นเป็นนอร์มัลฟอร์มที่ 2 อยู่แล้ว

1.2) รีเลชันนั้นไม่มีทรานซิทีฟดีเพนเดนซี

ตัวอย่างรีเลชัน การทำงานของพนักงาน ในภาพข้างบน เมื่อทำการแตกออกเป็นรีเลชันย่อยที่ไม่มีทรานซิทีฟดีเพนเดนซีแล้ว จะได้เป็นรีเลชันสองรีเลชัน คือรีเลชันพนักงานและรีเลชันตำแหน่งงาน ซึ่งอยู่ในรูปของนอร์มัลฟอร์มที่ 3 แล้วดังภาพข้างล่าง



ภาพที่ 2.10 แสดงตารางที่มีรีเลชันที่อยู่ในรูปนอร์มัลฟอร์มที่ 3

ที่มา : https://msit5.wordpress.com

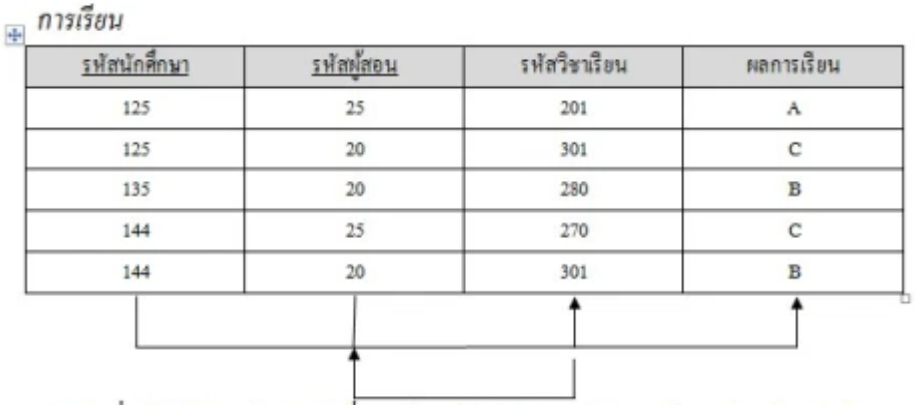
2.2.5.4 การแปลงให้อยู่ในรูปบอยซ์คอดด์นอร์มัลฟอร์ม (Boyce–Codd Normal Form: BCNF)

ในหนึ่งรีเลชันอาจจะประกอบด้วยหลายแคนดิเดตคีย์(Candidate Key) ทุกแอตทริบิวต์ในรีเลชันจะต้องขึ้นอยู่กับแคนดิเดตคีย์เสมอ เราสามารถกำหนดนิยามของ รีเลชันที่อยู่ในรูปของบอยซ์คอดด์นอร์มัลฟอร์ม ก็ต่อเมื่อรีเลชันมีคุณสมบัติตามเงื่อนไขดังต่อไปนี้

1) รีเลชันนั้นเป็นนอร์มัลฟอร์มที่ 3 อยู่แล้ว
2) ทุกแอตทริบิวต์ในรีเลชันจะต้องขึ้นกับแคนดิเดตคีย์

รีเลชันจะอยู่ในรูปบอยซ์คอดด์นอร์มัลฟอร์ม ถ้าทุกแอตทริบิวต์ขึ้นอยู่กับแคนดิเดตคีย์ (Candidate Key) ดังนั้นถ้าใน 1 รีเลชันมีแคนดิเดตคีย์เพียงตัวเดียวแล้วนอร์มัลฟอร์มที่ 3 และบอยซ์คอดด์นอร์มัลฟอร์ม จะเหมือนกัน โอกาสที่คุณสมบัติของบอยซ์คอดด์นอร์มัลฟอร์มจะถูกละเมิดนั้น เกิดขึ้นได้น้อย และจะเกิดได้กับรีเลชันที่มีแคนดิเดตคีย์มากกว่าหนึ่งเท่านั้น ดังตัวอย่างในภาพข้างล่าง รีเลชันการลงทะเบียนเรียน รีเลชันดังกล่าวอยู่ในรูปนอร์มัลฟอร์มที่ 3 แล้วแต่ก็ยังมีบางส่วนมีปัญหาอยู่ตรงจุดที่แอตทริบิวต์รหัสวิชาเรียน และผลการเรียนขึ้นอยู่กับคีย์นักศึกษา และคีย์ผู้สอน แต่ในขณะเดียวกันรหัสผู้สอนก็ขึ้นอยู่กับรหัสวิชาเรียน ทำให้ถ้าต้องการเปลี่ยนแปลงผู้สอนในวิชา 301 จะต้องมี

การเปลี่ยนแปลงถึง 2 ทัปเพิล ซึ่งผลลัพธ์ที่ได้อาจจะทำให้เกิดความผิดพลาดหากทำการแก้ไขไม่ครบถ้วน และถ้านักศึกษารหัส 135 ถอนการลงทะเบียนวิชา 280 ข้อมูลของผู้ที่สอนวิชานี้จะหายไปจากระบบเลย ถ้าเราลบข้อมูลนี้



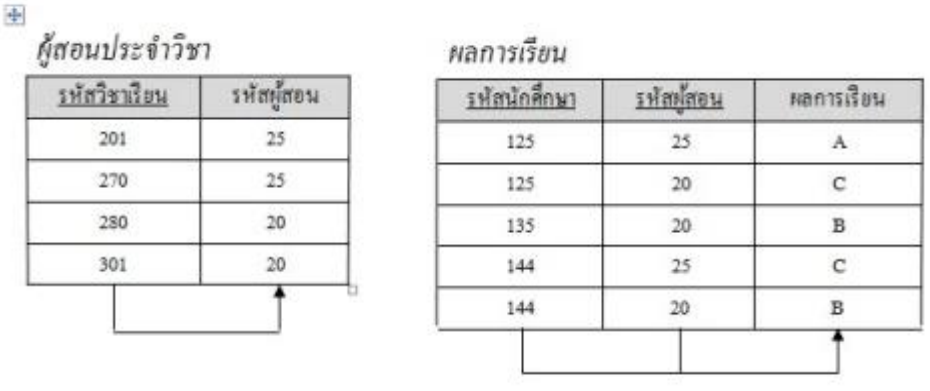
การเรียน

รหัสนักศึกษา	รหัสผู้สอน	รหัสวิชาเรียน	ผลการเรียน
125	25	201	A
125	20	301	C
135	20	280	B
144	25	270	C
144	20	301	B

ภาพที่ 2.11 แสดงตารางที่มีรีเลชันที่อยู่ในรูปนอร์มัลฟอร์มที่ 3 แต่ไม่อยู่ในรูปของบอยซ์

ที่มา : https://msit5.wordpress.com

เราสามารถทำการแตกตารางออกมาให้อยู่ในรูปของบอยซ์คอดด์นอร์มัลฟอร์มได้ โดยการแยกแอตทริบิวต์รหัสวิชาเรียนและรหัสผู้สอนซึ่งขึ้นอยู่กับแอตทริบิวต์รหัสวิชาเรียน ออกมาเป็นอีกหนึ่งรีเลชัน และแยกแอตทริบิวต์ รหัสนักศึกษา รหัสผู้สอน และผลการเรียนออกมาเป็นอีกหนึ่งรีเลชัน ดังแสดงในภาพข้างล่าง



ผู้สอนประจำวิชา

รหัสวิชาเรียน	รหัสผู้สอน
201	25
270	25
280	20
301	20

ผลการเรียน

รหัสนักศึกษา	รหัสผู้สอน	ผลการเรียน
125	25	A
125	20	C
135	20	B
144	25	C
144	20	B

ภาพที่ 2.12 แสดงตารางที่มีรีเลชันที่อยู่ในรูปนอร์มัลฟอร์มที่ 3 ที่อยู่ในรูปของบอยซ์

ที่มา : https://msit5.wordpress.com

2.2.5.5 การแปลงให้อยู่ในรูปนอร์มัลฟอร์มที่ 4 (Fourth Normal Form: 4NF)

ในขณะที่การทำให้อยู่ในรูปของนอร์มัลฟอร์มต่าง ๆ ที่ผ่านมาจะเกี่ยวข้องกับการขึ้นตรงต่อกันของข้อมูลในแต่ละแอตทริบิวต์หรือฟังก์ชันนัลดีเพนเดนซี แต่การทำให้อยู่ในรูปของนอร์มัลฟอร์มที่ 4 จะเกี่ยวข้องกับรูปแบบของการขึ้นตรงต่อกันของข้อมูลในระดับที่ซับซ้อนกว่า

1) มัลติแวลูดีเพนเดนซี (Multivalued Dependency)

ถ้าแต่ละแอตทริบิวต์ในหนึ่งรีเลชัน แบ่งออกเป็นกลุ่มของข้อมูลอิสระ เช่น แอตทริบิวต์ X, Y และ Z แบ่งออกเป็นกลุ่มข้อมูลของ X, Y และ Z ที่เป็นอิสระต่อกัน มัลติแวลูดีเพนเดนซี X –>> Y หมายถึงว่าค่า X หนึ่งค่าสามารถที่จะบอกค่า Y ได้หลาย ๆ (X Multi–Determinse Y) ไม่ว่า Z จะมีค่าเป็นอะไรก็ตาม

โดยปกติ ถ้า R ประกอบด้วย Attribute X, Y และ Z (Z = R – {XY}) ดังนั้น ถ้า X –>> Y แล้ว X –>> Z เสมอ สามารถเขียนใหม่เป็น X –>> Y I Z ถ้า Y เป็นสับเซทของ X หรือ X ยูเนี่ยน Y = R แล้ว เราเรียก X –>> Y ว่า ทริเวียลมัลติแวลูดีเพนเดนซี (Trivial Multivalued Dependency) ซึ่งจะต่างจากฟังก์ชันนัลดีเพนเดนซี X –> Y ที่ X จะสามารถบอกค่า Y ได้แค่เพียงค่าเดียว ดังตัวอย่างภาพข้างล่าง เนื่องจากแอตทริบิวต์ รหัสโครงการ รหัสบริษัท

โครงการก่อสร้าง

รหัสโครงการ	รหัสบริษัท	ที่ตั้งโครงการ
A001	B001	จันทบุรี
A001	B001	ระยอง
A001	B002	จันทบุรี
A001	B002	ระยอง

ภาพที่ 2.13 แสดงตารางที่มีรีเลชันที่อยู่ในรูปนอร์มัลฟอร์มที่ 4

ที่มา : https://msit5.wordpress.com

2) นิยามของนอร์มัลฟอร์มที่ 4

รีเลชันจะอยู่ในรูปของนอร์มัลฟอร์มที่ 4 ก็ต่อเมื่อมีคุณสมบัติตามเงื่อนไข

2.1) รีเลชันนั้นเป็นบอยซ์คอดด์นอร์มัลฟอร์มอยู่แล้ว

2.2) รีเลชันนั้นไม่มีทริเวียลมัลติแวลูดีเพนเดนซี

จากรีเลชันในภาพข้างบน เราสามารถขจัดทริเวียลมัลติแวลูดีเพนเดนซี โดยการแตกรีเลชันดังกล่าวออกเป็นรีเลชันย่อย 2 รีเลชัน ซึ่งจะทำให้ทั้งสองรีเลชันอยู่ในรูปของนอร์มัลฟอร์มที่ 4 ดังภาพข้างล่าง

บริษัทในโครงการ

รหัสโครงการ	รหัสบริษัท
A001	B001
A001	B002

ที่ตั้งโครงการ

รหัสโครงการ	ที่ตั้งโครงการ
A001	จันทบุรี
A001	ระยอง

ภาพที่ 2.14 แสดงตารางที่มีรีเลชันที่อยู่ในรูปนอร์มัลฟอร์มที่ 4

ที่มา : https://msit5.wordpress.com

2.2.5.6 การแปลงให้อยู่ในรูปนอร์มัลฟอร์มที่ 5 (Fifth Normal Form: 5NF)

การแปลงให้อยู่ในรูปของนอร์มัลฟอร์มที่ 5 จะพิจารณาถึงการขึ้นต่อกันของข้อมูลในการแยกข้อมูลในรีเลชันออกเป็นรีเลชันย่อย และประกอบรีเลชันย่อยกลับเป็นรีเลชันใหญ่เช่นเดิม ซึ่งเป็นการตรวจสอบว่าเมื่อรวมกันใหม่ด้วยวิธีการจอยน์แล้ว จะได้รีเลชันกลับมาเหมือนเดิมทุกประการหรือไม่

1) จอยน์โอเปอรชัน (Join Operation)

2) จอยน์ดีเพนเดนซี (Join Dependency)

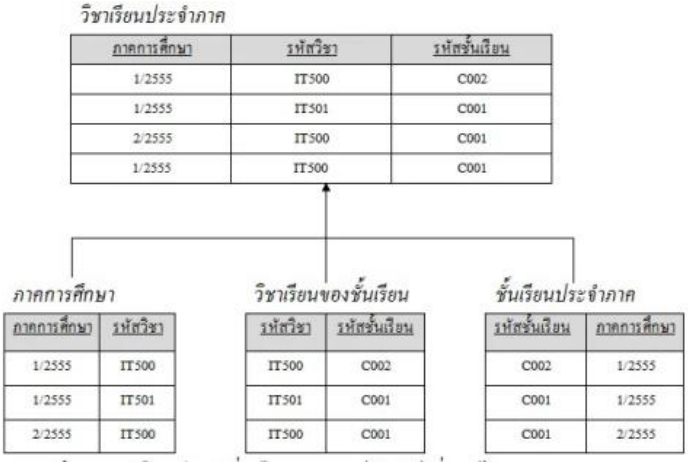
3) นิยามของ 5NF รีเลชันจะเป็น 5NF

3.1) รีเลชันนั้นเป็นนอร์มัลฟอร์มที่ 4 อยู่แล้ว

3.2) การแบ่งแยกรีเลชันมีคุณสมบัติจอยน์ดีเพนเดนซี

จากตัวอย่างในภาพข้างล่าง รีเลชัน วิชาเรียนประจำภาคอยู่ในรูปของนอร์มัลฟอร์มที่ 4 แล้ว เนื่องจากแอตทริบิวต์ภาคการศึกษาเป็นตัวกำหนดแอตทริบิวต์รหัสวิชาหลาย ค่าในขณะที่แอตทริบิวต์รหัสวิชา ก็เป็นตัวกำหนดแอตทริบิวต์รหัสชั้นเรียนหลายค่า รีเลชันนี้จึงไม่มีทริเวียลมัลติแวลูดีเพนเดนซี ต่อไปเราจึงทำการทดสอบคุณสมบัตินอร์มัลฟอร์มที่ 5 ของรีเลชันวิชาเรียนประจำภาค โดยเมื่อนำรีเลชันดังกล่าวมาทำการแตก

ย่อยออกเป็นสามรีเลชันคือ รีเลชันภาคการศึกษา รีเลชันวิชาเรียนของชั้นเรียน และรีเลชันชั้นเรียนประจำภาค และทำการจอยน์ทั้งสามรีเลชันรวมกลับเป็นหนึ่งรีเลชันอีกครั้งจะได้จำนวนข้อมูลเท่ากันกับรีเลชันก่อนที่จะมีแตกเป็นรีเลชันย่อยทุกประการ ซึ่งก็คือรีเลชันดังกล่าวมีคุณสมบัติจอยน์ดีเพนเดนซีและอยู่ในรูปของนอร์มัลฟอร์มที่ 5



ภาพที่ 2.15 แสดงตารางคุณสมบัติที่มีรีเลชันที่อยู่ในรูปนอร์มัลฟอร์มที่ 5

ที่มา : https://msit5.wordpress.com

2.3 เครื่องมือในการออกแบบ และวิเคราะห์ข้อมูล

2.3.1 แบบจำลองอนุกรมเวลา (Time series analysis) แบบจำลองอนุกรมเวลาเป็นแบบจำลองเศรษฐมิติที่ใช้ข้อมูลในอดีตอธิบายหรือทำนายตัวแปรนั้น ๆ ในอนาคต เช่น การส่งออก โดยแบบจำลองอนุกรมเวลาเป็นแบบจำลองที่นักเศรษฐศาสตร์ มีความเห็นตรงกันว่าเป็นแบบจำลองที่มีความแม่นยำในการคาดการณ์ระยะสั้นมากกว่าแบบจำลองเชิงโครงสร้าง แบบจำลองการพยากรณสามารถจำแนกได้เป็น 2 กลุ่ม

1) แบบจำลองอนุกรมเวลา (Time Series Model) เป็นแบบจำลองทางสถิติที่ใช้ข้อมูลอนุกรมเวลาในอดีตเพื่อคาดการณตัวแปรในอนาคต เช่น AR, ARIMA และ ARIMAX เป็นต้น

2) แบบจำลองเชิงโครงสร้าง (Structural Model) มีการกำหนดความสัมพันธ์ของตัวแปรทางเศรษฐกิจต่าง ๆ ในแบบจำลองตามทฤษฎีทางเศรษฐศาสตร์ไว้ เช่น DSGE Macro Model และ CGE

2.3.2 กระบวนการวิเคราะห์ข้อมูลด้วย (CRISP-DM) งานวิทยาการด้านข้อมูล (Data Science) ซึ่งกำลังมีบทบาทอย่างมากในโลกยุคปัจจุบัน และทวีความสำคัญยิ่งขึ้น

ในอนาคตก็มี CRISP–DM เป็นกระบวนการหลักในการจัดทำเหมืองข้อมูลเพื่อการวิเคราะห์และใช้ประโยชน์ในทางธุรกิจ กระบวนการวิเคราะห์ข้อมูล ด้วย CRISP–DM หรือ Cross Industry Standard Process for Data Mining พัฒนาขึ้นในปี ค.ศ. 1996 โดยความร่วมมือของ 3 บริษัท คือ Daimler Chrysler, SPSSและ NCR ประกอบด้วย 6 ขั้นตอนหลัก ได้แก่

1) รู้จักและเข้าใจในธุรกิจ (Business understanding) เป็นขั้นตอนแรกของกระบวนการ ที่มุ่งเน้นไปที่การทำความเข้าใจกระบวนการทางธุรกิจโดยรวม หัวหน้าโครงการหรือที่ปรึกษาด้านการวางระบบวิเคราะห์ข้อมูล จะต้องทำการสัมภาษณ์หรือรับฟังปัญหาความต้องการจากผู้บริหารองค์กรและหน่วยงานต่าง ๆ ที่จะนำผลการวิเคราะห์ข้อมูลไปใช้ประโยชน์ โดยความต้องการทั้งหมดจะนำมาจัดลำดับความสำคัญและกำหนดวัตถุประสงค์ที่จะนำไปสู่รูปแบบการวิเคราะห์ข้อมูลขององค์กร เช่น ผู้บริหารห้างสรรพสินค้า ต้องการรู้ว่าอะไรเป็นเหตุปัจจัยที่ทำให้ลูกค้าเป้าหมายตัดสินใจและเลือกที่จะเข้าห้าง ไม่ว่าจะเพื่อการจับจ่ายซื้อของ ใช้เป็นสถานที่นัดพบ/พักผ่อน หรือหาอาหารรับประทาน ร้านขายสินค้าออนไลน์อยากรู้ว่าผู้คนกำลังให้ความสนใจในสินค้า/บริการประเภทใดอยู่ แหล่งข้อมูลออนไลน์ใดที่มีอิทธิพลต่อการตัดสินใจเลือกซื้อสินค้าเป็นต้น

2) สร้างฐานข้อมูลให้ครบ (Data understanding) ขั้นตอนการจัดเก็บและรวบรวมข้อมูล ตลอดจนการพิจารณาตรวจสอบความถูกต้องของข้อมูลที่ได้รับ โดยเลือกว่าจะใช้ข้อมูลทั้งหมดหรือบางส่วนในการวิเคราะห์ให้สอดรับกับวัตถุประสงค์ที่กำหนดไว้ ในอดีตการศึกษาหาแนวโน้มความต้องการตลาด หรือพฤติกรรมผู้บริโภคในการตัดสินใจซื้อสินค้า เป็นเรื่องที่ยุ่งยากและต้องว่าจ้างบริษัทวิจัยสำรวจภาพรวม ควบคู่กับการพิจารณารายการสั่งซื้อสินค้าที่เก็บไว้ในฐานข้อมูลของบริษัท แต่ด้วยความก้าวหน้าทางด้านเทคโนโลยีในปัจจุบันและการทำธุรกรรมผ่านเครือข่ายอินเทอร์เน็ต ทำให้ข้อมูลมากมายมหาศาลวิ่งผ่านไปมาอยู่ในระบบเว็บไซต์หรือแอพที่เป็นช่องทางในการทำธุรกรรมต่าง ๆ จึงเป็นแหล่งข้อมูลสำคัญ อีกทั้งยังได้ข้อมูลความสนใจของคนที่พร้อมยอมให้อย่างเต็มใจจากห้องแชทต่าง ๆ ที่มีการพูดคุยหารือกัน ปัจจุบันการแกะรอยหรือสะกดรอยตามคนได้ดีที่สุดเกิดขึ้นได้ง่ายมากจากออนไลน์ ไม่ว่าจะเป็นพิกัดตำแหน่งที่อยู่ของเราที่อนุญาตให้แอพ ต่าง ๆ เข้าถึง

3) เตรียมข้อมูลให้พร้อมใช้ (Data preparation) ขั้นตอนการแปลงข้อมูลที่ได้รวบรวมมาและเลือกไว้ให้อยู่ในรูปแบบที่พร้อมสำหรับนำไปวิเคราะห์ในขั้นตอนต่อไปได้โดยการทำให้เป็นข้อมูลที่ถูกต้อง (*Data cleaning)* มักใช้เวลาค่อนข้างมาก ระบบการรับข้อมูลป้อนเข้าสู่ระบบที่ทันสมัยในปัจจุบันจะลดการคีย์ข้อมูลจากคนให้น้อยที่สุด แต่จะใช้วิธีการสแกน การติ๊กเลือก เพื่อลดความผิดพลาดให้น้อยที่สุด เพราะขั้นตอนใช้เวลามากกว่า 50%

ของเวลารวมทั้งหมด การลดข้อผิดพลาดของข้อมูลได้มากเท่าใดก็จะยิ่งมีประสิทธิภาพมากขึ้นเท่านั้น

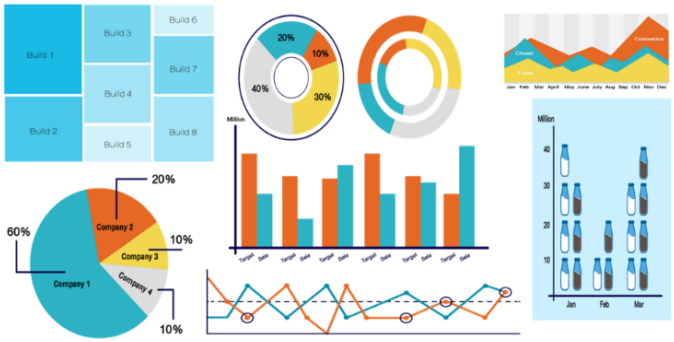
4) จัดทำและเลือกโมเดลที่ใช่ (Modeling) ขั้นตอนการสร้างตัวแบบทางคณิตศาสตร์และสถิติเพื่อการวิเคราะห์ข้อมูล โดยสามารถใช้เทคนิควิธีการต่าง ๆผสมผสานกัน อาทิ การจำแนก (Classification) การแบ่งกลุ่ม (Clustering) และการสร้างความสัมพันธ์ (Association rule) ในร้านสะดวกซื้อ จะนำข้อมูลการซื้อสินค้าของลูกค้าแต่ละรายมาหาความสัมพันธ์ เช่น คนที่ซื้อเครื่องดื่มแต่ละชนิดมักจะซื้อขนมหรือของกินอะไรร่วมอยู่ด้วย การใช้จ่ายของแต่ละคนจะอยู่ที่ประมาณกี่บาท คนส่วนใหญ่ที่เข้ามาจะซื้อสินค้ากี่ชิ้นต่อคน และเพื่อให้ทราบข้อมูลของผู้ซื้อ ร้านค้ามักจะใช้การออกบัตรเติมเงินที่จูงใจให้ใช้จากส่วนลดหรือสะสมแต้ม ทำให้สามารถติดตามประวัติการใช้จ่ายได้ง่ายขึ้น ซึ่งปัจจุบันมีการนำกล้องจับภาพผู้ซื้อในการแยกแยะเพศ อายุ และไลฟ์สไตล์ของคน

5) ประเมินผลก่อนตัดสินใจ (Evaluation) เป็นขั้นตอนก่อนนำผลลัพธ์ที่ได้จากขั้นตอนที่ 4 ไปใช้งาน ด้วยการวัดประสิทธิผลของผลลัพธ์ที่ได้กับวัตถุประสงค์ที่ตั้งไว้ในขั้นตอนแรก ว่ามีนัยสำคัญหรือความน่าเชื่อถือมากน้อยเพียงใด ทั้งนี้อาจต้องกลับไปทบทวนขั้นตอนที่ 2 – 4 ซ้ำอีกครั้ง ในกรณีที่ผลลัพธ์ไม่มีความน่าเชื่อถือเพียงพอหรือไม่สอดรับกับวัตถุประสงค์

6) เผยแพร่ผลวิเคราะห์ (Deployment) ขั้นตอนการนำผลลัพธ์ที่ได้ไปใช้งานเป็นการทั่วไป อาจจัดทำเป็นรูปแบบของรายงาน (*Report)* หรือแผนภาพ (Dashboard) ที่พร้อมให้ฝ่ายต่าง ๆ นำไปใช้ประโยชน์ในการวางแผน กำหนดกลยุทธ์ ในทางธุรกิจต่อไป

2.3.2 การแสดงผลแบบ (visualization)

2.3.2.1 การแสดงผลแบบ visualization ด้วย Tableau PublicTableau ทำหน้าที่แปลงข้อมูล (*Data)* ให้เป็นภาพ (*Visualization)* ตัวอย่างของการแปลงข้อมูลเป็นภาพเป็นสิ่งที่เราคุ้นเคย และมีมาเนิ่นนานแล้วก็คือการสร้างกราฟนั่นเอง จะเห็นได้ว่าเรามีการสร้าง ตาราง กราฟแท่ง กราฟเส้น กราฟวงกลม (pie) มาก่อนแล้วกราฟที่กล่าวมาถือเป็น Visualization อย่างหนึ่ง แต่ก็ยังมีกราฟรูปแบบอื่น ๆ อีกมากมายที่มีประสิทธิภาพและทรงพลัง *3* ในการสื่อสารมากกว่า รูปด้านล่างคือตัวอย่าง Visualization ที่สร้างโดย Tableau



ภาพที่ 2.16 กราฟในลักษณะต่าง ๆ

ที่มา : https://sites.google.com/

Tableau เป็นซอฟต์แวร์สำหรับทำ Data Visualization ที่ได้รับความนิยมอันดับต้นของโลก Gartner ได้จัดให้ Tableau อยู่ใน Magic Quadrant for Analytics and Business Intelligence Platform โดยได้คะแนนอยู่ในกลุ่มของผู้นำ (Leader) และมีคะแนน Ability to Execute สูงที่สุด Tableau ในมุมมองด้าน BI นั้น Tableau มีความโดดเด่นในการนำเสนอข้อมูลผ่าน Dashboard ที่สวยงามและมีประสิทธิภาพ โดยให้นิยามตัวเองว่าเป็น Modern BI ที่ส่งเสริมการทำ Self Service BI ให้สำเร็จได้ Tableau ในมุมมองด้าน Analytics นั้น Tableau มีความโดดเด่นในการทำ Descriptive Analytics และ Diagnostic Analytics ถ้าหากต้องการทำ Analytics ในระดับที่สูงขึ้นก็สามารถเชื่อมต่อกับภาษา R หรือ Python ได้

2.3.2.2 ลักษณะเด่นของ Tableau Public

1) ใช้งานง่ายด้วยหลัก Drag-and-Drop Interface โดยสามารถสร้าง Visualization จากการ Drag ข้อมูล (Dimension หรือ Measure) ไป Drop บน Shelf ต่าง ๆ (มีอยู่ประมาณ 10 Shelves เท่านั้น) แล้วโปรแกรม Tableau จะตีความ และสร้างเป็นภาพให้เห็นทันที

2) Tableau พัฒนาอย่างต่อเนื่อง เพื่อนำ User Experience มาปรับปรุงซอฟต์แวร์อยู่ตลอดเวลา โดยซอฟต์แวร์ มีการลงทุน R&D ในสัดส่วนที่สูงมาก และมีการ Upgrade เวอร์ชัน หลักเป็นประจำทุกปี ทำให้การแสดงผลทั้งใน Desktop, Mobile หรือแม้แต่ Browser แสดงผลได้ง่ายและสวยงาม

3) เพราะ Tableau ออกแบบมาเพื่อ Business User ไม่จำเป็นต้องเป็น IT User หรือผู้ที่เรียนจบมาทางด้านคอมพิวเตอร์ Tableau มองว่าผู้ที่เข้าใจข้อมูลน่าจะเป็นคนที่มีคำถามกับข้อมูล ไม่ใช่ว่ามีคำถามที่สงสัยแล้วต้องให้ฝ่าย IT เป็นคนวิเคราะห์หรือทำ Report ให้ Tableau สามารถทำแบบนี้ได้จำเป็นต้องทำให้ซอฟต์แวร์ใช้งานง่ายพอ Benchmark ที่สำคัญก็คือโปรแกรม Microsoft Excel (ซึ่งแทบจะใช้งานเป็นแทบทุกคน) Tableau ได้พิสูจน์และเป็นที่ยอมรับว่า Tableau ใช้ทำกราฟได้ง่ายกว่า เร็วกว่า และสวยกว่า Excel

4) ภาษา VizQL ผู้บิดทองหลังพระให้กับ Tableau VizQL คือ Visual Query Language ที่แปลงการ Drag-and-Drop เป็น Query การดึงข้อมูลแล้วแสดงผลเป็นภาพโดย ทำหน้าที่จัดการความซับซ้อนเบื้องหลังของการ Query และ Analysis VizQL ได้ถูกจดสิทธิบัตรโดย Tableau รูปด้านล่างแสดงการทำงานของ VizQL ผลของ VizQL ทำให้ผู้ใช้งานได้รับประสบการณ์อันน่าประทับใจในการสร้าง Visualization เพราะเราจะเห็นภาพทันทีที่เรา Drag-and-Drop ส่งผลให้เราสามารถตอบคำถามได้เร็วเท่าที่เราคิด นี่เป็นเหตุผลที่ Tableau สามารถทำ Data Exploration ได้โดดเด่นกว่าคู่แข่งมาก

2.4 วรรณกรรมที่เกี่ยวข้อง

กฤติยา พุตติ (2560) การศึกษานี้มีวัตถุประสงค์เพื่อศึกษาการขยายตัวของพื้นที่เมืองและอาคารประเภทที่อยู่อาศัยของเทศบาลนครนนทบุรีและศึกษาแนวโน้มการเกิดปริมาณขยะมูลฝอยในอนาคตจากการขยายตัวอาคารประเภทที่อยู่อาศัยในพื้นที่เทศบาลนครนนทบุรี ส่วนใหญ่มีประชากรเพิ่มมากขึ้นจึงทำให้ปริมาณการทิ้งขยะนั้นเพิ่มจำนวนขึ้น ซึ่งทำให้การปรับเปลี่ยนพฤติกรรมในการใช้ชีวิตประจำวัน เพื่อลดปริมาณขยะจะช่วยสามารถลดค่าใช้จ่าย และในอดีตมีการเกิดอุทกภัยขึ้นในแต่ละจังวัดก็จะส่งผลต่อสภาพเศรษฐกิจของประเทศไทยทำให้เกิดความเสียหายเพิ่มโดยปริมาณขยะมูลฝอยนั้นเพิ่มจำนวนมากจึงต้องมีความรู้เกี่ยวกับการพยากรณ์ปริมาณขยะมูลฝอยในประเทศไทยและมีการคัดแยกขยะ เพื่อให้มีการจัดการปริมาณขยะที่ลดลง ทำให้ขยะแต่ละบางประเภทมีการนำกลับมาใช้ใหม่ได้โดยที่ปริมาณขยะไม่เพิ่มมากขึ้น จึงทำให้เกิดการกำจัดขยะแบบถูกวิธี เช่น การฝังกลบ การนำกลับมาใช้ใหม่ จึงเล็งเห็นความสำคัญของการรวบรวมข้อมูลต่าง ๆ ของเรื่องปริมาณขยะมูลฝอย ในประเทศไทย โดยที่จะเชื่อมโยงกับฐานข้อมูลอื่น ๆ จากหน่วยงานที่เกี่ยวข้องเพื่อประกอบการตัดสินใจในการแก้ไขปัญหาและการลดปริมาณขยะโดยการนำข้อมูล

ปริมาณขยะมูลฝอยมาเข้าสู่กระบวนการวิเคราะห์ข้อมูล การรายงานผลของการพยากรณ์ผ่านเว็บไซต์ เพื่อที่จะคาดการณ์ได้อย่างแม่นยำ และเหมาะสมกับสถานการ์ณที่จะเกิดขึ้น โดยข้อมูลที่นำมาใช้ในการศึกษาเป็นข้อมูลจากกระทรวงทรัพยากรธรรมชาติและสิ่งแวดล้อม ตั้งแต่ปี 2553–2562 โดยใช้เครื่องมือที่เหมาะสมในการวิเคราะห์ข้อมูลทางดาต้า ไมน์นิ่ง เก็บรวบรวมข้อมูล เพื่อนำมาพยากรณ์และเปรียบเทียบข้อมูลด้วยเทคนิคทางสถิติ การเปรียบเทียบความคลาดเคลื่อนของการพยากรณ์ในวิธีการคาดประมาณจำนวนประชากรในพื้นที่ศึกษาโดยใช้สมการ exponential method คำนวณจำนวนประชากรในอนาคตในช่วง 10,20 และ 50 ปี ข้างหน้าด้วยสูตร Pt = P0ern [5] โดย Pt คือจำนวนประชากรในเวลาในอนาคตที่ต้องการฉายภาพ P0 คือ จำนวนประชากรฐานหรือประชากรในเวลาเริ่มต้นของการฉายภาพ r คือ อัตราเพิ่มประชากรต่อปี (ข้อมูลประชากร ปี พ.ศ. 2548 –2558) n คือ ระยะเวลาหรือจำนวนปีที่ต้องการฉายภาพในการโดยที่จะนำเสนอผ่านทางเว็บไซต์ เป็นต้น

สุรศักดิ์ โอสถิตพร (2558) ได้ทำการศึกษาเรื่อง รูปแบบการจัดการขยะมูลฝอยที่เหมาะสมกับเทศบาลตำบลดอยสะเก็ด อำเภอดอยสะเก็ด จังหวัดเชียงใหม่ โดยใช้แบบสอบถามและแบบสัมภาษณ์แบบมีโครงสร้าง กับประชากร 2 กลุ่ม คือประชาชนและผู้บริหารเทศบาล ได้แก่ นายกเทศมมนตรี สมาชิกสภาเทศบาล ผลการศึกษาพบว่าแนวคิดและวิธีการจัดการขยะมูลฝอยของประชาชนที่พักอาศัยอยู่ในเขตเทศบาลตำบลดอยสะเก็ดพบว่า กลุ่มตัวอย่างมีวิธีการจัดการลดขยะมูลฝอยภายในครัวเรือนตามหลักการจัดการขยะ 5R ซึ่งมีการทำมาโดยตลอด ด้านประชาชนให้ความเห็นว่า เทศบาลดอยสะเก็ดควรหาแนวทางประสานร่วมมือกับทุกภาคส่วนแบบมีส่วนร่วมโดยมีการวางแผนงานในส่วนของการจัดการปัญหาขยะมูลฝอยทั้งระบบซึ่งประกอบด้วยการศึกษาสภาพปัญหา การวางแผนแก้ไขปัญหามูลฝอย การมีส่วนร่วมในการติดตาม ตรวจสอบ และประเมินผลการจัดการขยะมูลฝอย ทั้งนี้ต้องมีความชัดเจนเกี่ยวกับนโยบายการจัดการขยะมูลฝอยในพื้นที่เทศบาลตำบลดอยสะเก็ด

ชลิดา ตระกูลสุนทร (2561) การศึกษาการวิจัยนี้มีวัตถุประสงค์เพื่อ 1.) พัฒนาตัวประมาณอัตราส่วนของพารามิเตอร์ค่าเฉลี่ยภายใต้เงื่อนไขค่ามากสุดและค่าน้อยสุด (yRMC) ภายใต้การสุ่มตัวอย่างอย่างง่ายแบบไม่คืนที่ 2.) เปรียบเทียบประสิทธิภาพ (MSE และ PRE) ของตัวประมาณที่พัฒนากับตัวประมาณพารามิเตอร์ค่าเฉลี่ย (y) ตัวประมาณอัตราส่วนของ

พารามิเตอร์ค่าเฉลี่ย (yr) ตัวประมาณพารามิเตอร์ค่าเฉลี่ยภายใต้เงื่อนไขค่ามากสุดและค่าน้อยสุด (y s) และตัวประมาณอัตราส่วนของ Subramani and Kumarapandiyan (y p) ทั้งทางทฤษฎีและการจาลองสถานการณ์ 150 สถานการณ์ 3.) ประมาณปริมาณขยะมูลฝอยที่จัดเก็บได้เฉลี่ยต่อปีของสานักงานเขตพื้นที่ กรุงเทพมหานคร ปีงบประมาณ 2562 ด้วยตัวประมาณที่พัฒนา โดยกำหนดเกณฑ์ค่าร้อยละความคลาดเคลื่อนสัมพัทธ์เฉลี่ยไม่เกินร้อยละ 10

จตุรงค์ พะยอมแย้ม (2559) การศึกษาในครั้งนี้มีวัตถุประสงค์เพื่อคาดการณ์การปริมาณขยะจากข้อมูลของการขยายตัวของพื้นที่เมือง ในเขตเทศบาลตำบล เจ้าพระยาสุรศักดิ์ จังหวัดชลบุรี โดยการหาขนาดของเมืองที่เปลี่ยนแปลงไปนั้นจะใช้การศึกษาการเปลี่ยนแปลงการใช้ประโยชน์ที่ดินปีตั้งแต่ พ.ศ.2559 ถึงปี พ.ศ.2570 ด้วยแบบจำลอง Cellular Automata Markov และนำข้อมูลขนาดของเมืองที่ได้มาทำการคาดการณ์ปริมาณขยะที่เกิดขึ้นและเปรียบเทียบกับการคาดการณ์การปริมาณขยะที่คำนวณจากการเพิ่มของประชากรเพื่อทำการวิเคราะห์และสรุปถึงผลของคาดการณ์การปริมาณขยะที่จะเกิดขึ้นในอนาคต โดยผลการศึกษาการเปลี่ยนแปลงการใช้ประโยชน์ที่ดินแสดงให้เห็นว่าพื้นที่เมืองมีขนาดเพิ่มขึ้น คือในปี พ.ศ.2559 พื้นที่เมืองมีขนาดอยู่ประมาณ 27.10 ตารางกิโลเมตร และในปีในปี พ.ศ.2570 มีขนาดเพิ่มขึ้นเป็นประมาณ 69.55 ตารางกิโลเมตร โดยผลจากการคาดการณ์ปริมาณขยะจากการคำนวณทั้งจากการเปลี่ยนแปลงของประชากรและจากการขยายตัวของเมือง โดยปริมาณขยะจากการคาดการณ์ทั้ง 2 แบบนั้นไม่แตกต่างกันมากนักและมีทิศทางที่เพิ่มขึ้นทุกปี จากผลการศึกษาสามารถสรุปได้ว่า ปริมาณขยะจากการคาดการณ์ด้วยจากการเปลี่ยนแปลงของประชากรและจากการขยายตัวของเมืองนั้นมีความใกล้เคียงกัน ซึ่งสามารถนำเอาผลจากการคาดการณ์ปริมาณขยะจากทั้ง 2 วิธีไปใช้ในการกำหนดนโยบายและแผนในการจัดการขยะมูลฝอยของเทศบาลตำบลเจ้าพระยาสุรศักดิ์ในอนาคตได้

สุกันยา บัวลาด (2560) การศึกษาการวิจัยมีวัตถุประสงค์เพื่อ 1.) ศึกษาระดับการจัดการขยะมูลฝอยที่มีประสิทธิภาพ 2.) ศึกษาปัจจัยที่ส่งผลต่อการจัดการขยะมูลฝอยที่มีประสิทธิภาพ และ 3.) ศึกษาข้อเสนอแนะต่อการจัดการขยะมูลฝอยที่มีประสิทธิภาพของประชาชนในเขตพื้นที่องค์การบริหารส่วนตำบลท่าสองคอน อำเภอเมือง จังหวัดมหาสารคาม โดยกลุ่มตัวอย่าง คือ หัวหน้าครัวเรือนหรือสมาชิกในครัวเรือนที่มีอายุ 18 ปี บริบูรณ์ขึ้นไปในเขตพื้นที่องค์การบริหารส่วนตำบลท่าสองคอน อำเภอเมือง จังหวัดมหาสารคาม

จำนวน 374 คน ได้มาโดยการสุ่มตัวอย่างแบบเป็นระบบผลการวิจัย พบว่าการจัดการขยะมูลฝอยของประชาชนในเขตพื้นที่องค์การบริหารส่วนตำบลท่าสองคอน อำเภอเมือง จังหวัดมหาสารคาม โดยรวมอยู่ในระดับมาก และเมื่อพิจารณาเป็นรายด้าน พบว่าอยู่ในระดับมากที่สุด คือ การคัดแยกขยะมูลฝอย อยู่ในระดับมาก คือ การเก็บรวบรวมขยะมูลฝอยและอยู่ในระดับน้อย คือ การทิ้งขยะมูลฝอย และการกำจัดขยะมูลฝอย ปัจจัยที่ส่งผลต่อการจัดการขยะมูลฝอยที่มีประสิทธิภาพของประชาชนในเขตพื้นที่องค์การบริหารส่วนตำบลท่าสองคอน อำเภอเมือง จังหวัดมหาสารคาม ได้แก่ จำนวนสถานะหรือบทบาทที่แสดงในชุมชน เจตคติเกี่ยวกับการจัดการขยะมูลฝอยในชุมชนการมีส่วนร่วมในการจัดการขยะมูลฝอยในชุมชนโดยตัวแปรดังกล่าวสามารถร่วมกันพยากรณ์การจัดการขยะมูลฝอยที่มีประสิทธิภาพของประชาชนในพื้นที่องค์การบริหารส่วน ตำบลท่าสองคอนได้ร้อยละ 33.2 อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.00 โดยมีความคลาดเคลื่อนมาตรฐานในการพยากรณ์ เท่ากับ ±0.377 การจัดการขยะมูลฝอยที่มีประสิทธิภาพของประชาชนในเขตพื้นที่องค์การบริหารส่วน ตำบลท่าสองคอน ได้แก่ ควรมีการให้ความรู้ประชาชนเกี่ยวกับการจัดการขยะมูลฝอยอย่างถูกวิธี สร้างความเข้าใจอย่างต่อเนื่องเกี่ยวกับผลกระทบของปัญหาขยะมูลฝอย

2.5 บทสรุป

วรรณกรรมที่เกี่ยวข้องทั้งหลายที่ได้กล่าวมาในข้างต้นทั้งหมดนั้น ผู้วิเคราะห์โครงงานได้เล็งเห็นความสำคัญของการเปรียบเทียบการพยากรณ์ปริมาณขยะมูลฝอยในประเทศไทย และเผยแพร่ข้อมูลสารสนเทศบน web browser เนื่องจากเทคโนโลยีปัจจุบันเข้ามาช่วยในการจัดการข้อมูลหรือดำเนินงานต่าง ๆ ให้มีความสะดวกสบายเป็นอย่างมาก อีกทั้งผู้ใช้งานยังเข้าถึงข้อมูลได้ง่าย และมีประสิทธิภาพ

บทที่ 3
วิธีการดำเนินโครงงาน

ผู้จัดทำได้วิเคราะห์ความสัมพันธ์ของแหล่งข้อมูลและการจัดการข้อมูลที่มีอยู่ให้เกิดประโยชน์และได้สารสนเทศที่สอดคล้องกับข้อมูลผู้จัดทำได้ศึกษาวิเคราะห์ข้อมูลเชิงสัมพันธ์เชิงสาเหตุและการจัดการข้อมูลได้อยู่ในรูปแบบที่ถูกต้องโดยได้วิเคราะห์และออกแบบข้อมูลใหม่โดยใช้เครื่องมือประกอบไปด้วย กระบวนการในการพัฒนาฐานข้อมูล CRISP-DM สร้างแบบฟอร์มข้อมูลและปรับโครงสร้างข้อมูลใหม่ เทคนิควิธีทางเหมืองข้อมูล วิเคราะห์ความสัมพันธ์ข้อมูล (Visual Analytics tableau) ในการวิเคราะห์ข้อมูลซึ่งในบทนี้จะแสดงถึงวิธีในการดำเนินงาน ดังนี้

3.1 แผนภาพกระแสข้อมูล Data Flow Diagram

แผนภาพกระแสข้อมูล Data Flow Diagram คือ แบบจำลองกระบวนการ (Process Model) เป็นประเภทหนึ่ง มักนำมาใช้ในการวิเคราะห์และออกแบบระบบเชิงโครงสร้าง โดยแสดงความสัมพันธ์ระหว่างกระบวนการกับข้อมูล ซึ่งเป็นการแสดงการเคลื่อนที่ของข้อมูลจากกระบวนการหนึ่งไปยังกระบวนการหนึ่งโดยไม่อิงเทคโนโลยีการจัดเก็บข้อมูล เช่น สื่อบันทึกข้อมูล ฯลฯ

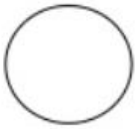
สัญลักษณ์ที่ใช้เป็นมาตรฐานในการแสดงแผนภาพกระแสข้อมูลมีหลายชนิด แต่ในที่นี้จะแสดงให้เห็นเพียง 2 ชนิด ได้แก่ ชุดสัญลักษณ์มาตรฐานที่พัฒนาโดย Gane and Sarson (1979) และชุดสัญลักษณ์มาตรฐานที่พัฒนาโดย DeMarco and Yourdon (DeMarco, 1979); Yourdon and Constantine, 1979) โดยมีสัญลักษณ์ดังต่อไปนี้

3.1.1 สัญลักษณ์ที่ใช้ในการวาดแผนภาพกระแสข้อมูล

สัญลักษณ์ที่นิมใช้ในการวาดแผนภาพกระแสข้อมูลมี 2 แบบ คือ

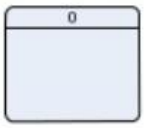
1) Gane and Sarson
2) DeMarco and Yourdon

ตารางที่ 3.1 สัญลักษณ์การไหลของข้อมูล แบบ Gane and Sarson

สัญลักษณ์	คำอธิบาย
	กระบวนการ (Process)
	การเคลื่อนที่ของข้อมูล (Data Flow)
	หน่วยภายนอก (External Entity)
	แฟ้มข้อมูล (Data Store/File)

ที่มา : https://www.gotoknow.org (2560)

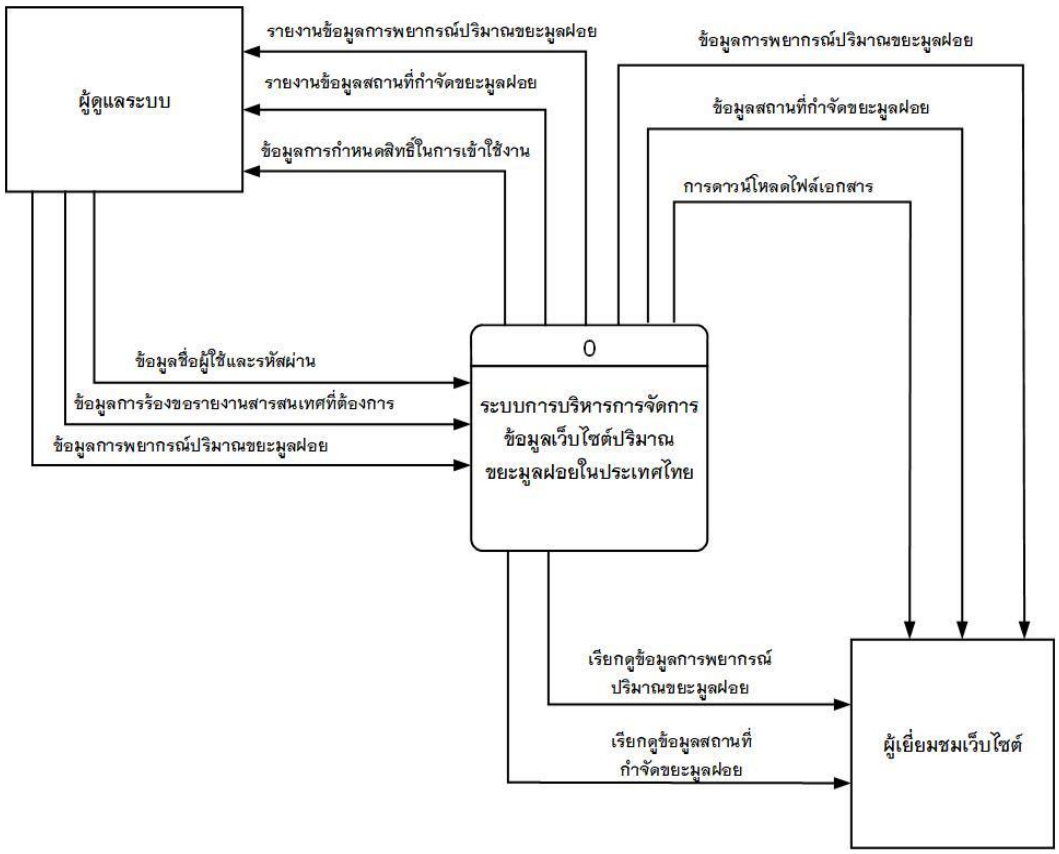
ตารางที่ 3.2 สัญลักษณ์การไหลของข้อมูล แบบ Yourdon

สัญลักษณ์	คำอธิบาย
	กระบวนการ (Process)
	การเคลื่อนที่ของข้อมูล (Data Flow)
	หน่วยภายนอก (External Entity)
	แฟ้มข้อมูล (Data Store/File)

ที่มา : https://www.gotoknow.org (2560)

3.1.2 แผนภาพบริบท Context Diagram

แผนภาพบริบท (Context Diagram) จะแสดงภาพโดยรวมระบบของการพัฒนาเว็บแอปพลิเคชันเพื่อการวิเคราะห์ข้อมูลปริมาณขยะมูลฝอยในประเทศไทย ดังภาพ 3.1

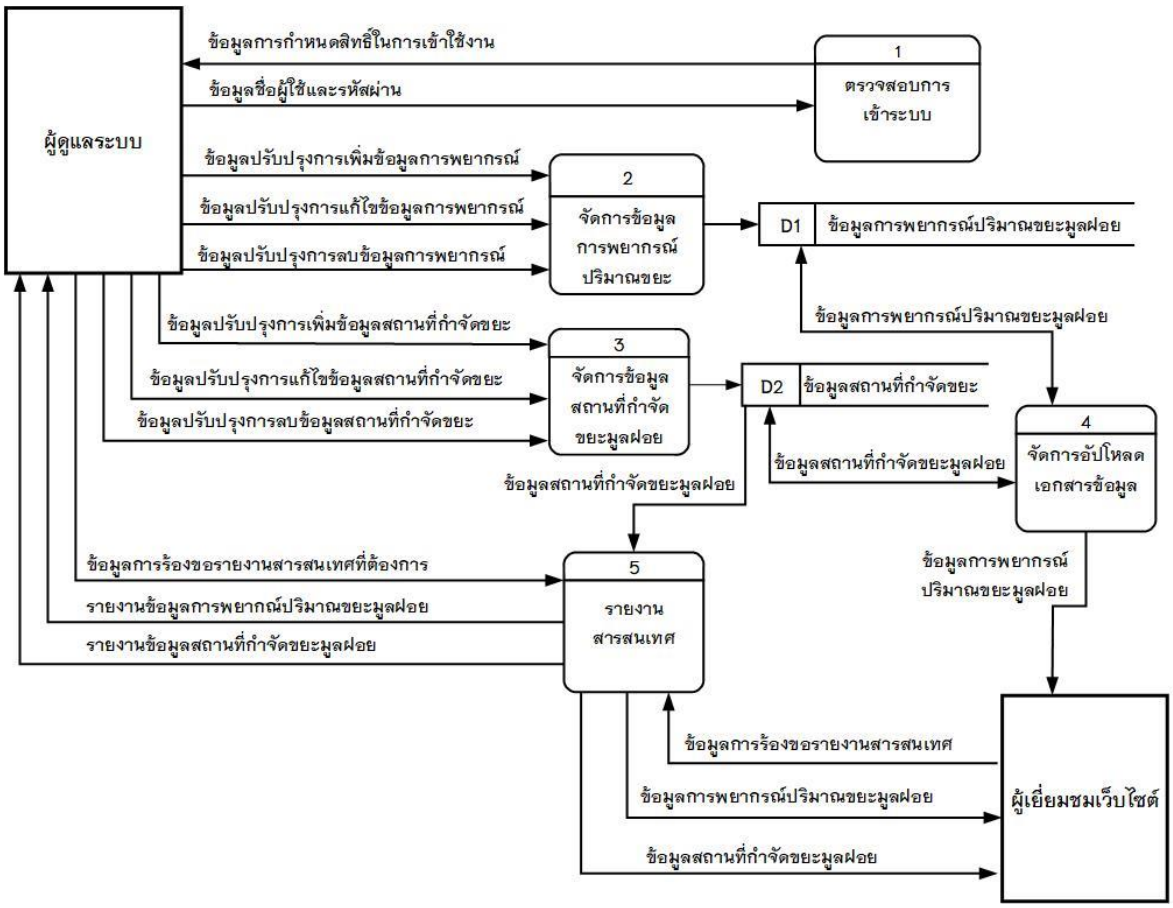


ภาพที่ 3.1 แผนภาพบริบท Context Diagram

จากภาพที่ 3.2 เป็นแผนผังระบบของการพัฒนาระบบบริหารจัดการพัฒนาเว็บแอปพลิเคชันเพื่อการวิเคราะห์ข้อมูลปริมาณขยะมูลฝอยในประเทศไทย
โดยสามารถแบ่งผู้ใช้ออกเป็น 2 ประเภทดังนี้

1) **ผู้ดูแลระบบ (System Administrator)** สามารถล็อกอินได้ สามารถออกจากระบบได้ สามารถจัดการข้อมูลผู้ใช้ระบบทั้งหมด

2) **ผู้ใช้ทั่วไป (User)** สามารถดูแหล่งการเรียนรู้ได้ สามารถดูการพยากรณ์ของปริมาณขยะ สามารถดูสถานที่กำจัดขยะ



ภาพที่ 3.2 แสดงแผนภาพกระแสข้อมูลระดับที่ 0 (Data Flow Diagram Level 0)

จากภาพที่ 3.2 แสดงแผนภาพกระแสข้อมูลระดับที่ 0 (Data Flow Diagram Level 0) ระบบบริหารฐานข้อมูลปริมาณขยะมูลฝอยในระเทศไทยและสถานที่กำจัดปริมาณขยะประกอบด้วย 5 กระบวนการหลักดังนี้

ตารางที่ 3.3 แสดงคำอธิบายกระบวนการ ตรวจสอบระบบ

Process Description	
System	ระบบบริหารจัดการพัฒนาเว็บแอปพลิเคชันเพื่อการวิเคราะห์ข้อมูลปริมาณขยะมูลฝอยในประเทศไทย
DFD Number	1
Process Name	ตรวจสอบการเข้าระบบ
Input Data Flow	ชื่อผู้ใช้และรหัสผ่าน ข้อมูลสิทธิ์การเข้าใช้ระบบ ข้อมูลผู้ดูแลระบบ
Output Data Flow	สิทธิ์ในการเข้าถึงระบบ สิทธิ์การเข้าใช้ระบบ ข้อมูลผู้ดูแลระบบ
Data Store Used	ข้อมูลผู้ใช้และรหัสผ่าน ข้อมูลผู้ดูแลระบบ
Description	เป็นกระบวนการสำหรับตรวจสอบ และกำหนดสิทธิ์ในการเข้าใจระบบ ผู้ใช้ถึงจะมีสิทธิ์เข้าใช้ระบบโดยมีผู้ใช้ระบบได้แก่ ผู้ดูแลระบบ

ตารางที่ 3.4 แสดงคำอธิบายกระบวนการ การจัดการข้อมูลการพยากรณ์ปริมาณขยะ

Process Description	
System	ระบบบริหารจัดการพัฒนาเว็บแอปพลิเคชันเพื่อการวิเคราะห์ข้อมูลปริมาณขยะมูลฝอยในประเทศไทย
DFD Number	2
Process Name	จัดการข้อมูลการพยากรณ์ปริมาณขยะมูลฝอยในประเทศไทย
Input Data Flow	ข้อมูลการพยากรณ์ปริมาณขยะมูลฝอยในประเทศไทย
Output Data Flow	ข้อมูลการพยากรณ์ปริมาณขยะมูลฝอยในประเทศไทย
Data Store Used	ข้อมูลการพยากรณ์ปริมาณขยะมูลฝอยในประเทศไทย
Description	เป็นกระบวนการสำหรับจัดการข้อมูลเกี่ยวกับการพยากรณ์ปริมาณขยะในประเทศไทย โดยผู้ใช้ระบบในส่วนนี้ได้แก่ การเพิ่ม ลบ แก้ไข ผู้มีสิทธิ์ใช้งานกระบวนการนี้ ได้แก่ ผู้ดูแลระบบ และออกสารสนเทศข้อมูลการพยากรณ์ปริมาณขยะมูลในประเทศไทย

ตารางที่ 3.5 แสดงคำอธิบายกระบวนการ จัดการข้อมูลสถานที่กำจัดขยะมูลฝอยในประเทศไทย

Process Description	
System	ระบบบริหารจัดการพัฒนาเว็บแอปพลิเคชันเพื่อการวิเคราะห์ข้อมูลปริมาณขยะมูลฝอยในประเทศไทย
DFD Number	3
Process Name	จัดการข้อมูลการสถานที่กำจัดขยะมูลฝอยในประเทศไทย
Input Data Flow	ข้อมูลการสถานที่กำจัดขยะมูลฝอยในประเทศไทย
Output Data Flow	ข้อมูลการสถานที่กำจัดขยะมูลฝอยในประเทศไทย
Data Store Used	ข้อมูลการสถานที่กำจัดขยะมูลฝอยในประเทศไทย
Description	เป็นกระบวนการสำหรับการแสดงสถานที่กำจัดในแต่ละจังหวัด โดยผู้ใช้ระบบในส่วนนี้ได้แก่ การเพิ่ม ลบ แก้ไข ผู้มีสิทธิ์ใช้งานกระบวนการนี้ ได้แก่ ผู้ดูแลระบบ และออกสารสนเทศข้อมูลสถานที่กำจัดขยะมูลในประเทศไทย

ตารางที่ 3.6 แสดงคำอธิบายกระบวนการจัดการอัปโหลดเอกสารข้อมูลต่างๆ

Process Description	
System	ระบบบริหารจัดการพัฒนาเว็บแอปพลิเคชันเพื่อการวิเคราะห์ข้อมูลปริมาณขยะมูลฝอยในประเทศไทย
DFD Number	4
Process Name	จัดการอัปโหลดเอกสารข้อมูลต่างๆ
Input Data Flow	ข้อมูลการพยากรณ์ปริมาณขยะมูลฝอยในประเทศไทย ข้อมูลสถานที่กำจัดขยะมูลฝอยในประเทศไทย
Output Data Flow	ข้อมูลการพยากรณ์ปริมาณขยะมูลฝอยในประเทศไทย และข้อมูลสถานที่กำจัดขยะมูลฝอยในประเทศไทย เป็นไฟล์ Excel หรือ PDF และเอกสารที่เกี่ยวข้องที่เป็นไฟล์ Excel หรือ PDF
Data Store Used	ข้อมูลการพยากรณ์ปริมาณขยะมูลฝอยในประเทศไทย และข้อมูลสถานที่กำจัดขยะมูลฝอยในประเทศไทย

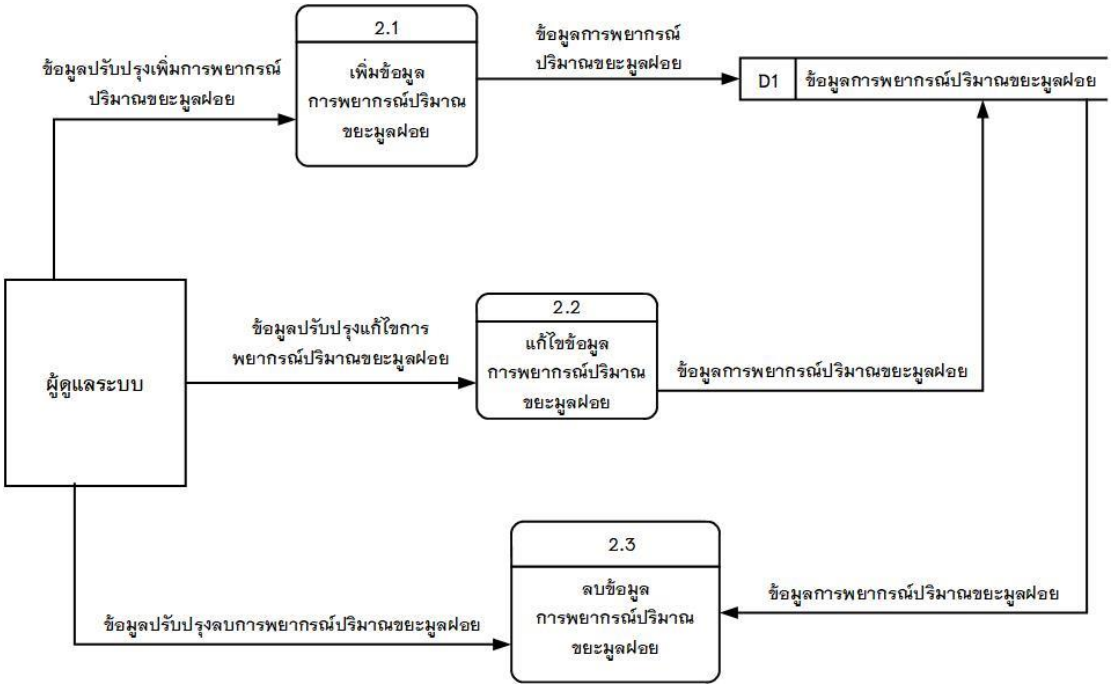
Description	เป็นกระบวนการสำหรับจัดการดาวน์โหลดเอกสารต่าง ๆเช่นข้อมูลการพยากรณ์ปริมาณขยะมูลฝอยในประเทศไทย และสถานที่กำจัดขยะมูลฝอยในประเทศไทยที่ผู้เยี่ยมชมเว็บไซต์สามารถดาวน์โหลดได้

ตารางที่ 3.7 แสดงคำอธิบายกระบวนการจัดการรายงานสารสนเทศ

Process Description	
System	ระบบบริหารจัดการพัฒนาเว็บแอปพลิเคชันเพื่อการวิเคราะห์ข้อมูลปริมาณขยะมูลฝอยในประเทศไทย
DFD Number	5
Process Name	จัดการรายงานสารสนเทศ
Input Data Flow	ข้อมูลการพยากรณ์ปริมาณขยะมูลฝอยในประเทศไทย,ข้อมูลสถานที่กำจัดขยะมูลฝอยในประเทศไทย
Output Data Flow	รายงานข้อมูลการพยากรณ์ขยะมูลฝอยในประเทศไทย และรายงานข้อมูลสถานที่กำจัดมูลฝอยในประเทศไทย
Data Store Used	ข้อมูลปริมาณการพยากรณ์ปริมาณขยะ, ข้อมูลสถานที่กำจัดขยะมูลฝอยในประเทศไทย
Description	เป็นกระบวนการสำหรับออกรายงานสารสนเทศ โดยผู้มีสิทธิ์เข้าใช้ระบบได้แก่ ผู้ดูแลระบบ

จากแผนภาพกระแสข้อมูล ระดับที่ 0 สามารถแยกย่อยเป็นกระบวนการย่อยระดับที่ 1 ได้ดังนี้

แผนภาพกระแสข้อมูล ระดับที่ 1 กระบวนการที่ 2 จัดการข้อมูล ประเด็นที่ 1 ปริมาณขยะมูลฝอยในประเทศไทย



ภาพที่ 3.3 แผนภาพกระแสข้อมูล ระดับที่ 1 กระบวนการที่ 2.1, 2.2, 2.3 จัดการข้อมูลการพยากรณ์ปริมาณขยะมูลฝอยในประเทศไทย

ตารางที่ 3.8 แสดงคำอธิบายกระบวนการที่ 2.1 เพิ่มข้อมูลการพยากรณ์ปริมาณขยะมูลฝอยในประเทศไทย

Process Description	
System	ระบบบริหารจัดการพัฒนาเว็บแอปพลิเคชันเพื่อการวิเคราะห์ข้อมูลปริมาณขยะมูลฝอยในประเทศไทย
DFD Number	2.1
Process Name	ระบบการเพิ่มข้อมูล
Input Data Flow	ข้อมูลการพยากรณ์ปริมาณขยะมูลฝอยในประเทศไทย
Output Data Flow	ข้อมูลการพยากรณ์ปริมาณขยะมูลฝอยในประเทศไทย

Data Store Used	ข้อมูลการพยากรณ์ปริมาณขยะมูลฝอยในประเทศไทย
Description	เป็นกระบวนการสำหรับการเพิ่มข้อมูลในการพยากรณ์ขยะมูลฝอย โดยผู้ใช้ระบบในส่วนนี้ได้แก่ ผู้ดูแลระบบ

ตารางที่ 3.9 แสดงคำอธิบายกระบวนการที่ 2.2 แก้ไขข้อมูลการพยากรณ์ปริมาณขยะมูลฝอยในประเทศไทย

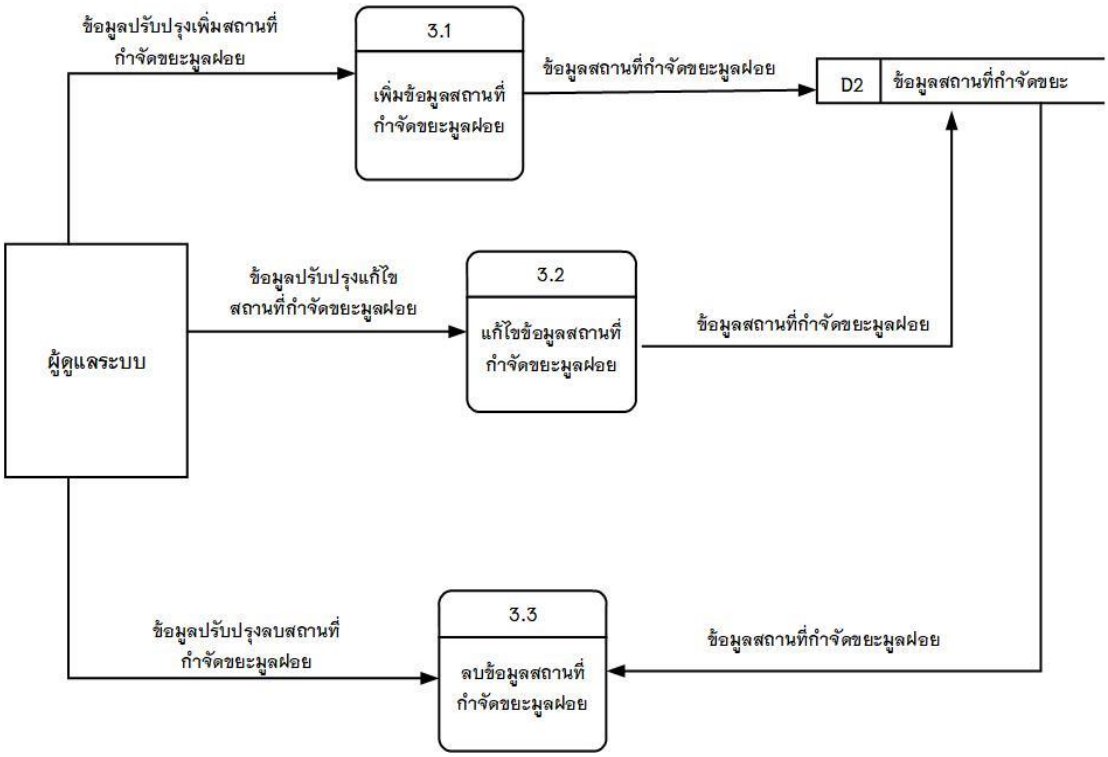
Process Description	
System	ระบบบริหารจัดการพัฒนาเว็บแอปพลิเคชันเพื่อการวิเคราะห์ข้อมูลปริมาณขยะมูลฝอยในประเทศไทย
DFD Number	2.2
Process Name	ระบบการเพิ่มข้อมูล
Input Data Flow	ข้อมูลการพยากรณ์ปริมาณขยะมูลฝอย
Output Data Flow	ข้อมูลการพยากรณ์ปริมาณขยะมูลฝอย
Data Store Used	ข้อมูลการพยากรณ์ปริมาณขยะมูลฝอย
Description	เป็นกระบวนการสำหรับการเพิ่มข้อมูลในการพยากรณ์ขยะมูลฝอย โดยผู้ใช้ระบบในส่วนนี้ได้แก่ ผู้ดูแลระบบ

ตารางที่ 3.10 แสดงคำอธิบายกระบวนการที่ 2.3 ลบข้อมูลการพยากรณ์ปริมาณมูลฝอยในประเทศไทย

Process Description	
System	ระบบบริหารจัดการพัฒนาเว็บแอปพลิเคชันเพื่อการวิเคราะห์ข้อมูลปริมาณขยะมูลฝอยในประเทศไทย
DFD Number	2.3
Process Name	ระบบการลบข้อมูล
Input Data Flow	ข้อมูลการพยากรณ์ปริมาณขยะมูลฝอย
Output Data Flow	ข้อมูลการพยากรณ์ปริมาณขยะมูลฝอย

Data Store Used	ข้อมูลการพยากรณ์ปริมาณขยะมูลฝอย
Description	เป็นกระบวนการสำหรับการลบข้อมูลในการพยากรณ์ขยะมูลฝอย โดยผู้ใช้ระบบในส่วนนี้ได้แก่ ผู้ดูแลระบบ

แผนภาพกระแสข้อมูล ระดับที่ 1 กระบวนการที่ 2 จัดการข้อมูล ประเด็นที่ 2 สถานที่กำจัดขยะในประเทศไทย



ภาพที่ 3.4 แผนภาพกระแสข้อมูล ระดับที่ 1 กระบวนการที่ 3 จัดการข้อมูลสถานที่กำจัดขยะมูลฝอยในประเทศไทย

ตารางที่ 3.11 แสดงคำอธิบายกระบวนการที่ 3.1 เพิ่มข้อมูลสถานที่กำจัดขยะมูลฝอยในประเทศไทย

Process Description	
System	ระบบบริหารจัดการพัฒนาเว็บแอปพลิเคชันเพื่อการวิเคราะห์ข้อมูลปริมาณขยะมูลฝอยในประเทศไทย

DFD Number	3.1
Process Name	การเพิ่มข้อมูลสถานที่กำจัดขยะมูลฝอยในประเทศไทย
Input Data Flow	ข้อมูลสถานที่กำจัดขยะมูลฝอยในประเทศไทย
Output Data Flow	ข้อมูลสถานที่กำจัดขยะมูลฝอยในประเทศไทย
Data Store Used	ข้อมูลสถานที่กำจัดขยะมูลฝอยในประเทศไทย
Description	เป็นกระบวนการสำหรับการเพิ่มข้อมูลสถานที่กำจัดขยะมูลฝอยในประเทศไทย โดยผู้ใช้ระบบในส่วนนี้ได้แก่ ผู้ดูแลระบบ

ตารางที่ 3.12 แสดงคำอธิบายกระบวนการที่ 3.2 แก้ไขสถานที่กำจัดขยะมูลฝอยในประเทศไทย

Process Description	
System	ระบบบริหารจัดการพัฒนาเว็บแอปพลิเคชันเพื่อการวิเคราะห์ข้อมูลปริมาณขยะมูลฝอยในประเทศไทย
DFD Number	3.2
Process Name	แก้ไขข้อมูลสถานที่กำจัดขยะมูลฝอยในประเทศไทย
Input Data Flow	ข้อมูลสถานที่กำจัดขยะมูลฝอยในประเทศไทย
Output Data Flow	ข้อมูลสถานที่กำจัดขยะมูลฝอยในประเทศไทย
Data Store Used	ข้อมูลสถานที่กำจัดขยะมูลฝอยในประเทศไทย
Description	เป็นกระบวนการสำหรับการเพิ่มข้อมูลข้อมูลสถานที่กำจัดขยะในประเทศไทย โดยผู้ใช้ระบบในส่วนนี้ได้แก่ ผู้ดูแลระบบ

ตารางที่ 3.13 แสดงคำอธิบายกระบวนการที่ 3.3 ลบข้อมูลสถานที่กำจัดขยะมูลฝอยในประเทศไทย

Process Description	
System	ระบบบริหารจัดการพัฒนาเว็บแอปพลิเคชันเพื่อการวิเคราะห์ข้อมูลปริมาณขยะมูลฝอยในประเทศไทย

DFD Number	3.3
Process Name	ลบข้อมูลสถานที่กำจัดขยะมูลฝอยในประเทศไทย
Input Data Flow	ข้อมูลสถานที่กำจัดขยะมูลฝอยในประเทศไทย
Output Data Flow	ข้อมูลสถานที่กำจัดขยะมูลฝอยในประเทศไทย
Data Store Used	ข้อมูลสถานที่กำจัดขยะมูลฝอยในประเทศไทย
Description	เป็นกระบวนการสำหรับการลบข้อมูลในสถานที่กำจัดขยะมูลฝอยในประเทศไทย โดยผู้ใช้ระบบในส่วนนี้ได้แก่ ผู้ดูแลระบบ

3.2 กระบวนการทำนอร์มัลไลเซชัน (Normalization)

การทำนอร์มัลไลเซชัน เป็นวิธีการในการกำหนดแอตทริบิวต์ให้กับแต่ละเอนทิตี เพื่อให้ได้โครงสร้างของตารางที่ดี สามารถควบคุมความซ้ำซ้อนของข้อมูลหลีกเลี่ยงความผิดปกติของข้อมูล โดยทั่วไปผลลัพธ์ของการนอร์มัลไลเซชัน จะได้ตารางที่มีโครงสร้างซับซ้อนน้อยลง แต่จำนวนของตารางจะมากขึ้น

การทำนอร์มัลไลเซชัน จะประกอบด้วยนอร์มัลฟอร์ม (Normal Form) แบบต่าง ๆ ที่มีเงื่อนไขของการทำให้อยู่ในรูปของนอร์มัลฟอร์มที่แตกต่างกันไป ขึ้นอยู่กับผู้ออกแบบฐานข้อมูลว่า ต้องการลดความซ้ำซ้อนในฐานข้อมูลให้อยู่ในระดับใด ซึ่งประกอบด้วยนอร์มัลฟอร์มแบบต่าง ๆ ดังต่อไปนี้

นอร์มัลฟอร์มที่ 1 (First Normal Form : 1NF)

นอร์มัลฟอร์มที่ 2 (Second Normal Form: 2NF)

นอร์มัลฟอร์มที่ 3 (Third Normal Form: 3NF)

บอยซ์คอดด์นอร์มัลฟอร์ม (Boyce–Codd Normal Form: BCNF)

นอร์มัลฟอร์มที่ 4 (Fourth Normal Form: 4NF)

นอร์มัลฟอร์มที่ 5 (Fifth Normal Form: 5NF)

กระบวนการทำนอร์มัลฟอร์มที่ 1 (First Normal Form : 1NF)

ค่าของแอททริบิวท์ต่าง ๆ ในแต่ละจะต้องมีค่าเพียงค่าเดียว และรีเลชันนั้นจะต้องไม่มีกลุ่มข้อมูลซ้ำ

ตารางที่ 3.14 ตารางข้อมูลก่อนที่จะทำการนอร์มัลฟอร์มที่ 1 ในประเด็นที่ 1

Province_id	Province_name	region_id	region	Year	Category_id	Category_name	garbage
1001	กำแพงเพชร	P01	ภาคเหนือ	2013	001	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	31,996.10
1002	เชียงราย	P01	ภาคเหนือ	2013	001	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	108,500.55
1003	เชียงใหม่	P01	ภาคเหนือ	2013	001	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	253,864.07
1004	ตาก	P01	ภาคเหนือ	2013	001	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	45,689.75

จากตารางที่ 3.14 อยู่ในรูปแบบ Unnormalized Form จะต้องทำให้อยู่ในรูปของ 1NF โดยสามารถทำให้อยู่ในรูปแบบ 1NF ได้

ตารางข้อมูลที่อยู่ในรูปของนอร์มัลฟอร์มที่ 1 ในประเด็นที่ 1 ปริมาณขยะมูลฝอยในประเทศไทย (First Normal Form : 1NF)

ตารางที่ 3.15 ตารางข้อมูลที่อยู่ในรูปของนอร์มัลฟอร์มที่ 1 ในประเด็นที่ 1

Province_id	Province_name	region_id	region	Year	Category_id	Category_name	garbage
1001	กำแพงเพชร	P01	ภาคเหนือ	2013	001	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	31,996.10
1001	กำแพงเพชร	P01	ภาคเหนือ	2013	002	ปริมาณขยะมูลฝอยที่ กำจัดไม่ถูกต้อง (ตัน)	127,181.69
1001	กำแพงเพชร	P01	ภาคเหนือ	2013	003	ปริมาณขยะมูลฝอยที่ นำไปใช้ประโยชน์ (ตัน)	63,282.05
1002	เชียงราย	P01	ภาคเหนือ	2013	001	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	108,500.55
1002	เชียงราย	P01	ภาคเหนือ	2013	002	ปริมาณขยะมูลฝอยที่ กำจัดไม่ถูกต้อง (ตัน)	240,928.64
1002	เชียงราย	P01	ภาคเหนือ	2013	003	ปริมาณขยะมูลฝอยที่ นำไปใช้ประโยชน์ (ตัน)	185,718.77
1003	เชียงใหม่	P01	ภาคเหนือ	2013	001	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	253,864.07
1003	เชียงใหม่	P01	ภาคเหนือ	2013	002	ปริมาณขยะมูลฝอยที่ กำจัดไม่ถูกต้อง (ตัน)	144,595.79
1003	เชียงใหม่	P01	ภาคเหนือ	2013	003	ปริมาณขยะมูลฝอยที่ นำไปใช้ประโยชน์ (ตัน)	228,944.71
1004	ตาก	P01	ภาคเหนือ	2013	001	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	45,689.75
1004	ตาก	P01	ภาคเหนือ	2013	002	ปริมาณขยะมูลฝอยที่ กำจัดไม่ถูกต้อง (ตัน)	142,227.62
1004	ตาก	P01	ภาคเหนือ	2013	003	ปริมาณขยะมูลฝอยที่ นำไปใช้ประโยชน์ (ตัน)	66,537.87

จากตารางที่ 3.15 จะปรับค่าให้เป็น 1NF โดยมีการทำ multi-value ให้กลายเป็น single value โดยเติมค่าของข้อมูลใน attribute ให้เต็ม โดยกำหนดตาราง Province (Province_id, Province_name, region, region_name, Year, Category_id, Category_name, garbage และในค่าข้อมูลของแต่ละ tuple จะไม่มีค่าที่ซ้ำกัน

กระบวนการทำนอร์มัลฟอร์มที่ 2 (Second Normal Form : 2NF)

รีเลชั่นใดๆ จะอยู่ในรูปแบบบรรทัดฐานระดับที่ 2 ก็ต่อเมื่อ รีเลชั่นนั้น ๆ อยู่ในรูปแบบบรรทัดฐานระดับที่ 1 และแอททริบิวต์ทุกตัวที่ไม่ได้เป็นคีย์หลัก จะต้องมีความสัมพันธ์ระหว่างค่าของแอททริบิวต์แบบ ฟังก์ชั่นกับคีย์หลัก (Fully Functional Dependency)

รีเลชั่นที่มีคุณสมบัติเป็นนอร์มัลฟอร์มระดับที่ 2 (Second Normal Form: 2NF)

1) จะต้องเป็นรีเลชั่นที่มีคุณสมบัติเป็น 1NF

2) แอททรีบิวท์ที่ไม่ใช่คีย์จะต้องขึ้นกับคีย์หลักอย่างแท้จริง

ตารางข้อมูลที่อยู่ในรูปของนอร์มัลฟอร์มที่ 2 ในประเด็นที่ 1 (Second Normal Form : 2NF)

ตารางที่ 3.16 ตารางข้อมูลที่อยู่ในรูปของนอร์มัลฟอร์มที่ 2 ในประเด็นที่ 1

Province_id	Province_name
1001	กำแพงเพชร
1001	กำแพงเพชร
1001	กำแพงเพชร
1002	เชียงราย
1002	เชียงราย
1002	เชียงราย
1003	เชียงใหม่
1003	เชียงใหม่
1003	เชียงใหม่
1004	ตาก
1004	ตาก
1004	ตาก

จากตารางที่ 3.16 จะมีการตัด Attribute ในตารางเดิม เมื่อนำไปใส่ตารางใหม่แล้วให้ตัดออกจากตารางเดิม ยกเว้น Attribute ที่เป็น PK ของตารางใหม่ ให้คงไว้ในตารางเดิม เพื่อใช้เป็น FK ระหว่างตารางเดิม โดยมีการกำหนดตารางใหม่ Province (Province_id, Province_name)

ตารางที่ 3.17 ตาราง region ที่อยู่ในรูปของนอร์มัลฟอร์มที่ 2 (Normal Form : 2NF)

region_id	ragion_name
P01	ภาคเหนือ
P01	ภาคเหนือ
P01	ภาคเหนือ
P01	ภาคเหนือ
P01	ภาคเหนือ
P01	ภาคเหนือ
P01	ภาคเหนือ
P01	ภาคเหนือ
P01	ภาคเหนือ
P01	ภาคเหนือ
P01	ภาคเหนือ
P01	ภาคเหนือ

ตารางที่ 3.18 ตาราง Category ที่อยู่ในรูปของนอร์มัลฟอร์มที่ 2 (Normal Form : 2NF)

Category_ID	Category
001	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)
002	ปริมาณขยะมูลฝอยที่ กำจัดไม่ถูกต้อง (ตัน)
003	ปริมาณขยะมูลฝอยที่ นำไปใช้ประโยชน์ (ตัน)
001	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)
002	ปริมาณขยะมูลฝอยที่ กำจัดไม่ถูกต้อง (ตัน)
003	ปริมาณขยะมูลฝอยที่ นำไปใช้ประโยชน์ (ตัน)
001	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)
002	ปริมาณขยะมูลฝอยที่ กำจัดไม่ถูกต้อง (ตัน)
003	ปริมาณขยะมูลฝอยที่ นำไปใช้ประโยชน์ (ตัน)
001	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)
002	ปริมาณขยะมูลฝอยที่ กำจัดไม่ถูกต้อง (ตัน)
003	ปริมาณขยะมูลฝอยที่ นำไปใช้ประโยชน์ (ตัน)

ตารางที่ 3.19 ตาราง Garbage_dumping ที่สมบูรณ์ในประเด็นที่ 1

Garbage_Dumping_id	Province_id	region_id	Category_id	Year	garbage
amg0001	1001	P01	001	2013	31,996.10
amg0002	1001	P01	002	2013	127,181.69
amg0003	1001	P01	003	2013	63,282.05
amg0004	1002	P01	001	2013	108,500.55
amg0005	1002	P01	002	2013	240,928.64
amg0006	1002	P01	003	2013	185,718.77
amg0007	1003	P01	001	2013	253,864.07
amg0008	1003	P01	002	2013	144,595.79
amg0009	1003	P01	003	2013	228,944.71
amg0010	1004	P01	001	2013	45,689.75
amg0011	1004	P01	002	2013	142,227.62
amg0012	1004	P01	003	2013	66,537.87

กระบวนการทำนอร์มัลฟอร์มที่ 1 ในประเด็นที่ 2 สถานที่กำจัดขยะ (First Normal Form : 1NF)

ค่าของแอททริบิวท์ต่าง ๆ ในแต่ละทัพเพิลจะต้องมีค่าเพียงค่าเดียว และรีเลชันนั้นจะต้องไม่มีกลุ่มข้อมูลซ้ำ

ตารางที่ 3.20 ตารางข้อมูลก่อนที่จะทำการนอร์มัลฟอร์มที่ 1 ในประเด็นที่ 2 สถานที่กำจัดขยะในประเทศไทย (First Normal Form : 1NF)

Province_id	Province_name	District_id	District_name	Sub-district_id	Sub-district_name	area size	Category_ID	Category_name	volume	year
1001	กำแพงเพชร	62170	-	S01	-	-	001	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	-	2013
							002			
1002	เชียงราย	57230	เทิง	S02	ปล้อง	10.5	001	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	383.7	2013
							002			
1003	เชียงใหม่	50130	แม่ออน	S03	ออนเหนือ	5	001	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	845.51	2013
							002			
1004	ตาก	63120	บ้านตาก	S04	เกาะตะเภา	10	001	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	15.00	2013
							002			

จากตารางที่ 3.20 อยู่ในรูปแบบ Unnormalized Form จะต้องทำให้อยู่ในรูปของ 1NF โดยสามารถทำให้อยู่ในรูปแบบ 1NF ได้

ตารางที่ 3.21 ตารางที่อยู่ในรูปของนอร์มัลฟอร์มที่ 1 ในประเด็นที่ 2

Province_id	Province_name	District_id	District_name	Sub-district_id	Sub-district_name	area size	Category_id	Category_name	volume	year
1001	กำแพงเพชร	-	-	-	-	-	001	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	-	2013
1001	กำแพงเพชร	62170	ลานกระบือ	S01	ลานกระบือ	13	002	ปริมาณขยะมูลฝอยที่ กำจัดไม่ถูกต้อง (ตัน)	93.14	2013
1002	เชียงราย	57230	เทิง	S02	ปล้อง	10.5	001	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	383.7	2013
1002	เชียงราย	57150	เชียงแสน	S03	เวียง	50	002	ปริมาณขยะมูลฝอยที่ กำจัดไม่ถูกต้อง (ตัน)	209.26	2013
1003	เชียงใหม่	50130	แม่ออน	S04	ออนเหนือ	5	001	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	845.51	2013
1003	เชียงใหม่	50150	แม่แตง	S05	แม่หอพระ	2	002	ปริมาณขยะมูลฝอยที่ กำจัดไม่ถูกต้อง (ตัน)	272.27	2013
1004	ตาก	63120	บ้านตาก	S06	เกาะตะเภา	10	001	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	15.00	2013
1004	ตาก	63160	พบพระ	S07	ช่องแคบ	5	002	ปริมาณขยะมูลฝอยที่ กำจัดไม่ถูกต้อง (ตัน)	331.3	2013

จากตารางที่ 3.21 จะปรับค่าให้เป็น 1NF โดยมีการทำ multi-value ให้กลายเป็น single value โดยเติมค่าของข้อมูลใน attribute ให้เต็ม โดยกำหนดตาราง Province (Province_id, Province_name, District_id, District_name, Sub-district_id, Sub-district_name, area size, Category_id, Category_name, Volume, Year และในค่าข้อมูลของแต่ละ tuple จะไม่มีค่าที่ซ้ำ

จากการพิจารณารีเลชันสถานที่กำจัดขยะ จะพบว่ามีฟังก์ชันการขึ้นต่อกัน ดังนี้ จะพบว่าในรีเลชันสถานที่กำจัดขยะจะมี Province_id, District_id, Sub-district_id, Category_id, ประกอบกันเป็นคีย์หลัก และจากการพิจารณาพบว่า จะไม่ตรงตามนิยามของรูปแบบนอมัลระดับที่ 2 เพราะแอตทริบิวต์ aera size, Volum, Year จะมีค่าเปลี่ยนแปลงตามแอตทริบิวส์ Province_id, District_id, Sub-district_id, Category_id ดังนั้นจึงต้องทำการแยกรีเลชันออกเป็น 5 รีเลชัน ดังนี้

ตารางที่ 3.22 ตาราง Province ที่อยู่ในรูปของนอร์มัลฟอร์มที่ 2 (Normal Form : 2NF)

Province_id	Province_name
1001	กำแพงเพชร
1001	กำแพงเพชร
1002	เชียงราย
1002	เชียงราย
1003	เชียงใหม่
1003	เชียงใหม่
1004	ตาก
1004	ตาก

ตารางที่ 3.23 ตาราง District ที่อยู่ในรูปของนอร์มัลฟอร์มที่ 2 (Normal Form : 2NF)

District_id	District_name
-	-
62170	ลานกระบือ
57230	เทิง
57150	เชียงแสน
50130	แม่ออน
50150	แม่แตง
63120	บ้านตาก
63160	พบพระ

ตารางที่ 3.24 ตาราง Sub-district ที่อยู่ในรูปของนอร์มัลฟอร์มที่ 2 (Normal Form : 2NF)

Sub-district_id	Sub-district
-	-
S01	ลานกระบือ
S02	ปล้อง
S03	เวียง
S04	ออนเหนือ
S05	แม่หอพระ
S06	เกาะตะเภา
S07	ช่องแคบ

ตารางที่ 3.25 ตารางข้อมูล Category ที่อยู่ในรูปของนอร์มัลฟอร์มที่ 2 (Normal Form: 2NF)

Category_id	category_name
001	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)
002	ปริมาณขยะมูลฝอยที่ กำจัดไม่ถูกต้อง (ตัน)
001	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)
002	ปริมาณขยะมูลฝอยที่ กำจัดไม่ถูกต้อง (ตัน)
001	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)
002	ปริมาณขยะมูลฝอยที่ กำจัดไม่ถูกต้อง (ตัน)
001	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)
002	ปริมาณขยะมูลฝอยที่ กำจัดไม่ถูกต้อง (ตัน)

ตารางที่ 3.26 ตารางข้อมูล Garbage_location ที่สมบูรณ์ในประเด็นที่ 2

Garbage_location_id	Province_id	District_id	Sub-district_id	Category_id	area size	volume	Year
log001	1001	-	-	001	-	-	2013
log002	1001	62170	S01	002	13	93.14	2013
log003	1002	57230	S02	001	10.5	383.7	2013
log004	1002	57150	S03	002	50	209.26	2013
log005	1003	50130	S04	001	5	845.51	2013
log006	1003	50150	S05	002	2	272.27	2013
log007	1004	63120	S06	001	10	15.00	2013
log008	1004	63160	S07	002	5	331.3	2013

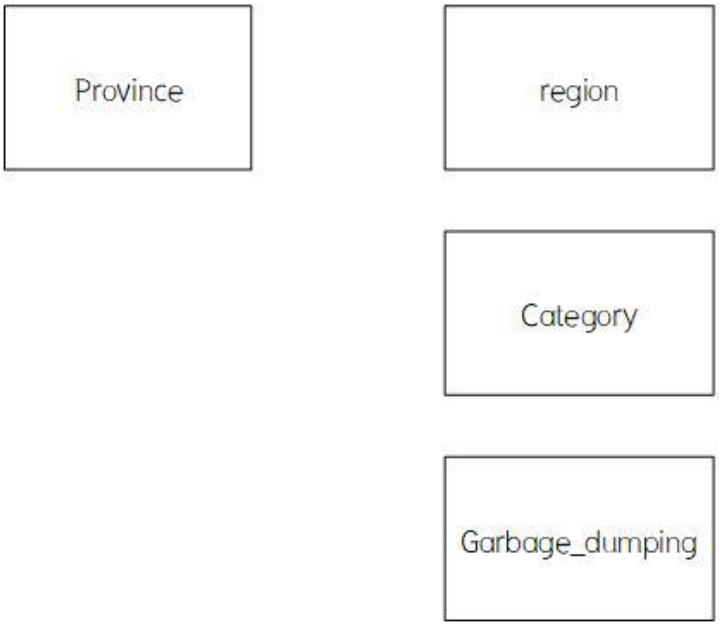
จากการที่ผู้วิเคราะห์ได้ออกแบบเป็นฐานข้อมูลไว้ มาปรับปรุงโดยใช้วิธีการทำนอร์มัลไลเซชัน เพื่อปรับเปลี่ยนโครงสร้างของรีเลชันต่าง ๆ ทำให้สามารถแก้ปัญหาลดความซ้ำซ้อนของข้อมูลในลักษณะต่าง ๆ และทำให้รีเลชันมีโครงสร้างที่เหมาะสม ในการใช้งานต่อไป โดยผู้วิเคราะห์ได้ทำนอร์มัลไลเซชันได้ 2 และระดับขั้นตอนโดยประเด็นที่ 1 จะได้ตารางทั้งหมด 4 ตารางประกอบไปด้วยตารางที่ 1 ชื่อรีเลชัน Province ตารางที่ 2 ชื่อรีเลชัน reion ตารางที่ 3 ชื่อรีเลชัน Category ตารางที่ 4 ชื่อรีเลชัน Garbage_Dumping และในระดับขั้นตอนโดยประเด็นที่ 2 จะได้ตารางทั้งหมด 5 ตารางประกอบไปด้วยตารางที่ 1 ชื่อรีเลชัน Province ตารางที่ 2 ชื่อรีเลชัน District ตารางที่ 3 ชื่อรีเลชัน Sub-district_id ตารางที่ 4 ชื่อรีเลชัน Category และตารางที่ 5 ชื่อรีเลชัน Garbage_location จากข้อมูลของปริมาณขยะมูลฝอยในประเทศไทยและสถานที่กำจัดขยะในประเทศไทยที่ได้นำมาผ่านกระบวนการนอร์มัลไลเซชันทั้ง 2 ขั้นตอนก็เพียงพอต่อการใช้งานและตอบสนองความต้องการของผู้ใช้งานแล้ว

3.3 ความสัมพันธ์ของฐานข้อมูล (ER-Diagram)

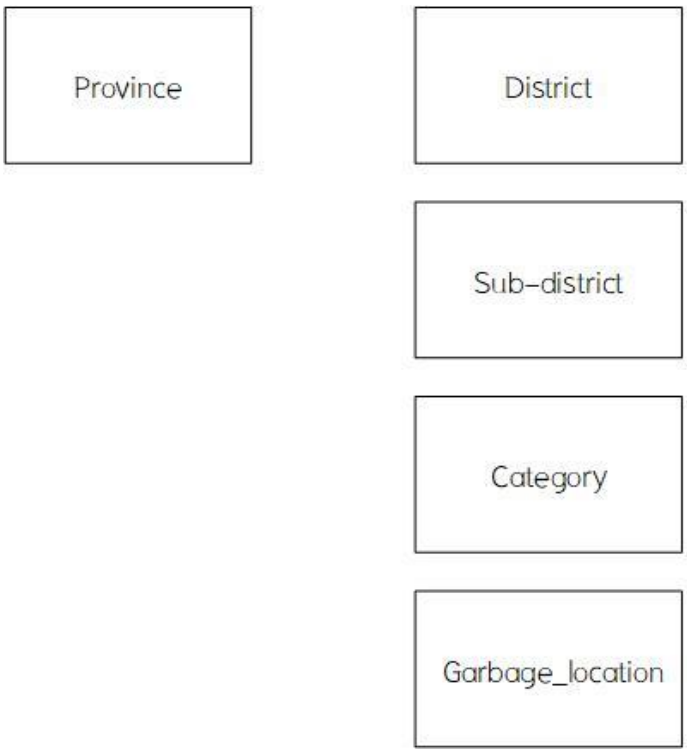
การออกแบบฐานข้อมูลโดยใช้แบบจำลองความสัมพันธ์ระหว่างข้อมูล (ER Model) มีขั้นตอนที่เกี่ยวข้องหลายขั้นตอน และต้องใช้ความรอบคอบในการออกแบบเพื่อให้ได้ ER-Diagram ที่ถูกต้องเหมาะสมกับระบบงาน ในขั้นตอนการออกแบบฐานข้อมูลนี้ใช้ตัวอย่างฐานข้อมูลลงทะเบียนสามารถนำมาสร้างแบบจำลองข้อมูลด้วย ER Model โดยมีขั้นตอนในการพัฒนาดังนี้

3.3.1 การกำหนดเอนทิตี้

การกำหนดเอนทิตี้เป็นการกำหนดสิ่งที่สนใจและต้องการจัดเก็บข้อมูลที่มีในระบบงาน โดยดูจากลักษณะหน้าที่ของระบบงานว่ามีรายละเอียดการทำงานอย่างไร โดยจะต้องพิจารณาด้วยว่าเป็นเอนทิตี้ประเภทใด เช่น เอนทิตี้แบบปกติ (Regular Entity) หรือเอนทิตี้แบบอ่อนแอ (Weak Entity) เช่น ฐานข้อมูลของปริมาณขยะในประเด็นที่ 1 ประกอบด้วย เอนทิตี้ Province เอนทิตี้ region เอนทิตี้ category เอนทิตี้ Garbage_dumping และฐานข้อมูลของสถานที่กำจัดขยะในประเด็นที่ 2 ประกอบด้วย เอนทิตี้ Province เอนทิตี้ District เอนทิตี้ Sub-district เอนทิตี้ Category เอนทิตี้ Garbage_location



ภาพที่ 3.5 การกำหนดเอนทิตี้ในประเด็นที่ 1 ปริมาณขยะมูลฝอยในประเทศไทย

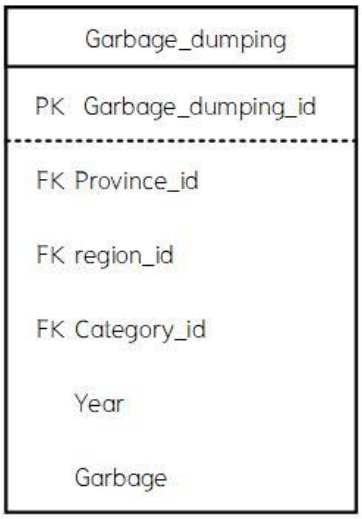


ภาพที่ 3.6 การกำหนดเอนทิตี้ในประเด็นที่ 2 สถานที่กำจัดขยะมูลฝอยในประเทศไทย

3.3.2 การกำหนดแอตทริบิวต์ของเอนทิตี

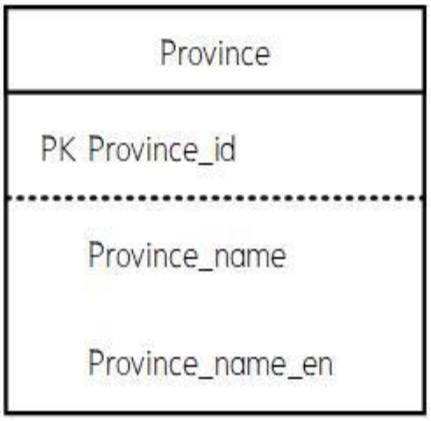
การกำหนดแอตทริบิวต์ของเอนทิตีเป็นการกำหนดคุณลักษณะต่าง ๆ ของแต่ละเอนทิตีรวมทั้งพิจารณาแอตทริบิวต์ที่จะทำหน้าที่เป็นคีย์หลักของเอนทิตีด้วย ดังตัวอย่างประเด็นที่ 1 ปริมาณขยะ ประกอบด้วย

1) เอนทิตี้ Garbage_dumping ประกอบด้วย Garbage_dumping_id, Province_id, region_id, Category_id, Year, Garbage



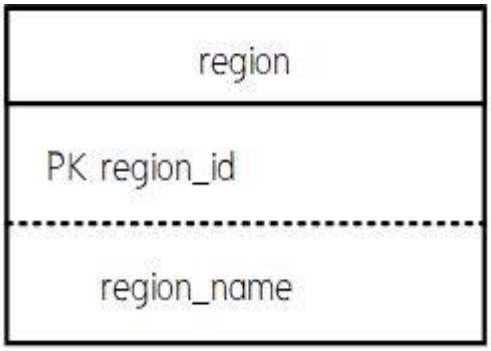
ภาพที่ 3.7 การกำหนดแอตทริบิวต์ Garbage_dumping ในประเด็นที่ 1 ปริมาณขยะมูลฝอยในประเทศไทย

2) เอนทิตี้ Province ประกอบด้วย Province_id, Province_name, Province_name_en



ภาพที่ 3.8 การกำหนดแอตทริบิวต์ Province ในประเด็นที่ 1 ปริมาณขยะมูลฝอย

3) เอนทิตี้ region ประกอบด้วย region_id, region_name



ภาพที่ 3.9 การกำหนดแอตทริบิวต์ region ในประเด็นที่ 1 ปริมาณขยะมูลฝอยในประเทศไทย

4) เอนทิตี้ Category ประกอบด้วย Category_id, Category_name

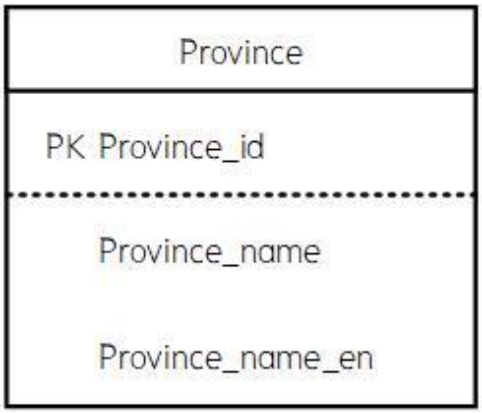
ภาพที่ 3.10 การกำหนดแอตทริบิวต์ Category ในประเด็นที่ 1 ปริมาณขยะมูลฝอย ในประเทศไทย

ดังตัวอย่างประเด็นที่ 2 สถานที่กำจัดขยะ ประกอบด้วย

1) เอนทิตี้ Garbage_location ประกอบด้วย Garbage_location_id, Province_id, District_id, Sub–distrct_id, Category_id, Area size, Volume, year

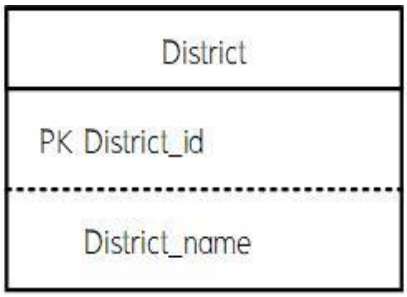
ภาพที่ 3.11 การกำหนดแอตทริบิวต์ Garbage_location ในประเด็นที่ 2 สถานที่กำจัดขยะมูลฝอยในประเทศไทย

2) เอนทิตี้ Province ประกอบด้วย Province_id, Province_name, Province_name_en



ภาพที่ 3.12 การกำหนดแอตทริบิวต์ Province ในประเด็นที่ 2 สถานที่กำจัดขยะมูลฝอยในประเทศไทย

3) เอนทิตี District ประกอบด้วย District_id, District_name



ภาพที่ 3.13 การกำหนดแอตทริบิวต์ District ในประเด็นที่ 2
สถานที่กำจัดขยะมูลฝอยในประเทศไทย

4) เอนทิตี Sub-District ประกอบด้วย Sub-District_id, Sub-District _Name

ภาพที่ 3.14 การกำหนดแอตทริบิวต์ Sub-District ในประเด็นที่ 2
สถานที่กำจัดขยะมูลฝอยในประเทศไทย

5) เอนทิตี Category ประกอบด้วย Category_id, Category_name

ภาพที่ 3.15 การกำหนดแอตทริบิวต์ Category ในประเด็นที่ 2
สถานที่กำจัดขยะมูลฝอยในประเทศไทย

3.3.3 การกำหนดความสัมพันธ์ระหว่างเอนทิตี้ (Relationships)

การกำหนดแอตทริบิวต์ของเอนทิตี้เป็นการกำหนดคุณลักษณะต่าง ๆ ของแต่ละเอนทิตี้รวมทั้งพิจารณาแอตทริบิวต์ที่จะทำหน้าที่เป็นคีย์หลักของเอนทิตี้ด้วย โดยแบ่งหลักได้อยู่ 3 ประเภท

1) ความสัมพันธ์แบบหนึ่งต่อหนึ่ง (One-to-One Relationship)

เป็นการแสดงความสัมพันธ์ของข้อมูลเอนทิตี้หนึ่งว่ามีความสัมพันธ์กับข้อมูล อย่างมากหนึ่งข้อมูลกับอีกเอนทิตี้หนึ่งในลักษณะที่เป็นหนึ่งต่อหนึ่ง

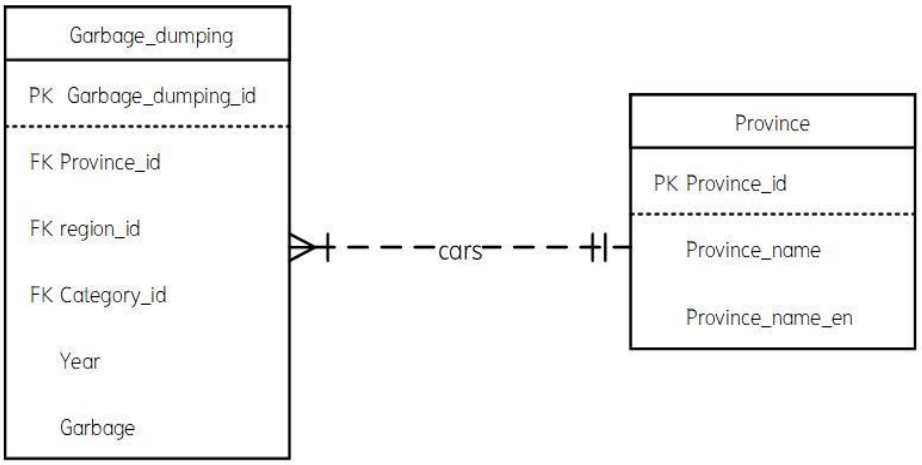
2) ความสัมพันธ์แบบหนึ่งต่อกลุ่ม (One-to-Many Relationship)

เป็นการแสดงความสัมพันธ์ระหว่างข้อมูลของเอนทิตี้หนึ่งว่ามีความสัมพันธ์กับข้อมูลของอีกเอนทิตี้หนึ่ง

3) ความสัมพันธ์แบบกลุ่มต่อกลุ่ม (Many-to-Many Relationship)

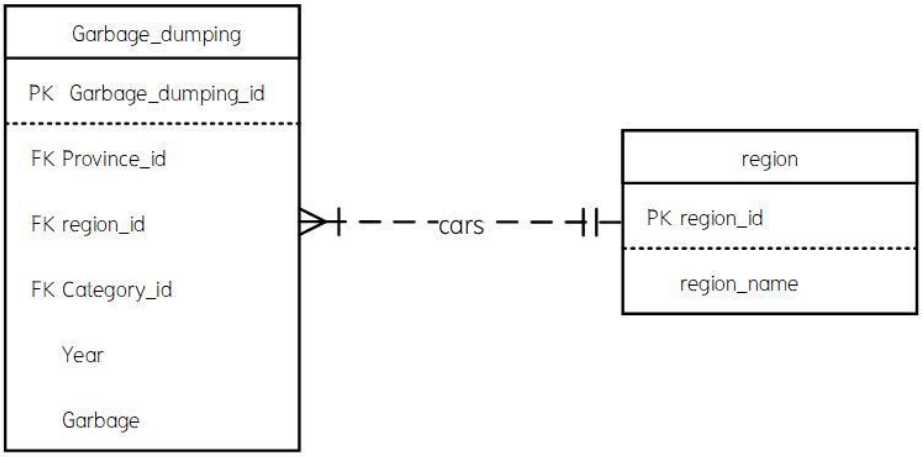
เป็นการแสดงความสัมพันธ์ของข้อมูลระหว่างเอนทิตี้ในลักษณะกลุ่มต่อกลุ่ม เช่น ความสัมพันธ์ของ แผนกกับพนักงาน

3.1) ตาราง Garbage Dumping ในประเด็นที่ 1 ปริมาณขยะมูลฝอยในประเทศไทย ที่มีความสัมพันธ์ระหว่างเอนทิตี้กับตาราง Province แบบหนึ่งต่อกลุ่ม (1 : N) หมายความว่า ปริมาณขยะ 1 ปริมาณสามารถมีได้หลายจังหวัด



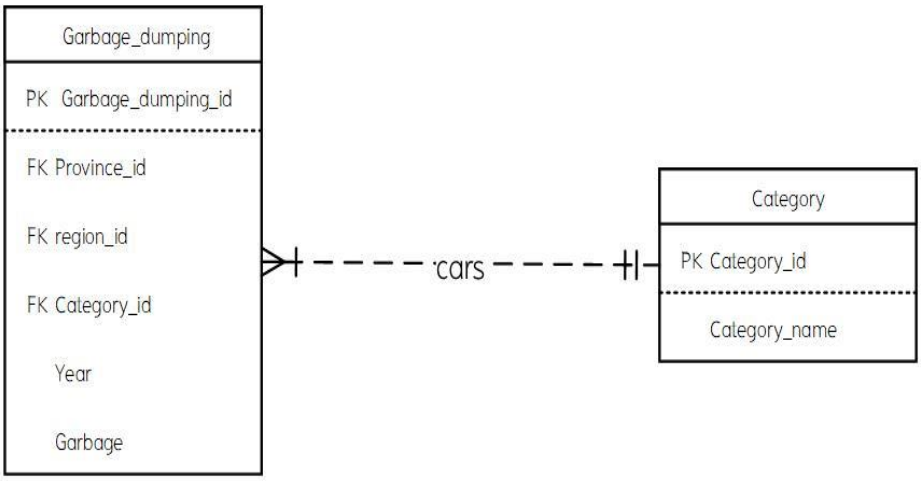
ภาพที่ 3.16 แสดงความสัมพันธ์ระหว่างตาราง Garbage_dumping กับตาราง Province แบบหนึ่งต่อกลุ่ม (1 : N)

3.2) ตาราง Garbage_dumping ในประเด็นที่ 1 ปริมาณขยะมูลฝอยในประเทศไทยที่มีความสัมพันธ์ระหว่างเอนทิตี้กับตาราง region แบบหนึ่งต่อกลุ่ม (1: N) หมายความว่า ปริมาณขยะ 1 ปริมาณสามารถมีได้หลายๆภาค



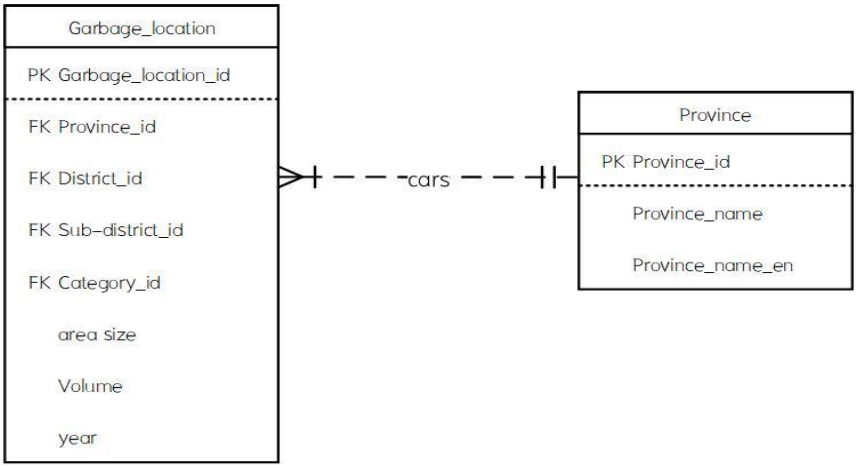
ภาพที่ 3.17 แสดงความสัมพันธ์ระหว่างตาราง Garbage_dumping กับตาราง region แบบหนึ่งต่อกลุ่ม (1 : N)

3.3) ตาราง Garbage_dumping ในประเด็นที่ 1 ปริมาณขยะมูลฝอยในประเทศไทยที่มีความสัมพันธ์ระหว่างเอนทิตี้กับตาราง Category แบบหนึ่งต่อกลุ่ม (1 : N) หมายความว่า ปริมาณขยะ 1 ปริมาณสามารถมีประเภทในการกำจัดได้ขยะได้หลายประเภท



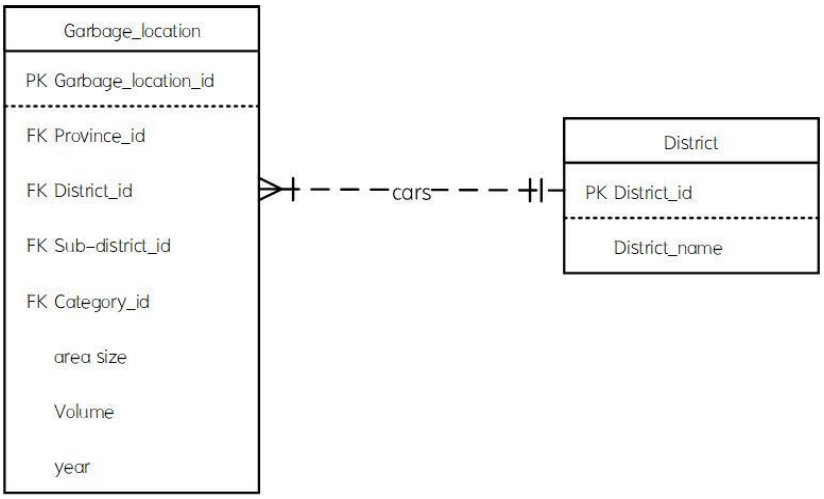
ภาพที่ 3.18 แสดงความสัมพันธ์ระหว่างตาราง Garbage_dumping กับตาราง Category แบบหนึ่งต่อกลุ่ม (1 : N)

3.4) ตาราง Garbage_location ในประเด็นที่ 2 มีความสัมพันธ์ระหว่างเอนทิตี้กับตาราง Province แบบหนึ่งต่อกลุ่ม (1 : N) หมายความว่า สถานที่กำจัด 1 สถานที่สามารถมีจังหวัดได้หลายจังหวัด



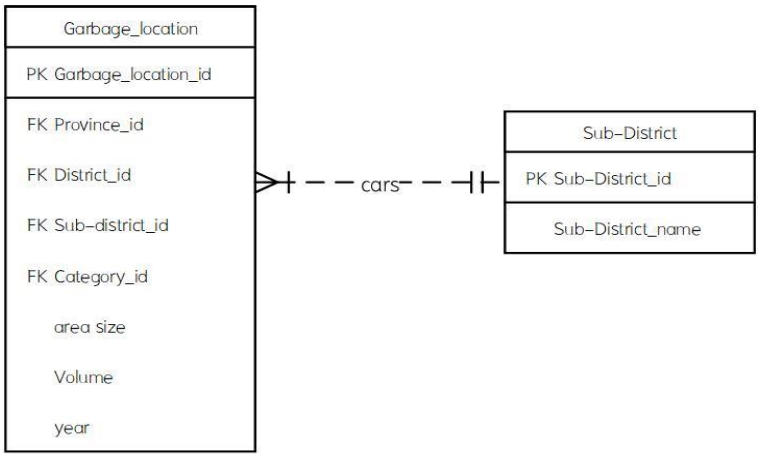
ภาพที่ 3.19 แสดงความสัมพันธ์ระหว่างตาราง Garbage_location กับตาราง Province แบบหนึ่งต่อกลุ่ม (1 : N)

3.5) ตาราง Garbage_location ในประเด็นที่ 2 มีความสัมพันธ์ระหว่างเอนทิตี้กับตาราง District แบบหนึ่งต่อกลุ่ม (1 : N) หมายความว่า สถานที่กำจัด 1 สถานที่กำจัดสามารถมีได้หลายอำเภอ



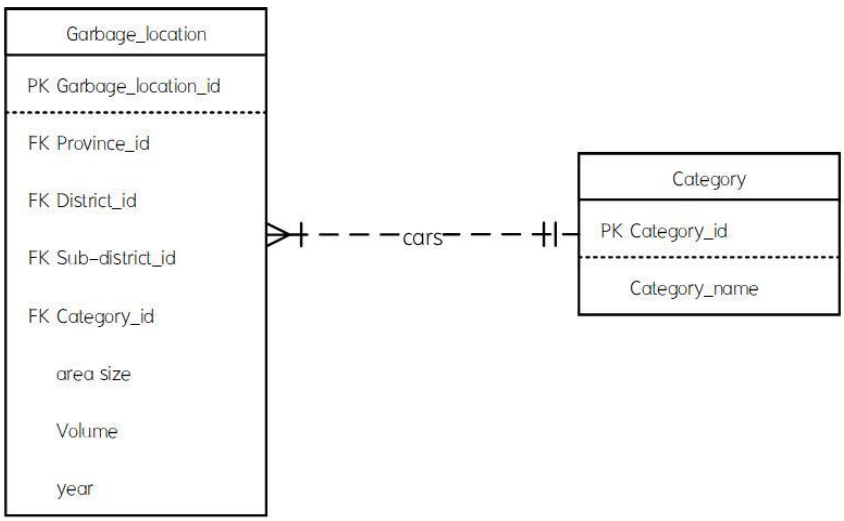
ภาพที่ 3.20 แสดงความสัมพันธ์ระหว่างตาราง Garbage_location กับตาราง District แบบหนึ่งต่อกลุ่ม (1 : N)

3.6) ตาราง Garbage_location ในประเด็นที่ 2 มีความสัมพันธ์ระหว่างเอนทิตี้กับตาราง Sub-District แบบหนึ่งต่อกลุ่ม (1 : N) หมายความว่า สถานที่กำจัด 1 สถานที่กำจัดสามารถมีได้หลายตำบล



ภาพที่ 3.21 แสดงความสัมพันธ์ระหว่างตาราง Garbage_location กับตาราง Sub-District แบบหนึ่งต่อกลุ่ม (1 : N)

3.7) ตาราง Garbage location ในประเด็นที่ 2 มีความสัมพันธ์ระหว่างเอนทิตี้กับตาราง Category แบบหนึ่งต่อกลุ่ม (1 : N) หมายความว่า สถานที่กำจัด 1 สถานที่กำจัดสามารถมีประเภทได้หลายประเภท

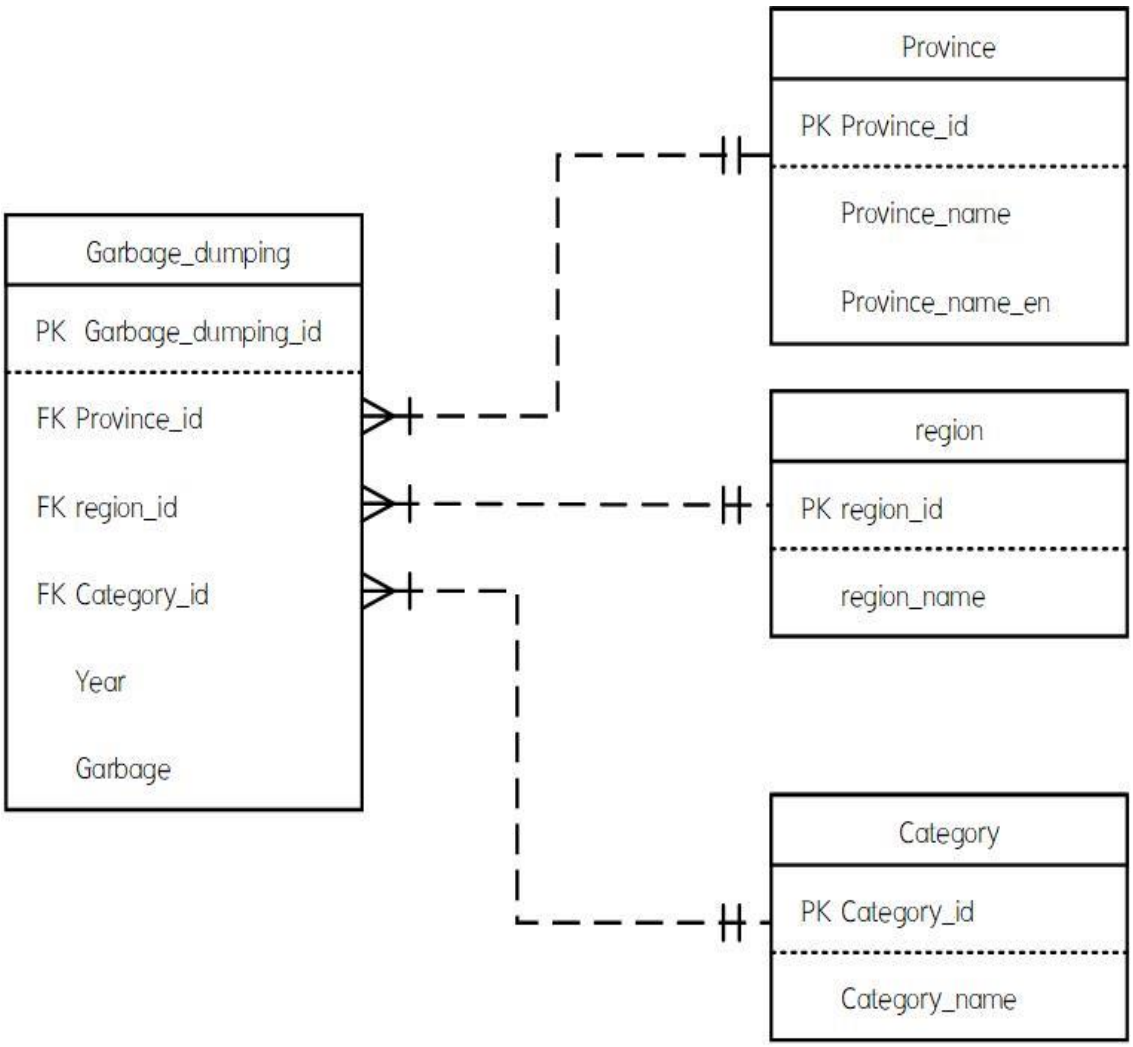


ภาพที่ 3.22 แสดงความสัมพันธ์ระหว่างตาราง Garbage location กับตาราง Category แบบหนึ่งต่อกลุ่ม (1 : N)

3.3.4 การเขียนเป็นแบบจำลองแผนภาพ Entity Relationship Diagram (ERD)

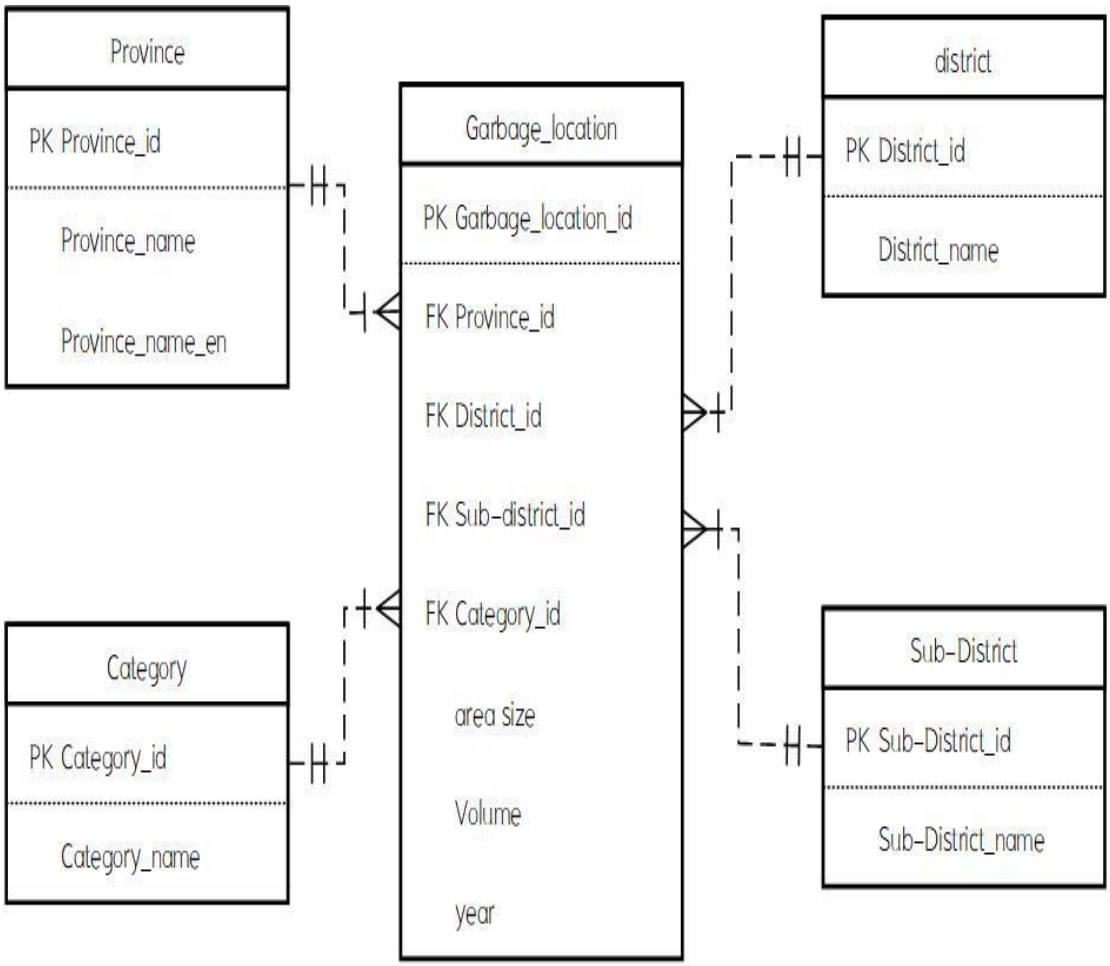
การนำรายละเอียดการออกแบบทั้งหมดในแต่ละขั้นตอน นำมาวาดแบบจำลองแผนภาพ E-R Diagram กำหนดให้ความสัมพันธ์ระหว่างเอนทิตี้ ดังต่อไปนี้

1) การเขียนเป็นแบบจำลองแผนภาพ ในประเด็นที่ 1 Entity Relationship Diagram



ภาพที่ 3.23 แสดงความสัมพันธ์ของฐานข้อมูลแบบ Crow's Foot Model ในประเด็นที่ 1 ปริมาณขยะมูลฝอยในประเทศไทย

2) การเขียนเป็นแบบจำลองแผนภาพ ในประเด็นที่ 2 Entity Relationship Diagram



ภาพที่ 3.24 แสดงความสัมพันธ์ของฐานข้อมูลแบบ Crow's Foot Model ในประเด็นที่ 2 สถานที่กำจัดขยะมูลฝอยในประเทศไทย

3.4 พจนานุกรมข้อมูล (Data Dictionary)

จากการออกแบบฐานข้อมูลซึ่งมีการจัดการระบบฐานข้อมูลให้กับระบบที่ประกอบไปด้วยตารางข้อมูลต่าง ๆ ดังตารางที่ 3.27

ตารางที่ 3.27 แสดงชื่อตารางทั้งหมดของระบบฐานข้อมูล ในประเด็นที่ 1 ปริมาณขยะมูลฝอย

ลำดับ	ชื่อตาราง	ประเภท	รายละเอียด
1	Garbage_dumping	Transaction	เก็บข้อมูลปริมาณขยะ
2	Garbage_location	Transaction	เก็บข้อมูลสถานที่กำจัดขยะ
3	Province	Reference	เก็บข้อมูลจังหวัด
4	region	Reference	เก็บข้อมูลภาค
5	Category	Reference	เก็บข้อมูลประเภทการกำจัดขยะ
6	District	Reference	เก็บข้อมูลอำเภอ
7	Sub-district	Reference	เก็บข้อมูลตำบล

คำอธิบาย ประเภทของตาราง ได้แก่

Transaction หมายถึง ตารางที่มีการเปลี่ยนแปลงของข้อมูล

Reference หมายถึง ตารางที่มีการอ้างอิงถึงข้อมูล

จากตารางที่ 3.28 สามารถแสดงรายละเอียดของแต่ละตารางได้ดังนี้

ตารางที่ 3.28 แสดงรายละเอียดของตาราง Garbage_dumping

ชื่อตาราง : Garbage_dumping
ประเภทตาราง : Transaction
คำอธิบาย : เก็บข้อมูลปริมาณขยะ
คีย์หลัก : Garbage_dumping_id
คีย์รอง : Province_id, region_id, Category_id

เขตข้อมูล	ชนิดและขนาด	ความหมาย	ตัวอย่าง
Garbage_dumping_id	int (11)	รหัสปริมาณขยะ	amg0001
Province_id	Int (11)	รหัสจังหวัด	1001
region_id	VARCHAR (5)	รหัสภาค	P01

Category_id	Int (11)	รหัสประเภทการกำจัดขยะ	001
Year	Year (4)	ปี	2013
garbage	Double	ปริมาณขยะ	31,996.10

ตารางที่ 3.29 แสดงรายละเอียดของตาราง Garbage_location

ชื่อตาราง : Garbage_location
ประเภทตาราง : Transaction
คำอธิบาย : เก็บข้อมูลสถานที่กำจัดขยะ
คีย์หลัก : Garbage_location_id
คีย์รอง : Province_id, District_id, Sub-district_id, Category_id

เขตข้อมูล	ชนิดและขนาด	ความหมาย	ตัวอย่าง
Garbage_location_id	int (11)	รหัสปริมาณขยะ	amg0001
Province_id	int (11)	รหัสจังหวัด	1001
District_id	int (11)	รหัสอำเภอ	001
Sub-district_id	int (11)	รหัสตำบล	2013
Category_id	int (11)	รหัสประเภทการกำจัดขยะ	31,996.10
Area size	Double	ขนาดพื้นที่	13
Volume	Double	ปริมาณขยะ	93.14
Year	Year (4)	ปี	2013

ตารางที่ 3.30 แสดงรายละเอียดของตาราง Province

ชื่อตาราง : Province
ประเภทตาราง : Reference
คำอธิบาย : เก็บข้อมูลจังหวัด
คีย์หลัก : Province_id
คีย์รอง : -

เขตข้อมูล	ชนิดและขนาด	ความหมาย	ตัวอย่าง
Province_id	int (10)	รหัสจังหวัด	1001

Province_name	VARCHAR (100)	ชื่อจังหวัด	กำแพงเพชร
Province_name_en	VARCHAR (100)	ชื่อจังหวัด ภาษาอังกฤษ	Chiang Mai

ตารางที่ 3.31 แสดงรายละเอียดของตาราง District

ชื่อตาราง : District ประเภทตาราง : Reference คำอธิบาย : เก็บข้อมูลประเภทการกำจัดขยะ คีย์หลัก : District_id คีย์รอง : –			
เขตข้อมูล	**ชนิดและขนาด**	**ความหมาย**	**ตัวอย่าง**
District_id	int (11)	รหัสอำเภอ	62170
District_Name	VARCHAR (100)	ชื่ออำเภอ	ลานกระบือ

ตารางที่ 3.32 แสดงรายละเอียดของตาราง Sub-district

ชื่อตาราง : Sub-district ประเภทตาราง : Reference คำอธิบาย : เก็บข้อมูลประเภทการกำจัดขยะ คีย์หลัก : Sub-district_id คีย์รอง : –			
เขตข้อมูล	**ชนิดและขนาด**	**ความหมาย**	**ตัวอย่าง**
Sub-district_id	int (11)	รหัสตำบล	S01
Sub-district_Name	VARCHAR (100)	ชื่อตำบล	ลานกระบือ

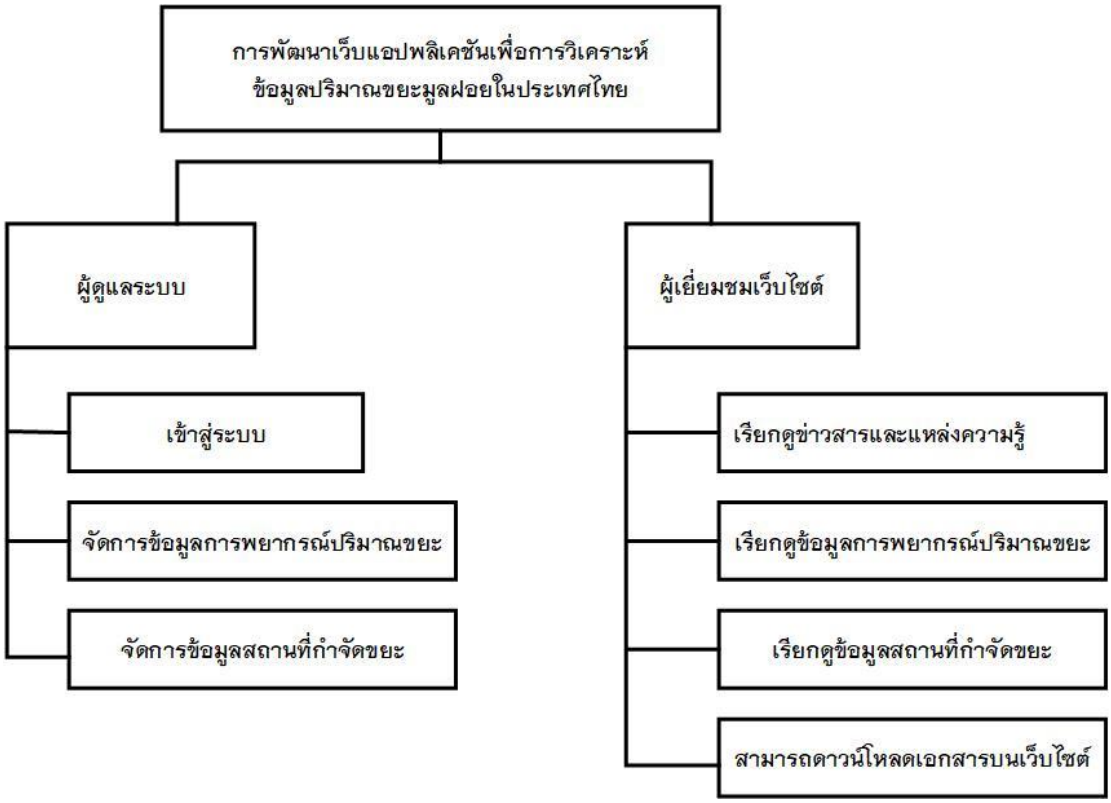
ตารางที่ 3.33 แสดงรายละเอียดของตาราง Category

ชื่อตาราง : Category
ประเภทตาราง : Reference
คำอธิบาย : เก็บข้อมูลประเภทการกำจัดขยะ
คีย์หลัก : Category_id
คีย์รอง : –

เขตข้อมูล	ชนิดและขนาด	ความหมาย	ตัวอย่าง
Category_id	int (5)	รหัสประเภทการกำจัดขยะ	001
Category_Name	VARCHAR (100)	ชื่อประเภทการกำจัดขยะ	ปริมาณขยะมูลฝอยที่กำจัดถูกต้อง (ตัน)

3.5 โครงสร้างระบบ (System Structure)

ลักษณะภายในของโครงสร้างระบบแสดงถึงโครงสร้างการใช้งานในระบบฐานข้อมูลปริมาณขยะมูลฝอยในประเทศไทย ส่วนที่เป็นส่วนสำคัญโดยตรง ได้แก่ ผู้ดูแลระบบและผู้เยี่ยมชมเว็บไซต์ ดังภาพที่ 3.25



ภาพที่ 3.25 แผนผังโครงสร้างการใช้งานในระบบฐานข้อมูลปริมาณขยะและสถานที่กำจัดขยะ

จากภาพที่ 3.26 เพื่อความสะดวกในการออกแบบฐานข้อมูลผู้วิเคราะห์จึงจัดแบ่งระบบออกเป็นส่วนย่อย ๆ มีทั้งหมด 2 ส่วนด้วยกัน และให้แต่ละส่วนมีหน้าที่รับผิดชอบการทำงานใน แต่ละด้าน โดยไม่คาบเกี่ยวกันแต่สัมพันธ์กัน ส่วนแรก คือส่วนของผู้ดูแลระบบ จะเป็นส่วนที่ สามารถจัดการข้อมูลทั้งหมดในระบบได้และสามารถอัปโหลดข้อมูลบนเว็บไซต์ ส่วนที่สอง คือส่วนผู้เยี่ยมชมเว็บไซต์ เป็นส่วนที่บุคคลภายนอกสามารถเข้ามาเยี่ยมชมข้อมูลบนเว็บไซต์ได้และสามารถดาวน์โหลดข้อมูลบนเว็บไซต์ได้

3.6 กระบวนการ CRISP–DM (Cross Industry Standard Process for Data Mining)

CRISP–DM เป็นกระบวนการหลักในการจัดทำเหมืองข้อมูลเพื่อการวิเคราะห์และใช้ประโยชน์ในทางธุรกิจกระบวนวิเคราะห์ข้อมูลด้วย CRISP–DM หรือ (Cross Industry Standard Process for Data Mining) ประกอบด้วย 6 ขั้นตอน แต่ละขั้นตอนในรูปจะเป็นขั้นตอนที่ต่อเนื่องกันนั่นคือขั้นตอนถัดไปจะรอผลลัพธ์จากขั้นตอนก่อนหน้าซึ่งแสดงด้วยลูกศรที่เชื่อมระหว่างกล่องสี่เหลี่ยมแต่ละกล่อง ตัวอย่างเช่นเมื่อได้ผลลัพธ์จากขั้นตอนการเตรียมข้อมูล (Data Preparation) แล้วจะนำไปสร้างโมเดลจำแนกประเภทข้อมูลในชั้น Modeling และหลังจากนั้นอาจจะย้อนกลับมาเปลี่ยนแปลงข้อมูลให้ถูกต้องมากขึ้นเพื่อหวังว่าโมเดลที่ให้ความถูกต้องมากขึ้นก็ได้ เป็นต้น

ภาพที่ 3.26 แสดงกระบวนการ CRISP–DM

ที่มา : dataminingtrend.com

ในกระบวนการนี้ประกอบด้วย 6 ขั้นตอน

1) Business Understanding เป็นขั้นตอนในกระบวนการ เป็นการแปลงปัญหาที่ได้ให้อยู่ในรูปโจทย์ของการวิเคราะห์ข้อมูล Data Mining พร้อมทั้งวางแผนในการดำเนินการ

2) Data Understanding เริ่มจากการเก็บรวบรวมข้อมูล หลังจากนั้นก็เป็นการตรวจสอบข้อมูลที่ได้ทำการรวบรวมมา เพื่อดูความถูกต้อง และพิจารณาว่าใช้ข้อมูลทั้งหมดหรือจำเป็นต้องเลือกข้อมูลบางส่วนมาใช้ในการวิเคราะห์

3) Data Preparation เป็นขั้นตอนที่ทำการแปลงข้อมูลที่ได้ทำการเก็บรวบรวมมาให้กลายเป็นข้อมูลที่สามารถนำไปวิเคราะห์ในขั้นถัดไปได้ โดยการแปลงข้อมูลนี้อาจจะต้องมี

การทำข้อมูลให้ถูกต้อง เช่น แปลงข้อมูลให้อยู่ในช่วงเดียวกัน หรือการเติมข้อมูลข้อมูลที่ขาดหายไป เป็นขั้นตอนที่ใช้เวลามากที่สุดของกระบวนการ CRISP-DM

4) Modeling เป็นขั้นตอนการวิเคราะห์ข้อมูลด้วยเทคนิคทาง Data Mining ที่ได้แนะนำไปแล้ว เช่นการจำแนกประเภทข้อมูล หรือการแบ่งกลุ่มข้อมูล

5) Evaluation ได้ผลการวิเคราะห์ข้อมูลด้วยเทคนิคทาง Data Mining แล้วแต่ก่อนที่จะนำผลลัพธ์ที่ได้ไปใช้งานต้องมีการวัดประสิทธิภาพของผลลัพธ์ที่ได้ว่าตรงกับวัตถุประสงค์ ที่ได้ตั้งไว้หรือไม่ มีความน่าเชื่อถือเพียงใด

6) Deployment มีการนำความรู้ที่ได้จากการได้ผลลัพธ์ด้วยเทคนิค Data Mining ไปใช้ประโยชน์ต่อในองค์กร

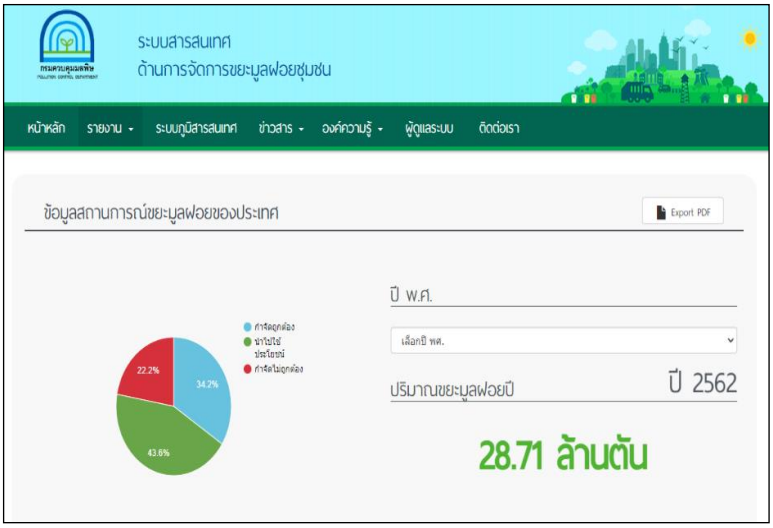
3.1.1 ความเข้าใจในธุรกิจ (Business Understanding)

ผู้วิเคราะห์ทำความเข้าใจกระบวนการทางข้อมูล และรับฟังปัญหารวมถึงการมองหาปัญหาจากเรื่องต่าง ๆ เพื่อที่นำผลการวิเคราะห์ข้อมูลไปใช้ให้เกิดประโยชน์ โดยเรียงลำดับความสำคัญ และกำหนดวัตถุประสงค์ที่จะนำไปสู่รูปแบบการวิเคราะห์ข้อมูล ให้อยู่ในรูปโจทย์ของการวิเคราะห์ฐานข้อมูล และวางแผนในการดำเนินการโดยการวิเคราะห์ข้อมูลรูปแบบความสัมพันธ์ที่ซ่อนอยู่ในข้อมูลของข้อมูลปริมาณขยะมูลฝอย ในประเทศไทย เพื่อที่จะนำผลการวิเคราะห์ข้อมูลไปใช้ประโยชน์ตามความต้องการ และตามวัตถุประสงค์ที่จะนำไปสู่รูปแบบการวิเคราะห์ข้อมูล

3.1.2 การทำความเข้าใจข้อมูล (Data Understanding)

ผู้วิเคราะห์ข้อมูลทำการจัดเก็บรวบรวมข้อมูล และพิจารณาตรวจสอบข้อมูลที่ได้รับดูความถูกต้อง มีความน่าเชื่อถือของข้อมูล โดยพิจารณาว่าจะใช้ข้อมูลทั้งหมดหรือจำเป็นต้องเลือกข้อมูลบางส่วนมาใช้ในการวิเคราะห์ ให้สอดคล้องกับวัตถุประสงค์ที่กำหนดไว้ ผู้วิเคราะห์ได้สำรวจภาพรวมก่อนนำข้อมูลมาใช้ในการวิเคราะห์ข้อมูลจากเว็บไซต์ https://thaimsw.pcd.go.th ซึ่งเป็นระบบสารสนเทศของกรมควบคุมมลพิษ ผู้วิเคราะห์ข้อมูล จึงได้ดำเนินการตามขั้นตอนดังนี้

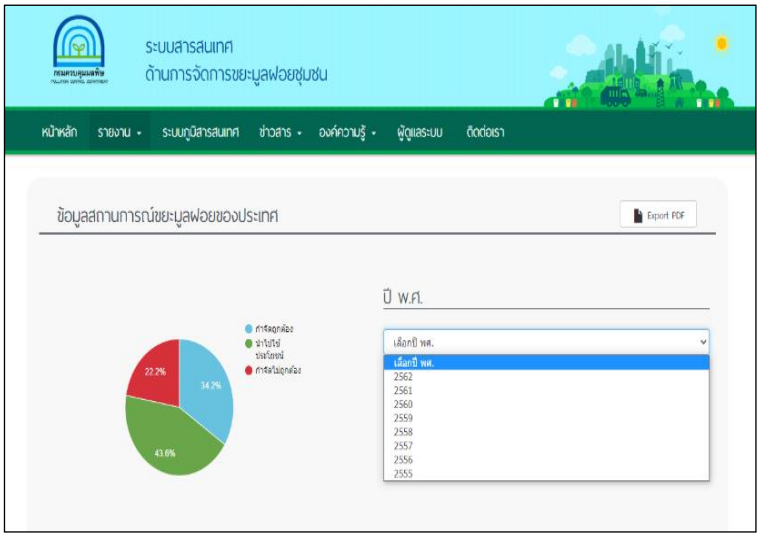
1) ผู้วิเคราะห์รวบรวมข้อมูลจาก จากเว็บไซต์ https://thaimsw.pcd.go.th ผู้วิเคราะห์ทำการเลือกที่รายงาน ข้อมูลสถานการ์ณขยะมูลฝอยของประเทศ



ภาพที่ 3.27 แสดงขั้นตอนการหาข้อมูลจากเว็บไซต์

ที่มา : https://thaimsw.pcd.go.th

2) ผู้วิเคราะห์ทำการจัดเก็บ และรวบรวมข้อมูลทั้งหมด ของข้อมูลปริมาณขยะมูลฝอยในช่วงปี พ.ศ. 2556 – 2562



ภาพที่ 3.28 แสดงขั้นตอนการดาวน์โหลดข้อมูล

ที่มา : https://thaimsw.pcd.go.th

3) ตรวจสอบความถูกต้อง และความน่าเชื่อถือของข้อมูล ซึ่ง 1 ไฟล์ จะประกอบด้วย 5 แอตทริบิวท์ คือ จังหวัด, ปริมาณขยะมูลฝอยที่เกิดขึ้น, ปริมาณขยะมูลฝอยที่นำไปใช้ประโยชน์, ปริมาณขยะมูลฝอยที่กำจัดถูกต้อง, ปริมาณขยะมูลฝอยที่กำจัดไม่ถูกต้อง

จังหวัด	ปริมาณขยะมูลฝอยที่ เกิดขึ้น (ตัน)	ปริมาณขยะมูลฝอยที่ นำไปใช้ไประโยชน์ (ตัน)	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	ปริมาณขยะมูลฝอยที่ กำจัดไม่ถูกต้อง (ตัน)
กระบี่	200,593.05	47,628.85	50,534.25	102,429.95
กรุงเทพมหานคร	7,446,000.00	478,150.00	3,244,850.00	3,723,000.00
กาญจนบุรี	250,495.85	7,921.96	8,068.69	234,505.20
กาฬสินธุ์	348,216.45	126,883.49	30,638.10	190,694.86
กำแพงเพชร	256,739.86	55,247.64	89,498.00	111,994.22
ขอนแก่น	682,514.91	182,613.53	49,869.95	450,031.43
จันทบุรี	220,843.93	21,520.51	78,777.95	120,545.46
ฉะเชิงเทรา	314,406.09	35,349.63	6,424.00	272,632.46
ชลบุรี	907,884.66	79,607.49	397,120.00	431,157.17
ชัยนาท	129,691.80	54,987.25	27,353.10	47,351.45
ชัยภูมิ	396,644.01	108,705.68	3,650.00	284,288.32
ชุมพร	133,221.35	3,051.40	10,658.00	119,511.95
ตรัง	242,670.25	88,476.00	7,730.70	146,463.55
ตราด	97,715.20	18,583.80	30,039.50	49,091.90
ตาก	233,519.00	48,384.73	10,220.00	174,914.27
นครนายก	93,570.36	20,221.15	34,025.30	39,323.92

ภาพที่ 3.29 ไฟล์ข้อมูลการบันทึกปริมาณขยะมูลฝอยรายปี

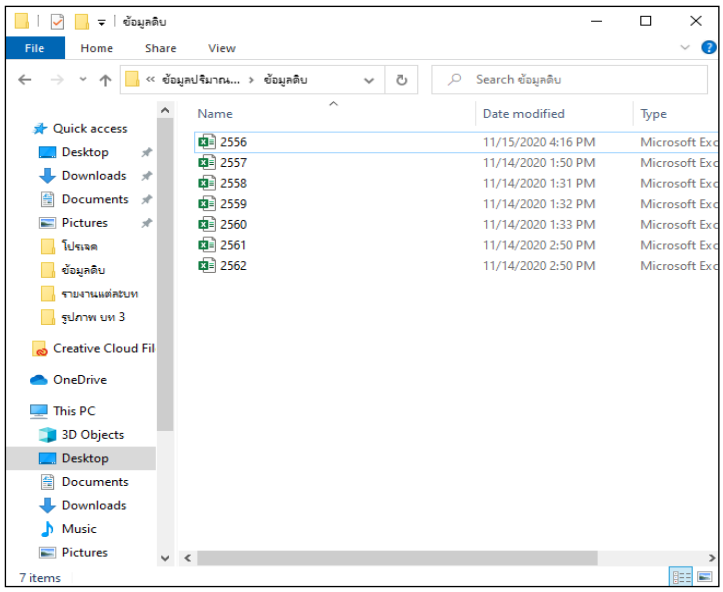
ที่มา : https://thaimsw.pcd.go.th

ขั้นตอนการแปลงข้อมูลที่ได้รวบรวมมา ให้อยู่ในรูปแบบที่พร้อมสำหรับนำไปวิเคราะห์ในขั้นตอนต่อไปได้ โดยการทำให้เป็นข้อมูลที่ถูกต้อง (Data cleaning) มักใช้เวลาค่อนข้างมากระบบการรับข้อมูลป้อนเข้าสู่ระบบในปัจจุบันจะลดการคีย์ข้อมูลจากคนให้น้อยที่สุด แต่จะใช้วิธีการสแกน การติ๊กเลือก เพื่อลดความผิดพลาดให้น้อยที่สุด เพราะขั้นตอนใช้เวลามากการลดข้อผิดพลาดของข้อมูลได้มากเท่าใด ก็จะยิ่งมีประสิทธิภาพมากขึ้นเท่านั้น

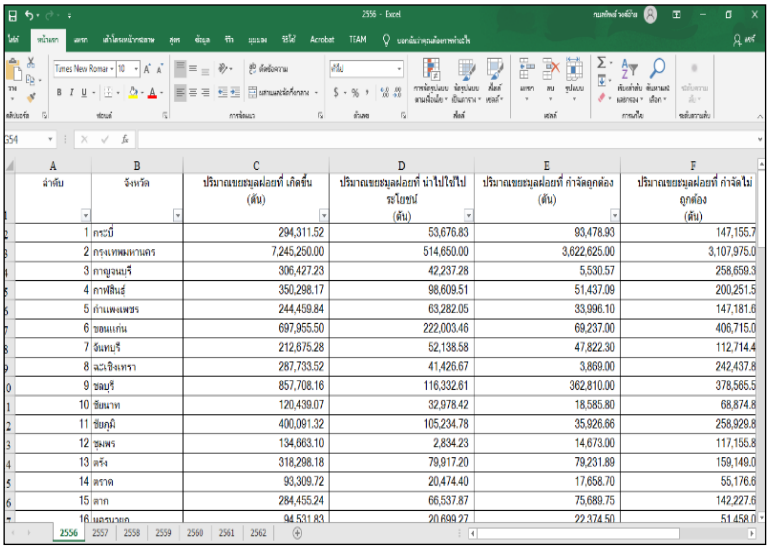
3.1.3 การเตรียมข้อมูล (Data Preparation)

ผู้วิเคราะห์ข้อมูลทำการแปลงข้อมูลที่ได้ทำการเก็บรวบรวม และทำการจัดการข้อมูลที่ได้ให้อยู่ในรูปแบบที่สามารถนำไปวิเคราะห์ในขั้นถัดไปได้ เพื่อลดความผิดพลาดการทำข้อมูลให้ถูกต้อง โดยใช้กระบวนการ data cleaning เพื่อให้ได้ข้อมูลอยู่ในรูปแบบและมาตรฐานเดียวกัน และกำจัดข้อมูลเสียออก ผู้วิเคราะห์ข้อมูลจึงได้ดำเนินการตามขั้นตอนดังนี้

1) การปรับมาตรฐาน (Standardizing) ผู้วิเคราะห์ข้อมูลทำการรวบแปลงไฟล์ข้อมูลให้เป็นไฟล์เดียวกันอยู่ในรูปแบบเดียวกันพร้อมทั้งสร้างแอททริบิวต์เพื่อกำหนดคีย์หลักของแต่ละไฟล์ขึ้นมาเพื่อเป็นตัวที่ใช้ระบุความสัมพันธ์ของข้อมูลที่จะแสดงปริมาณขยะมูลฝอยในประเทศไทย



ภาพที่ 3.30 ไฟล์ข้อมูลปริมาณขยะมูลฝอยทั้งหมดที่ได้จากเว็บไซต์

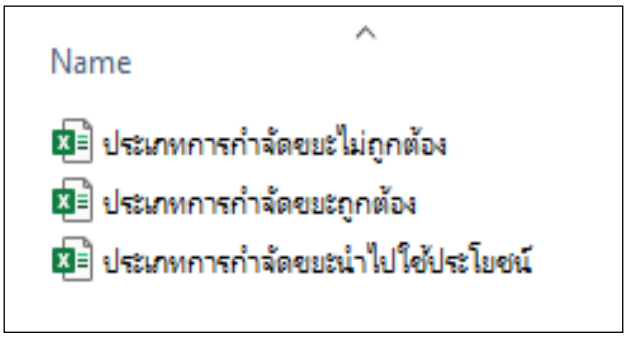


	A	B	C	D	E	F
	ลำดับ	จังหวัด	ปริมาณขยะมูลฝอยที่ เกิดขึ้น (ตัน)	ปริมาณขยะมูลฝอยที่ นำไปใช้ไป รอโยชน์ (ตัน)	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	ปริมาณขยะมูลฝอยที่ กำจัดไม่ ถูกต้อง (ตัน)
2	1	กระบี่	294,311.52	53,676.83	93,478.93	147,155.7
3	2	กรุงเทพมหานคร	7,245,250.00	514,650.00	3,622,625.00	3,107,975.0
4	3	กาญจนบุรี	306,427.23	42,237.28	5,530.57	258,659.3
5	4	กาฬสินธุ์	350,298.17	98,609.51	51,437.09	200,251.5
6	5	กำแพงเพชร	244,459.84	63,282.05	33,996.10	147,181.6
7	6	ขอนแก่น	697,955.50	222,003.46	69,237.00	406,715.0
8	7	จันทบุรี	212,675.28	52,138.58	47,822.30	112,714.4
9	8	ฉะเชิงเทรา	287,733.52	41,426.67	3,869.00	242,437.8
10	9	ชลบุรี	857,708.16	116,332.61	362,810.00	378,565.5
11	10	ชัยนาท	120,439.07	32,978.42	18,585.80	68,874.8
12	11	ชัยภูมิ	400,091.32	105,234.78	35,926.66	258,929.8
13	12	ชุมพร	134,663.10	2,834.23	14,673.00	117,155.8
14	13	เชียง	318,298.18	79,917.20	79,231.89	159,149.0
15	14	ตราด	93,309.72	20,474.40	17,658.70	55,176.6
16	15	ตาก	284,455.24	66,537.87	75,689.75	142,227.6
17	16	นครนายก	94,531.83	20,699.27	22,374.50	51,458.0

ภาพที่ 3.31 ข้อมูลที่ปรับมาตรฐานรวมให้เป็นรายปีในแต่ละจังหวัดในประเทศไทย

2) การจัดหมวดหมู่ (Transfrom) ผู้วิเคราะห์นำ dataset ที่ผ่านการรวบรวมรายปีมาจัดหมวดหมู่เพื่อความถูกต้อง ผู้วิเคราะห์ดำเนินการจัดกลุ่มข้อมูลแบ่งตามภูมิภาค โดยการสร้างแอตทริบิวท์เพิ่มขึ้น 1 ตัว คือ แอตทริบิวท์ region (ภูมิภาค) เพื่อแบ่งหมวดหมู่และแอตทริบิวท์ที่กำหนดให้ Province_name เป็นคีย์หลักเพื่อเป็นตัวที่ใช้ระบุความสัมพันธ์ของข้อมูลที่จะแสดงข้อมูลภาคต่าง ๆ ในแต่ละปี เตรียมความพร้อมการนำข้อมูลไปแสดงบนโปรแกรม Minitab และโปรแกรม Tableau Public โดยดำเนินตามขั้นตอนดังนี้

2.1) การจัดกลุ่มตามจังหวัดโดยแบ่งเป็นประเภทของการกำจัดขยะ



ภาพที่ 3.32 ไฟล์ข้อมูลที่ต้องการจัดกลุ่มให้รวมอยู่ในแต่ละประเภทของการกำจัดขยะมูลฝอย

2.2) นำข้อมูลที่ผ่านการรวบรวมเป็นรายจังหวัดที่แบ่งเป็นตามประเภท

	C1-T	C2-T	C3-T	C4-T	C5
	Province_Name	Ragion	Year	Category	garbage
1	กำแพงเพชร	ภาคเหนือ	2013	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	43996
2	กำแพงเพชร	ภาคเหนือ	2014	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	62123
3	กำแพงเพชร	ภาคเหนือ	2015	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	77498
4	กำแพงเพชร	ภาคเหนือ	2016	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	55159
5	กำแพงเพชร	ภาคเหนือ	2017	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	53151
6	กำแพงเพชร	ภาคเหนือ	2018	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	52239
7	กำแพงเพชร	ภาคเหนือ	2019	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	57772

Sheet1

ภาพที่ 3.33 แสดงการจัดกลุ่มรายประเภทการกำจัดขยะมูลฝอย

จากรูปภาพที่ 3.26 ประกอบไปด้วย 5 แอตทริบิวท์ คือ

1) Province_name เก็บข้อมูลชื่อจังหวัด

2) region เก็บข้อมูลภาค

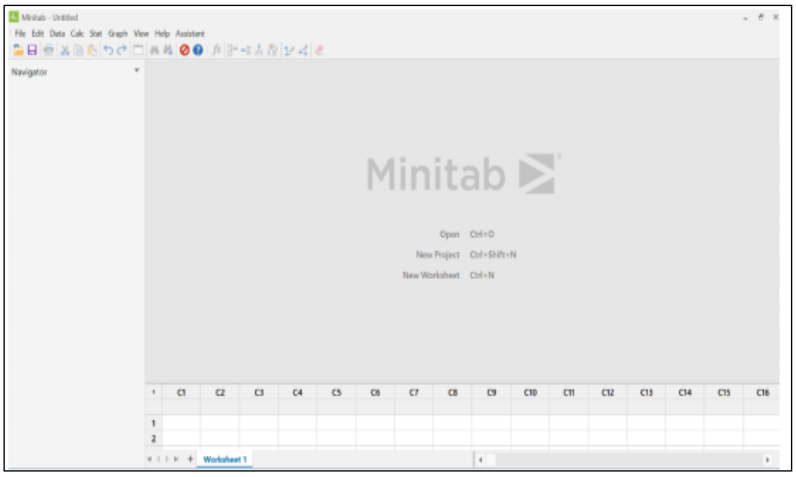
3) Year เก็บข้อมูลรายปี

4) Category เก็บข้อมูลประเภทการกำจัดขยะ

5) garbage เก็บข้อมูลปริมาณขยะ

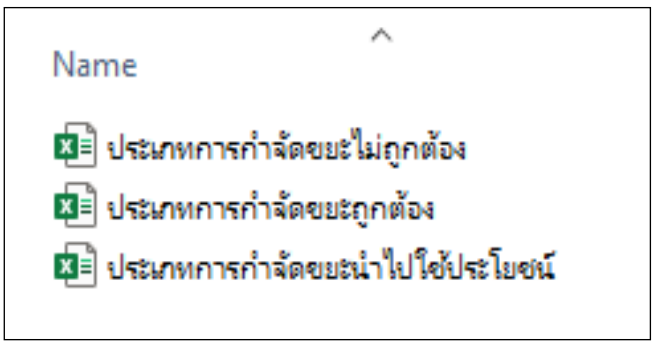
3) การอัปโหลดข้อมูล (Loading) โดยใช้ Minitab ดำเนินกระบวนการการโหลดข้อมูลเข้าระบบฐานข้อมูลของ Minitab เพื่อเป็นการพยากรณ์ของข้อมูล โดยมีขั้นตอนดำเนินการดังนี้

3.1) โหลดข้อมูลที่ได้เตรียมไว้แล้วลงโปรแกรม Minitab



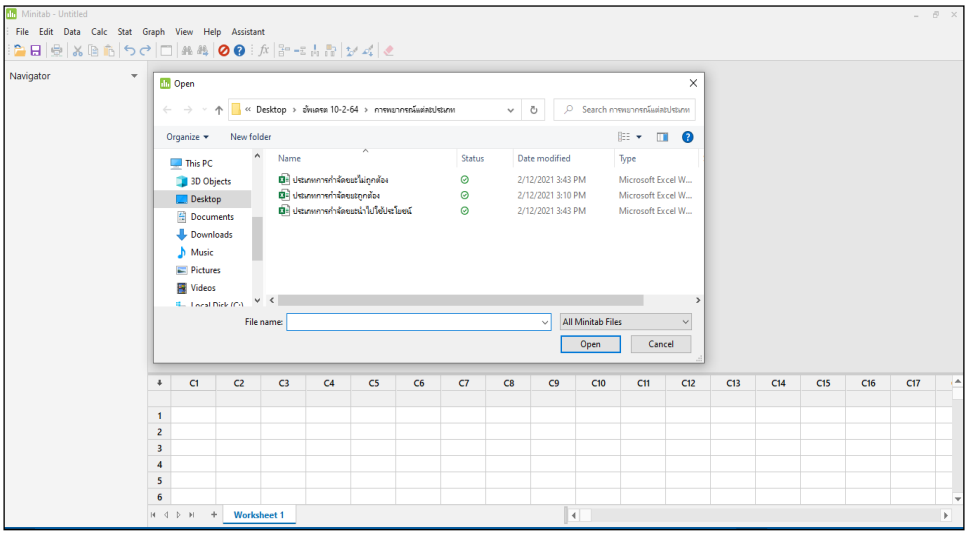
ภาพที่ 3.34 หน้าต่างพื้นที่การทำงานของโปรแกรม Minitab

3.2) จัดเตรียมข้อมูลที่นำมาใช้ในการพยากรณ์



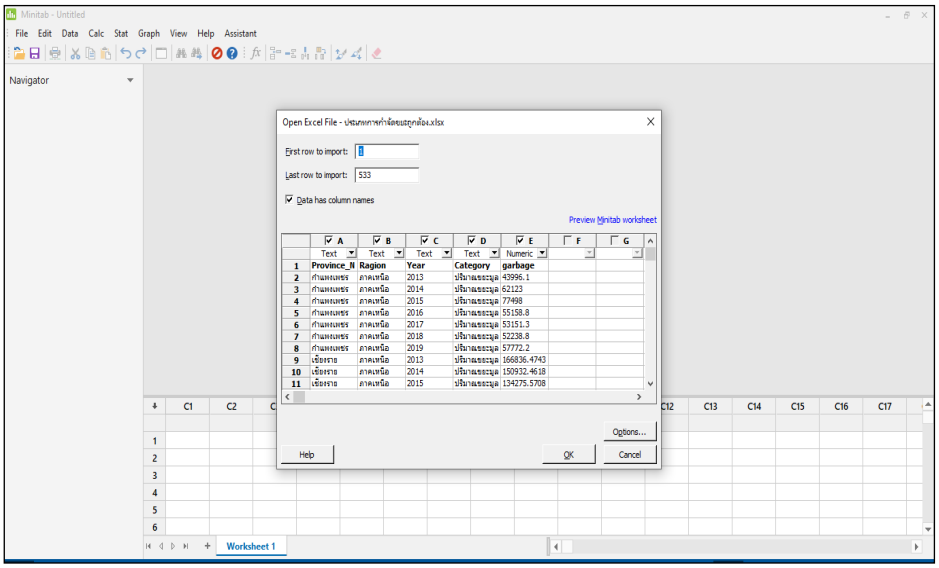
ภาพที่ 3.35 ข้อมูลที่นำมาใช้ในการพยากรณ์แต่ละประเภท

3.3) ใช้ฟังก์ชัน File open เพื่อทำการนำข้อมูลเข้าฐานข้อมูลรูปแบบของ Microsoft Excel และเลือกข้อมูลที่ผ่านการจัดกลุ่มข้อมูลแล้ว และนำข้อมูลเข้าโปรแกรม



ภาพที่ 3.36 เลือกไฟล์ที่ต้องการนำข้อมูลมาพยากรณ์

3.4) ทำการ Import ข้อมูลทั้งหมดของแต่ละภาคเข้าโปรแกรมเพื่อทำการหาค่าผลลัพธ์ของการพยากรณ์



ภาพที่ 3.37 Import ข้อมูลเพื่อหาค่าการพยากรณ์

3.5) ข้อมูลที่ถูก Import ลงโปรแกรมจะมีทั้งหมด 5 แอตทริบิวท์ ได้แก่ Province_Name, region, year, Category และจะนำแอตทริบิวท์ garbage มาคำนวณหาค่าพยากรณ์ด้วยวิธีต่างๆ

	C1-T	C2-T	C3-T	C4-T	C5
	Province_Name	Ragion	Year	Category	garbage
1	กำแพงเพชร	ภาคเหนือ	2013	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	43996
2	กำแพงเพชร	ภาคเหนือ	2014	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	62123
3	กำแพงเพชร	ภาคเหนือ	2015	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	77498
4	กำแพงเพชร	ภาคเหนือ	2016	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	55159
5	กำแพงเพชร	ภาคเหนือ	2017	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	53151
6	กำแพงเพชร	ภาคเหนือ	2018	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	52239
7	กำแพงเพชร	ภาคเหนือ	2019	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	57772

Sheet1

ภาพที่ 3.38 แอตทริบิวท์ที่นำมาคำนวณหาค่าพยากรณ์

3.1.4 การสร้างแบบจำลอง (Modeling)

ผู้วิเคราะห์ข้อมูลจะวิเคราะห์ด้วยเทคนิคทางสถิติ ทำการเรียนรู้จากข้อมูลเดิมของปริมาณขยะมูลฝอย การเปรียบเทียบความคลาดเคลื่อนของวิธีการพยากรณ์ โดยนำโมเดล Time series analysis มาช่วยในการหาค่าความคลาดเคลื่อนของการพยากรณ์ในวิธีต่าง ๆ โดยมีขั้นตอนการดำเนินการดังนี้

ในการพยากรณ์โดยใช้รูปแบบอนุกรมเวลาจะเป็นการพยากรณ์ข้อมูลในอนาคตจากข้อมูลในอดีตเท่านั้น ตัวแปรอื่น ๆ จะไม่นำมาพิจารณา ผู้วิเคราะห์ข้อมูลได้เลือกข้อมูลในรูปแบบมีแนวโน้ม (Trend)

	C1-T	C2-T	C3-T	C4-T	C5
	Province_Name	Ragion	Year	Category	garbage
1	กำแพงเพชร	ภาคเหนือ	2013	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	43996
2	กำแพงเพชร	ภาคเหนือ	2014	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	62123
3	กำแพงเพชร	ภาคเหนือ	2015	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	77498
4	กำแพงเพชร	ภาคเหนือ	2016	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	55159
5	กำแพงเพชร	ภาคเหนือ	2017	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	53151
6	กำแพงเพชร	ภาคเหนือ	2018	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	52239
7	กำแพงเพชร	ภาคเหนือ	2019	ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	57772

Sheet1

ภาพที่ 3.39 ตารางข้อมูลปริมาณขยะมูลฝอยโดยแบ่งตามจังหวัดและประเภทการกำจัดขยะ

การสร้างโมเดล Time Series Models จะทำการเลือกแอททริบิวต์ข้อมูลเชิงปริมาณที่จัดเก็บไว้ มาคำนวณหาค่าความเป็นไปได้เชิงพยากรณ์ โดยการคำนวณจากทางคณิตศาสตร์ดังนี้

ขั้นตอนที่ 1 คำนวณหาค่า Smoothing Factor จากสมการ

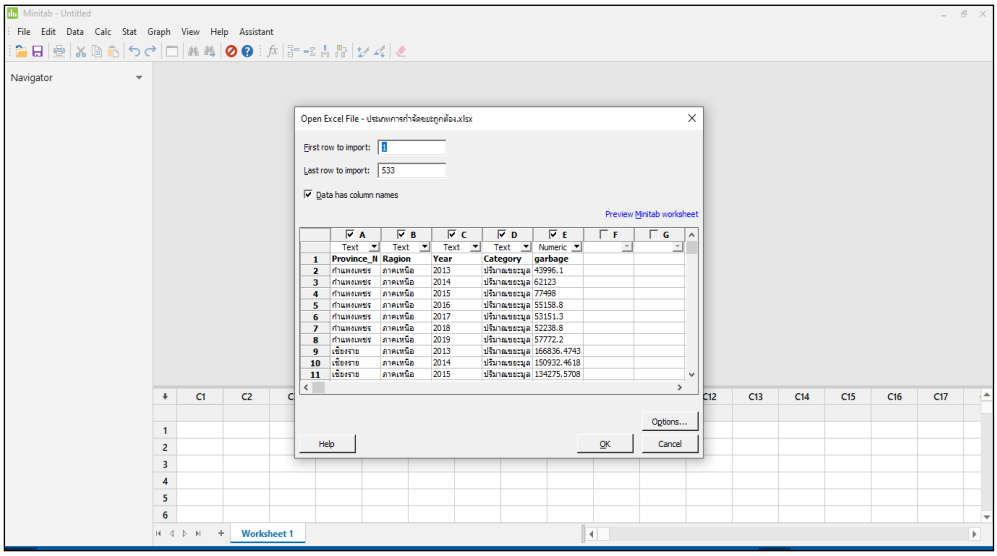
$$2 / (n + 1)$$

ขั้นตอนที่ 2 คำนวณ EMA

$$EMA_n = aP_n + EMA_{n-1}(1 - a)$$

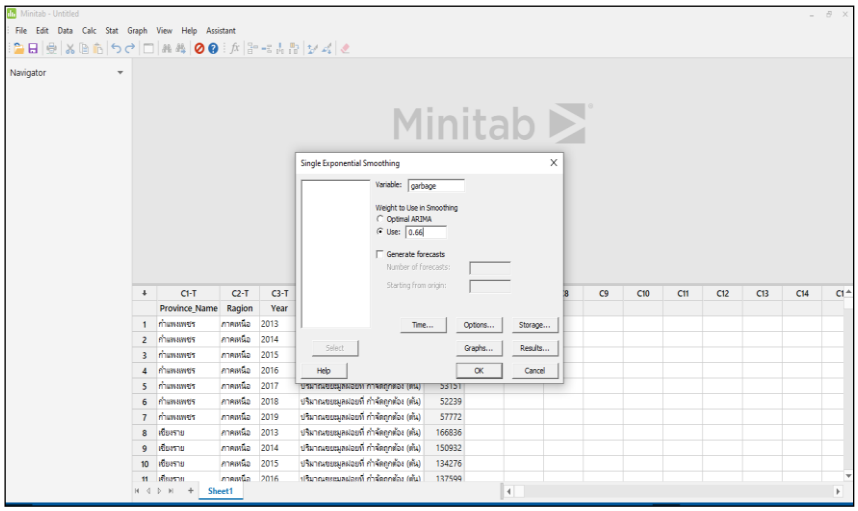
1) การใช้เทคนิคการวิเคราะห์อนุกรมเวลา ด้วยวิธีการพยากรณ์ค่าเฉลี่ยเคลื่อนที่ถ่วงน้ำหนัก โดยผลลัพธ์จากการพยากรณ์ผู้วิเคราะห์ข้อมูลได้ข้อมูลที่ผ่านการวิเคราะห์ข้อมูลจากการสร้างโมเดล Time Series Models จึงนำข้อมูลดังกล่าว มาทดลองกับโปรแกรม Minitab ซึ่งมีขั้นตอนการทำงานดังนี้

ขั้นตอนที่ 1 ใช้โปรแกรม Minitab เลือกคำสั่ง file –> open –>เลือกไฟล์ที่จะนำเข้า



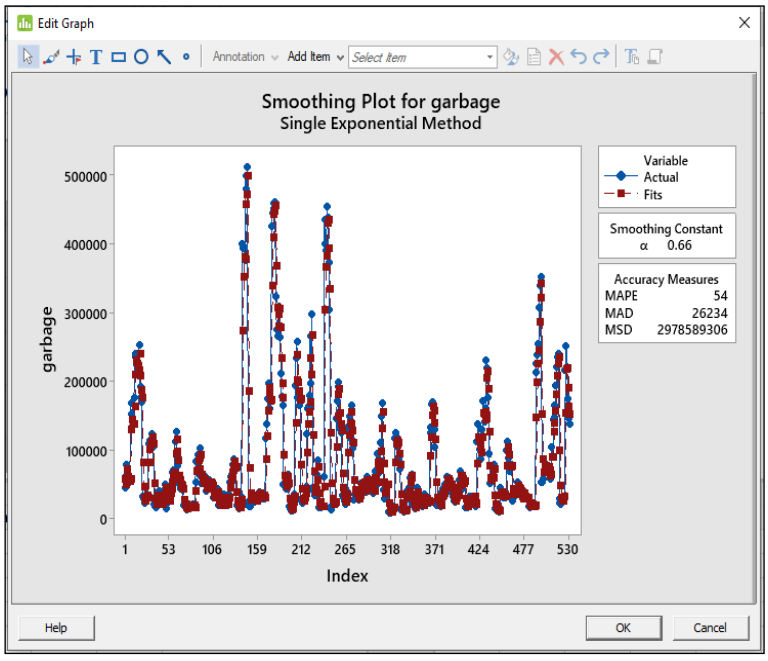
ภาพที่ 3.40 แสดงขั้นตอนการนำข้อมูลเข้าโปรแกรม Minitab

ขั้นตอนที่ 2 คลิกเลือกคำสั่งที่แถบเมนู stat –> time series –> single Exponential Smoothing –> ok



ภาพที่ 3.41 แสดงขั้นตอนการกำหนดรูปแบบการพยากรณ์วิธี Single EMA

ขั้นตอนที่ 3 ผลลัพธ์ที่แสดงการพยากรณ์แล้วทำการแทนค่าสมการตามสูตรอนุกรมเวลาวิธีการ Single Exponential Moving Average ในการพยากรณ์จะมีความคลาดเคลื่อน MAPE 54% ในภาพของแต่ละประเภทการกำจัดขยะมูลฝอย



ภาพที่ 3.42 แสดงผลลัพธ์กราฟวิธี Single EMA

การพยากรณ์โดยการวิเคราะห์อนุกรมเวลาของการทำเหมืองข้อมูล ผ่านโปรแกรม โดยผลลัพธ์ที่ได้จากการพยากรณ์ล่วงหน้าและหลักเกณฑ์ในการเลือกค่าตรวจสอบนั้น มีหลักเกณฑ์ปฏิบัติคือ ค่า MAPE ส่วนใหญ่แล้วใช้เปรียบเทียบความแม่นยำของค่าพยากรณ์ของข้อมูลทางธุรกิจ ทั้งนี้ค่า MAPE เป็นค่าวัดเปรียบเทียบดังนั้นค่า MAPE จะนิยมกว่าค่า MAD ซึ่งจากการศึกษาพบว่าถ้าค่า MAPE ที่สามารถเชื่อถือได้ดังตารางที่ 3.20

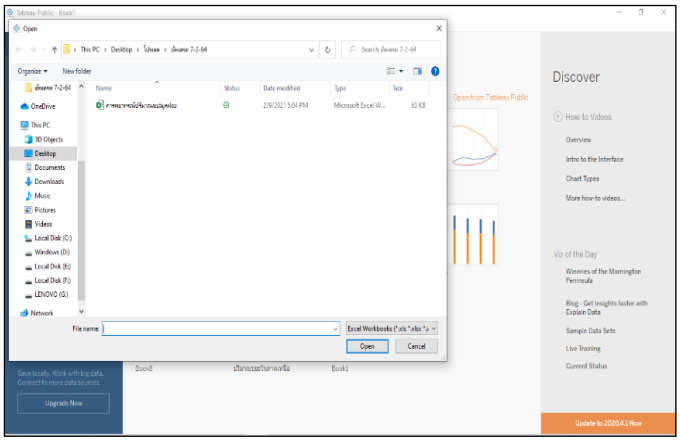
ตารางที่ 3.34 ตารางเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE)

ค่า MAPE	**ความน่าเชื่อถือ**
น้อยกว่า 10%	จัดว่าการพยากรณ์ค่อนข้างแม่นยำ
10% ถึง 20%	จัดว่าการพยากรณ์ใช้ได้ดี
20% ถึง 50%	จัดว่าการพยากรณ์พอใช้
มากกว่า 50%	จัดว่าการพยากรณ์ไม่แม่นยำ

ที่มา : รศ.สุพรรณี อิ้งปัญสัตวงศ์ (2563) , 12 ธันวาคม 2563.

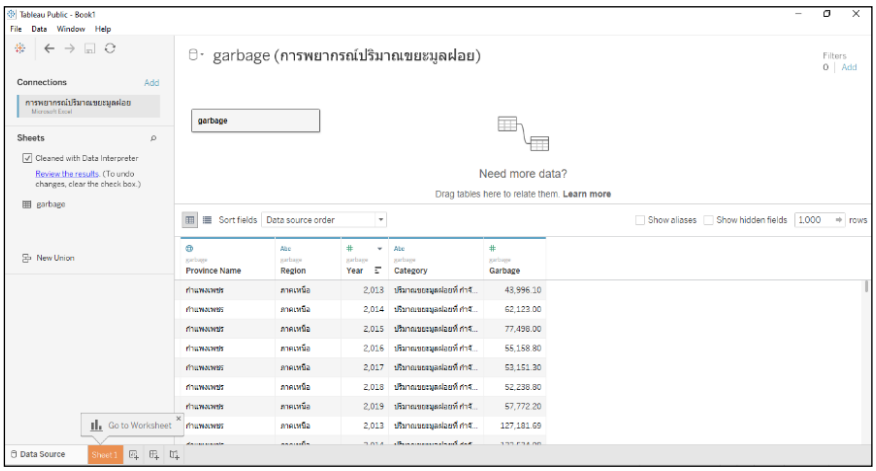
2) การพยากรณ์โดยการวิเคราะห์อนุกรมเวลาของข้อมูล (Data mining time series analysis) ผ่านโปรแกรม Tableau Public โดยผลลัพธ์จากการพยากรณ์ล่วงหน้าผู้วิเคราะห์ข้อมูลได้ข้อมูลที่ผ่านการวิเคราะห์ข้อมูลจึงนำข้อมูลดังกล่าวมาทดสอบกับโปรแกรม Tableau Public ซึ่งมีขั้นตอนการทำงานดังนี้

2.1) หน้าต่างที่ทำการเชื่อมต่อฐานข้อมูลไฟล์ Excel เพื่อนำมาวิเคราะห์



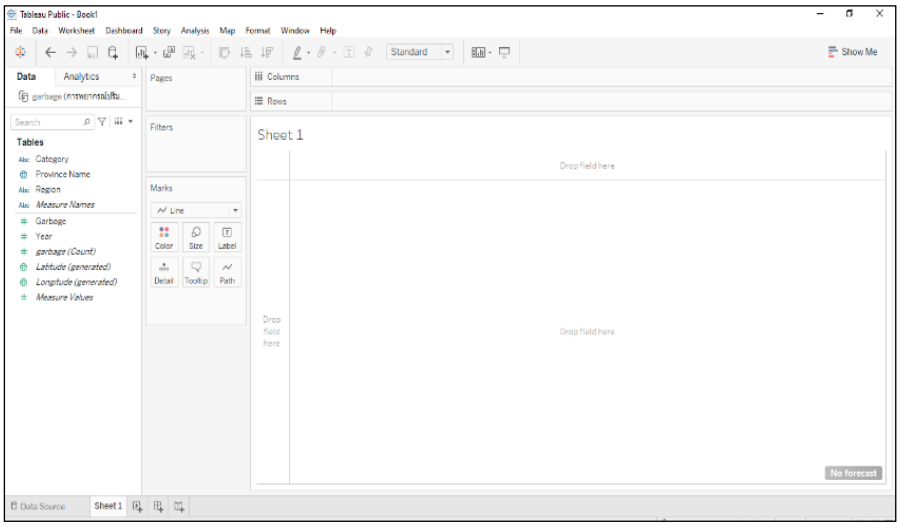
ภาพที่ 3.43 หน้าต่างที่ทำการเชื่อมต่อฐานข้อมูลไฟล์ Excel เพื่อนำมาวิเคราะห์

2.2) ทำการใช้เครื่องมือทำเหมืองข้อมูล (Data Mining) ของโปแกรม Tableau Public ในการพยากรณ์ข้อมูล



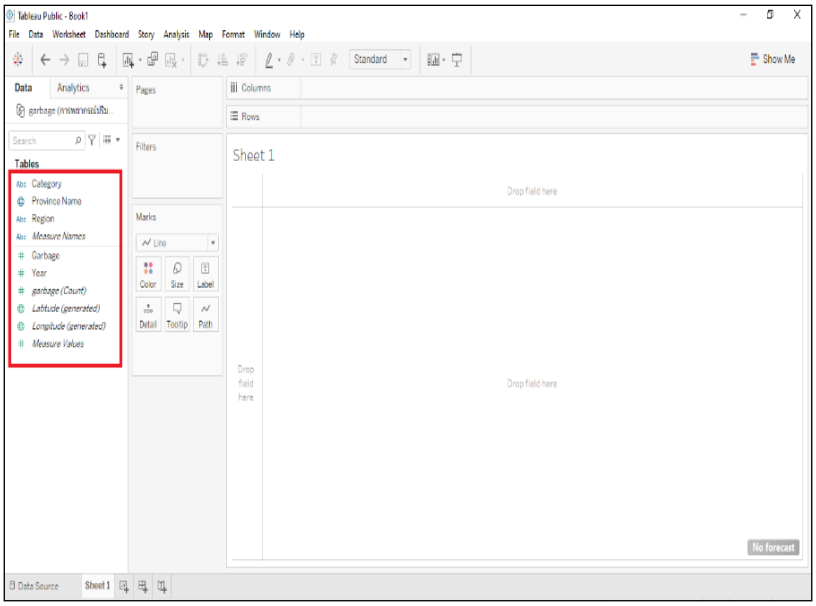
ภาพที่ 3.44 Import แสดงผลการโหลดฐานข้อมูลเข้าสู่โปรแกรม Tableau Public

2.3) หน้าต่างพื้นที่การทำงานของโปรแกรม Tableau Public



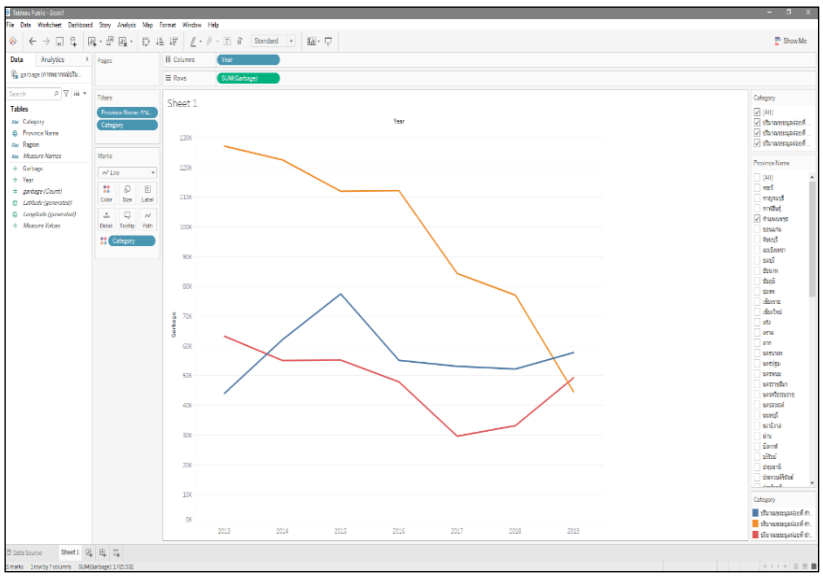
ภาพที่ 3.45 หน้าต่างพื้นที่ทำงานของโปรแกรม tableau

2.4) หน้าต่างที่แสดงข้อมูลที่ Import ข้อมูลเข้ามาแล้วจะมีข้อมูลอยู่แถบด้านซ้ายมือ



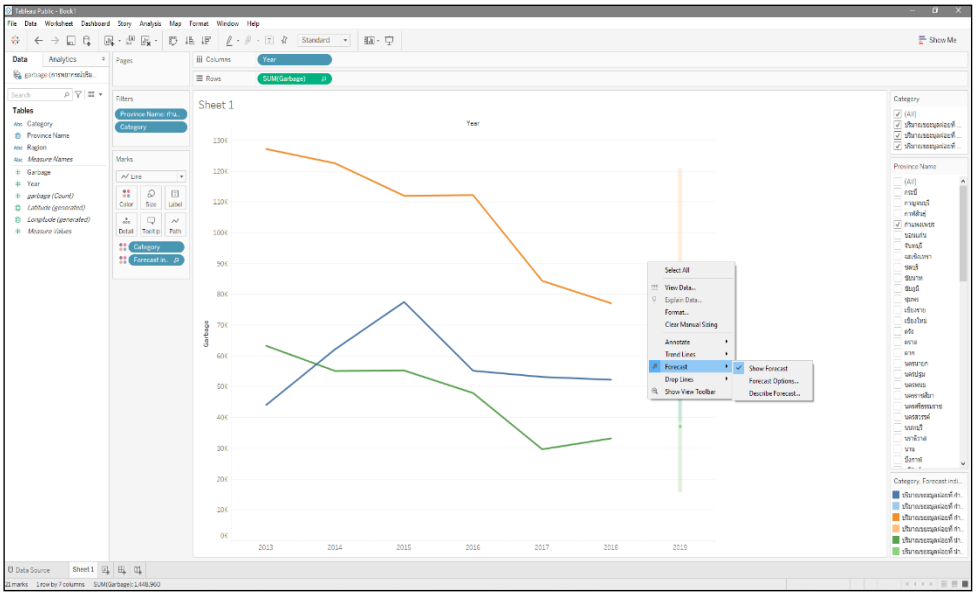
ภาพที่ 3.46 หน้าต่างที่แสดงข้อมูลที่ Import ข้อมูล

2.5) การเตรียมกราฟข้อมูลก่อนการวิเคราะห์ข้อมูล



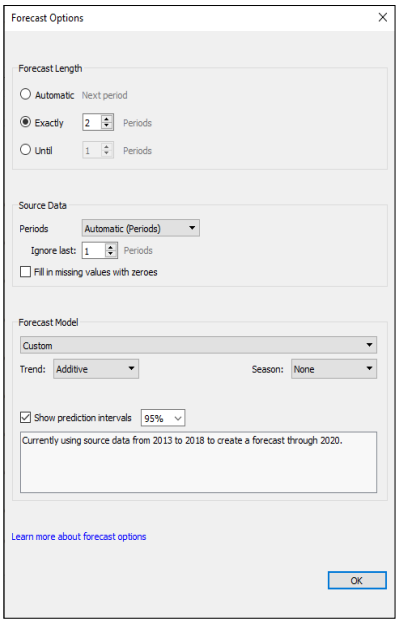
ภาพที่ 3.47 แสดงการเตรียมกราฟข้อมูลก่อนการวิเคราะห์ข้อมูล

2.6) เครื่องมือในการวิเคราะห์ข้อมูลของโปรแกรม Tableau Public



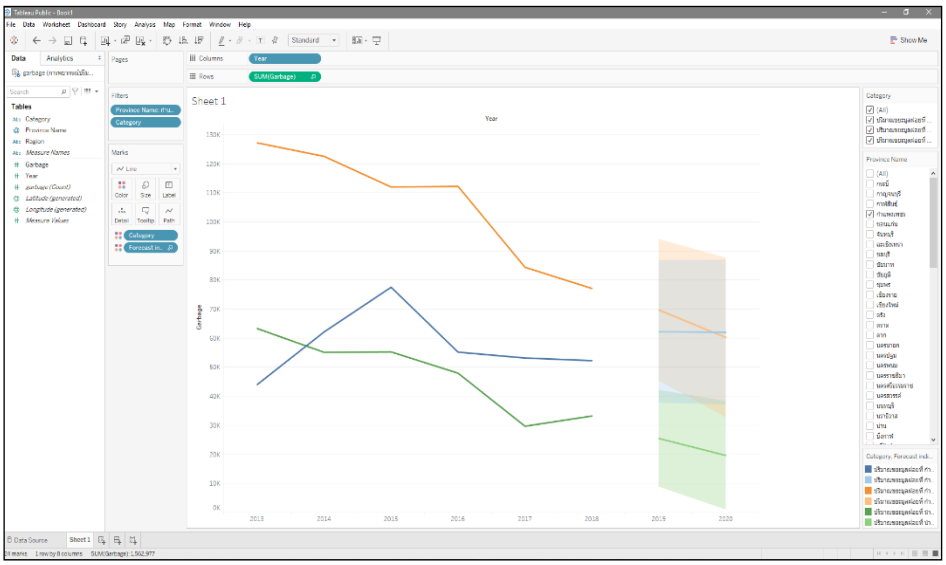
ภาพที่ 3.48 การแสดงเครื่องมือในการวิเคราะห์

2.7) การใช้เครื่องมือทำเหมืองข้อมูล (Data Mining) ของโปแกรม Tableau Public เพื่อพยากรณ์ข้อมูลปริมาณขยะมูลฝอยในช่วงเวลา 1 ปีถัดไปโดยใช้ Exactly ในการกำหนดค่าการพยากรณ์



ภาพที่ 3.49 การแสดงเครื่องมือในการพยากรณ์

2.8) ผลลัพธ์ในการพยากรณ์ช่วงเวลา 2 ปีถัดไป



ภาพที่ 3.50 การแสดงผลลัพธ์ในการพยากรณ์ช่วงเวลา 2 ปีถัดไป

2.9) ผลการคำนวณค่าความคลาดเคลื่อน (MAPE)

Describe Forecast

Summary | Models

All forecasts were computed using exponential smoothing.

Avg. Garbage

Color	Model			Quality Metrics					Smoothing Coefficients		
Category	Level	Trend	Season	RMSE	MAE	MASE	MAPE	AIC	Alpha	Beta	Gamma
ปริมาณขยะมูลฝอยที่ นำไปใช้ประโยชน์ (ตัน)	Additive	Additive	None	9,095	7,814	1.19	10.7%	119	0.500	0.000	0.000
ปริมาณขยะมูลฝอยที่ กำจัดไม่ถูกต้อง (ตัน)	Additive	Additive	None	11,351	9,904	0.94	5.9%	122	0.500	0.013	0.000
ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	Additive	Additive	None	2,405	2,207	0.75	2.7%	103	0.072	0.500	0.000

ภาพที่ 3.51 ผลการคำนวณค่าความคลาดเคลื่อน (MAPE)

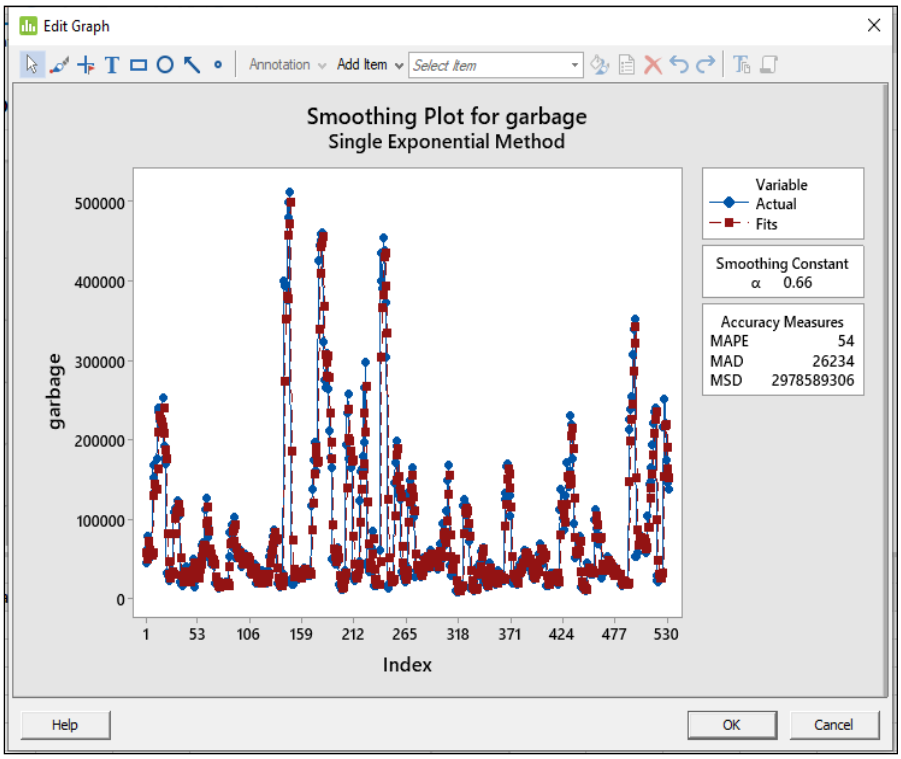
3.1.5 การประเมินประสิทธิภาพ (Evaluation)

เป็นขั้นตอนก่อนนำผลลัพธ์ที่ได้จากขั้นตอนที่ 4 ไปใช้งานผู้วิเคราะห์ข้อมูลวัดประสิทธิภาพของผลลัพธ์ได้กับวัตถุประสงค์ที่ตั้งไว้ในขั้นตอนแรก ว่ามีนัยสำคัญหรือความน่าเชื่อถือมากน้อยเพียงใด โดยทำการทดลองโมเดลเทคนิค Time Series Models

ด้วยการวิเคราะห์อนุกรมเวลาของข้อมูล (Data mining time series analysis) ผ่าน 2 โปรแกรม มีค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ยดังนี้

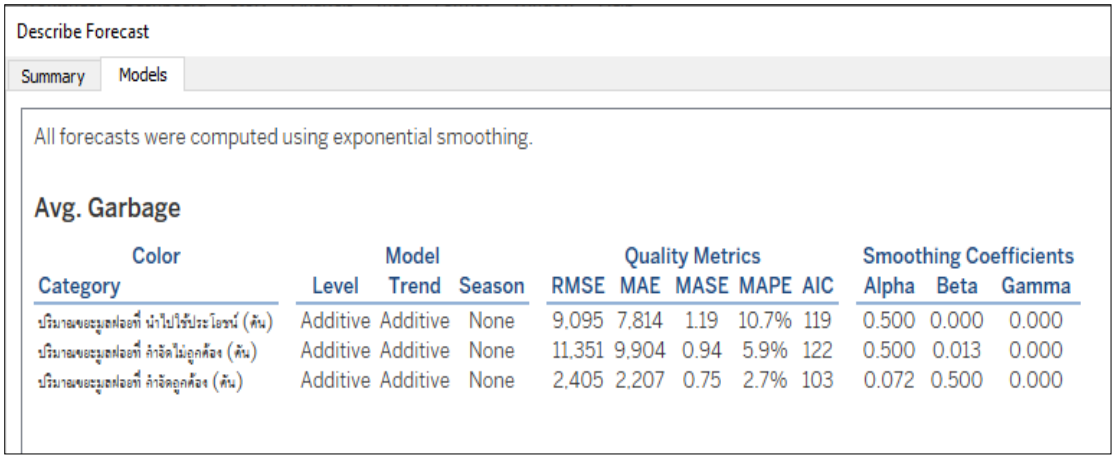
1) การวัดประสิทธิภาพของผลลัพธ์ด้วยโปรแกรม Minitab

ผลการตรวจสอบค่าเปอร์เซ็นต์ความคลาดเคลื่อนสมบูรณ์เฉลี่ย (Mean Absolute Percentage Error : MAPE) จากการทดลองโมเดลเทคนิค Time Series Models ด้วยการ วิเคราะห์อนุกรมเวลาของข้อมูล (Data mining time series analysis) ผ่านโปรแกรม Minitab ได้ผลลัพธ์ถึง 54%



ภาพที่ 3.52 ผลลัพธ์การวัดประสิทธิภาพของผลลัพธ์โปรแกรม Minitab

2) การวัดประสิทธิภาพของผลลัพธ์ด้วยโปรแกรม Tableau Public



Describe Forecast

Summary | Models

All forecasts were computed using exponential smoothing.

Avg. Garbage

Color	Model			Quality Metrics					Smoothing Coefficients		
Category	Level	Trend	Season	RMSE	MAE	MASE	MAPE	AIC	Alpha	Beta	Gamma
ปริมาณขยะมูลฝอยที่ นำไปใช้ประโยชน์ (ตัน)	Additive	Additive	None	9,095	7,814	1.19	10.7%	119	0.500	0.000	0.000
ปริมาณขยะมูลฝอยที่ กำจัดไม่ถูกต้อง (ตัน)	Additive	Additive	None	11,351	9,904	0.94	5.9%	122	0.500	0.013	0.000
ปริมาณขยะมูลฝอยที่ กำจัดถูกต้อง (ตัน)	Additive	Additive	None	2,405	2,207	0.75	2.7%	103	0.072	0.500	0.000

ภาพที่ 3.53 ผลลัพธ์การวัดประสิทธิภาพของผลลัพธ์ด้วยโปรแกรม Tableau Public

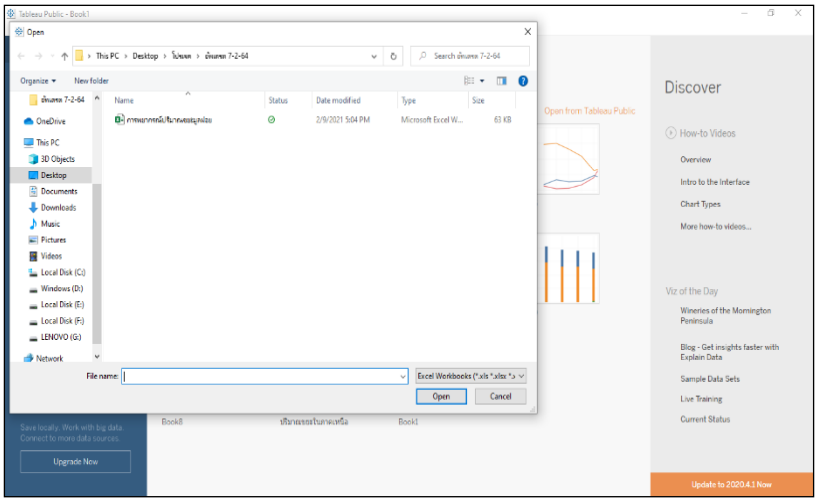
ผลการตรวจสอบค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percentage Error : MAPE) จากการทดลองโมเดลเทคนิค Time Series Models ด้วยการวิเคราะห์อนุกรมเวลาของข้อมูล (Data mining time series analysis) ผ่านโปรแกรม Tableau Public ได้ผลลัพธ์ต่ำสุดที่แบ่งเป็นประเภท อยู่ 3 ประเภท 1.) ปริมาณขยะมูลฝอยที่นำไปใช้ประโยชน์ 10.7% 2.) ปริมาณขยะมูลฝอยที่กำจัดไม่ถูกต้อง 5.9% 3.) ปริมาณขยะมูลฝอย ที่กำจัดถูกต้อง 2.7%

จากการทดสอบประสิทธิภาพโมเดล และการประเมินค่าความคลาดเคลื่อน MAPE ที่ค่าคงที่เกิดยิ่งน้อยยิ่งเป็นผลดีแสดงว่าค่าพยากรณ์ออกมานั้นมีความคลาดเคลื่อนจากค่าจริง เพียงใด ผู้วิเคราะห์จึงเลือกใช้โมเดลการวิเคราะห์อนุกรมเวลาของข้อมูล (Data mining time series analysis) ผ่านโปรแกรม Tableau Public ให้ผลลัพธ์ที่น่าพึงพอใจอยู่ในระดับใช้ได้ดี สามารถนำโมเดลไปใช้งานได้ ผู้วิเคราะห์จึงเลือกใช้โมเดลอนุกรมเวลาของ Tableau Public ในการพยากรณ์ปริมาณขยะมูลฝอยในประเทศไทย เพื่อประโยชน์ในการตัดสินให้สอดคล้องกับความต้องการต่อไป

3.1.6 การนำผลลัพธ์ไปใช้งาน (Deployment)

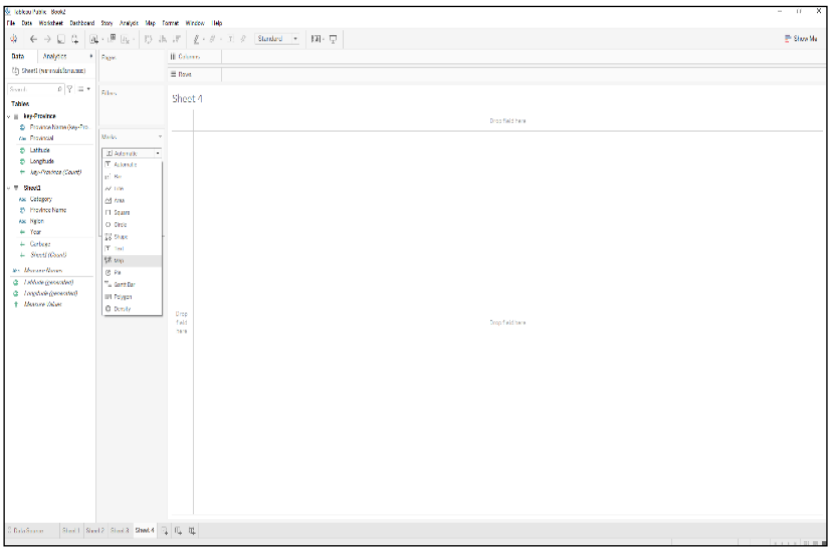
ผู้วิเคราะห์ข้อมูลนำผลองค์ความรู้ที่ได้เหล่านี้ไปนำเสนอข้อมูลแบบ visualization ด้วยการแสดงผลข้อมูลในรูปแบบของภาพโดยใช้โปรแกรม Tableau Public โดยดำเนินการตามขั้นตอนดังนี้

1) ทำการ Import ข้อมูลปริมาณขยะมูลฝอยในประเทศไทยเข้าโปรแกรม Tableau Public



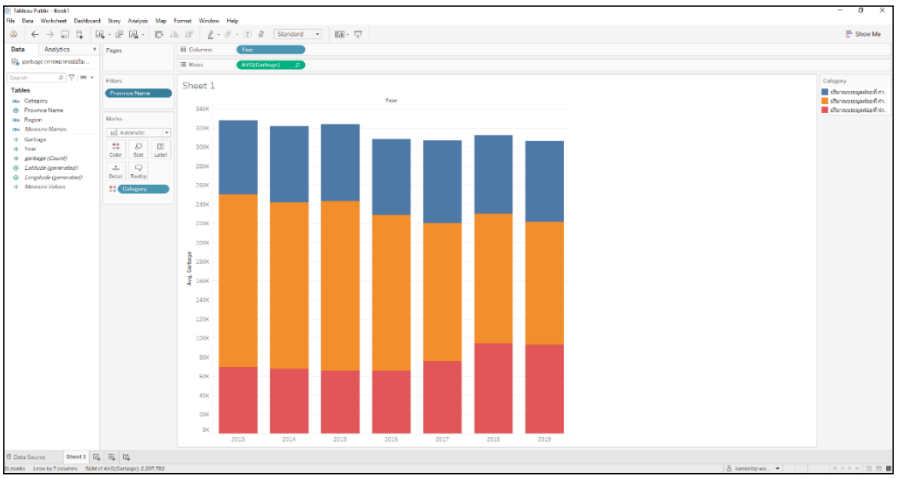
ภาพที่ 3.54 Import ข้อมูลปริมาณขยะมูลฝอยในประเทศไทยเพื่อหาผลลัพธ์

2) เข้าหน้า Dash Board การทำงานข้อมูลที่ Import จะอยู่แถบด้านซ้าย



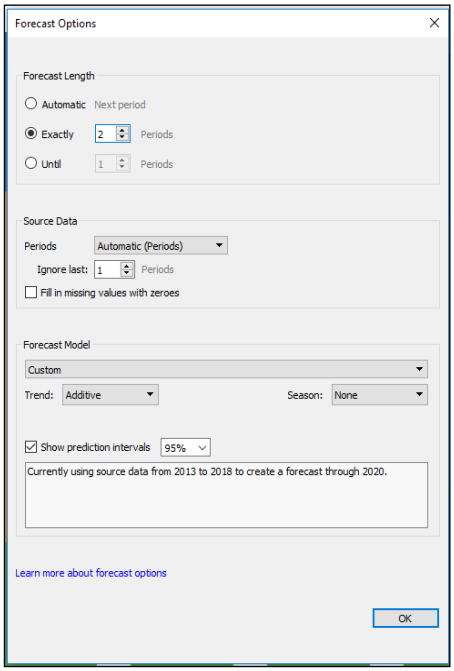
ภาพที่ 3.55 แสดงขั้นตอนการปรับ Mark ให้เป็น Bar

3) กำหนด Columns เป็น year กำหนด Rows เป็น SUM(Garbage), กำหนด category_name ใส่ใน Color และกำหนด Province ใส่ใน filters

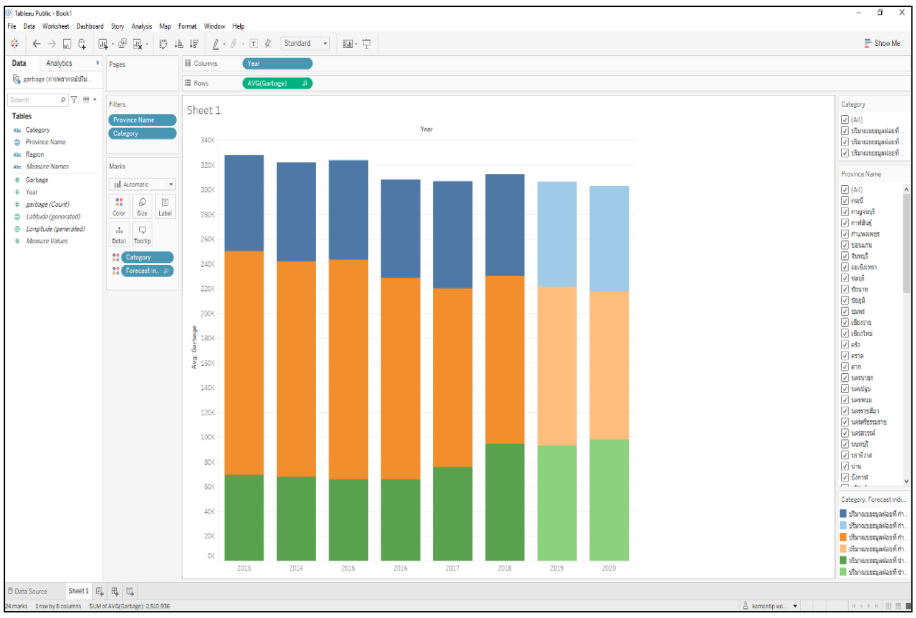


ภาพที่ 3.56 แสดงขั้นตอนการแสดง category บน Map

4) การกำหนดค่าการพยากรณ์โดยรูปแบบกราฟแท่ง

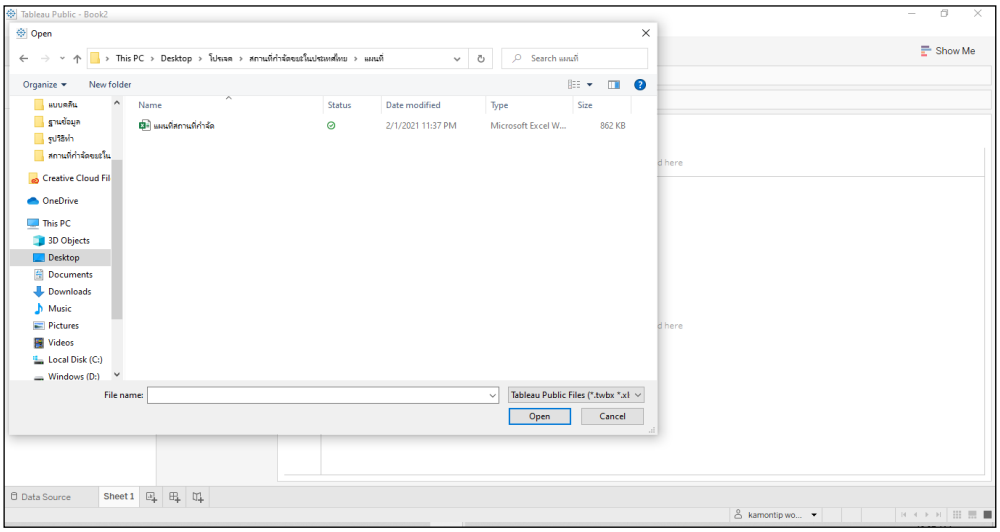


ภาพที่ 3.57 แสดงขั้นตอนการกำหนดค่าการพยากรณ์ในรูปแบบกราฟแท่ง



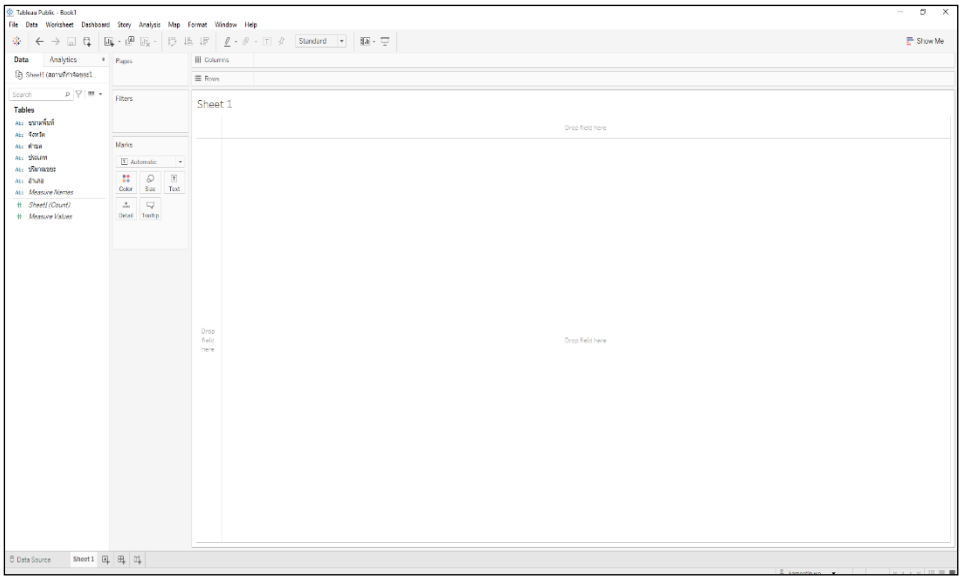
ภาพที่ 3.58 แสดงผลลัพธ์การพยากรณ์เป็นสีและรายละเอียดต่างๆ

5) หน้า Dash Board ในประเด็นที่ 2 สถานที่กำจัดขยะ การทำงานข้อมูลที่ Import จะอยู่แถบด้านซ้าย



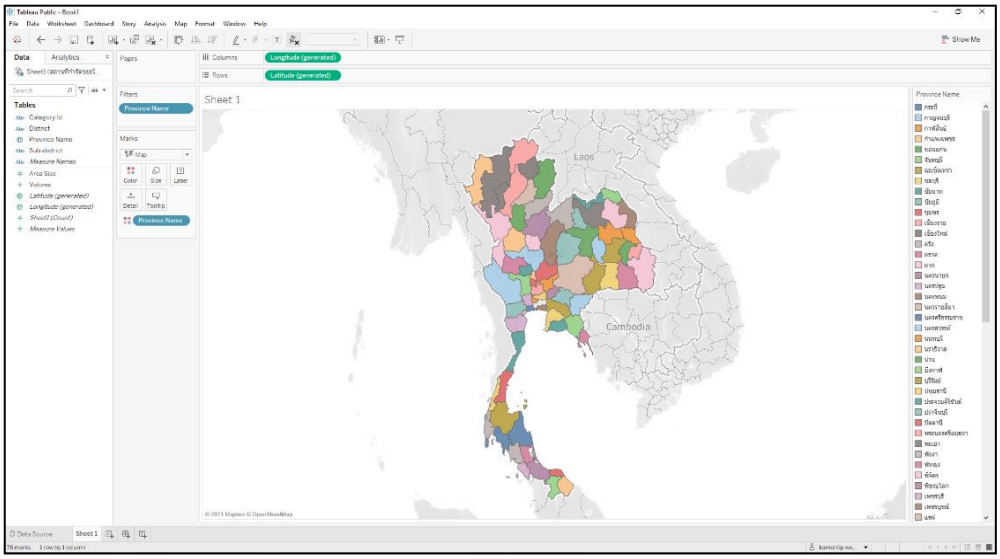
ภาพที่ 3.59 Import ข้อมูลสถานที่กำจัดขยะมูลฝอยในประเทศไทยเพื่อหาผลลัพธ์

6) เข้าหน้า Dash Board การทำงานข้อมูลที่ Import จะอยู่แถบด้านซ้าย

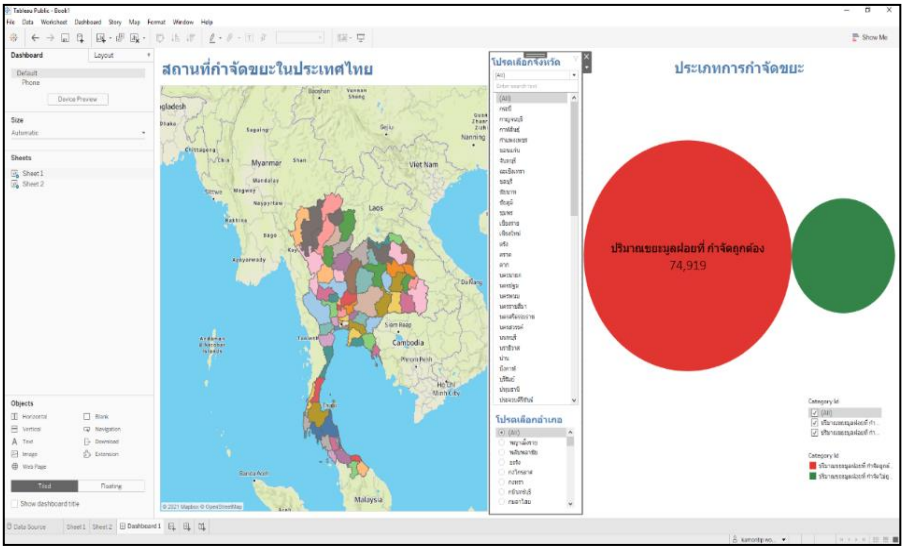


ภาพที่ 3.60 แสดงขั้นตอนการปรับ Mark ให้เป็น Map

7) กำหนด Columns เป็น Longitude (generated) กำหนด Rows เป็น Latitude (generated) และกำหนด Province Name ใส่ใน Color, กำหนด District ใส่ใน Detail และกำหนด SUM (Area Size) ใส่ใน Detail และกำหนด Province Name และกำหนด District ใส่ใน Filters



ภาพที่ 3.61 แสดงขั้นตอนการแสดงสถานที่กำจัดขยะในประเทศไทย บน Map



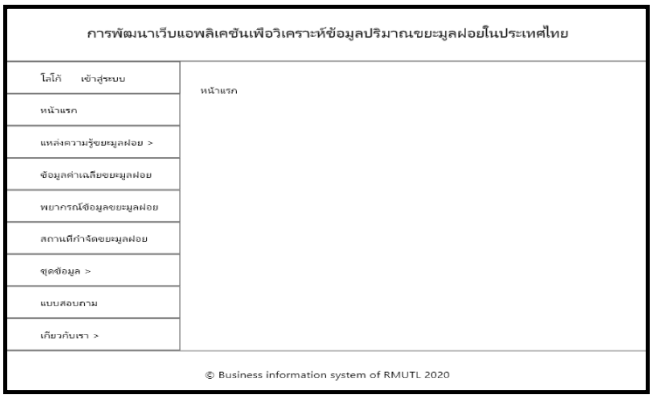
ภาพที่ 3.62 แสดงผลลัพธ์สถานที่กำจัดขยะมูลฝอยในประเทศไทย

3.6 การออกแบบและพัฒนาเว็บไซต์

การวางแผนการจัดลำดับ เนื้อหาสาระของเว็บไซต์ ออกเป็นหมวดหมู่ เพื่อจัดทำเป็นโครงสร้างในการจัดวางหน้าเว็บเพจทั้งหมด ทำให้เห็นโครงสร้างทั้งหมดของเว็บไซต์และการออกแบบโครงสร้างหรือจัดระเบียบของข้อมูลที่ชัดเจน แยกย่อยเนื้อหาออกเป็นหาออกเป็นส่วนต่าง ๆ ที่สัมพันธ์และให้มีอยู่ในมาตรฐานเดียวกัน จะช่วยให้นำมาใช้งานและง่ายต่อการเข้าอ่านเนื้อหาของผู้ใช้เว็บไซต์

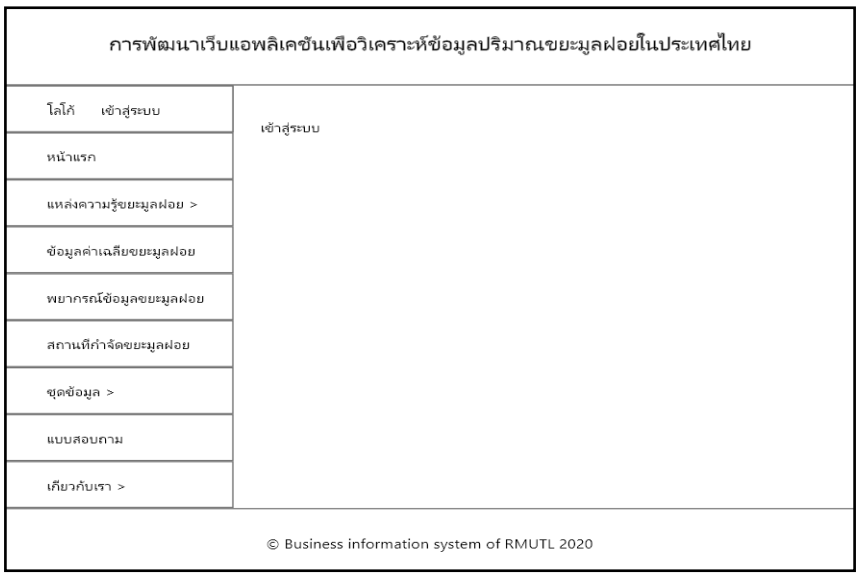
3.6.1 การออกแบบ Wireframe หน้าจอเว็บไซต์

1) หน้าแรกของเว็บไซต์ แสดงเมนูต่าง ๆ ของหน้าเว็บไซต์



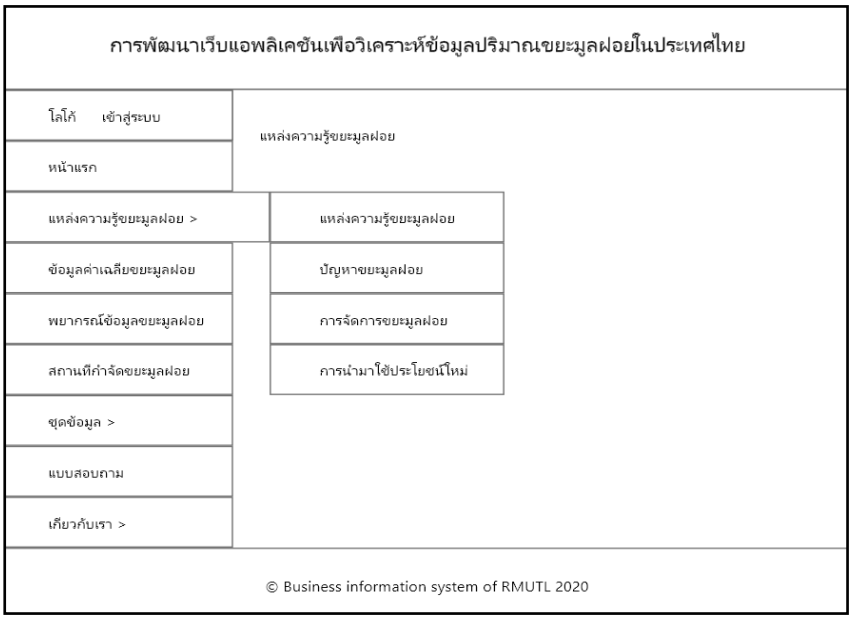
ภาพที่ 3.63 หน้าแรกของเว็บไซต์ แสดงเมนูต่าง ๆ ของหน้าเว็บไซต์

2) หน้าเข้าสู่ระบบของผู้ดูแลระบบ



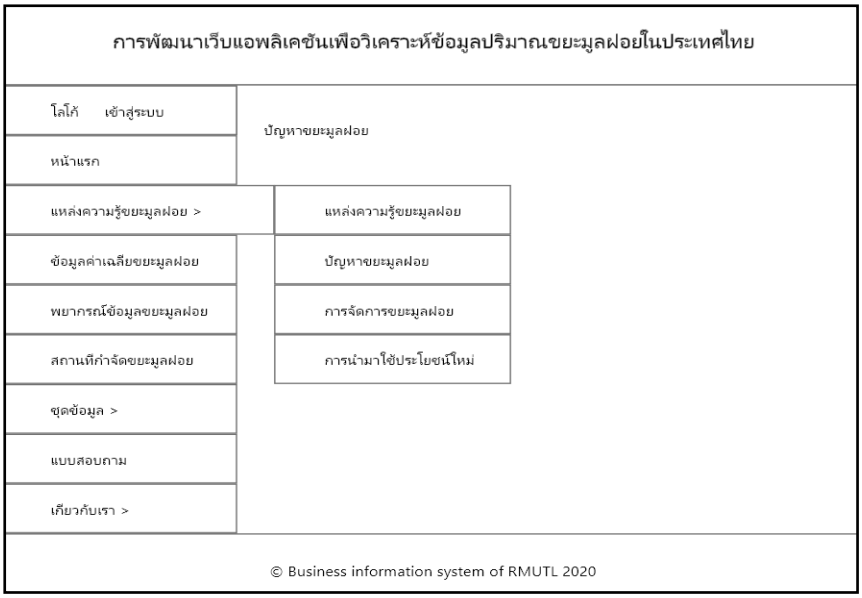
ภาพที่ 3.64 หน้าเข้าสู่ระบบของผู้ดูแลระบบ

3) หน้าแหล่งความรู้ขยะมูลฝอย



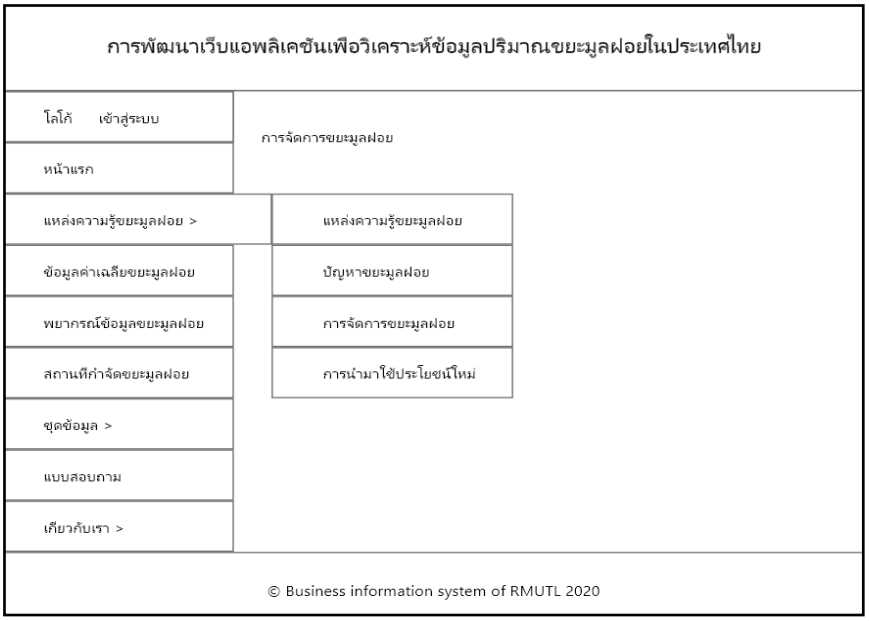
ภาพที่ 3.65 หน้าแหล่งความรู้ขยะมูลฝอย

4) หน้าความรู้เกี่ยวกับปัญหาขยะมูลฝอย



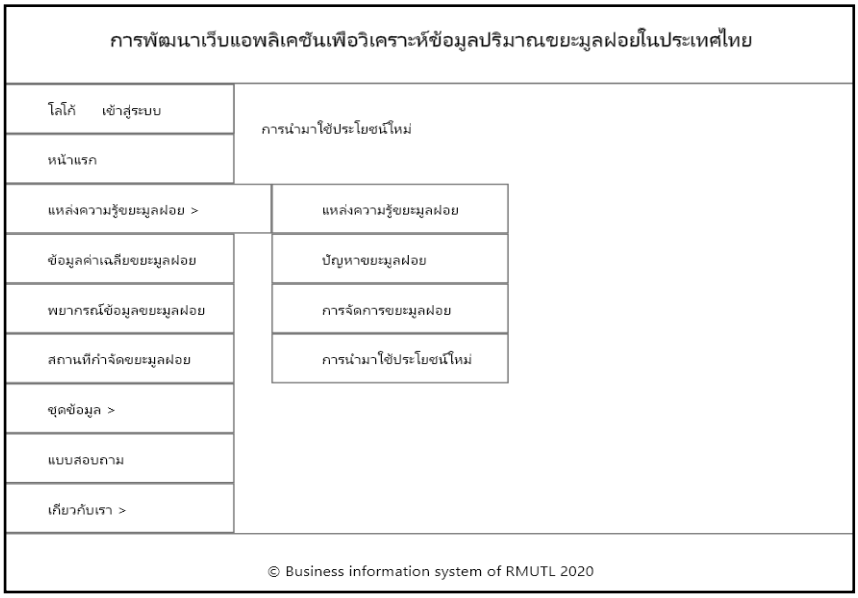
ภาพที่ 3.66 หน้าความรู้เกี่ยวกับปัญหาขยะมูลฝอย

5) หน้าการจัดการขยะมูลฝอย



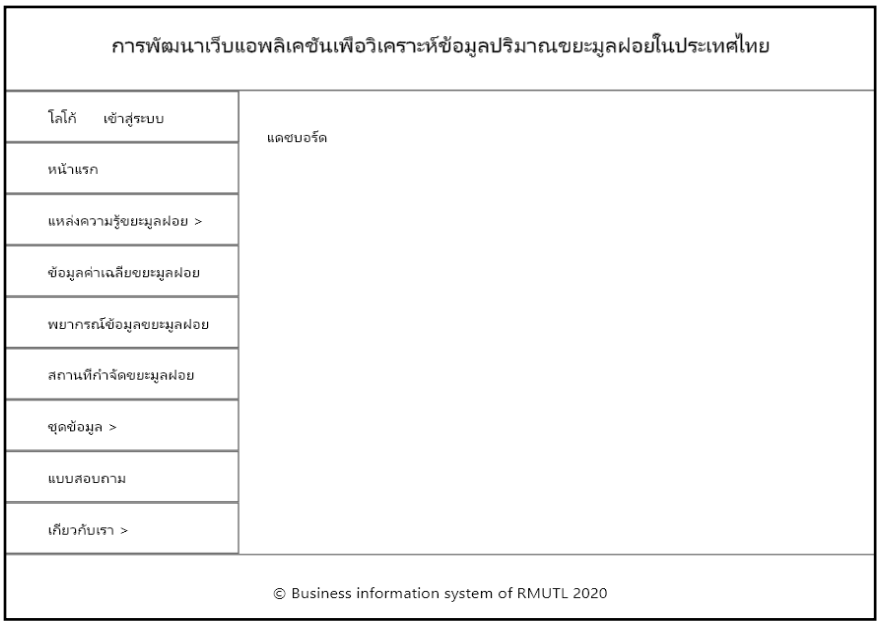
ภาพที่ 3.67 หน้าการจัดการขยะมูลฝอย

6) หน้าการนำขยะมาใช้ประโยชน์



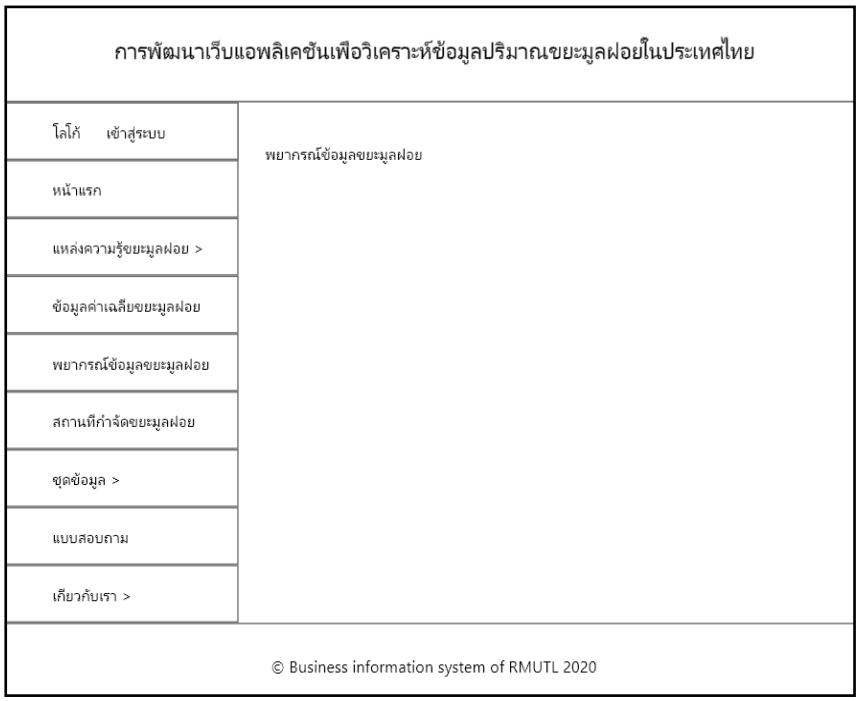
ภาพที่ 3.68 หน้าการนำขยะมาใช้ประโยชน์

7) หน้าการแสดงแดชบอร์ด



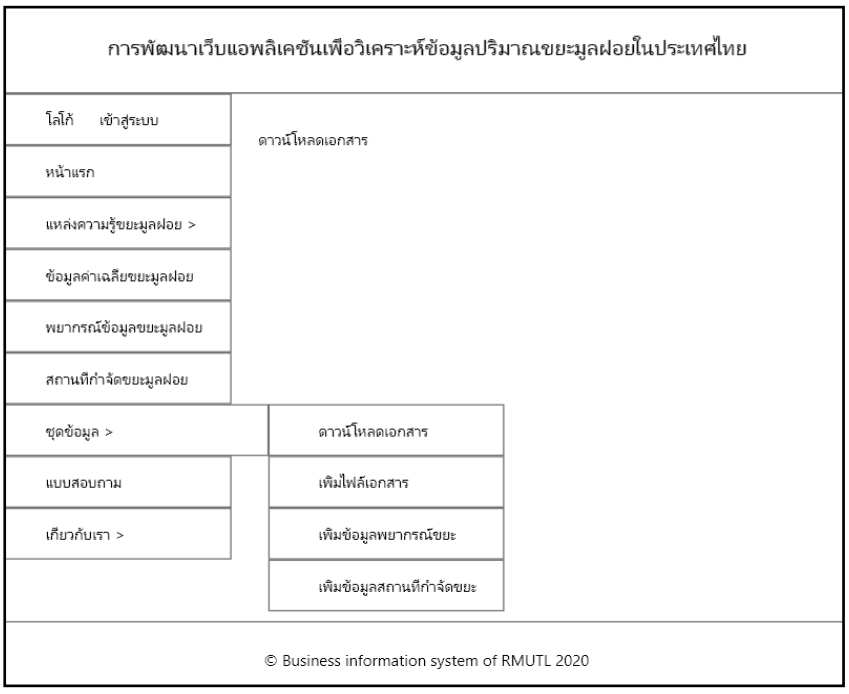
ภาพที่ 3.69 หน้าการแสดงแดชบอร์ด

8) หน้าการแสดงพยากรณ์ปริมาณขยะมูลฝอย



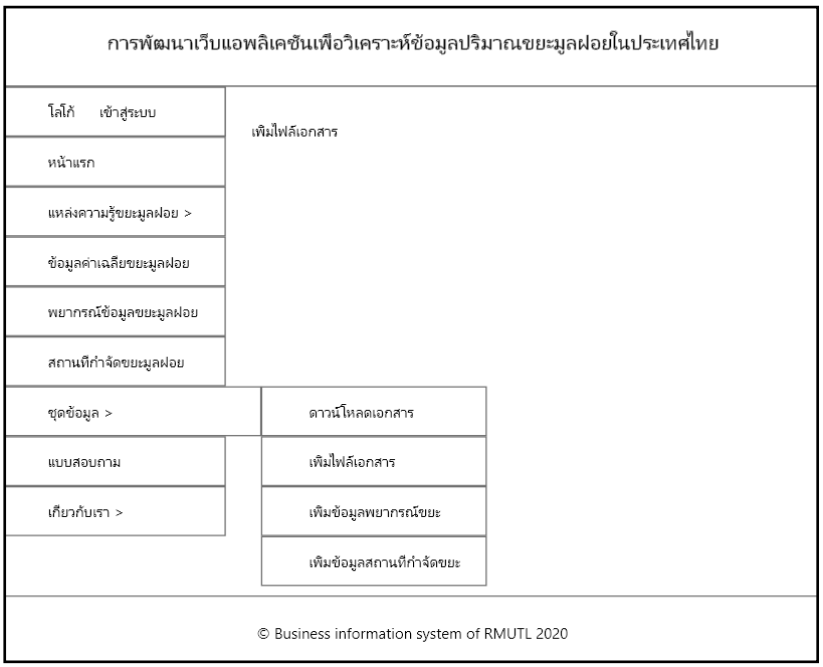
ภาพที่ 3.70 หน้าการแสดงพยากรณ์ปริมาณขยะมูลฝอย

9) หน้าการแสดงดาวน์โหลดเอกสารต่างๆ



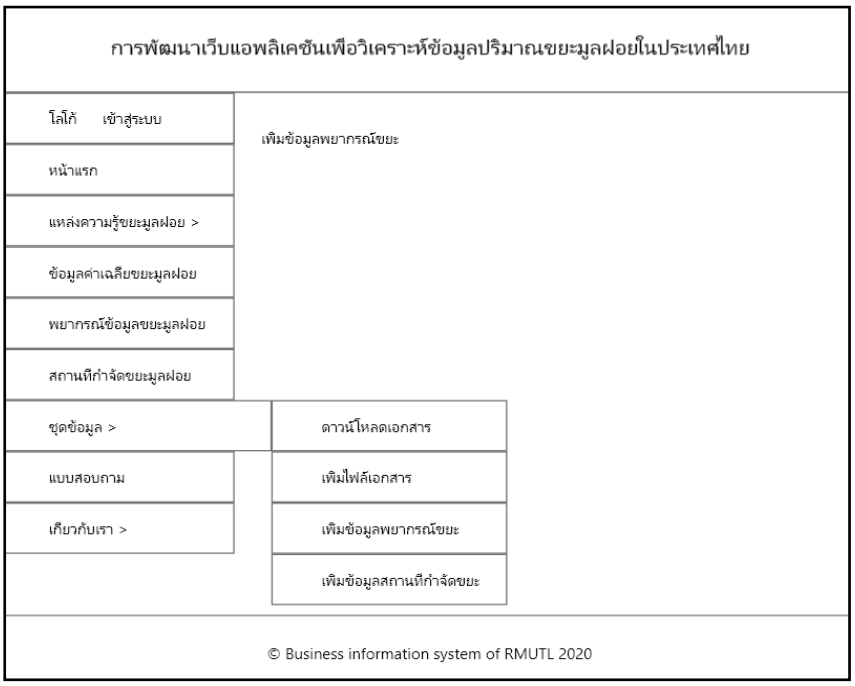
ภาพที่ 3.71 หน้าดาวน์โหลดเอกสารต่างๆ

10) หน้าเพิ่มไฟล์เอกสาร



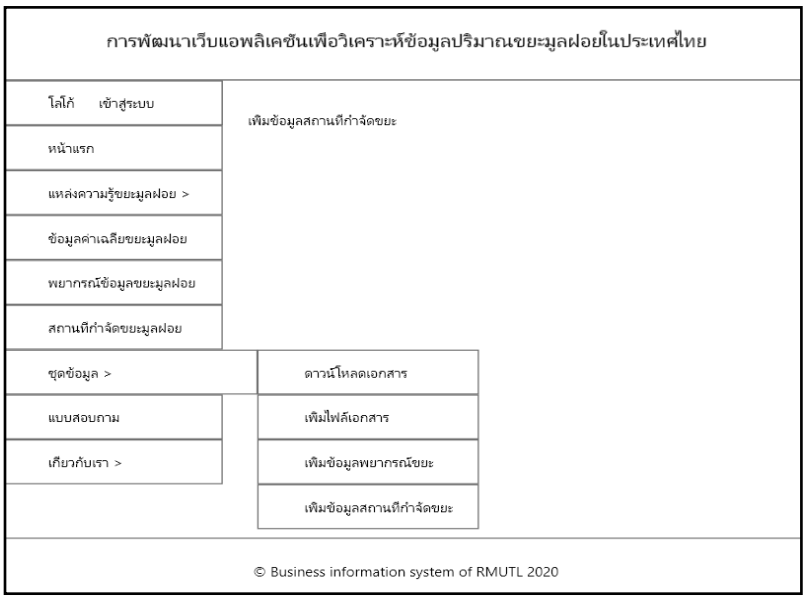
ภาพที่ 3.72 หน้าเพิ่มไฟล์เอกสาร

11) หน้าการเพิ่มข้อมูลการพยากรณ์ขยะมูลฝอย



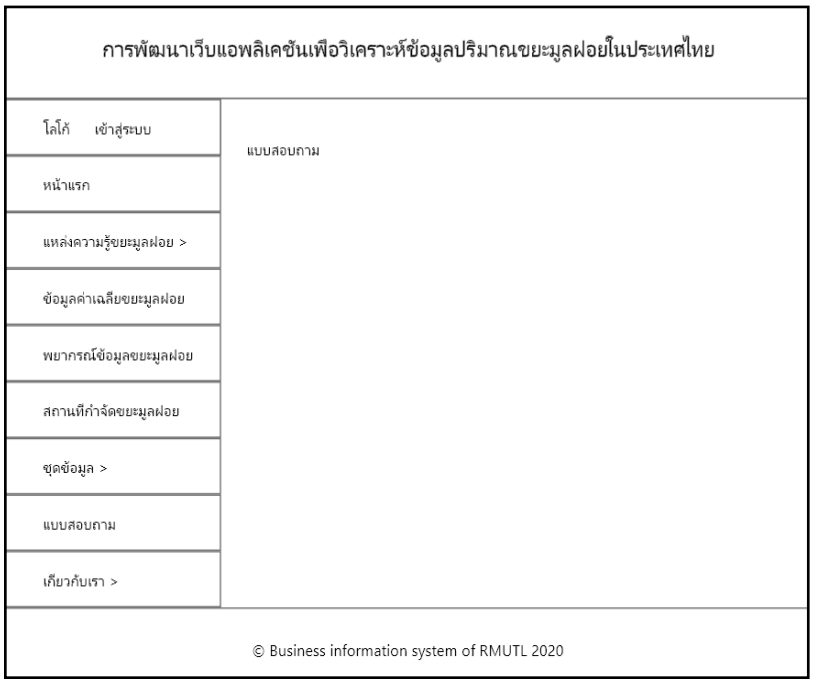
ภาพที่ 3.73 หน้าเพิ่มข้อมูลการพยากรณ์ขยะมูลฝอย

12) หน้าการเพิ่มข้อมูลสถานที่กำจัดขยะมูลฝอย



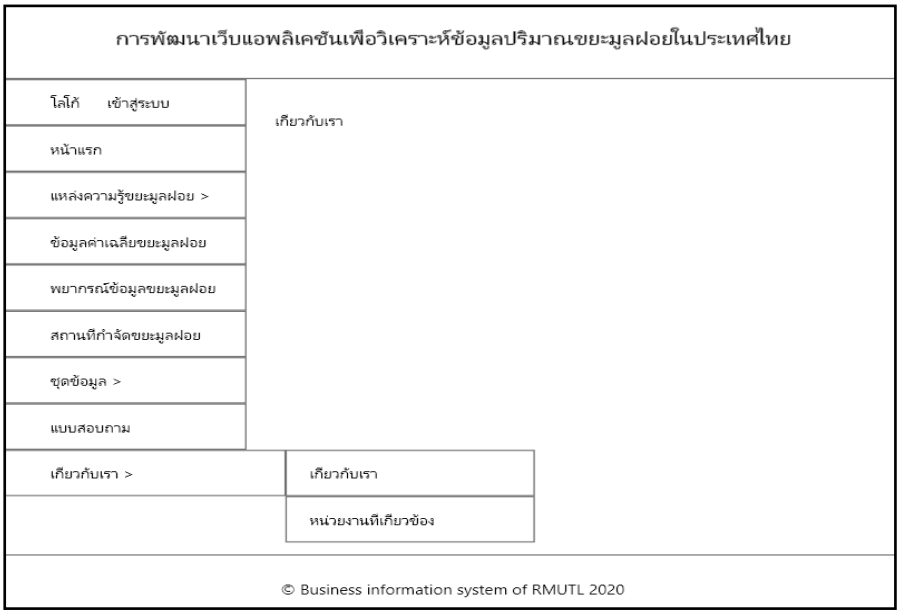
ภาพที่ 3.74 หน้าการเพิ่มสถานที่กำจัดขยะมูลฝอย

13) หน้าการแสดงแบบสอบถาม



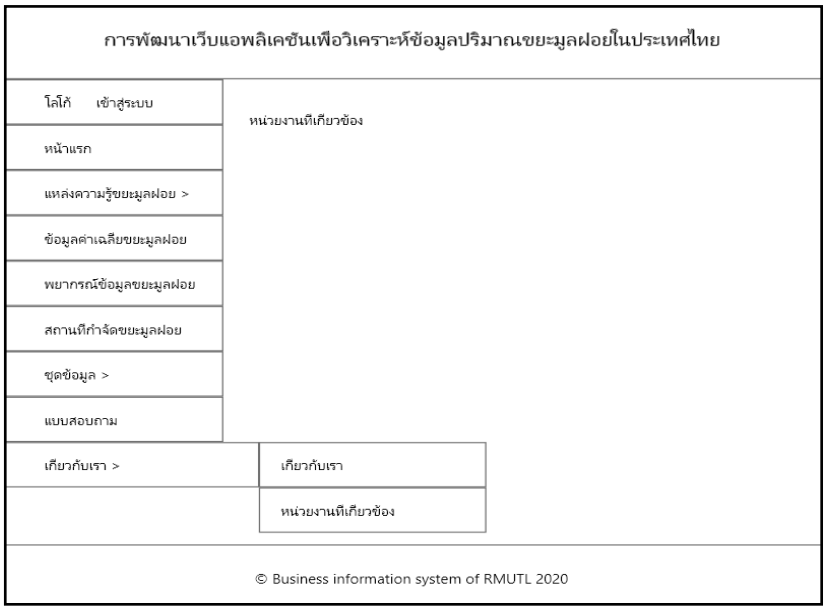
ภาพที่ 3.75 หน้าการแสดงแบบสอบถาม

14) หน้าการแสดงเกี่ยวกับเรา



ภาพที่ 3.76 หน้าการแสดงเกี่ยวกับเรา

15) หน้าการแสดงหน่วยงานที่เกี่ยวข้อง



ภาพที่ 3.77 หน้าการแสดงหน่วยงานที่เกี่ยวข้อง

3.7 บทสรุป

จากขั้นตอนการดำเนินงานผู้วิเคราะห์ได้แสดงกระบวนการวิเคราะห์การพยากรณ์ข้อมูลปริมาณขยะมูลฝอยในประเทศไทย ด้วยกระบวนการ Data Flow Diagram กระบวนการทำนอร์มัลไลเซชัน (Normalization) และวิเคราะห์ข้อมูล CRISP–DM มาใช้ในการรวบรวมข้อมูลวิเคราะห์ข้อมูลที่เกี่ยวข้องกับงาน เพื่อให้ได้สารสนเทศของงานอย่างเพียงพอในการนำไปใช้ประโยชน์ โดยผู้วิเคราะห์ได้วิเคราะห์ข้อมูลเป็นหลักสำคัญที่จะช่วยให้เข้าใจในงานแต่ละส่วน จากการวิเคราะห์ข้อมูลด้วยกระบวนการวิเคราะห์ข้อมูล CRISP–DM รวมถึงการสร้างโมเดล Time series analysis โดยเลือกใช้โปรแกรม Tableau Public ในการพยากรณ์ค่าปริมาณขยะมูลฝอยในประเทศไทยล่วงหน้าของปีถัดไปในช่วงเวลาเดียวกัน และนำผลการพยากรณ์ที่ได้ไปประเมินประสิทธิภาพของโมเดล ด้วยวิธีการคำนวณหาค่าเฉลี่ยเปอร์เซ็นต์ความคลาดเคลื่อนสมบูรณ์ (MAPE) ที่ต่ำที่สุด ผู้วิเคราะห์ได้นำข้อมูลสารสนเทศมาทำการแสดงผลและเผยแพร่ข้อมูลสารสนเทศผ่านเว็บแอปพลิเคชันจะพัฒนาโดยใช้ภาษา HTML, PHP, JavaScript และชุดคำสั่ง CSS3 เพื่อนำเข้าวิเคราะห์ข้อมูลด้วยโมเดลที่เลือกใช้และผ่านการทดสอบประสิทธิภาพของการวิเคราะห์ข้อมูลแล้ว และนำเสนอข้อมูลภาพด้วยเฟรมเวิร์คการแสดงข้อมูลภาพชื่อ highcharts.js เพื่อสร้างกราฟและ dashboard สำหรับแสดงผลข้อมูลภาพจากข้อมูลที่ผ่านการวิเคราะห์แล้วภายในเว็บแอปพลิเคชันที่พัฒนาขึ้น

บทที่ 4
ผลการดำเนินงาน

การจัดทำโครงงานเรื่อง การพัฒนาเว็บแอปพลิเคชันเพื่อการวิเคราะห์ข้อมูลปริมาณขยะมูลฝอยในประเทศไทย โดยมีวัตถุประสงค์ เพื่อการวิเคราะห์และพยากรณ์ข้อมูลปริมาณขยะมูลฝอยในประเทศไทย เพื่อเป็นข้อมูลความรู้ก่อให้เกิดประโยชน์ โดยนำข้อมูลที่ได้มาจัดแสดงผลเป็นสารสนเทศผ่านทางเว็บไซต์ มีรายละเอียดเกี่ยวกับข้อมูล ปริมาณขยะมูลฝอยที่เกี่ยวข้อง อีกทั้งยังมีข้อมูลปริมาณขยะมูลฝอยย้อนหลังของแต่ละภาค ผู้วิเคราะห์ได้ทำการวิเคราะห์ข้อมูลตามวัตถุประสงค์ที่ได้กำหนดไว้จนสำเร็จลุล่วงตามเป้าหมายและมีผลการ ดำเนินงานดังนี้

4.1 ผลการดำเนินงาน

จากการศึกษาและจัดทำการวิเคราะห์ข้อมูลปริมาณขยะมูลฝอยจากกระทรวงทรัพยากรธรรมชาติและสิ่งแวดล้อม เพื่อเผยแพร่ข้อมูลบนเว็บไซต์ โดยมีวัตถุประสงค์เพื่อวิเคราะห์ข้อมูลปริมาณขยะมูลฝอยจากกระทรวงทรัพยากรธรรมชาติและสิ่งแวดล้อม เพื่อเผยแพร่ข้อมูลบนเว็บไซต์ ให้สามารถนำไปเผยแพร่ต่อสาธารณะชนเพื่อเป็นข้อมูลความรู้ก่อให้เกิดประโยชน์ โดยนำข้อมูลที่ได้มาจัดแสดงผลเป็นสารสนเทศผ่านเว็บไซต์ ซึ่งในส่วนการดำเนินงานพบว่า ค่าภายในข้อมูลนั้นมีการเปลี่ยนแปลงตามช่วงเวลาต่าง ๆ ตามฤดูกาลที่เปลี่ยนแปลงไป เพื่อให้ข้อมูลเป็นไปตามข้อเท็จจริงจึงต้องทำความเข้าใจและวิเคราะห์ข้อมูลโดย ดำเนินการรับข้อมูลด้วยกระบวนการวิเคราะห์ข้อมูลของ CRISP-DM หรือ (Cross Industry Standard Process for Data Mining) มาใช้กับข้อมูลปริมาณขยะมูลฝอยจากการศึกษาพบว่าเทคนิคการ อนุกรมเวลา (Time Series Analysis) ที่นำมาใช้งานเหมาะสมกับการวิเคราะห์ข้อมูลได้อย่างดีจึงใช้ เทคนิคทางเหมืองข้อมูล (Data Mining) ในรูปแบบของอนุกรมเวลา (Time series) ทำการพยากรณ์ ทำนายผลของปริมาณขยะมูลฝอยล่วงหน้า โดยทดสอบกับหลายโปรแกรมจนได้ผลการพยากรณ์ที่ดี ที่สุด ด้วยการนำเสนอข้อมูลสารสนเทศในรูปแบบให้เห็นถึง Visualization ส่งผลทำให้เกิดความ เข้าใจในข้อมูลได้อย่างง่ายสะดวกรวดเร็วต่อการเรียนรู้และศึกษาข้อมูล และง่ายต่อการแสดงผล รวมถึงลดความผิดพลาดของข้อมูลและความซ้ำซ้อนจากข้อมูลจำนวนมาก ผู้วิเคราะห์ข้อมูลจึงได้ ทำการวิเคราะห์ข้อมูลสร้างเป็น

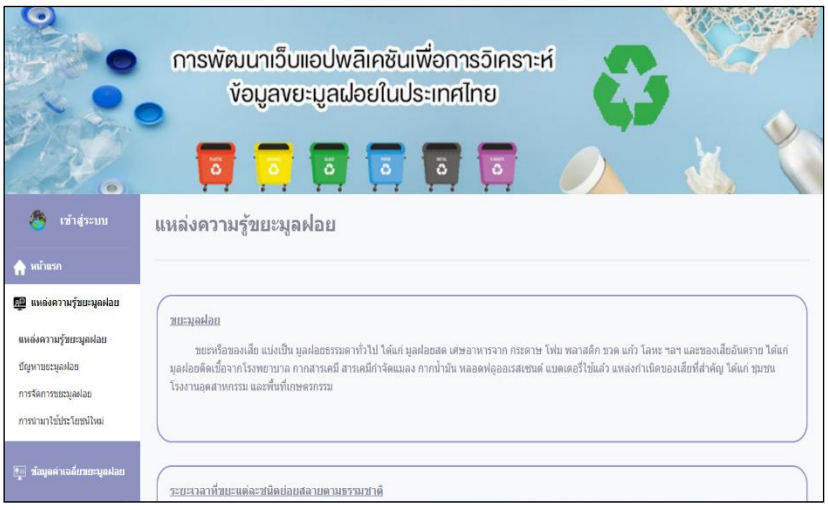
สารสนเทศแล้วนำมาพัฒนาจัดแสดงบนเว็บไซต์ตามวัตถุประสงค์ที่ ได้กำหนดไว้จนสำเร็จลุล่วงดังนี้

เว็บไซต์การพัฒนาเว็บแอปพลิเคชันเพื่อการวิเคราะห์ข้อมูลปริมาณขยะมูลฝอยในประเทศไทย

1) แถบเมนูสามารถใช้งานได้ความต้องการของการใช้งาน ซึ่งเป็นหน้าสำหรับผู้ใช้งานทั่วไปสามารถเลือกดู และศึกษาค้นคว้าเข้าไปใช้งานส่วนต่าง ๆ ของเว็บไซต์

ภาพที่ 4.1 หน้าแรกของเว็บไซต์

2) หน้าเว็บไซต์แสดงแหล่งความรู้เกี่ยวกับขยะมูลฝอยในประเทศไทย ระยะเวลาที่ขยะแต่ละชนิดย่อยสลายตามธรรมชาติ และชนิดและภาชนะรองรับขยะมูลฝอย



ภาพที่ 4.2 หน้าแหล่งความรู้ขยะมูลฝอย

2.1) หน้าเว็บไซต์แสดงแหล่งความรู้เกี่ยวกับปัญหาขยะมูลฝอยในประเทศไทย อธิบายสาเหตุและปัญหาต่าง ๆ ที่ทำให้เกิดขยะมูลฝอย

ภาพที่ 4.3 หน้าปัญหาขยะมูล

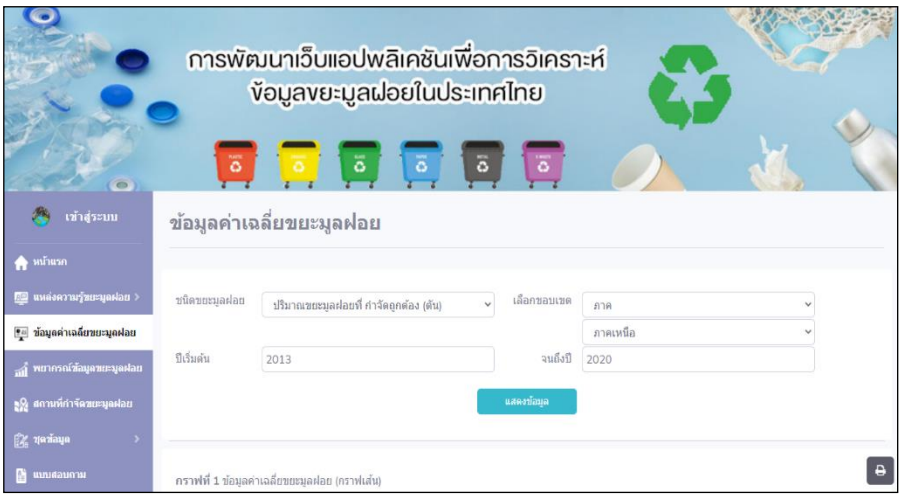
2.2) หน้าเว็บไซต์แสดงแหล่งความรู้เกี่ยวกับการจัดการขยะมูลฝอยในประเทศไทย แนวทางจัดการขยะมูลฝอย และวิธีการจัดการขยะแบบไหนถูกต้อง

ภาพที่ 4.4 หน้าการจัดการขยะมูลฝอย

2.3) หน้าเว็บไซต์แสดงแหล่งความรู้เกี่ยวกับการการลด และการนำขยะมูลฝอยกลับมาใช้ประโยชน์ใหม่ในประเทศไทย

ภาพที่ 4.5 หน้าการนำมาใช้ประโยชน์ใหม่

3) หน้าเว็บไซต์แสดงข้อมูลปริมาณขยะมูลฝอยย้อนหลังของแต่ละภาค ทั้งหมด 4 ภาค ได้แก่ ภาคเหนือ ภาคตะวันออกเฉียงเหนือ ภาคกลาง ภาคใต้ มีการอธิบายกราฟเส้นและกราฟแท่ง



ภาพที่ 4.6 หน้าข้อมูลค่าเฉลี่ยขยะมูลฝอย

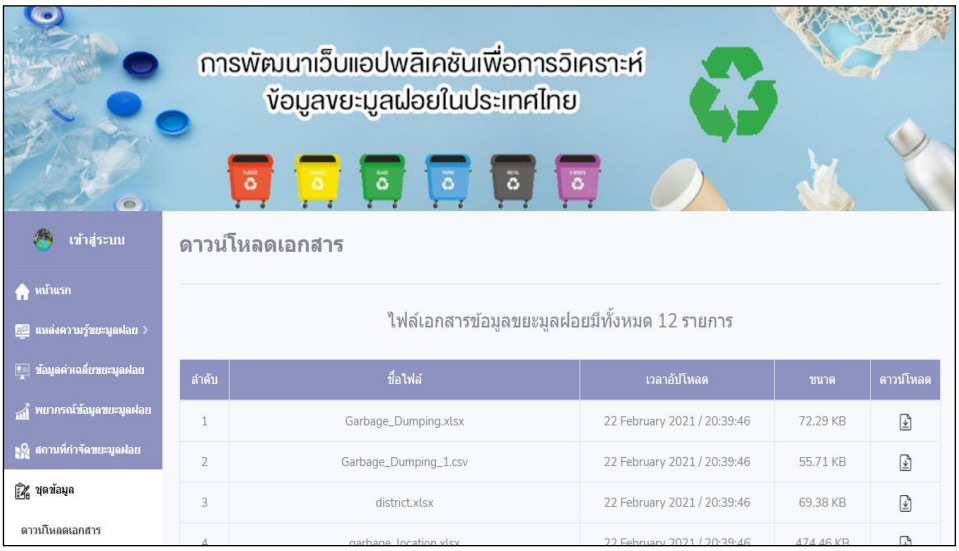
4) หน้าเว็บไซต์แสดงผลการวิเคราะห์ข้อมูลปริมาณขยะมูลฝอยในประเทศไทย ผ่านโปรแกรม Tableau Public เพื่อให้ได้ผลลัพธ์ต่ำสุด ใช้ในการวิเคราะห์ข้อมูลปริมาณขยะมูลฝอยในประเทศไทย โดยแสดงแต่ละประเภทของการกำจัดขยะมูลฝอย

ภาพที่ 4.7 หน้าพยากรณ์ข้อมูลขยะมูลฝอย

5) หน้าเว็บไซต์แสดงสถานที่กำจัดขยะมูลฝอย โดยแสดงแผนที่ในประเทศไทย และตารางแสดงข้อมูลสถานที่กำจัดขยะมูลฝอย

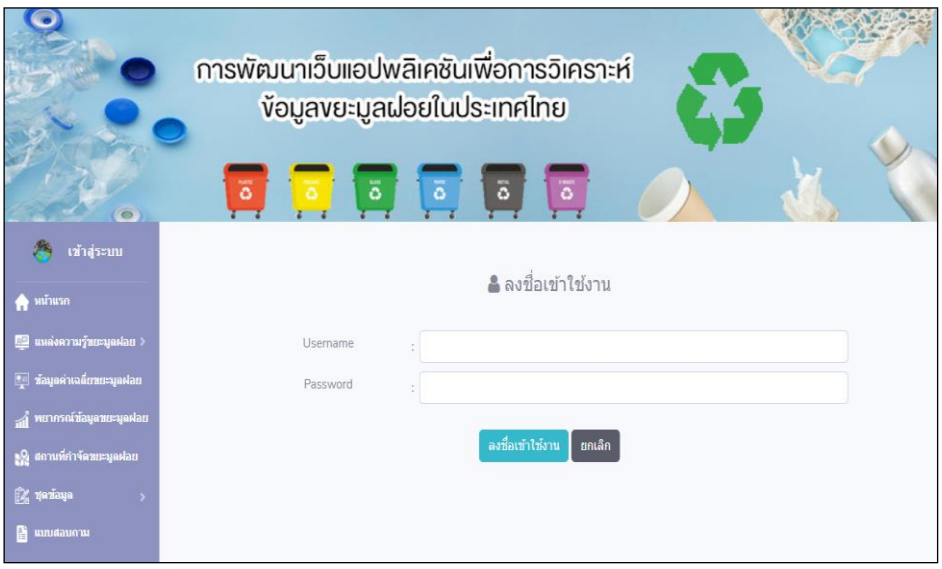
ภาพที่ 4.8 หน้าสถานที่กำจัดขยะมูลฝอย

6) หน้าเว็บไซต์แสดงชุดข้อมูลดาวน์โหลดเอกสาร ซึ่งเป็นหน้าสำหรับผู้ใช้งานทั่วไปสามารถดาวน์โหลดชุดข้อมูลเอกสารได้



ภาพที่ 4.9 หน้าดาวน์โหลดเอกสาร

6.1) หน้าเว็บไซต์การเข้าสู่ระบบ สำหรับผู้ดูแลเว็บไซต์ อัปโหลดข้อมูล สำหรับผู้ดูแลเว็บไซต์ที่สามารถ เพิ่ม ลบ แก้ไข ข้อมูล ได้



ภาพที่ 4.10 หน้าเข้าสู่ระบบ

6.2) หน้าเว็บไซต์แบบสอบถามความพึงพอใจเกี่ยวกับเว็บไซต์ และเนื้อหาของเว็บไซต์ โดยอ้างอิงคำถามแบบสอบถามจากเว็บ SurveyCan

ภาพที่ 4.11 หน้าเว็บไซต์แบบสอบถามความพึงพอใจของเว็บไซต์

7) หน้าเว็บไซต์เกี่ยวกับเราจะแสดงข้อมูลเกี่ยวกับเว็บไซต์ วัตถุประสงค์ เป้าหมายของเว็บไซต์ และคณะผู้จัดทำ

ภาพที่ 4.12 หน้าเว็บไซต์ข้อมูลเกี่ยวกับเว็บไซต์

4.2 การอธิปรายผล

4.2.1 อภิปรายการดำเนินโครงงาน จากวัตถุประสงค์เพื่อการวิเคราะห์และพยากรณ์ข้อมูลปริมาณขยะมูลฝอย ผู้วิเคราะห์ข้อมูลได้ทำการศึกษาถึงปัญหาเก็บรวบรวมข้อมูล และจัดการกับข้อมูลให้อยู่ในรูปแบบที่พร้อมต่อการวิเคราะห์เพื่อใช้ในการพยากรณ์วิเคราะห์ผล จากผลดำเนินการพบว่า

1) ข้อมูลปริมาณขยะมูลฝอยที่ได้จากเว็บไซต์ กระทรวงทรัพยากรธรรมชาติและสิ่งแวดล้อม ชุดข้อมูลมีจำนวนมากทำให้การศึกษาและการวิเคราะห์ข้อมูลใช้เวลามาก

2) ข้อมูลที่เผยแพร่แก่บุคคลภายนอกนั้นยังขาดความชัดเจนและติดตามผลย้อนหลัง ส่งผลให้กับการรับรู้และตระหนักถึงปัญหาปริมาณขยะมูลฝอยนั้นยังมีประสิทธิภาพไม่เพียงพอ โดยการวิเคราะห์ข้อมูลได้ดำเนินการแก้ไขปัญหาดังกล่าว ผู้วิเคราะห์ข้อมูลได้ดำเนินการ กับข้อมูลในการใช้การวิเคราะห์ข้อมูลตามกระบวนการของกระบวนการวิเคราะห์ข้อมูลของ CRISP–DM หรือ (Cross Industry Standard Process for Data Mining) มาใช้กับข้อมูลปริมาณขยะมูลฝอย ซึ่งผู้วิเคราะห์ได้ดำเนินการตามกระบวน CRISP–DM หรือ (Cross Industry Standard Process for Data Mining) ดังนี้

1) ผู้วิเคราะห์ทำความเข้าใจกระบวนการทางข้อมูล และรับฟังปัญหารวมถึงการมองหาปัญหาจากเรื่องต่าง ๆ เพื่อที่นำผลการวิเคราะห์ข้อมูลไปใช้ให้เกิดประโยชน์ และวางแผนในการดำเนินการโดยการวิเคราะห์ข้อมูลของข้อมูลปริมาณขยะมูลฝอยจากกระทรวงทรัพยากรธรรมชาติและสิ่งแวดล้อมจากช่วงเวลาที่มี ความถี่หรืออัตราที่เกิดขึ้นซ้ำ ๆ กัน เพื่อที่จะนำผลการวิเคราะห์ข้อมูลไปใช้ประโยชน์ และตามวัตถุประสงค์ที่จะนำไปสู่รูปแบบการวิเคราะห์ข้อมูล

2) ผู้วิเคราะห์ข้อมูลทำการจัดเก็บรวบรวมข้อมูล และพิจารณาข้อมูลที่ได้รับดูความถูกต้อง มีความน่าเชื่อถือโดยพิจารณาว่าจะใช้ข้อมูลทั้งหมดหรือจำเป็นต้องเลือกข้อมูลบางส่วน มาใช้ในการวิเคราะห์ให้สอดคล้องกับวัตถุประสงค์ที่กำหนดไว้ ผู้วิเคราะห์ได้สำรวจภาพรวมก่อน นำข้อมูลมาใช้ในการวิเคราะห์ข้อมูล จากเว็บไซต์กระทรวงทรัพยากรธรรมชาติและสิ่งแวดล้อม ซึ่งเป็นศูนย์กลางการให้บริการข้อมูล

3) ผู้วิเคราะห์ข้อมูลทำการแปลงข้อมูลที่ได้ทำการเก็บรวบรวม และทำการจัดการข้อมูลที่ได้ให้อยู่ในรูปแบบที่สามารถนำไปวิเคราะห์ในขั้นถัดไปได้ เพื่อลดความผิดพลาด

การทำ ข้อมูลให้ถูกต้อง โดยใช้กระบวนการ data cleaning เพื่อให้ได้ข้อมูลอยู่ในรูปแบบและมาตรฐาน เดียวกัน และกำจัดข้อมูลเสียออก

4) ผู้วิเคราะห์ข้อมูลจะวิเคราะห์ด้วยเทคนิคทางสถิติ ทำการเรียนรู้จากข้อมูลเดิมของปริมาณขยะมูลฝอยเพื่อให้ได้คำตอบที่ดีที่สุด การเปรียบเทียบความคลาดเคลื่อนของวิธีการพยากรณ์ โดยนำโมเดล Time series analysis มาช่วยในการหาค่าความคลาดเคลื่อนของการพยากรณ์ในวิธี ต่าง ๆ มาเปรียบเทียบตัวโปรแกรมว่าโปรแกรมไหนให้ความถูกต้องแม่นยำมากที่สุดนำมาใช้ในการ พยากรณ์การเกิดขยะมูลฝอย ผู้วิเคราะห์ข้อมูลได้เลือกใช้เทคนิครูปแบบอนุกรมเวลา Time series Models ด้วยชุดข้อมูลปริมาณขยะมูลฝอยรายปีของแต่ละภูมิภาค

5) ผู้วิเคราะห์ได้ดำเนินการนำผลลัพธ์ที่ได้จากขั้นตอนก่อนหน้าไปวัดประสิทธิผลของ ผลลัพธ์ที่ได้โดยการนำไปตรวจสอบค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percentage Error : MAPE) จากการทดลองโมเดลเทคนิค Time Series Models ด้วยการวิเคราะห์อนุกรมเวลาของข้อมูล (Data mining time series analysis) ผ่านโปรแกรม Minitab และ โปรแกรม Tableau Public ให้ผลลัพธ์ค่าความคลาดเคลื่อน MAPE ออกมาต่างกัน ซึ่งโปรแกรม Tableau Public ให้ผลลัพธ์ค่าความคลาดเคลื่อนของข้อมูลต่ำที่สุด 9.3% โดยในที่นี้ โปรแกรมแต่ละโปรแกรมมีสูตรการคำนวณที่เฉพาะขึ้นอยู่กับข้อมูลหรือค่าว่างทั้งนี้ ผู้วิเคราะห์ข้อมูลพบว่าเทคนิคทางเหมือนข้อมูล Time series Forecasting ในโปรแกรม Tableau public ออกมาดีที่สุด

6) ผู้วิเคราะห์ข้อมูลนำผลองค์ความรู้ที่ได้เหล่านี้ไปนำเสนอข้อมูลแบบ visualization ด้วยการแสดงผลข้อมูลในรูปแบบของภาพโดยใช้โปรแกรม Tableau Public ผู้วิเคราะห์ นำผลลัพธ์ที่ได้มาจัดแสดงแบบการนำเสนอออกเป็นภาพ จินตภาพ (visualization) ส่งผลทำให้เกิด ความเข้าใจในข้อมูลได้ง่าย สะดวก ลดขั้นตอนการทำงานและตอบสนองต่อความต้องการของผู้ใช้ สามารถกรองและเจาะลึกถึงรายละเอียดของข้อมูลได้ เพิ่มประสิทธิภาพให้กับการศึกษา และค้นคว้าข้อมูลโดยจัดทำในรูปแบบเว็บไซต์ (Website) เพื่อให้ง่ายต่อการแสดงผลข้อมูล สารสนเทศ และเลือกใช้งานตามต้องการเพราะมีเมนูการใช้งานที่แบ่งแยกชัดเจน ทำให้การดำเนินงานมีประสิทธิภาพมากยิ่งขึ้นซึ่งตรงตามวัตถุประสงค์ และเป้าหมายหลักของ เว็บไซต์ ที่ต้องการเผยแพร่ข้อมูลสารสนเทศให้ผู้ใช้งาน

4.2.2 อภิปรายผลสำรวจจากแบบสอบถาม

โดยส่วนแสดงผลหน้าเว็บไซต์ได้มีส่วนของแบบสอบถามเพื่อให้ผู้ที่เข้ามาเยี่ยมชมเว็บไซต์สามารถตอบแบบสอบถามและแสดงความคิดเห็นเกี่ยวกับเว็บไซต์ได้ ผู้วิเคราะห์จึงได้ใช้ ประโยชน์จากส่วนนี้ในการประเมินผลการใช้งานของเว็บไซต์จากกลุ่มตัวอย่าง 52 คน จากการกรอกแบบสอบถามความพึงพอใจในการใช้งานเกี่ยวกับเว็บไซต์ บริการแหล่งความรู้ปริมาณขยะมูลฝอย ที่ผู้วิเคราะห์สร้างขึ้น

จากการทดสอบการทำงานและแบบสอบถาม ได้พบว่าเทคนิคการอนุกรมเวลา (Time Series Analysis) ที่นำมาใช้งานเหมาะสมกับการวิเคราะห์ข้อมูลปริมาณขยะมูลฝอยของเราได้เป็นอย่างดี สารสนเทศและเว็บไซต์ที่นำมาแสดง สำหรับการวิเคราะห์ข้อมูลปริมาณขยะมูลฝอยจาก กระทรวงทรัพยากรธรรมชาติและสิ่งแวดล้อม

ความรู้ด้านปริมาณขยะมูลฝอยจากกระทรวงทรัพยากรธรรมชาติและสิ่งแวดล้อม พบว่าสามารถนำไปใช้งานให้เกิดประโยชน์ได้จริง สามารถแสดงข้อมูลที่มีอยู่ให้เข้าถึงง่าย สะดวกต่อความเข้าใจ และนอกจากจะเป็นประโยชน์ต่อส่วนรวมแล้ว ยังเกิดประโยชน์ช่วยในการตัดสินใจที่เกิดจากการพยากรณ์ข้อมูลที่มีอยู่ทางผู้วิเคราะห์ข้อมูลคิดว่าหากมีโอกาสที่ได้พัฒนาข้อมูลเหล่านี้ให้มีมากยิ่งขึ้นไปอีกก็จะนำมาซึ่งประโยชน์ และองค์ความรู้อีกมากมายที่จะส่งเสริมการเตรียมตัวและเฝ้าระวังจากการเกิด ปริมาณขยะมูลฝอยให้ดียิ่งขึ้นต่อไป

ผู้จัดทำได้แบ่งการประเมินเป็น 3 ตอน ประกอบไปด้วย

1) ข้อมูลทั่วไปของผู้เข้าชมเว็บไซต์
2) เกี่ยวกับการวิเคราะห์ข้อมูลการพยากรณ์ และการออกแบบหน้าเว็บไซต์
3) ข้อเสนอแนะ

ในตอนที่ 1 ข้อมูลทั่วไปของผู้เข้าชมเว็บไซต์ ประกอบไปด้วย ข้อมูลทั่วไป ได้แก่ เพศ สถานะ ของ ผู้ตอบแบบสอบถาม โดยสามารถสรุปข้อมูล จากกลุ่มตัวอย่าง 52 คน แบ่งเป็น

เพศหญิงจำนวน 28 คน คิดเป็นร้อยละ 54%

เพศชายจำนวน 24 คน คิดเป็นร้อยละ 46%

อาจารย์จำนวน 7 คน คิดเป็นร้อยละ 14%

นักศึกษาจำนวน 24 คน คิดเป็นร้อยละ 46%

บุคคลทั่วไปจำนวน 21 คน คิดเป็นร้อยละ 40%

ในตอนที่ 2 เกี่ยวกับการวิเคราะห์ข้อมูลการพยากรณ์ และการออกแบบหน้าเว็บไซต์

1) การประเมินด้านเนื้อหาของเว็บไซต์
2) การประเมินด้านการออกแบบและการจัดรูปแบบ
3) การประเมินด้านด้านประโยชน์และการนำไปใช้

การประเมินความพึงพอใจของผู้ตอบแบบสอบในภาพรวมด้วยค่าเฉลี่ยคะแนนความพึงพอใจ โดยมีการคิดคะแนนและเกณฑ์ระดับความพึงพอใจดังนี้

ระดับความพึงพอใจ ควรปรับปรุง มีค่าคะแนน 1
ระดับความพึงพอใจ น้อย มีค่าคะแนน 2
ระดับความพึงพอใจ ปานกลาง มีค่าคะแนน 3
ระดับความพึงพอใจ ดี มีค่าคะแนน 4
ระดับความพึงพอใจ ดีมาก มีค่าคะแนน 5

เมื่อนำคำตอบของผู้ตอบแบบสอบถามมาแจกแจงความถี่และหาค่าเฉลี่ย แล้วกำหนดระดับค่าเฉลี่ยความพึงพอใจเป็นดังนี้

$$\text{ความกว้างของอันตรภาคชั้น} = \frac{\text{คะแนนสูงสุด} - \text{คะแนนต่ำสุด}}{\text{จำนวนชั้น}}$$

จากเกณฑ์ดังกล่าว สามารถแปลความหมายของความพึงพอใจได้ดังนี้

คะแนนค่าเฉลี่ย 1.00 – 1.80 หมายถึง ระดับความพึงพอใจ ควรปรับปรุง
คะแนนค่าเฉลี่ย 1.81 – 2.60 หมายถึง ระดับความพึงพอใจ น้อย
คะแนนค่าเฉลี่ย 2.61 – 3.40 หมายถึง ระดับความพึงพอใจ ปานกลาง
คะแนนค่าเฉลี่ย 3.41 – 4.20 หมายถึง ระดับความพึงพอใจ ดี
คะแนนค่าเฉลี่ย 4.21 – 5.00 หมายถึง ระดับความพึงพอใจ ดีมาก

ผู้วิจัยจึงนำค่าคะแนนคำนวณสถิติ คือ ค่าคะแนนเฉลี่ย ($\bar{x}$) และค่าส่วนเบี่ยงเบนมาตรฐาน (S.D.) ผลการประเมินแสดงรายละเอียดดังนี้ การประเมินด้านเนื้อหาแสดงดังตารางที่ 4.1 การประเมินด้าน การออกแบบและการจัดรูปแบบเว็บไซต์ แสดงดังตารางที่ 4.2 และการประเมินด้านประโยชน์ และการนำไปใช้แสดงดังตารางที่ 4.3

ตารางที่ 4.1 แสดงข้อมูลการประเมินผลด้านเนื้อหา

ลำดับ	หัวข้อคำถาม	จำนวนผู้ตอบแบบสอบถาม	คะแนนเฉลี่ย $(\overline{x})$	S.D.	ระดับความพึงพอใจ
1	การวิเคราะห์และพยากรณ์ข้อมูลมี ความชัดเจนน่าเชื่อถือ	52	4.28	0.61	ดีมาก
2	แสดงรายละเอียดโมเดลการพยากรณ์ได้ครบถ้วน	52	4.12	0.72	ดี
3	ข้อมูลที่นำเสนอบนเว็บไซต์มีขนาดและองค์ประกอบที่เหมาะสม	52	4.34	0.56	ดีมาก
4	การจัดลำดับเนื้อหาเป็นขั้นตอนมี ความต่อเนื่องอ่านแล้วเข้าใจ	52	4.3	0.65	ดีมาก
5	รูปแบบ Visualization มีความเหมาะสมกับข้อมูล	52	4.18	0.66	ดี
ผลรวม			4.24	0.64	ดีมาก

จากตารางที่ 4.1 พบว่า 1.) การวิเคราะห์และพยากรณ์ข้อมูลมีความชัดเจน น่าเชื่อถือ ค่าเฉลี่ยสูงสุด เท่ากับ 4.28 (S.D. = 0.61) อยู่ในเกณฑ์ระดับดีมาก 2.) แสดงรายละเอียดโมเดลการพยากรณ์ได้ครบถ้วน ค่าเฉลี่ยสูงสุด เท่ากับ 4.12 (S.D. = 0.72) อยู่ในเกณฑ์ระดับดี 3.) ข้อมูลที่นำเสนอบนเว็บไซต์มีขนาดและองค์ประกอบที่เหมาะสม ค่าเฉลี่ยสูงสุด เท่ากับ 4.34 (S.D. = 0.56) อยู่ในเกณฑ์ระดับดีมาก 4.) การจัดลำดับเนื้อหาเป็นขั้นตอน มีความต่อเนื่อง อ่านแล้วเข้าใจ ค่าเฉลี่ยสูงสุด เท่ากับ 4.3 (S.D. = 0.65) อยู่ในเกณฑ์ระดับดีมาก 5.) รูปแบบ Visualization มีความเหมาะสมกับข้อมูล ค่าเฉลี่ยสูงสุด เท่ากับ 4.18 (S.D. = 0.66) อยู่ในเกณฑ์ ระดับดี ตามลำดับ และผลรวมของการประเมินผลเว็บไซต์บริการแหล่งความรู้ข้อมูลปริมาณขยะมูลฝอยในประเทศไทยด้านเนื้อหา จากจำนวนผู้ตอบแบบสอบถามทั้งหมด 52 คน ค่าเฉลี่ยเท่ากับ 4.24 (S.D. = 0.64) อยู่ในเกณฑ์ระดับดีมาก

ตารางที่ 4.2 การประเมินด้านการออกแบบและการจัดรูปแบบเว็บไซต์

ลำดับ	หัวข้อคำถาม	จำนวนผู้ตอบแบบสอบถาม	คะแนนเฉลี่ย ($\bar{x}$)	S.D.	ระดับความพึงพอใจ
1	การจัดรูปแบบในเว็บไซต์ง่ายต่อการ อ่านและสะดวกต่อการใช้งาน	52	4.3	0.71	ดีมาก
2	เว็บไซต์มีความสวยงาม มีความ ทันสมัย น่าสนใจ	52	4.24	0.66	ดีมาก
3	สีสันในการออกแบบเว็บไซต์มีความ เหมาะสม	52	4.26	0.75	ดีมาก
4	สีพื้นหลังกับสีตัวอักษรมีความเหมาะสมต่อการอ่าน	52	4.4	0.70	ดีมาก
5	ขนาดตัวอักษร และรูปแบบตัวอักษร มีความสวยงามและอ่านได้ง่าย	52	4.26	0.66	ดีมาก
ผลรวม			4.30	1.30	ดีมาก

จากตารางที่ 4.2 พบว่า 1.) การจัดรูปแบบในเว็บไซต์ง่ายต่อการอ่านและสะดวกต่อการใช้งาน ค่าเฉลี่ยสูงสุด เท่ากับ 4.3 (S.D. = 0.71) อยู่ในเกณฑ์ระดับดีมาก 2.) เว็บไซต์มีความสวยงาม มีความทันสมัย น่าสนใจ ค่าเฉลี่ยสูงสุด เท่ากับ 4.24 (S.D. = 0.66) อยู่ในเกณฑ์ระดับดีมาก 3.) สีสันในการออกแบบเว็บไซต์มีความเหมาะสม ค่าเฉลี่ยสูงสุด เท่ากับ 4.26 (S.D. = 0.75) อยู่ในเกณฑ์ระดับดี 4.) สีพื้นหลังกับสีตัวอักษรมีความเหมาะสมต่อการอ่าน ค่าเฉลี่ยสูงสุด เท่ากับ 4.4 (S.D. = 0.70) อยู่ในเกณฑ์ระดับดีมาก 5.) ขนาดตัวอักษร และรูปแบบตัวอักษร มีความสวยงาม และอ่านได้ง่าย ค่าเฉลี่ยสูงสุด เท่ากับ 4.26 (S.D. = 0.66) อยู่ในเกณฑ์ระดับ ดีมาก ตามลำดับ และผลรวมการประเมินผลเว็บไซต์บริการแหล่งความรู้ข้อมูลปริมาณขยะมูลฝอยในประเทศไทยด้านการออกแบบและการจัดรูปแบบเว็บไซต์ จากจำนวนผู้ตอบแบบสอบถามทั้งหมด 52 คน ค่าเฉลี่ยเท่ากับ 4.30 (S.D. = 1.30) อยู่ในเกณฑ์ระดับดีมาก

ตารางที่ 4.3 การประเมินด้านประโยชน์และการนำไปใช้

ลำดับ	หัวข้อคำถาม	จำนวนผู้ตอบแบบสอบถาม	คะแนนเฉลี่ย ($\bar{x}$)	S.D.	ระดับความพึงพอใจ
1	เนื้อหามีรายละเอียดและประโยชน์ เพียงพอสามารถนำไปประยุกต์ใช้ได้	52	4.14	0.78	ดีมาก
2	สื่อ ข่าวสาร และการพยากรณ์เป็น ประโยชน์และสามารถนำไปใช้ได้	52	4.24	0.77	ดีมาก
3	สามารถใช้เป็นแหล่งความรู้และ แนวทางในการพัฒนาต่อไปได้	52	4.16	0.58	ดีมาก
4	แหล่งข้อมูลที่ตรงกับความต้องการ ของผู้ใช้งาน	52	4.34	0.66	ดีมาก
ผลรวม			4.22	0.70	ดีมาก

จากตารางที่ 4.3 พบว่า 1.) เนื้อหามีรายละเอียดและประโยชน์เพียงพอสามารถนำไปประยุกต์ใช้ได้ ค่าเฉลี่ยสูงสุด เท่ากับ 4.14 (S.D. = 0.78) อยู่ในเกณฑ์ระดับดีมาก 2.) สื่อ ข่าวสาร และการพยากรณ์ เป็นประโยชน์และสามารถนำไปใช้ได้ ค่าเฉลี่ยสูงสุด เท่ากับ 4.24 (S.D. = 0.77) อยู่ในเกณฑ์ระดับดีมาก 3) สามารถใช้เป็นแหล่งความรู้และแนวทางในการ พัฒนาต่อไปได้ ค่าเฉลี่ยสูงสุด เท่ากับ 4.16 (S.D. = 0.58) อยู่ในเกณฑ์ระดับดีมาก 4.) แหล่งข้อมูลที่ตรงกับความต้องการของผู้ใช้งาน ค่าเฉลี่ยสูงสุด เท่ากับ 4.34 (S.D. = 0.66) อยู่ในเกณฑ์ระดับดี ตามลำดับ และผลรวมการประเมินผลเว็บไซต์บริการแหล่งความรู้ข้อมูล ปริมาณน้ำฝนด้านการประเมินด้านประโยชน์และการนำไปใช้ จากจำนวนผู้ตอบแบบสอบถาม ทั้งหมด 52 คน ค่าเฉลี่ยเท่ากับ 4.22 (S.D. = 0.70) อยู่ในเกณฑ์ระดับดีมาก

4.3 บทสรุป

เนื้อหาในบทที่ 4 เป็นผลจาการดำเนินงานโครงงานในข้างต้นเนื้อหาในส่วนของผลการดำเนินงานการวิเคราะห์ข้อมูลข้อมูลประมาณขยะมูลฝอย ผู้วิเคราะห์ได้ทดสอบโปรแกรมที่ใช้ในการพยากรณ์และเลือกผลลัพธ์ที่ดีที่สุดนำมาจัดแสดงในรูปแบบการนำเสนอออกเป็นภาพผ่านบนหน้า เว็บไซต์ โดยมีส่วนแสดงการพยากรณ์ปริมาณขยะมูลฝอยล่วงหน้า ส่วนของการแสดงข้อมูลสารสนเทศปริมาณขยะมูลฝอย ส่วนของการแสดงแหล่งความรู้เกี่ยวกับการกำจัดขยะมูลฝอยที่ถูกต้อง, การกำจัดไม่ถูกต้อง และการนำไปใช้ประโยชน์ ส่วนการติดต่อสื่อสารกับผู้วิเคราะห์ข้อมูลและเว็บไซต์ ส่วนแสดงหน่วยงานที่เกี่ยวข้อง ส่วนแบบสอบถามเพื่อให้ผู้ที่เข้ามาเยี่ยมชมเว็บไซต์สามารถทำแบบสอบถามและแสดงความคิดเกี่ยวกับเว็บไซต์ได้ ส่วนของการดาวน์โหลดชุดข้อมูลที่สามารถให้ผู้เยี่ยมชมเว็บไซต์เข้ามาดาวน์โหลดชุดข้อมูลเพื่อไปพัฒนาต่อไปได้ และส่วนของการอัปโหลดข้อมูลผู้วิเคราะห์ข้อมูลได้ทำการวิเคราะห์ และนำผลที่ได้จากการวิเคราะห์มาแสดงผ่านหน้าเว็บไซต์ตามวัตถุประสงค์ที่ได้กำหนดไว้จนสำเร็จลุล่วงตาม เป้าหมายเพื่อที่จะได้นำผลการดำเนินงานไปสรุปและเป็นข้อเสนอแนะในบทที่ 5 ต่อไป

บทที่ 5
สรุปและข้อเสนอแนะ

การวิเคราะห์ข้อมูลปริมาณขยะมูลฝอยจากกระทรวงทรัพยากรธรรมชาติและสิ่งแวดล้อม ทางผู้วิเคราะห์ข้อมูลได้ทำการศึกษาและวิเคราะห์ข้อมูลจากแหล่งข้อมูลที่ได้มาโดยนำมาเป็นวิทยาการทั้งทางด้านโปรแกรมต่าง ๆ โดยมีการใช้ในการวิเคราะห์ข้อมูลตามกระบวนการของกระบวนการวิเคราะห์ข้อมูล ของ CRISP–DM หรือ (Cross Industry Standard Process for Data Mining), กระบวนการ นอมัลไลเซชัน หรือ (Normalization) และกระบวนการ E–R Diagram

5.1 บทสรุปผลโครงงาน

ผู้วิเคราะห์ข้อมูลดำเนินการรวบรวมข้อมูลจากแหล่งข้อมูลต่าง ๆ วิเคราะห์ข้อมูลที่ได้มานำไปหาองค์ความรู้ที่ซ่อนอยู่ภายใต้เทคนิคทางเหมืองข้อมูล (Data Mining) พบว่าเทคนิคการวิเคราะห์อนุกรมเวลา (Time Series Analysis) ที่นำมาใช้งานเหมาะสมและสอดคล้องกับข้อมูลได้ เป็นอย่างดี จึงใช้เทคนิคทางเหมืองข้อมูล (Data Mining) ทำการพยากรณ์ปริมาณขยะมูลฝอย ล่วงหน้าโดยทดสอบสองโปรแกรมเพื่อวัดประสิทธิภาพของโมเดลในแต่ละโปรแกรมแล้วเลือก โปรแกรมที่ให้ผลลัพธ์ที่ดีที่สุดนำมาสร้างเป็นสารสนเทศ และจัดแสดงผลผ่านทางเว็บไซต์ โดยผู้วิเคราะห์ได้ทำการศึกษาความรู้พื้นฐานทางด้านภาษาที่ใช้ในการพัฒนาระบบ คือ ภาษาเอชทีเอ็ม แอล (HTML) พีเอชพี (PHP), ซีเอสเอส (CSS3) และจาวาสคริปต์ (javascript) เมื่อทำการศึกษาความรู้พื้นฐานในภาษาแล้วจึง ทำออกแบบหน้าต่างของเว็บไซต์

ผู้วิเคราะห์ข้อมูลดำเนินการรวบรวมแบบสอบถามที่ผู้เข้าชมเว็บไซต์ที่ตอบแบบสอบถามและแสดงความคิดเห็นเกี่ยวกับเว็บไซต์ ผู้วิเคราะห์จึงได้ใช้ประโยชน์จากส่วนนี้ในการประเมินผล การใช้งานของเว็บไซต์จากกลุ่มตัวอย่าง 52 คนจากการกรอกแบบสอบถามความพึงพอใจในการ ใช้งานเกี่ยวกับเว็บไซต์ บริการแหล่งความรู้ปริมาณขยะมูลฝอยในประเทศไทยที่ผู้วิเคราะห์สร้างขึ้น โดยแบ่งออกเป็น 3 ด้าน ดังนี้ ผลรวมของการประเมินผลเว็บไซต์บริการแหล่งความรู้ข้อมูลปริมาณขยะมูลฝอยในประเทศไทย ด้านเนื้อหา ค่าเฉลี่ยเท่ากับ 4.24 (S.D. = 0.64) อยู่ในเกณฑ์ระดับดีมาก ผลรวมการประเมินผลเว็บไซต์บริการ แหล่งความรู้ข้อมูลปริมาณขยะมูลฝอยในประเทศไทยด้านการออกแบบและการจัดรูปแบบเว็บไซต์

ค่าเฉลี่ยเท่ากับ 4.30 (S.D. = 1.30) อยู่ในเกณฑ์ระดับดีมาก ผลรวมการประเมินผลเว็บไซต์บริการแหล่งความรู้ ข้อมูลปริมาณขยะมูลฝอยในประเทศไทยด้านการประเมินด้านประโยชน์และการนำไปใช้ ค่าเฉลี่ยเท่ากับ 4.22 (S.D. = 0.70) อยู่ในเกณฑ์ระดับดีมาก

5.2 ข้อจำกัดของเว็บไซต์

5.2.1 ผู้ใช้งานไม่สามารถปรับเปลี่ยนรูปแบบของกราฟที่แสดงข้อมูลปริมาณขยะมูลฝอยในประเทศไทยของเว็บไซต์ได้ เช่น เป็นกราฟวงกลม กราฟแท่ง หรือแผนภูมิโดนัท

5.2.2 ผู้ใช้งานไม่สามารถรับรู้การแจ้งเตือน การปรับปรุงข้อมูลภายในเว็บไซต์ได้

5.2.3 ผู้ใช้งานไม่สามารถ เพิ่ม ลบ แก้ไข และอัปโหลดข้อมูลภายในเว็บไซต์ได้

5.3 ปัญหาและอุปสรรคของโครงงาน

5.3.1 ไม่มีความเข้าใจเกี่ยวกับ Dataset ข้อมูลมีความซับซ้อนจึงต้องใช้เวลาในการศึกษาทำให้เกิดความล่าช้าในการวิเคราะห์และออกแบบ

5.3.2 มีแหล่งข้อมูลที่ใช้ในการวิเคราะห์เป็นข้อมูลที่ไม่สมบูรณ์ทำให้มีค่าซ้ำซ้อน หรือจัดอยู่ในรูปแบบที่ไม่สามารถนำไปวิเคราะห์ได้ ทำให้การเตรียมข้อมูลเพิ่มมากขึ้น

5.3.3 มีโปรแกรมบางประเภทที่ไม่สามารถใช้ฟรีได้ จึงทำให้ผู้วิเคราะห์ได้ทำการเลือกใช้โปรแกรมฟรีที่บางครั้งที่ไม่เหมาะสมตรงตามข้อมูลที่กำหนดไว้

5.3.4 มีการสร้างงานมัลติมีเดียเป็นงานที่ละเอียดและซับซ้อนทำให้ต้องใช้เวลาในการจัดทำที่นานพอสมควร

5.4 ข้อเสนอแนะ

การวิเคราะห์ข้อมูลปริมาณขยะมูลฝอยในประเทศไทยเพื่อเปรียบเทียบการพยากรณ์ที่วิเคราะห์เสร็จแล้วนี้แม้ จะสามารถทำงานได้ถูกต้องตรงตามวัตถุประสงค์และขอบเขตที่ตั้งไว้ แต่ก็ยังมีข้อจำกัดหลายประการ ซึ่งหากจะวิเคราะห์ข้อมูลให้เว็บไซต์สามารถทำงานได้อย่างมีประสิทธิภาพ และประสิทธิผลมากขึ้น ผู้วิเคราะห์ข้อมูลควรจะต้องปรับปรุงในประเด็นต่าง ๆ ต่อไปนี้

5.4.1 ควรมีการเพิ่มชุดข้อมูลให้มากขึ้น เพื่อการวิเคราะห์ข้อมูลที่มีความแม่นยำมากยิ่งขึ้น ลดความคลาดเคลื่อนของข้อมูล โดยหากชุดข้อมูลนี้ได้ศึกษาร่วมกับข้อมูลประชากรของ

ประเทศไทยของกรมการปกครอง ก็อาจสร้างความน่าเชื่อถือ หรือการพยากรณ์ทำนายผลที่มีความแม่นยำมากขึ้น

5.4.2 ควรมีการพัฒนาด้านการแสดงผลในด้านต่าง ๆ เพิ่มเติม เช่น ข้อมูลจำนวนประชากร เพื่อให้ครอบคลุมปัจจัยในการเกิดฝน และเพิ่มประสิทธิภาพในการนำเสนอข้อมูล

บรรณานุกรม

กฤติยา พุตติ (2560) **การคาดการณ์การเปลี่ยนแปลงของปริมาณขยะมูลฝอยจากการขยายตัวของอาคารที่อยู่อาศัยในอนาคต : กรณีศึกษาเทศบาลนครนนทบุรี.** สืบค้น 4 พฤศจิกายน 2556, จาก : https://li01.tcithaijo.org/index.php

สุรศักดิ์ โอสถิตพร (2558) **รูปแบบการจัดการขยะมูลฝอยที่เหมาะสมกับเทศบาลตำบลดอยสะเก็ด อำเภอดอยสะเก็ด จังหวัดเชียงใหม่.** สืบค้น 8 พฤศจิกายน 2563 , จาก :https://dric.nrct.go.th/index.php

จตุรงค์ พะยอมแย้ม (2559) **การคาดการณ์ปริมาณขยะ, การขยายตัวของเมือง.** สืบค้น 13 พฤศจิกายน 2563 , จาก :https://li01.tcithaijo.org/index.php

กรมควบคุมมลพิษ (2558). **ข้อมูลปริมาณขยะมูลฝอยของประเทศไทย.** กรมควบคุมมลพิษ. สืบค้นเมื่อ 1 พฤจิกายน 2563 , จาก : https://thaimsw.pcd.go.th/search.php

กรมควบคุมมลพิษ (2558). **ข้อมูลสถานที่กำจัดปริมาณขยะมูลฝอยในประเทศไทย.** กรมควบคุมมลพิษ. สืบค้นเมื่อ 10 ตุลาคม 2563 , จาก : https://thaimsw.pcd.go.th/search.php

ชาคริต กุลไกรศรี (2013). **นอร์มัลไลเซชัน (Normalization).** สืบค้น 16 มกราคม 2564 , จาก : https://1th.me/97HQ4

สุพรรณี อึ้งปัญสัตวงค์. (2551). **การวิเคราะห์ข้อมูลอนุกรมเวลา.** สืบค้น 16 มกราคม 2564 , จาก : https://1th.me/YSkY3

ภาคผนวก ก
คู่มือการใช้งานเว็บไซต์

ภาคผนวก ก
คู่มือการใช้งานเว็บไซต์

จากการดำเนินงานการวิเคราะห์ข้อมูลปริมาณขยะมูลฝอยจากกระทรวงทรัพยากรธรรมชาติและสิ่งแวดล้อม เพื่อนำเสนอข้อมูลสารสนเทศผ่านเว็บไซต์ มีการใช้งานทั้งหมด 9 ส่วน คือ ส่วนการแบ่งภูมิภาคของประเทศไทย ส่วนของแหล่งความรู้ ส่วนการแสดงข้อมูลปริมาณขยะมูลฝอยย้อนหลังของแต่ละภาค ส่วนของการใช้โมเดลในการพยากรณ์ ขั้นตอนในการพยากรณ์ ส่วนของผลลัพธ์ในการพยากรณ์ ปริมาณขยะมูลฝอยล่วงหน้า ของแต่ละภาค ส่วนของหน่วยงานที่เกี่ยวข้อง ส่วนของการอัปโหลดข้อมูล ส่วนแบบสอบถามความพึงพอใจของผู้เข้าชมเว็บไซต์ และส่วนของเกี่ยวกับเว็บไซต์กับผู้จัดทำโครงงาน

ก.1 คู่มือการใช้งานสำหรับผู้เยี่ยมชมหน้าแรก

1) หน้าเข้าสู่ระบบเว็บไซต์ จะแบ่งส่วนการใช้งานออกเป็น 9 ส่วน ดังภาพ ก.1

ภาพที่ ก.1 หน้าเข้าสู่ระบบเว็บไซต์

หมายเลข 1 หน้าเข้าสู่ระบบ และการออกระบบของผู้ดูแลเว็บไซต์

หมายเลข 2 หน้าแรกของเว็บไซต์

หมายเลข 3 หน้าเกี่ยวกับแหล่งความรู้ขยะมูลฝอย

หมายเลข 4 หน้าเกี่ยวกับข้อมูลค่าเฉลี่ยขยะมูลฝอย

หมายเลข 5 หน้าเกี่ยวกับการพยากรณ์ข้อมูลขยะมูลฝอย

หมายเลข 6 หน้าเกี่ยวกับสถานที่กำจัดขยะมูลฝอย

หมายเลข 7 หน้าเกี่ยวกับชุดข้อมูลขยะมูลฝอย

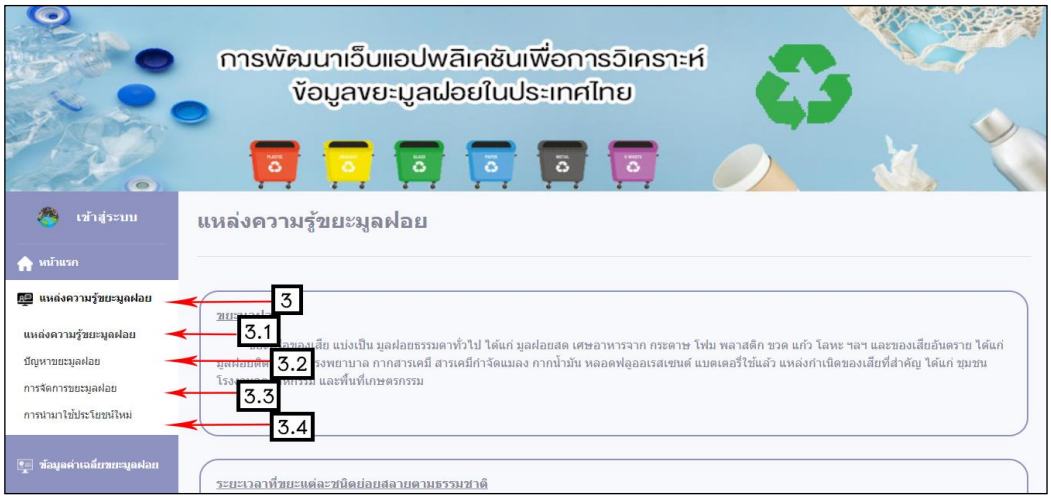
หมายเลข 8 หน้าแบบสอบถามความพึงพอใจของผู้เข้าชมเว็บไซต์

หมายเลข 9 หน้าของเกี่ยวกับเว็บไซต์กับผู้จัดทำโครงงาน

2) หน้าแรกของเว็บไซต์ ดังภาพที่ ก.2

ภาพที่ ก.2 หน้าแรกของเว็บไซต์

3) หน้าแหล่งความรู้เกี่ยวกับขยะมูลฝอย ดังภาพที่ ก.3



ภาพที่ ก.3 หน้าแหล่งความรู้เกี่ยวกับขยะมูลฝอย

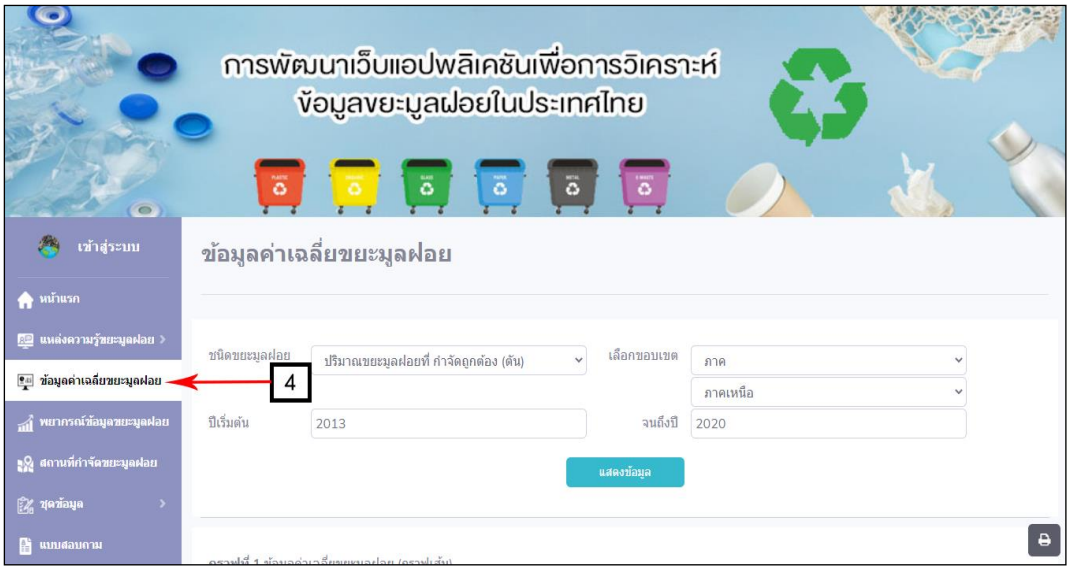
หมายเลข 3.1 ส่วนของแหล่งความรู้เกี่ยวกับขยะมูลฝอย

หมายเลข 3.2 ส่วนของแหล่งความรู้เกี่ยวกับปัญหาขยะมูลฝอย

หมายเลข 3.3 ส่วนของแหล่งความรู้เกี่ยวกับการจัดการขยะมูลฝอย

หมายเลข 3.4 ส่วนของแหล่งความรู้เกี่ยวกับการนำมาใช้ประโยชน์ใหม่

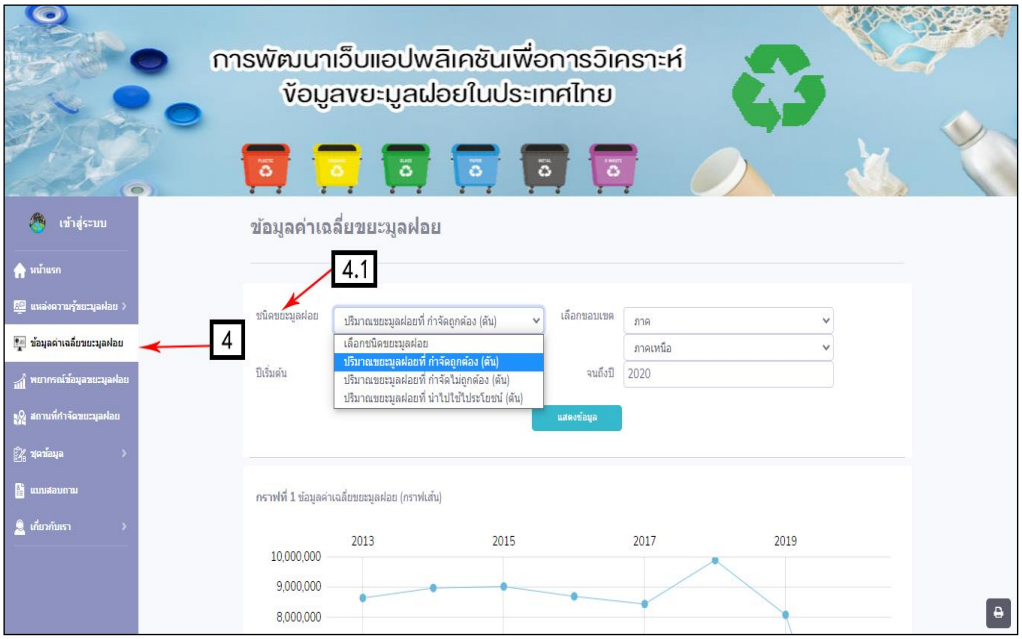
4) หน้าข้อมูลค่าเฉลี่ยขยะมูลฝอย ดังภาพที่ ก.4



ภาพที่ ก.4 หน้าข้อมูลค่าเฉลี่ยขยะมูลฝอย

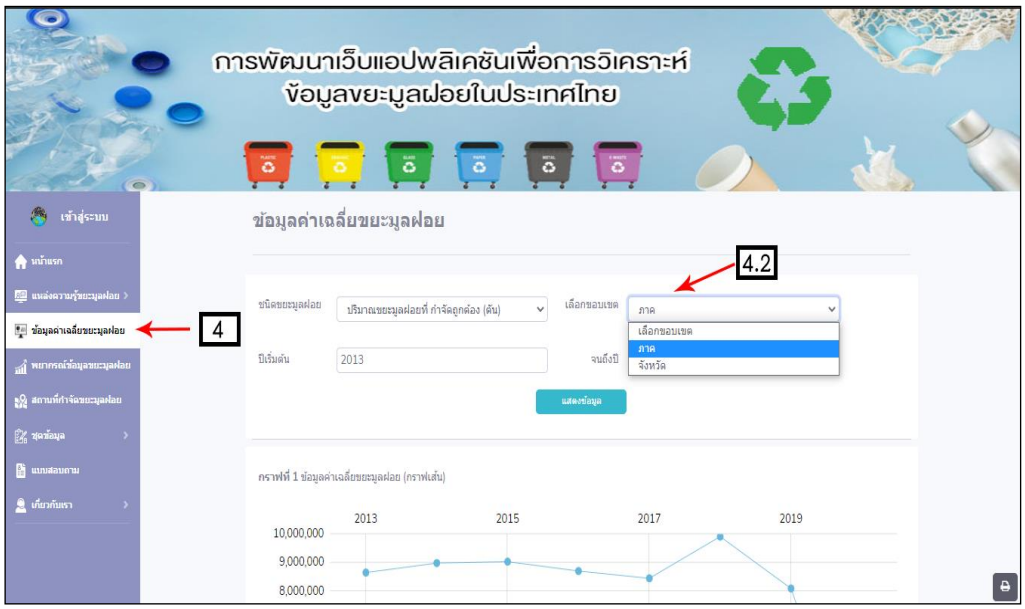
4.1 หน้าการเลือกแสดงชนิดขยะมูลฝอยที่จะแสดงข้อมูล โดยสามารถเลือกชนิดของข้อมูลขยะมูลฝอยได้

- ปริมาณยะมูลฝอยที่ กำจัดถูกต้อง(ตัน)
- ปริมาณยะมูลฝอยที่ กำจัดไม่ถูกต้อง(ตัน)
- ปริมาณยะมูลฝอยที่ นำไปใช้ประโยชน์(ตัน)



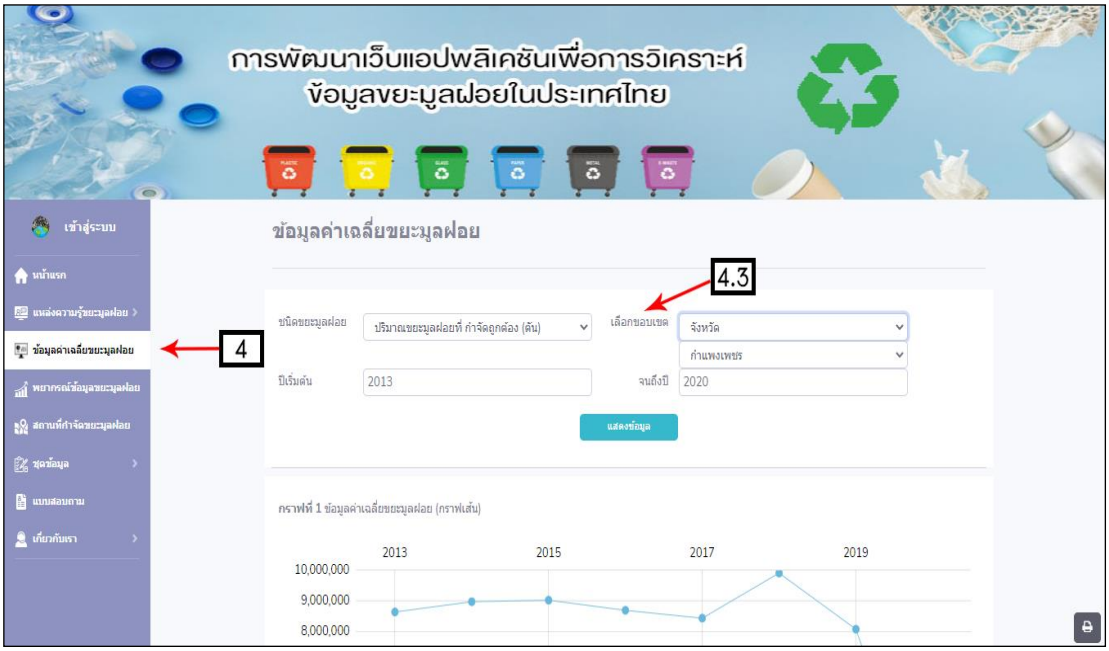
ภาพที่ ก.4.1 หน้าการเลือกแสดงชนิดขยะมูลฝอยที่จะแสดงข้อมูล

4.2 หน้าการเลือกขอบเขตที่จะแสดงข้อมูลขยะมูลฝอย โดยเลือกขอบเขตของภาคจะมีทั้งหมดอยู่ 4 ภาคด้วยกัน คือ ภาคเหนือ ภาคกลาง ภาคตะวันออกเฉียงเหนือ และภาคใต้



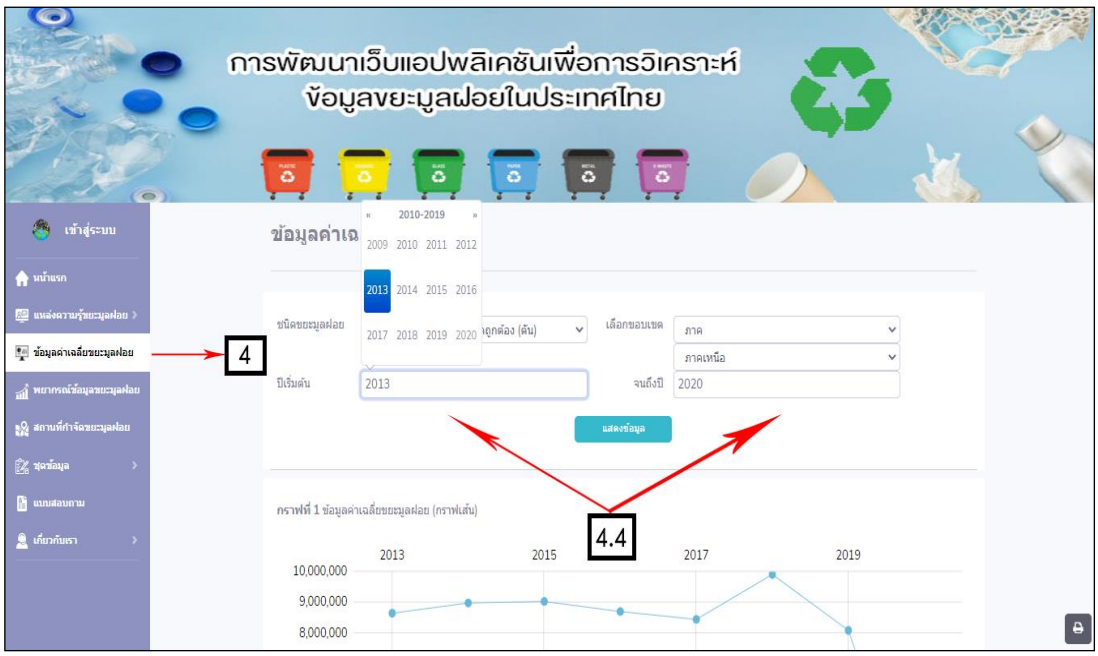
ภาพที่ ก.4.2 หน้าการเลือกขอบเขตของภาค

4.3 หน้าการเลือกขอบเขตที่จะแสดงข้อมูลขยะมูลฝอย โดยเลือกขอบเขตของจังหวัด จะมีทั้งหมด 76 จังหวัด



ภาพที่ ก.4.3 หน้าการเลือกขอบเขตของจังหวัด

4.4 หน้าแสดงปีเริ่มต้น และปีสิ้นสุดที่จะให้แสดงข้อมูลขยะมูลฝอย โดยสามารถปีที่ต้องการที่จะให้แสดงได้

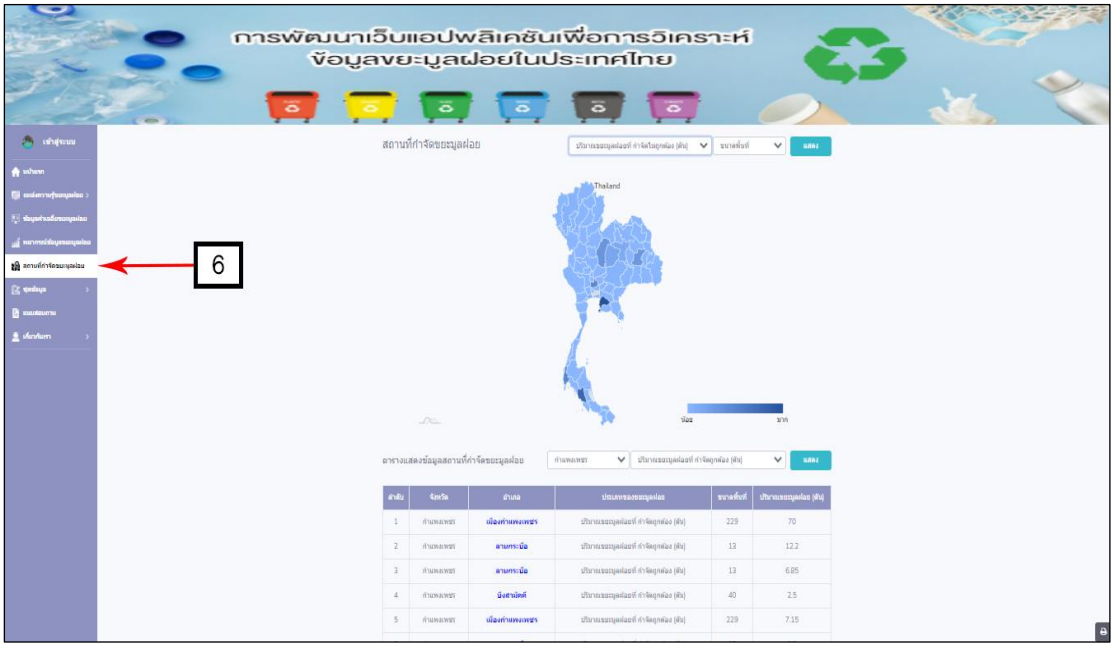


ภาพที่ ก.4.4 หน้าแสดงปีเริ่มต้น และปีสิ้นสุดที่จะให้แสดงข้อมูลขยะมูลฝอย

5) หน้าการพยากรณ์ข้อมูลขยะมูลฝอย ดังภาพที่ ก.5

ภาพที่ ก.5.1 หน้าการพยากรณ์ข้อมูลขยะมูลฝอย

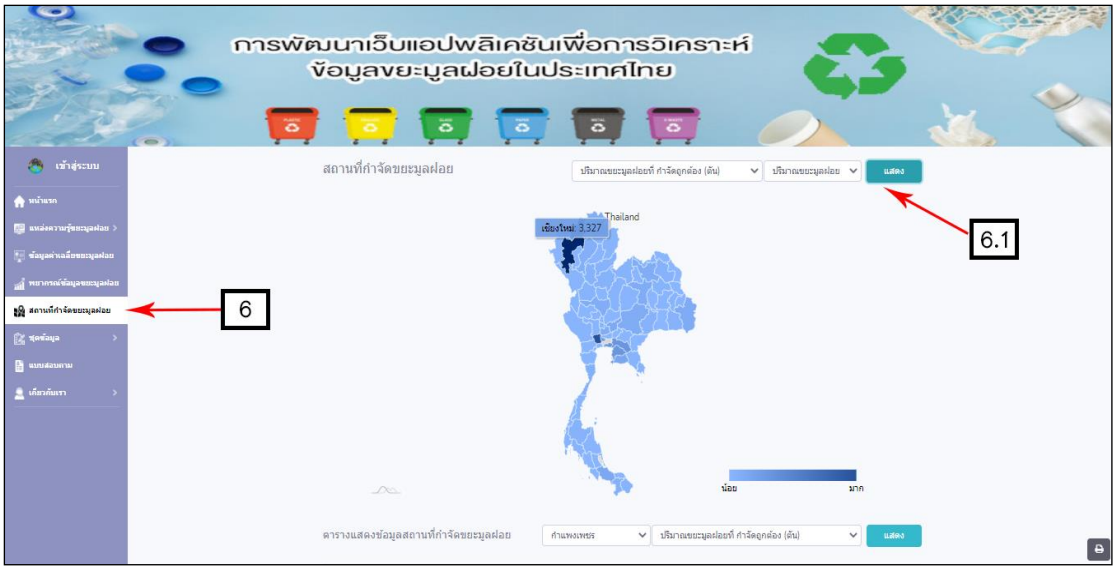
6) หน้าแสดงแผนที่จำกัดขยะมูลฝอย ดังภาพที่ ก.6



ภาพที่ ก.6 หน้าแสดงแผนที่จำกัดขยะมูลฝอย

6.1 หน้าแสดงแผนที่จำกัดขยะมูลฝอย โดยเลือกปริมาณขยะมูลฝอยได้ว่าต้องการกำจัดแบบถูกต้อง(ตัน) หรือกำจัดไม่ถูกต้อง(ตัน) และเลือกปริมาณขยะมูลฝอย เพื่อที่จะแสดงข้อมูลตัวเลขบนแผนที่ในแต่ละจังหวัด ดังตัวอย่าง

- เลือกปริมาณขยะมูลฝอยกำจัดถูต้อง(ตัน)
- เลือกปริมาณขยะมูลฝอย แล้วกดแสดง



ภาพที่ ก.6.1 หน้าแสดงปริมาณขยะมูลฝอยบนแผนที่

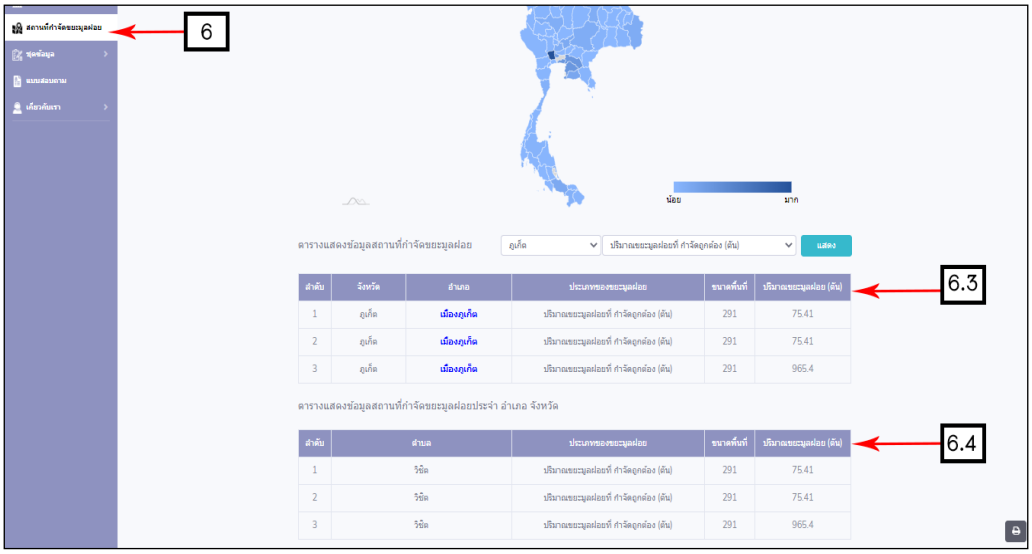
6.2 หน้าแสดงแผนที่จำกัดขยะมูลฝอย โดยเลือกปริมาณขยะมูลฝอยได้ว่าต้องการกำจัดแบบถูกต้อง(ตัน) หรือกำจัดไม่ถูกต้อง(ตัน) และเลือกขนาดพื้นที่ เพื่อที่จะแสดงข้อมูลตัวเลข บนแผนที่ในแต่ละจังหวัด ดังตัวอย่าง

- เลือกปริมาณขยะมูลฝอยกำจัดถูต้อง(ตัน)
- เลือกขนาดพื้นที่ แล้วกดแสดง

ภาพที่ ก.6.2 หน้าแสดงขนาดพื้นที่บนแผนที่

6.3 หน้าแสดงตารางข้อมูลสถานที่กำจัดขยะมูลฝอยในประเทศไทย จะแสดงค่าทั้งตามที่เราเลือกจังหวัด และวิธีการกำจัดของขยะมูลฝอยที่ต้องการจะให้แสดง จากตารางสามารถเลือกดูเป็นรายจังหวัดได้ โดย การเลือกที่ฟิวเตอร์ Province_Name และสามารถเลือกดูจังหวัดนั้น ๆ เป็นรายปี โดยเลือก ฟิวเตอร์ Category_Name ดังตัวอย่าง

- Province_Name เป็น จังหวัดภูเก็ต
- Category_Name เป็น ปริมาณขยะมูลฝอยที่จำกัดถูกต้อง(ตัน)

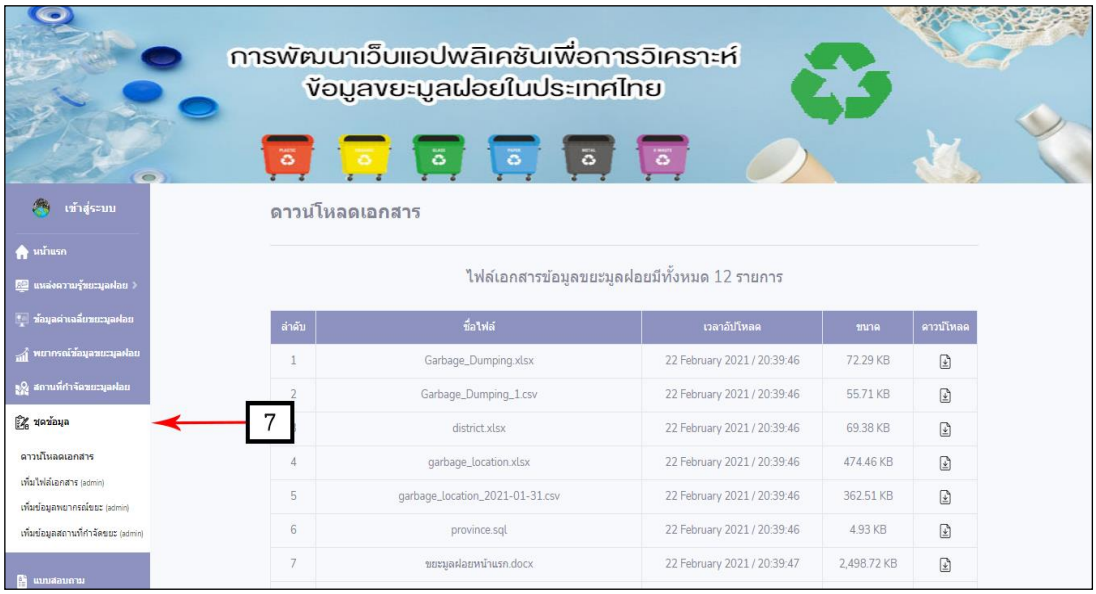


ภาพที่ ก.6.2 หน้าแสดงตารางข้อมูลสถานที่กำจัดขยะมูลฝอย

หมายเลข 6.3 ส่วนของตารางอำเภอในแต่ละจังหวัด

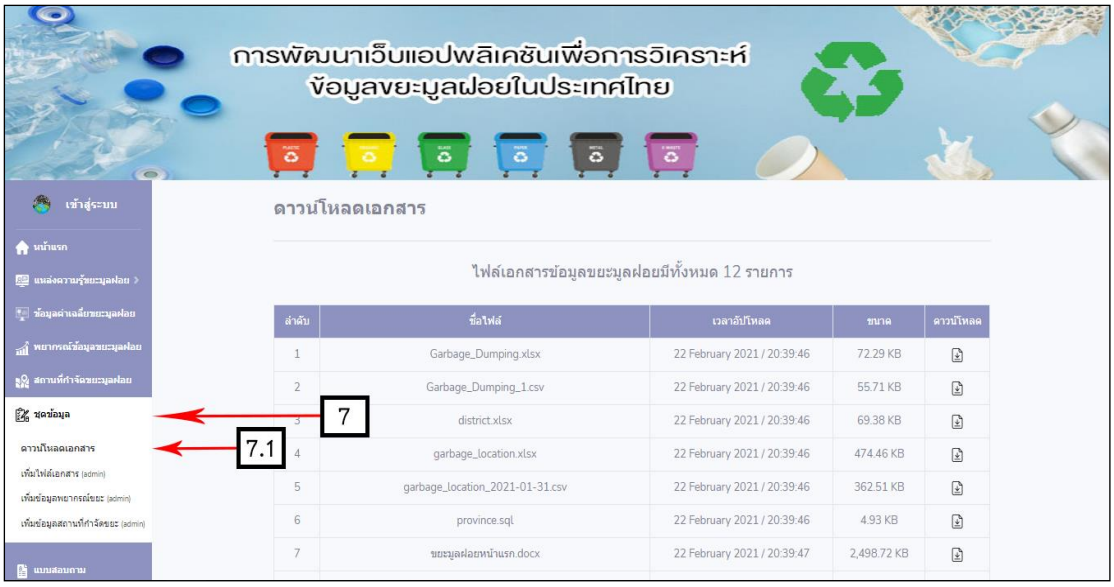
หมายเลข 6.4 ส่วนของตารางตำบลในแต่ละจังหวัด

7) หน้าแสดงชุดข้อมูลปริมาณขยะมูลฝอย ดังภาพที่ ก.7



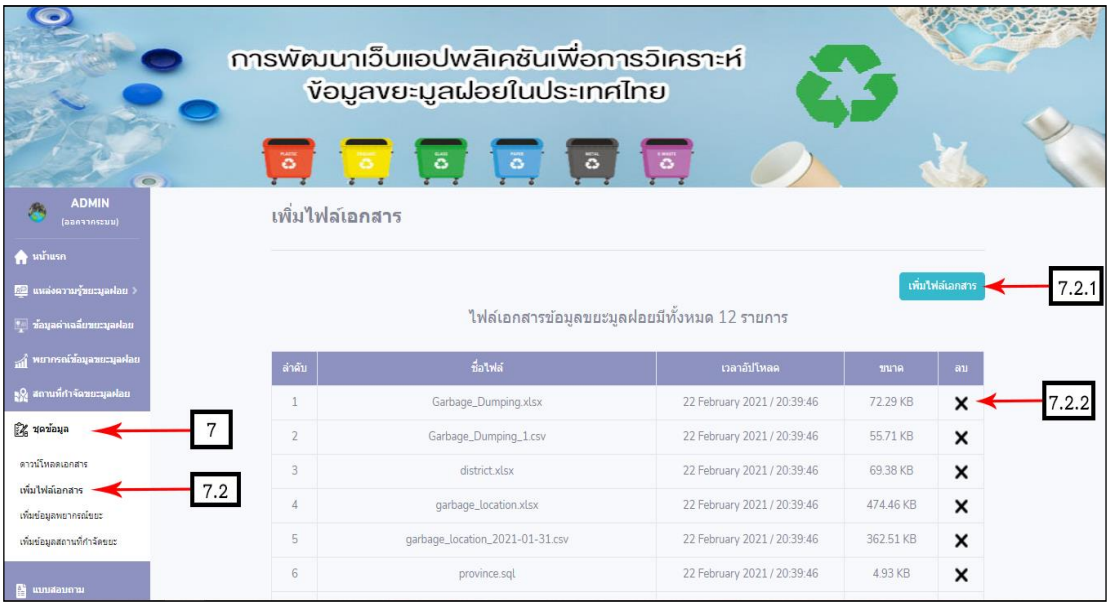
ภาพที่ ก.7 หน้าแสดงชุดข้อมูลปริมาณขยะมูลฝอย

7.1 หน้าดาวน์โหลดชุดข้อมูลปริมาณขยะมูลฝอย ชุดข้อมูลนี้เป็นชุดข้อมูลสำหรับผู้ใช้งานทั่วไปที่สามารถดาวน์โหลดชุดข้อมูลเพื่อนำไปประยุกต์ใช้ได้



ภาพที่ ก.7.1 หน้าดาวน์โหลดชุดข้อมูลปริมาณขยะมูลฝอย

7.2 หน้าเพิ่มไฟล์เอกสารขยะมูลฝอย ชุดข้อมูลนี้เป็นชุดข้อมูลสำหรับผู้ดูแลระบบเท่านั้นที่สามารถเพิ่มไฟล์ชุดข้อมูลได้ และลบชุดข้อมูลได้

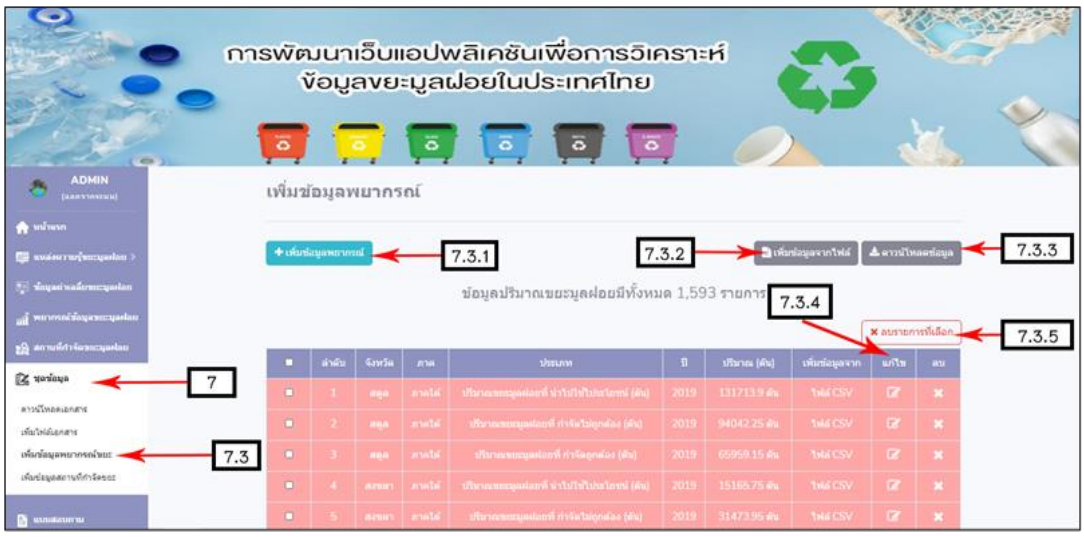


ภาพที่ ก.7.2 หน้าเพิ่ม และลบไฟล์เอกสารขยะมูลฝอย

หมายเลข 7.2.1 เพิ่มไฟล์เอกสาร

หมายเลข 7.2.2 ลบไฟล์เอกสาร

7.3 หน้าเพิ่มไฟล์ชุดข้อมูลการพยากรณ์ปริมาณขยะมูลฝอย ชุดข้อมูลนี้ เป็นชุดข้อมูลสำหรับผู้ดูแลระบบเท่านั้นที่สามารถเพิ่มไฟล์ชุดข้อมูลการพยากรณ์ปริมาณขยะมูลฝอยได้



ภาพที่ ก.7.3 หน้าเพิ่มไฟล์ชุดข้อมูลการพยากรณ์ปริมาณขยะมูลฝอย

หมายเลข 7.3.1 เพิ่มข้อมูลพยากรณ์

หมายเลข 7.3.2 เพิ่มข้อมูลจากไฟล์

หมายเลข 7.3.3 ดาวน์โหลดข้อมูล

หมายเลข 7.3.4 แก้ไขข้อมูล

หมายเลข 7.3.5 ลบข้อมูลตามรายการที่เลือก

7.4 หน้าเพิ่มไฟล์ชุดข้อมูลสถานที่จำกัดขยะมูลฝอย ชุดข้อมูลนี้เป็นชุดข้อมูลสำหรับผู้ดูแลระบบเท่านั้นที่สามารถเพิ่มไฟล์ชุดข้อมูลสถานที่จำกัดขยะมูลฝอยได้

ภาพที่ ก.7.4 หน้าเพิ่มไฟล์ชุดข้อมูลสถานที่จำกัดขยะมูลฝอย

หมายเลข 7.4.1 เพิ่มข้อมูลสถานที่กำจัดขยะมูลฝอย

หมายเลข 7.4.2 เพิ่มข้อมูลจากไฟล์

หมายเลข 7.4.3 ดาวน์โหลดข้อมูล

หมายเลข 7.4.4 แก้ไขข้อมูล

หมายเลข 7.4.5 ลบข้อมูลตามรายการที่เลือก

8) หน้าแบบสอบถามความพึงพอใจของผู้เข้าชมเว็บไซต์ ดังภาพที่ ก.8

ภาพที่ ก.8 หน้าแบบสอบถามความพึงพอใจของผู้เข้าชมเว็บไซต์

9) หน้าแสดงรายละเอียดเกี่ยวกับเว็บไซต์และผู้จัดทำโครงงาน ดังภาพที่ ก.9

ภาพที่ ก.9 หน้าแสดงรายละเอียดเกี่ยวกับเว็บไซต์

หมายเลข 9.1 ข้อมูลทั่วไปของเว็บไซต์
หมายเลข 9.2 วัตถุประสงค์ของเว็บไซต์
หมายเลข 9.3 เป้าหมายของเว็บไซต์
หมายเลข 9.4 อาจารย์ที่ปรึกษา
หมายเลข 9.5 คณะผู้จัดทำ
หมายเลข 9.6 หน่วยงานที่เกี่ยวข้อง

ภาคผนวก ข
แบบฟอร์มและเอกสารที่ใช้ในโครงงาน

ภาคผนวก ข
แบบฟอร์มและเอกสารที่ใช้ในโครงงาน

ผู้จัดทำได้ปรับปรุงแก้ไขคำถามในแบบสอบถามที่อ้างอิงมากจากเว็บไซต์ Google forms ให้สอดคล้องกับเนื้อหาและข้อมูลของเว็บไซต์ เพื่อสอบถามความพึงพอใจของผู้เข้าชมเว็บไซต์

แบบสอบถาม

โครงงานเรื่อง : การพัฒนาเว็บแอปพลิเคชันเพื่อการวิเคราะห์ข้อมูลปริมาณขยะมูลฝอยในประเทศไทย

คำชี้แจง : แบบสอบถามประกอบไปด้วยชุดคำถาม 3 ตอน ดังนี้

ตอนที่ 1 ข้อมูลทั่วไป

ตอนที่ 2 เกี่ยวกับการวิเคราะห์ข้อมูลการพยากรณ์ และการออกแบบหน้าเว็บไซต์

ตอนที่ 3 ข้อเสนอแนะ

ตอนที่ 1 ข้อมูลทั่วไป

คำชี้แจง ใส่เครื่องหมาย ✓ ลงใน ☐ หน้าคำตอบที่ตรงกับข้อมูลของผู้ตอบแบบสอบถามเพียงข้อเดียว

1. เพศ

☐หญิง ☐ชาย

2. สถานะ

☐อาจารย์ ☐นักศึกษา ☐บุคคลทั่วไป

ตอนที่ 2 เกี่ยวกับการวิเคราะห์ข้อมูลการพยากรณ์ และการออกแบบหน้าเว็บไซต์

คำชี้แจง ใส่เครื่องหมาย ✓ ลงในช่องที่ผู้ตอบแบบสอบถามพึงพอใจมากที่สุด

เกณฑ์การประเมิน ระดับ 5 หมายถึง ดีมาก

ระดับ 4 หมายถึง ดี

ระดับ 3 หมายถึง ปานกลาง

ระดับ 2 หมายถึง น้อย

ระดับ 1 หมายถึง ควรปรับปรุง

รายการ	ระดับความพึงพอใจ				
	5	4	3	2	1
ด้านเนื้อหา					
1. การวิเคราะห์และการพยากรณ์ข้อมูลมีความชัดเจนน่าเชื่อถือ					
2. แสดงรายละเอียดโมเดลการพยากรณ์ได้ครบถ้วน					
3. ข้อมูลที่นำเสนอบนเว็บไซต์มีขนาดและองค์ประกอบที่เหมาะสม					
4. การจัดลำดับเนื้อหาเป็นขั้นตอน มีความต่อเนื่องอ่านแล้วเข้าใจ					
5. รูปแบบ Visualization มีความเหมาะสมกับข้อมูล					
ด้านการออกแบบและการจัดรูปแบบเว็บไซต์					
1. การจัดรูปแบบในเว็บไซต์ง่ายต่อการอ่านและสะดวกต่อการใช้งาน					
2. เว็บไซต์มีความสวยงาม มีความทันสมัย น่าสนใจ					
3. สีสันในการออกแบบเว็บไซต์มีความเหมาะสม					
4. สีพื้นหลังกับสีตัวอักษรมีความเหมาะสมต่อการอ่าน					
5. ขนาดตัวอักษร และรูปแบบตัวอักษร มีความสวยงามและอ่านได้ง่าย					
ด้านประโยชน์และการนำไปใช้					
1. เนื้อหามีรายละเอียดและประโยชน์เพียงพอสามารถนำไปประยุกต์ใช้ได้					
2. สื่อ ข่าวสาร และการพยากรณ์ เป็นประโยชน์และสามารถนำไปใช้ได้					
3. สามารถใช้เป็นแหล่งความรู้และแนวทางในการพัฒนาต่อไปได้					
4. แหล่งข้อมูลที่ตรงกับความต้องการของผู้ใช้งาน					

ข้อเสนอแนะ

.....

.....

.....

ภาคผนวก ค
แบบบันทึกรายละเอียดการเข้าพบอาจารย์ที่ปรึกษา

FM-BIS-10

แบบบันทึกรายละเอียดการเข้าพบอาจารย์ที่ปรึกษา

วิชา Computer Information System Project

หลักสูตรระบบสารสนเทศทางคอมพิวเตอร์ สาขาบริหารธุรกิจ

คณะบริหารธุรกิจและศิลปศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เชียงใหม่

ชื่อโครงการ การพัฒนาเว็บแอปพลิเคชันเพื่อการวิเคราะห์ข้อมูลปริมาณขยะมูลฝอยในประเทศไทย

ผู้จัดทำ 1. นางสาว กมลทิพย์ วงค์อ้าย รหัสนักศึกษา 61521207084-6

2. นางสาว นาฏยา สุริยะมณี รหัสนักศึกษา 61521207094-5

อาจารย์ที่ปรึกษา ผู้ช่วยศาสตราจารย์ ดร. ชัชฎาพร ปุกแก้ว

ครั้งที่	วัน/เดือน/ปี	รายละเอียดการเข้าพบอาจารย์ที่ปรึกษา	ความคืบหน้าของผลงาน	ลายมืออาจารย์ที่ปรึกษา
1	2/12/63	ตรวจดู บทที่ 1 - 2	5%	
2	14/12/63	ตรวจ ดูข้อมูล ปริมาณขยะ และ สถานที่กำจัดขยะ	5%	
3	13/01/64	ตรวจความก้าวหน้า บทที่ 3 และดูหน้าเว็บไซต์	5%	
4	20/01/64	ตรวจความก้าวหน้าบทที่ 3 และดูเว็บไซต์	5%	
5	10/02/64	แก้ไข DFD และ ER	5%	
6	24/02/64	ตรวจเอกสาร บทที่ 1-5 และเว็บไซต์	5%	
7	03/03/64	ตรวจเว็บ และเอกสารรูปเล่ม	5%	

ลงชื่อ น.ส. กมลทิพย์ วงค์อ้าย (นักศึกษา) ลงชื่อ น.ส. นาฏยา สุริยะมณี (นักศึกษา)

(กมลทิพย์ วงค์อ้าย) (นาฏยา สุริยะมณี)

ประวัติผู้จัดทำ

ชื่อ-สกุล นางสาว กมลทิพย์ วงค์อ้าย

วันเดือนปีเกิด 8 กรกฎาคม 2539

ภูมิลำเนา 153 ต.สันกลาง อ.สันป่าตอง จ.เชียงใหม่

E-mail befern.01@gmail.com

ประวัติการศึกษา

- ระดับมัธยมต้น โรงเรียนบ้านโป่งน้อย จังหวัดเชียงใหม่ สำเร็จการศึกษา 2555
- ระดับมัธยมปลาย โรงเรียนหอพระ จังหวัดเชียงใหม่ สำเร็จการศึกษา 2558
- ระบบประกาศนียบัตรชั้นสูง สาขาวิชาคอมพิวเตอร์ธุรกิจ วิทยาลัยเทคโนโลยีโปลิลานนาเชียงใหม่ สำเร็จ 2561
- ระดับปริญญาตรี สาขาวิชาระบบสารสนเทศทางธุรกิจ มหาวิทยาลัย เทคโนโลยีราชมงคลล้านนา จังหวัดเชียงใหม่ สำเร็จการศึกษาปี 2564

ประวัติผู้จัดทำ

ชื่อ-สกุล นางสาวนาฏติยา สุริยะมณี

วันเดือนปีเกิด 24 พฤศจิกายน 2540

ภูมิลำเนา 141/453 ต.ต้นเปา อ.สันกำแพง จ.เชียงใหม่

E-mail nardtiya249@gmail.com

ประวัติการศึกษา

- ระดับมัธยมต้น โรงเรียนขาลสุวรรณอนุสรณ์ จ.เชียงใหม่
- ระดับประกาศนียบัตรวิชาชีพ สาขาคอมพิวเตอร์ธุรกิจ วิทยาลัยเทคโนโลยีเอเชีย จ.เชียงใหม่
- ระบบประกาศนียบัตรชั้นสูง สาขาคอมพิวเตอร์ธุรกิจ วิทยาลัยเทคโนโลยีเอเชีย จ.เชียงใหม่
- ระดับปริญญาตรี สาขาวิชาระบบสารสนเทศทางธุรกิจ มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา จังหวัดเชียงใหม่ สำเร็จการศึกษาปี 2564