บทที่ 3
วิธีการดำเนินโครงการ

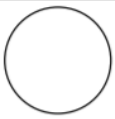
ผู้วิเคราะห์ได้วิเคราะห์ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย เพื่อเผยแพร่ข้อมูลบนเว็บไซต์ และการจัดการข้อมูลที่มีอยู่ให้เกิดประโยชน์ และได้สารสนเทศที่สอดคล้องกับข้อมูล ผู้วิเคราะห์ได้วิเคราะห์ข้อมูล และออกแบบฐานข้อมูลโดยใช้เครื่องมือที่ประกอบไปด้วย แผนภาพกระแสข้อมูล (Data Flow Diagram) กระบวนการทำนอร์มัลไลเซชัน (Normalization) เพื่อลดการซ้ำซ้อนของข้อมูล และกระบวนในการพัฒนาฐานข้อมูลโดยใช้กระบวนการ CRISP-DM ทำการวิเคราะห์ข้อมูลแบบอนุกรมเวลาโดยใช้โปรแกรม tableau public โปรแกรม Minitab และโปรแกรม IBM SPSS เพื่อเปรียบเทียบค่าด้วยวิธี Single Exponential Smoothing ในการวิเคราะห์ข้อมูล ซึ่งในบทนี้จะแสดงถึงวิธีในการดำเนินงาน ดังนี้

3.1 แผนภาพกระแสข้อมูล Data Flow Diagram

แผนภาพกระแสข้อมูล (Data Flow Diagram: DFD) หรือเรียกอีกอย่างหนึ่งว่า แผนภาพการไหลของข้อมูล เป็นเครื่องมือที่ใช้เพื่อแสดงการไหลของข้อมูลและการประมวลผลต่าง ๆ ในระบบ มีความสัมพันธ์กับแหล่งเก็บข้อมูลที่ใช้โดยแผนภาพนี้จะเป็นสื่อที่ช่วยให้การวิเคราะห์ข้อมูลให้เป็นไปได้โดยง่าย และมีความเข้าใจตรงกันระหว่างผู้วิเคราะห์ระบบด้วยกัน ระหว่างผู้วิเคราะห์ระบบกับโปรแกรมเมอร์ หรือระหว่างผู้วิเคราะห์ระบบกับผู้ใช้ระบบ

สัญลักษณ์ที่ใช้ในการเขียนแผนภาพกระแสข้อมูลนั้น ประกอบด้วย 4 สัญลักษณ์ที่แสดงถึงการประมวลผล การไหลของข้อมูล สวนที่ใช้เก็บข้อมูล และสิ่งที่อยู่นอกระบบ โดยไดมีการศึกษาคิดค้นพัฒนาวิธีการอยูหลายแบบ แต่ที่เป็นมาตรฐานมี 2 กลุ่ม คือกลุ่มที่คิดค้นโดย Gane and Sarson (1979) และกลุ่มของ DeMarco and Yourdon (SeMarco, 1979) ถึงแม้สัญลักษณ์บางอย่างของสององค์กรนี้จะต่างกัน แต่องค์ประกอบของแผนภาพและหลักการเขียนแผนภาพไมไดแตกต่างกัน ดังตารางที่ 3.1

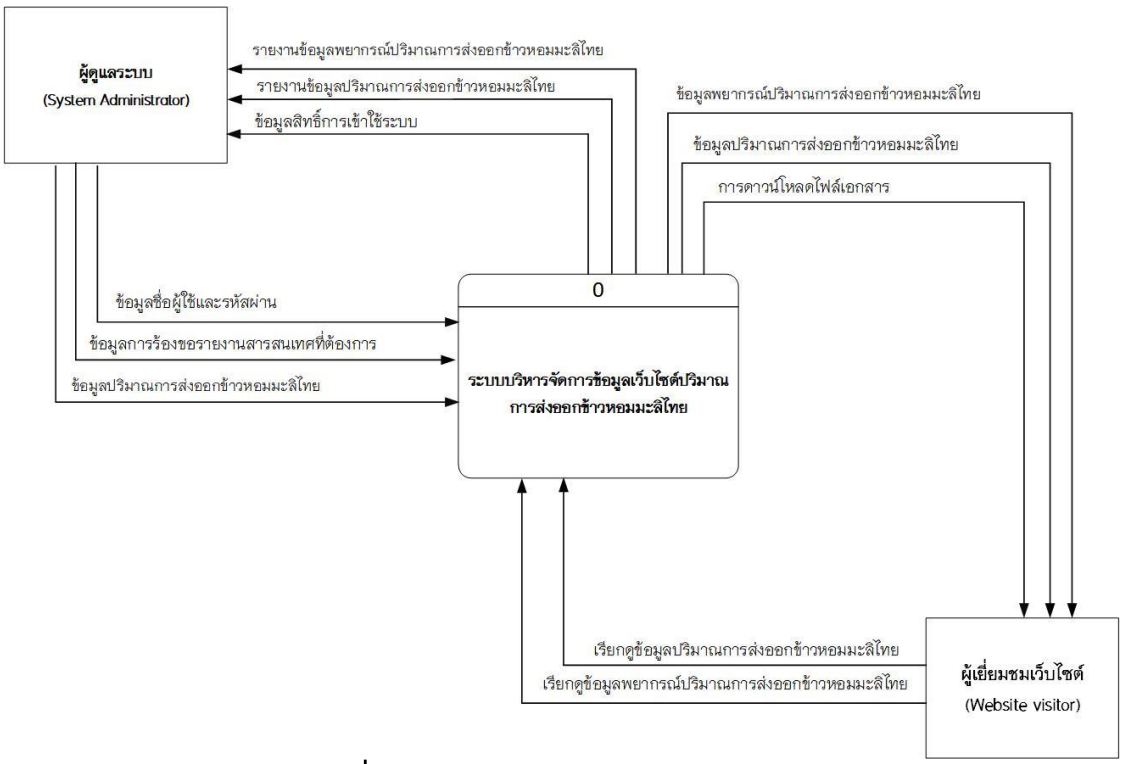
ตารางที่ 3.1 สัญลักษณ์ที่ใช้ในการเขียนแผนการไหลของข้อมูล

ชื่อสัญลักษณ์	DeMarco & Yourdon symbols	Gane & Sarson symbols
การประมวลผล (Process)		
แหล่งเก็บข้อมูล (Data Store)		
กระแสข้อมูล (Data Flow)		
สิ่งที่อยู่ภายนอก (External Entity)		

ที่มา : myweb.cmu.ac.th(2562) , 12 ธันวาคม 2563.

3.1.1 แผนภาพบริบท Context Diagram

แผนภาพบริบท (Context Diagram) จะแสดงภาพโดยรวมระบบของการพัฒนาระบบบริหารจัดการฐานข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย ทั้งหมดดังภาพที่ 3.1



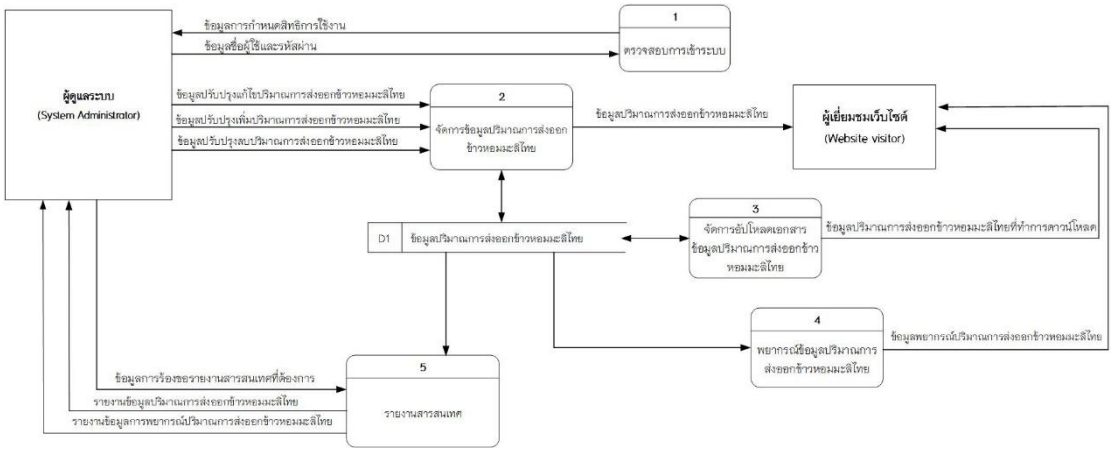
ภาพที่ 3.1 แผนภาพบริบท Context Diagram

จากรูปภาพที่ 3.1 เป็นแผนภาพบริบทระบบของระบบบริหารจัดการฐานข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย โดยสามารถแบ่งผู้ใช้ออกเป็น 2 ประเภทดังนี้

1) **ผู้ดูแลระบบ (System Administrator)** สามารถลงชื่อเข้าใช้และรหัสผ่านได้ สามารถเรียกดูข้อมูลพยากรณ์ปริมาณการส่งออกข้าวหอมมะลิไทยได้ สามารถเรียกดูข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทยได้ และสามารถจัดการข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทยได้

2) **ผู้เยี่ยมชมเว็บไซต์ (Website visitor)** สามารถเรียกดูข้อมูลการพยากรณ์ปริมาณการส่งออกข้าวหอมมะลิไทยได้ สามารถเรียกดูข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทยได้ และสามารถดาวน์โหลดเอกสารบนเว็บไซต์ได้

3.1.2 แผนภาพกระแสข้อมูลระดับที่ 0 (Data Flow Diagram Level 0)



ภาพที่ 3.2 แผนภาพกระแสข้อมูลระดับที่ 0 (Data Flow Diagram Level 0)

จากภาพที่ 3.2 แสดงแผนภาพกระแสข้อมูลระดับที่ 0 (Data Flow Diagram Level 0) ระบบบริหารฐานข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย ประกอบด้วย 5 กระบวนการหลัก ดังนี้

ตารางที่ 3.2 คำอธิบายกระบวนการ ตรวจสอบการเข้าระบบ

Process Description	
System	ระบบบริหารจัดการฐานข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย
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 และกำหนดสิทธิ์ในการเข้าใช้ระบบ โดยระบบจะตรวจสอบชื่อผู้ใช้และรหัสผ่าน ผู้มีสิทธิ์เข้าใช้ระบบโดยมีชื่อผู้ใช้ระบบได้แก่ ผู้ดูแลระบบ

ตารางที่ 3.3 คำอธิบายกระบวนการ จัดการข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย

Process Description	
System	ระบบบริหารจัดการฐานข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย
DFD Number	2
Process Name	จัดการข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย
Input Data Flow	ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย
Output Data Flow	ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย
Data Store Used	ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย
Description	เป็นกระบวนการสำหรับจัดการข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทยได้แก่ การเพิ่ม ลบ แก้ไข ผู้มีสิทธิใช้งานกระบวนการนี้ ได้แก่ ผู้ดูแลระบบ และออกสารสนเทศข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทยให้กับผู้เยี่ยมชมเว็บไซต์

ตารางที่ 3.4 คำอธิบายกระบวนการ จัดการอัปโหลดเอกสารข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย

Process Description	
System	ระบบบริหารจัดการฐานข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย
DFD Number	3
Process Name	จัดการอัปโหลดเอกสารข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย
Input Data Flow	ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย
Output Data Flow	ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทยเป็นไฟล์ Excel หรือ pdf และข้อมูลเอกสารที่เกี่ยวข้องที่เป็นไฟล์ Excel หรือ pdf
Data Store Used	ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย
Description	เป็นกระบวนการสำหรับจัดการดาวน์โหลดเอกสารต่าง ๆ เช่นข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทยที่ผู้เยี่ยมชมเว็บไซต์สามารถดาวน์โหลดได้

ตารางที่ 3.5 คำอธิบายกระบวนการ พยากรณ์ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย

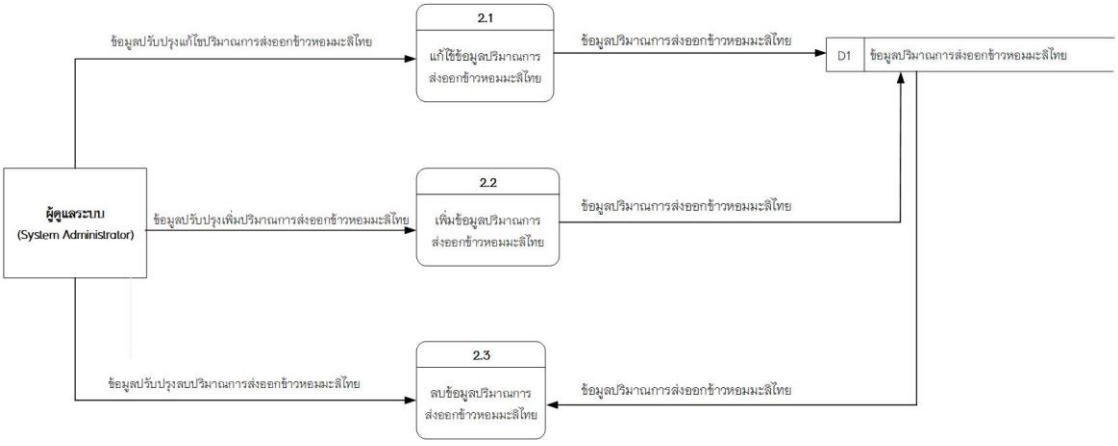
Process Description	
System	ระบบบริหารจัดการฐานข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย
DFD Number	4
Process Name	พยากรณ์ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย
Input Data Flow	ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย
Output Data Flow	ข้อมูลพยากรณ์ปริมาณการส่งออกข้าวหอมมะลิไทย
Data Store Used	ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย
Description	เป็นกระบวนการพยากรณ์ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย เพื่อเผยแผ่ข้อมูลออกสู่เว็บไซต์

ตารางที่ 3.6 คำอธิบายกระบวนการ แสดงรายงานสารสนเทศ

Process Description	
System	ระบบบริหารจัดการฐานข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย
DFD Number	5
Process Name	รายงานสารสนเทศ
Input Data Flow	ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย, ข้อมูลพยากรณ์ปริมาณการส่งออก
Output Data Flow	รายงานข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย และรายงานข้อมูลพยากรณ์ปริมาณการส่งออกข้าวหอมมะลิไทย
Data Store Used	ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย, ข้อมูลพยากรณ์ปริมาณการส่งออกข้าวหอมมะลิไทย
Description	เป็นกระบวนการสำหรับออกรายงานสารสนเทศ โดยผู้มีสิทธิ์เข้าใช้ระบบได้แก่ ผู้ดูแลระบบ

จากแผนภาพกระแสข้อมูล ระดับที่ 0 กระบวนการที่ 2 สามารถแยกย่อยเป็นกระบวนการย่อยระดับที่ 1 ได้ดังนี้

แผนภาพกระแสข้อมูล ระดับที่ 1 กระบวนการที่ 2.1, 2.2, 2.3 จัดการข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย



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

ตารางที่ 3.7 คำอธิบายกระบวนการที่ 2.1 แก้ไข้ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย

Process Description	
System	ระบบบริหารจัดการฐานข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย
DFD Number	2.1
Process Name	แก้ไข้ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย
Input Data Flow	ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย
Output Data Flow	ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย
Data Store Used	ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย
Description	เป็นกระบวนการสำหรับแก้ไขข้อมูลข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย โดยผู้มีสิทธิ์เข้าใช้ระบบได้แก่ผู้ดูแลระบบ

ตารางที่ 3.8 แสดงคำอธิบายกระบวนการที่ 2.2 เพิ่มข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย

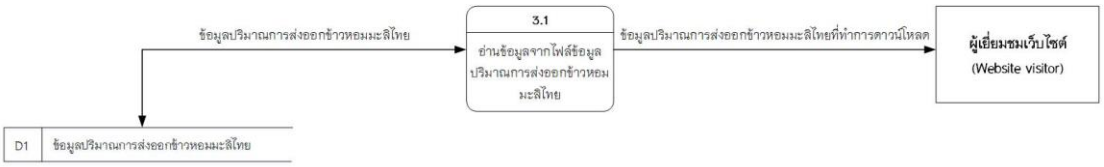
Process Description	
System	ระบบบริหารจัดการฐานข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย
DFD Number	2.2
Process Name	เพิ่มข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย
Input Data Flow	ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย
Output Data Flow	ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย
Data Store Used	ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย
Description	เป็นกระบวนการเพิ่มข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย โดยผู้มีสิทธิ์เข้าใช้ระบบได้แก่ผู้ดูแลระบบ

ตารางที่ 3.9 คำอธิบายกระบวนการที่ 2.3 ลบข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย

Process Description	
System	ระบบบริหารจัดการฐานข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย
DFD Number	2.3
Process Name	ลบข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย
Input Data Flow	ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย
Output Data Flow	ข้อมูลพยากรณ์ปริมาณการส่งออกข้าวหอมมะลิไทยที่ถูกลบออกจากระบบ
Data Store Used	ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย
Description	เป็นกระบวนการสำหรับลบข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทยโดยผู้มีสิทธิ์เข้าใช้ระบบได้แก่ผู้ดูแลระบบ

จากแผนภาพกระแสข้อมูล ระดับที่ 0 กระบวนการที่ 3 สามารถแยกย่อยเป็นกระบวนการย่อยระดับที่ 1 ได้ดังนี้

แผนภาพกระแสข้อมูล ระดับที่ 1 กระบวนการที่ 3.1 อ่านข้อมูลจากไฟล์ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย



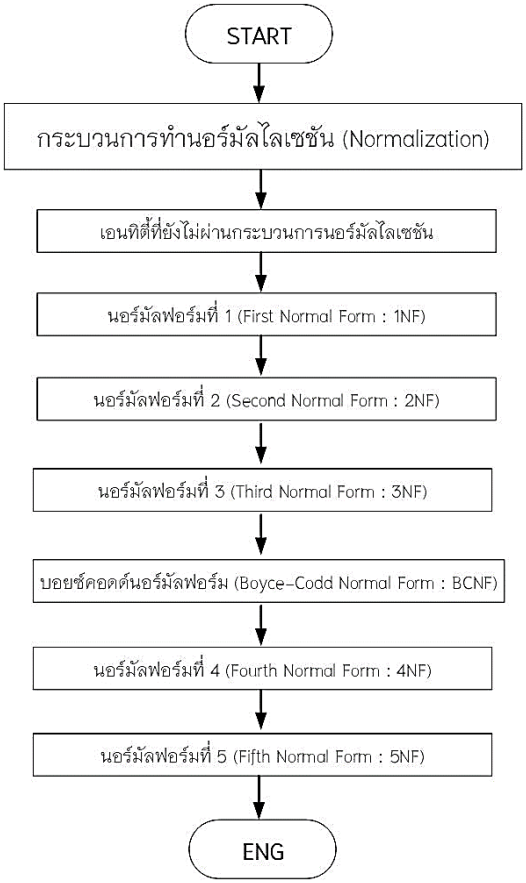
ภาพที่ 3.4 แผนภาพกระแสข้อมูล ระดับที่ 1 กระบวนการที่ 3.1 อ่านข้อมูลจากไฟล์ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย

ตารางที่ 3.10 แสดงคำอธิบายกระบวนการที่ 3.1 อ่านข้อมูลจากไฟล์ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย

Process Description	
System	ระบบบริหารจัดการฐานข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย
DFD Number	3.1
Process Name	อ่านข้อมูลจากไฟล์ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย
Input Data Flow	ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย
Output Data Flow	ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย
Data Store Used	ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย
Description	เป็นกระบวนการสำหรับการอ่านไฟล์ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย เพื่อบันทึกข้อมูลไฟล์ลงดาต้าสโตร์

3.2 กระบวนการทำนอร์มัลไลเซชัน (Normalization)

การทำนอร์มัลไลเซชัน เป็นวิธีการในการกำหนดแอตทริบิวต์ให้กับแต่ละเอนทิตี้ เพื่อให้ได้โครงสร้างของตารางที่ดี สามารถควบคุมความซ้ำซ้อนของข้อมูล โดยทั่วไปการทำนอร์มัลไลเซชันจะประกอบด้วยนอร์มัลฟอร์ม (Normal Form) แบบต่าง ๆ ที่มีเงื่อนไขของการทำให้อยู่ในรูปของนอร์มัลฟอร์มที่แตกต่างกันไป ถึงแม้ว่าการนอร์มัลไลเซชันจะเป็นสิ่งสำคัญและจำเป็นสำหรับการออกแบบฐานข้อมูล แต่ก็ไม่ได้หมายความว่าจะต้องทำการนอร์มัลไลเซชันจนถึงระดับนอร์มัลฟอร์มที่ 6 แต่จะขึ้นอยู่กับผู้ออกแบบฐานข้อมูลว่าต้องการลดความซ้ำซ้อนในฐานข้อมูลให้อยู่ในระดับใด การออกแบบฐานข้อมูลที่ดีจึงต้องพิจารณาถึงความต้องการของผู้ใช้และต้องสามารถตอบสนองได้อย่างรวดเร็ว เพราะฉะนั้นในบางกรณีจึงมีการลดระดับการนอร์มัลไลเซชันในบางส่วนของการออกแบบฐานข้อมูล เพื่อให้ระบบสามารถตอบสนองได้ตามความต้องการของผู้ใช้



ภาพที่ 3.5 กระบวนการทำนอร์มัลไลเซชัน (Normalization)

ที่มา : slideplayer.in.th/slide/16099894/ (2558), 12 ธันวาคม 2563.

ซึ่งการทำนอร์มัลไลเซชันประกอบด้วยนอร์มัลฟอร์มแบบต่าง 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)

3.2.1 กระบวนการทำนอร์มัลฟอร์มที่ 1 (First Normal Form: 1NF)

เป็นระดับที่ใช้สำหรับการปรับโครงสร้างของข้อมูลของรีเลชัน ให้มีคุณสมบัติตามนิยามคือ รีเลชันใด ๆ จะมีคุณสมบัติอยู่ในรูปแบบนอร์มัลฟอร์มระดับที่ 1 ก็ต่อเมื่อ ทุกแอตทริบิวต์ในแต่ละตารางค่าของข้อมูลเพียงค่าเดียว คือต้องไม่มีค่ากลุ่มข้อมูลที่ไม่ซ้ำกัน

ตารางที่ 3.12 ตารางข้อมูลก่อนการทำนอร์มัลฟอร์มที่ 1 (First Normal Form : 1NF)

category_id	category_name	continent_id	continent_name	country_id	country_name	year	volume	value	Ton / Baht
cate01	ข้าวข้าว100%ชั้น1	con01	เอเชีย	cou81	ญี่ปุ่น	2013	11	408474.00	37,134
				cou60	มาเลเซีย		22	751718.00	34,169
				cou971	สหรัฐอาหรับเอมิเรตส์		24	847440.00	35,310
				cou86	สาธารณรัฐประชาชนจีน		28	900223.20	31698
				cou84	สาธารณรัฐสังคมนิยมเวียดนาม		14	578452.00	41,318
				cou852	ฮ่องกง		33	1808169.00	54,793
cate02	ข้าวขาวหักเอวันเลิศ	con02	ยุโรป	cou47	นอร์เวย์		387	9437382.00	24,386
				cou31	เนเธอร์แลนด์		451	12225257.00	27,107
				cou32	เบลเยียม		556	16110656.00	28,976
				cou351	โปรตุเกส		582	8238792.00	14,156
				cou33	ฝรั่งเศส		623	12246934.00	19,658
				cou358	ฟินแลนด์		741	17834388.00	24,068

จากตารางที่ 3.12 ข้อมูลอยู่ในรูปแบบ Unnormalized Form จะต้องทำให้อยู่ในรูปของ 1NF โดยสามารถทำให้อยู่ในรูปแบบ 1NF ได้โดยการกระจายข้อมูลดังนี้

ตารางที่ 3.13 ตารางข้อมูลที่อยู่ในรูปของนอร์มัลฟอร์มที่ 1 (First Normal Form : 1NF)

category_id	category_name	continent_id	continent_name	country_id	country_name	year	volume	value	Ton / Baht
cate01	ข้าวข้าว100%ชั้น1	con01	เอเชีย	cou81	ญี่ปุ่น	2013	11	408474.00	37,134
cate01	ข้าวข้าว100%ชั้น1	con01	เอเชีย	cou60	มาเลเซีย	2013	22	751718.00	34,169
cate01	ข้าวข้าว100%ชั้น1	con01	เอเชีย	cou971	สหรัฐอาหรับเอมิเรตส์	2013	24	847440.00	35,310
cate01	ข้าวข้าว100%ชั้น1	con01	เอเชีย	cou86	สาธารณรัฐประชาชนจีน	2013	28	900223.20	31698
cate01	ข้าวข้าว100%ชั้น1	con01	เอเชีย	cou84	สาธารณรัฐสังคมนิยมเวียดนาม	2013	14	578452.00	41,318
cate01	ข้าวข้าว100%ชั้น1	con01	เอเชีย	cou852	ฮ่องกง	2013	33	1808169.00	54,793
cate02	ข้าวขาวหักเอวันเลิศ	con02	ยุโรป	cou47	นอร์เวย์	2013	387	9437382.00	24,386
cate02	ข้าวขาวหักเอวันเลิศ	con02	ยุโรป	cou31	เนเธอร์แลนด์	2013	451	12225257.00	27,107
cate02	ข้าวขาวหักเอวันเลิศ	con02	ยุโรป	cou32	เบลเยียม	2013	556	16110656.00	28,976
cate02	ข้าวขาวหักเอวันเลิศ	con02	ยุโรป	cou351	โปรตุเกส	2013	582	8238792.00	14,156
cate02	ข้าวขาวหักเอวันเลิศ	con02	ยุโรป	cou33	ฝรั่งเศส	2013	623	12246934.00	19,658
cate02	ข้าวขาวหักเอวันเลิศ	con02	ยุโรป	cou358	ฟินแลนด์	2013	741	17834388.00	24,068

จากข้อมูลข้างต้น นำมาทำกระบวนการนอร์มัลฟอร์มที่ 1 (First Normal Form : 1NF) จะได้ตารางดังนี้ rice_export (category_name, category_id, continent_name, continent_id, country_name, country_id, year, volume, value, Ton/Baht) ซึ่งทุก ๆ field ในแต่ละ record จะเป็น single value นั่นคือ ในตารางหนึ่ง ๆ จะไม่มีค่าของกลุ่มข้อมูลที่ไม่ซ้ำกัน (Repeating Group)

3.2.1 กระบวนการทำนอร์มัลฟอร์มที่ 2 (Second Normal Form: 2NF)

เป็นการแก้ไขปัญหาความซ้ำซ้อนของข้อมูล ที่ปรากฏในรีเลชันที่ผ่านการนอร์มัลไลซ์ ระดับที่ 1 โดยรูปแบบนอร์มัลฟอร์มระดับที่ 2 จะต้องมีคุณสมบัติดังนี้คือ รีเลชันใด ๆ จะมีคุณสมบัติอยู่ในรูปแบบนอร์มัลฟอร์มระดับที่ 2 ก็ต่อเมื่อ

1. รีเลชันนั้นมีคุณสมบัติอยู่ในรูปแบบนอร์มัลฟอร์มระดับที่ 1

2. ทุกแอตทริบิวต์ที่ไม่มีคีย์หลัก ต้องมีความสัมพันธ์กับแอตทริบิวต์ที่เป็นคีย์หลัก หรือทุกแอตทริบิวต์ที่ประกอบกันเป็นคีย์หลัก มิใช่ส่วนหนึ่งส่วนใดของคีย์หลัก (เป็นกรณีที่คีย์หลักนั้น มีคีย์ร่วม (Composition Key) คือมีหลายแแอตทริบิวต์ประกอบกันเป็นคีย์หลัก) จากการพิจารณาจะได้ตารางข้อมูลที่อยู่ในรูปของนอร์มัลฟอร์มที่ 2 ดังนี้

จากตารางที่ 3.13 จะพบว่าในรีเลชันการส่งออกข้าวหอมมะลิไทยจะมี category_id, continent_id และ country_id ประกอบกันเป็นคีย์หลัก และจากการพิจารณาพบว่า จะไม่ตรงตามนิยามของรูปแบบนอร์มัลฟอร์มละดับที่ 2 เพราะแอตทริบิวต์ year, volume, value, และ Ton/Baht จะมีค่าเปลี่ยนแปลงตามแอตทริบิวส์ category_id, continent_id และ country_id จึงต้องทำการแยกรีเลชันออกเป็น 4 รีเลชันโดยให้ตารางที่ 1 ชื่อรีเลชัน category, ตารางที่ 2 ชื่อรีเลชัน continent, ตารางที่ 3 ชื่อรีเลชัน country, และตารางที่ 4 ชื่อรีเลชัน rice_export เพื่อให้ตารางข้างต้นมีคุณสมบัติเป็นนอร์มัลฟอร์มที่ 2 จะต้องทำการแตกตารางออกมา ตามความสัมพันธ์ของฟังก์ชันการขึ้นต่อกัน เป็น 4 ตาราง ดังนี้

ตารางที่ 3.14 ตาราง category ที่อยู่ในรูปของนอร์มัลฟอร์มที่ 2 (Second Normal Form : 2NF)

category_id	category_name
cate01	ข้าวข้าว100%ชั้น1
cate01	ข้าวข้าว100%ชั้น1
cate01	ข้าวข้าว100%ชั้น1
cate01	ข้าวข้าว100%ชั้น1
cate01	ข้าวข้าว100%ชั้น1
cate01	ข้าวข้าว100%ชั้น1
cate02	ข้าวขาวหักเอวันเลิศ
cate02	ข้าวขาวหักเอวันเลิศ
cate02	ข้าวขาวหักเอวันเลิศ
cate02	ข้าวขาวหักเอวันเลิศ
cate02	ข้าวขาวหักเอวันเลิศ
cate02	ข้าวขาวหักเอวันเลิศ

ตารางที่ 3.15 ตาราง continent ที่อยู่ในรูปของนอร์มัลฟอร์มที่ 2 (Normal Form : 2NF)

continent_id	continent_name
con01	เอเชีย
con01	เอเชีย
con01	เอเชีย
con01	เอเชีย
con01	เอเชีย
con01	เอเชีย
con02	ยุโรป
con02	ยุโรป
con02	ยุโรป
con02	ยุโรป
con02	ยุโรป
con02	ยุโรป

ตารางที่ 3.16 ตาราง country ที่อยู่ในรูปของนอร์มัลฟอร์มที่ 2 (Normal Form : 2NF)

country_id	country_name
cou81	ญี่ปุ่น
cou60	มาเลเซีย
cou971	สหรัฐอาหรับเอมิเรตส์
cou86	สาธารณรัฐประชาชนจีน
cou84	สาธารณรัฐสังคมนิยมเวียดนาม
cou852	ฮ่องกง
cou47	นอร์เวย์
cou31	เนเธอร์แลนด์
cou32	เบลเยียม
cou351	โปรตุเกส
cou33	ฝรั่งเศส
cou358	ฟินแลนด์

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

rice_id	category_id	continent_id	country_id	year	volume	value	Ton / Baht
R1	cate01	con01	cou81	2013	11	408474.00	37,134
R2	cate01	con01	cou60	2013	22	751718.00	34,169
R3	cate01	con01	cou971	2013	24	847440.00	35,310
R4	cate01	con01	cou86	2013	28	900223.20	31698
R5	cate01	con01	cou84	2013	14	578452.00	41,318
R6	cate01	con01	cou852	2013	33	1808169.00	54,793
R7	cate02	con02	cou47	2013	387	9437382.00	24,386
R8	cate02	con02	cou31	2013	451	12225257.00	27,107
R9	cate02	con02	cou32	2013	556	16110656.00	28,976
R10	cate02	con02	cou351	2013	582	8238792.00	14,156
R11	cate02	con02	cou33	2013	623	12246934.00	19,658
R12	cate02	con02	cou358	2013	741	17834388.00	24,068

จากตารางที่ 3.17 จะเห็นได้ว่า Attribute ทุกตัวไม่มี Transitive Dependency จึงอยู่ในรูปแบบของ 3NF แล้ว และไม่มี Attribute อื่นใน Relation ที่สามารถระบุค่าของ Attribute ที่เป็นคีย์หลักหรือส่วนใดของคีย์หลักในกรณีที่คีย์หลักเป็นคีย์ผสม ตารางดังกล่าวจึงอยู่ในรูปแบบของ BCNF แล้ว

จากการที่ผู้วิเคราะห์ได้ออกแบบตารางข้อมูลไว้แล้วนั้น ได้นำตารางข้อมูลมาปรับปรุงใหม่โดยใช้วิธีการทำนอร์มัลไลเซชัน เพื่อปรับเปลี่ยนโครงสร้างของรีเลชันต่าง ๆ ช่วยลดปัญหาความซ้ำซ้อนของข้อมูลในลักษณะต่าง ๆ และทำให้รีเลชันมีโครงสร้างที่เหมาะสมในการใช้งานต่อไป โดยผู้วิเคราะห์ได้ทำการนอร์มัลไลเซชันในระดับที่ 1 และ 2 โดยจะได้ตารางทั้งหมด 4 ตาราง ประกอบไปด้วยตารางที่ 1 ชื่อรีเลชัน category, ตารางที่ 2 ชื่อรีเลชัน continent, ตารางที่ 3 ชื่อรีเลชัน country, และตารางที่ 4 ชื่อรีเลชัน rice_export ที่ได้จากข้อมูลการส่งออกข้าวหอมมะลิไทย นำมาผ่านกระบวนการนอร์มัลไลเซชันทั้ง 2 ขั้นตอนนั้น สามารถนำตารางมาใช้เพื่อจัดเก็บข้อมูลลงในฐานข้อมูลและใช้วิเคราะห์ข้อมูลในขั้นตอนต่อไปได้อย่างเหมาะสมแล้ว

3.3 ความสัมพันธ์ของฐานข้อมูล (ER-Diagram)

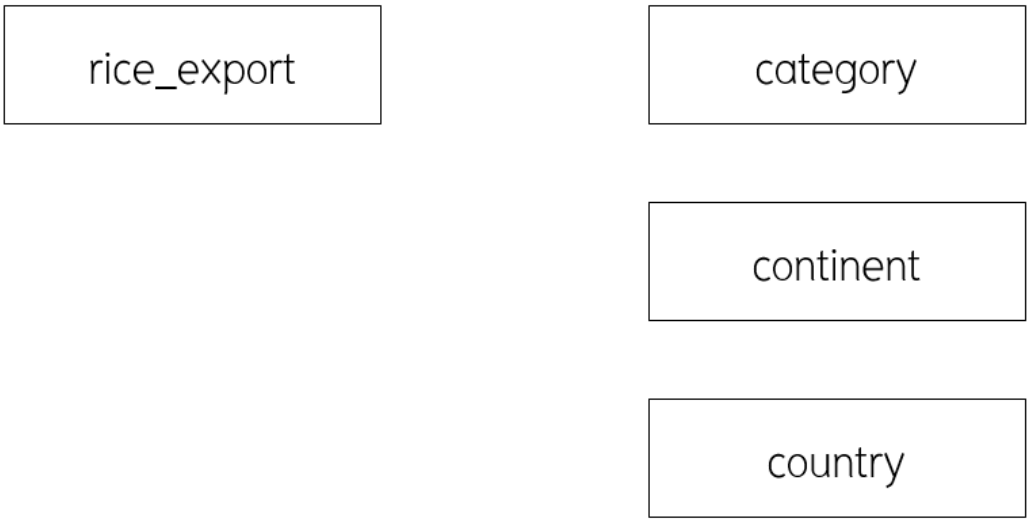
ER Diagram คือ แบบจำลองที่ใช้อธิบายโครงสร้างของฐานข้อมูลซึ่งเขียนออกมาในลักษณะของรูปภาพ การอธิบายโครงสร้างและความสัมพันธ์ของข้อมูล (Relationship) ประกอบด้วย

1) เอนทิตี้ (Entity) เป็นวัตถุ หรือสิ่งของที่เราสนใจในระบบงานนั้น ๆ
2) แอททริบิว (Attribute) เป็นคุณสมบัติของวัตถุที่เราสนใจ
3) ความสัมพันธ์ (Relationship) คือ ความสัมพันธ์ระหว่างเอนทิตี้

ER Diagram มีความสำคัญต่อการพัฒนาระบบงานฐานข้อมูล Application ต่าง ๆ ที่ต้องการการเก็บข้อมูลอย่างมีระบบ มีโครงสร้าง ดังนั้น ER Diagram จึงใช้เพื่อเป็นเอกสารในการสื่อสารระหว่าง นักออกแบบระบบ และนักพัฒนาระบบ เพื่อให้สื่อสารอย่างตรงกัน และเป็นสากล โดยมีขั้นตอนในการพัฒนาดังนี้

3.3.1 การกำหนดเอนทิตี้

การกำหนดเอนทิตี้เป็นการกำหนดสิ่งที่สนใจและต้องการจัดเก็บข้อมูลที่มีในระบบงาน โดยดูจากลักษณะหน้าที่ของระบบงานว่ามีรายละเอียดการทำงานอย่างไร ผู้วิเคราะห์ได้กำหนดเอนทิตี้ของฐานข้อมูลการส่งออกข้าวหอมมะลิไทย ประกอบด้วย เอนทิตี้ rice_export, เอนทิตี้ category, เอนทิตี้ continent, เอนทิตี้ country



ภาพที่ 3.6 การกำหนดเอนทิตี้ของฐานข้อมูลการส่งออกข้าวหอมมะลิไทย

3.3.2 การกำหนดแอตทริบิวต์ของเอนทิตี้

การกำหนดแอตทริบิวต์ของเอนทิตี้เป็นการกำหนดคุณลักษณะต่าง ๆ ของแต่ละเอนทิตี้รวมทั้งพิจารณาแอตทริบิวต์ที่จะทำหน้าที่เป็นคีย์หลักของเอนทิตี้ด้วย ประกอบไปด้วย

เอนทิตี้ rice_export (rice_id, category_name, continent_name, country_name, year, volue, value, ton_baht)

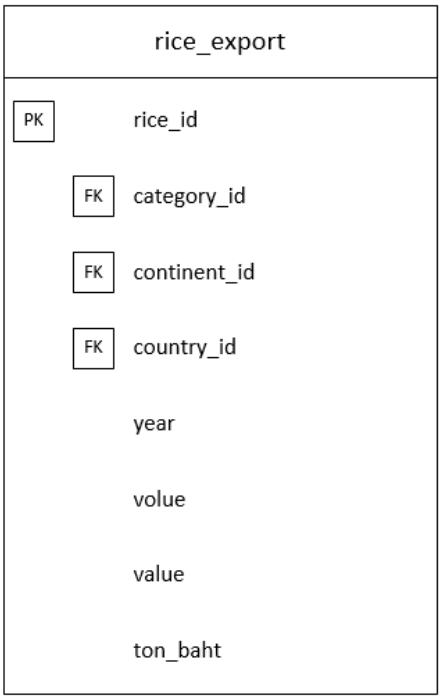
เอนทิตี้ category(category _id, category _name

เอนทิตี้ continent(continent_id, continent_name)

เอนทิตี้ country(country_id, country_name)

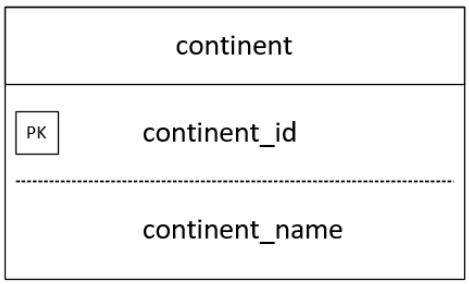
การกำหนดคีย์หลัก (Primary key) เป็นการกำหนดคีย์ของแต่ละเอนทิตี้ว่าในแต่ละเอนทิตี้ใช้แอตทริบิวต์ใดเป็นคีย์หลัก โดยการใช้สัญลักษณ์ PK หน้าชื่อแอตทริบิวต์ที่เป็นคีย์หลักดังต่อไปนี้

ภาพที่ 3.7 เป็นการกำหนดแอตทริบิวต์ของเอนทิตี้ rice_export ประกอบด้วยแอตทริบิวต์ rice_id, แอตทริบิวต์ category_name, แอตทริบิวต์ continent_name, แอตทริบิวต์ country_name, แอตทริบิวต์ year, แอตทริบิวต์ volue, แอตทริบิวต์ value, แอตทริบิวต์ton_baht โดยมีคีย์หลักคือ rice_id และคีย์นอก (Foreign Key) คือแอตทริบิวต์ category_id, continent_id, country_id



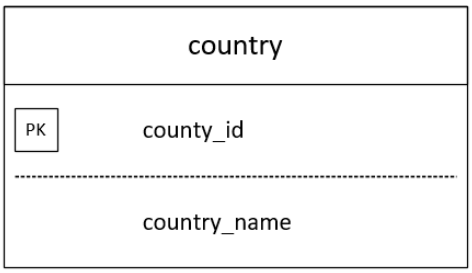
ภาพที่ 3.7 การกำหนดแอตทริบิวต์ของเอนทิตี้ rice_export

ภาพที่ 3.8 การกำหนดแอตทริบิวต์ของเอนทิตี้ continent ประกอบด้วยแอตทริบิวต์ continent_id, แอตทริบิวต์ continent_name โดยมีคีย์หลัก (Primary key) คือแอตทริบิวต์ continent_id



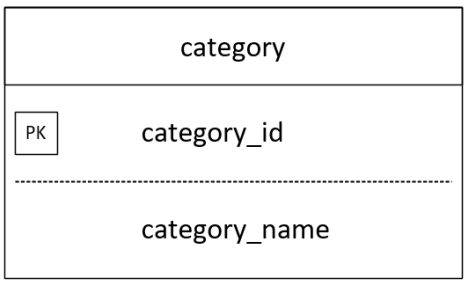
ภาพที่ 3.8 การกำหนดแอตทริบิวต์ของเอนทิตี้ continent

ภาพที่ 3.9 การกำหนดแอตทริบิวต์ของเอนทิตี้ country ประกอบด้วยแอตทริบิวต์ country_id, แอตทริบิวต์ country_name โดยมีคีย์หลัก (Primary key) คือแอตทริบิวต์ countrytb_id



ภาพที่ 3.9 การกำหนดแอตทริบิวต์ของเอนทิตี้ country

ภาพที่ 3.10 การกำหนดแอตทริบิวต์ของเอนทิตี้ category ประกอบด้วยแอตทริบิวต์ category_id, แอตทริบิวต์ category_name โดยมีคีย์หลัก (Primary key) คือแอตทริบิวต์ category_id



ภาพที่ 3.10 การกำหนดแอตทริบิวต์ของเอนทิตี้ category

3.3.3 การกำหนดความสัมพันธ์ระหว่างเอนทิตี้ (Relationships)

การสร้างความสัมพันธ์ระหว่างเอนทิตี้ในฐานข้อมูลเชิงสัมพันธ์กระทำได้ โดยการกำหนดให้เอนทิตี้ที่มีความสัมพันธ์กันมีแอททริบิวต์ที่เหมือนกัน และใช้ค่าของแอททริบิวต์ที่เหมือนกันเป็นตัวระบุข้อมูลในเอนทิตี้ที่มีความสัมพันธ์กัน ความสัมพันธ์ระหว่างเอนทิตี้แบ่งออกเป็น 3 ประเภท ดังนี้

1. ความสัมพันธ์แบบหนึ่งต่อหนึ่ง (One-to-one Relationship)

ความสัมพันธ์แบบหนึ่งต่อหนึ่งเป็นการแสดงความสัมพันธ์ของข้อมูลในเอนทิตี้หนึ่งว่ามีความสัมพันธ์กับข้อมูลของอีกเอนทิตี้หนึ่ง ในลักษณะที่เป็นหนึ่งต่อหนึ่ง เช่น ประชาชน 1 คน จะต้องมีหมายเลขรหัสประจำตัวประชาชน 1 หมายเลข ซึ่งไม่ซ้ำกัน

2. ความสัมพันธ์แบบหนึ่งต่อกลุ่ม (One-to-many Relationship)

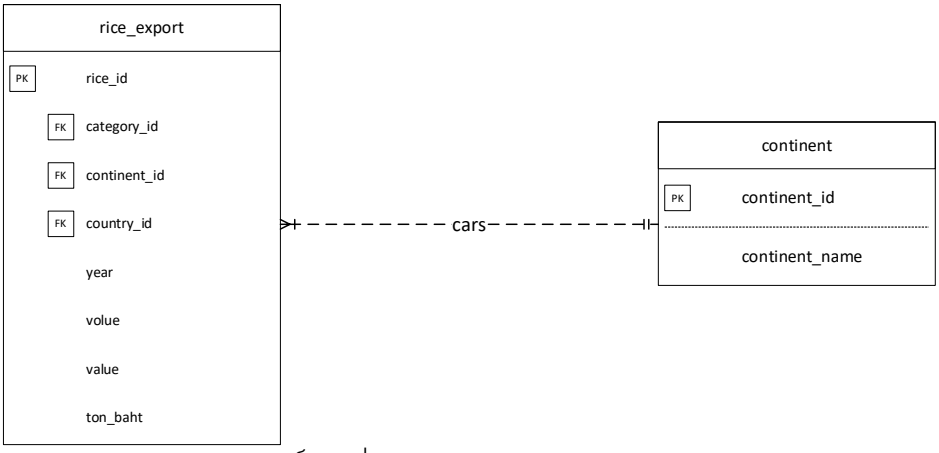
ความสัมพันธ์แบบหนึ่งต่อกลุ่ม เป็นการแสดงความสัมพันธ์ของข้อมูลของเอนทิตี้หนึ่ง ว่ามีความสัมพันธ์กับข้อมูลหลายข้อมูลกับอีกเอนทิตี้หนึ่ง เช่น แผนกแต่ละแผนกจะประกอบไปด้วยพนักงานที่สังกัดอยู่ในแผนกหลายคน

3. ความสัมพันธ์แบบกลุ่มต่อกลุ่ม (Many-to-many Relationship)

ความสัมพันธ์แบบกลุ่มต่อกลุ่ม เป็นการแสดงความสัมพันธ์ของข้อมูลของ 2 เอนทิตี้ ในลักษณะแบบกลุ่มต่อกลุ่ม เช่น นักศึกษาหลายคน อาจเรียนอยู่ในหลายหลักสูตร

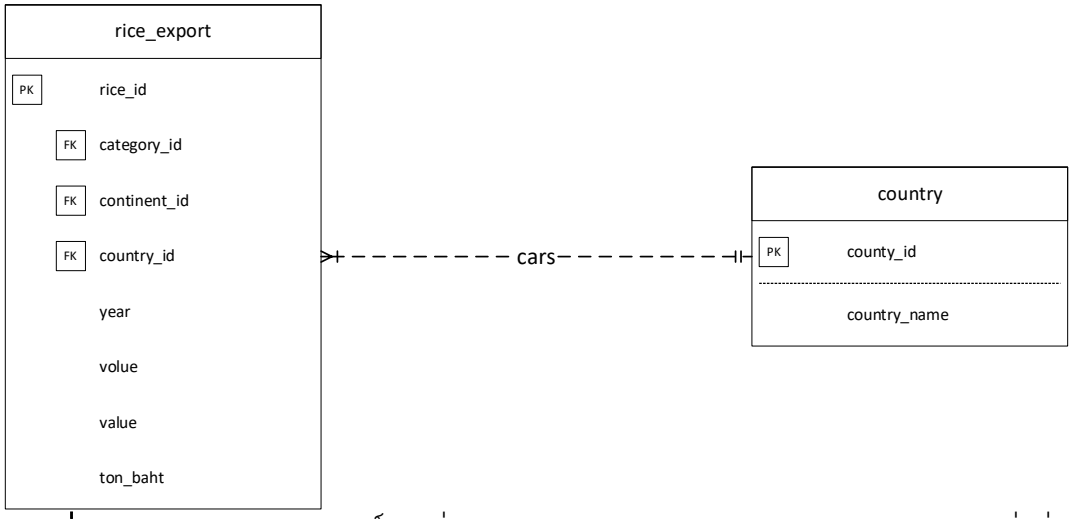
จากการศึกษาผู้วิเคราะห์ได้กำหนดความสัมพันธ์ระหว่างเอนทิตี้ของทั้ง 4 ตาราง ประกอบไปด้วยตารางที่ 1 รีเลชัน category, ตารางที่ 2 รีเลชัน continent, ตารางที่ 3 รีเลชัน country, และตารางที่ 4 รีเลชัน rice_export มีความสัมพันธ์ระหว่างเอนทิตี้ดังนี้

1) ตาราง rice_export มีความสัมพันธ์ระหว่างเอนทิตี้กับตาราง continent แบบหนึ่งต่อกลุ่ม (1 : N) หมายความว่า การส่งออกข้าว 1 ครั้ง สามารถส่งออกข้าวได้หลายทวีป



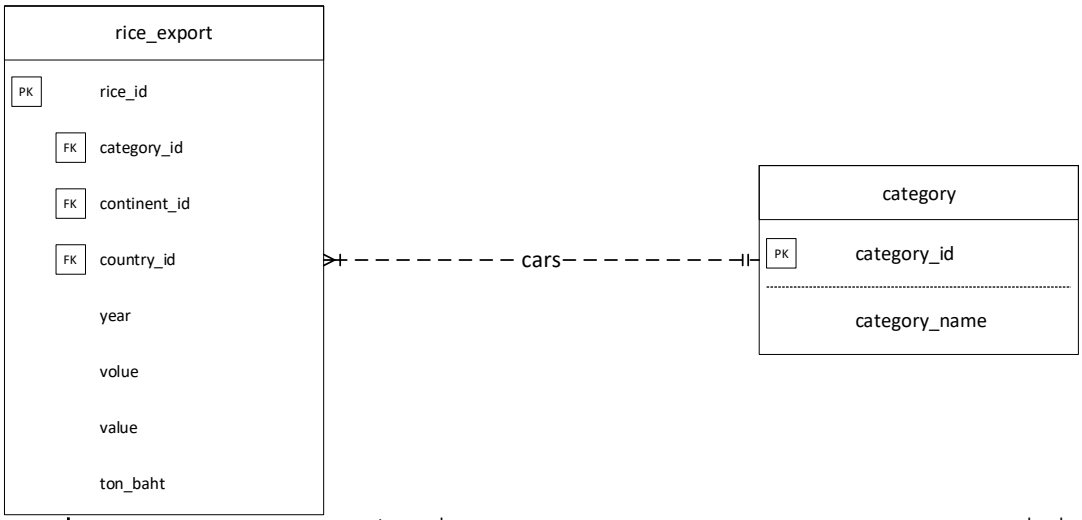
ภาพที่ 3.11 แสดงความสัมพันธ์ระหว่างตาราง rice_export กับตาราง continent แบบหนึ่งต่อกลุ่ม (1 : N)

2) ตาราง rice_export มีความสัมพันธ์ระหว่างเอนทิตี้กับตาราง country แบบหนึ่งต่อกลุ่ม (1 : N) หมายความว่า การส่งออกข้าว 1 ครั้ง สามารถส่งออกข้าวได้หลายประเทศ



ภาพที่ 3.12 แสดงความสัมพันธ์ระหว่างตาราง rice_export กับตาราง country แบบหนึ่งต่อกลุ่ม (1 : N)

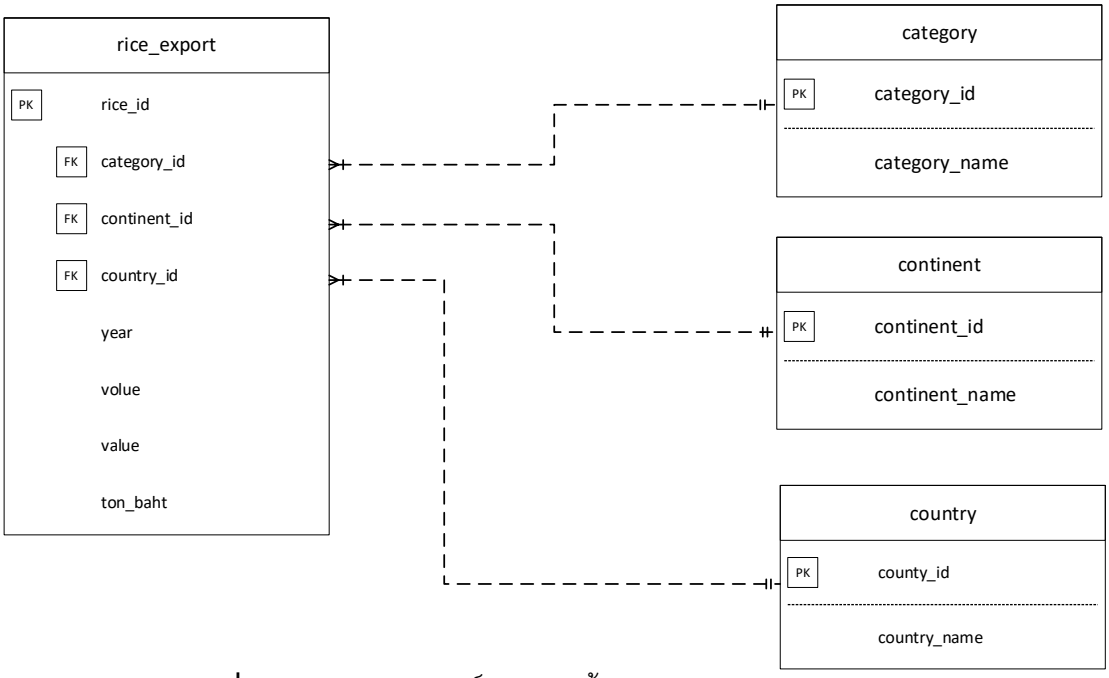
3.) ตาราง rice_export มีความสัมพันธ์ระหว่างเอนทิตี้กับตาราง category แบบหนึ่งต่อกลุ่ม (1 : N) หมายความว่า การส่งออกข้าว 1 ครั้ง สามารถส่งออกข้าวได้หลายประเภทข้าว



ภาพที่ 3.13 แสดงความสัมพันธ์ระหว่างตาราง rice_export กับตาราง category แบบหนึ่งต่อกลุ่ม (1 : N)

3.3.4 การเขียนเป็นแบบจำลองแผนภาพ Entity Relationship Diagram (ERD)

การนำรายละเอียดการออกแบบทั้งหมดตั้งแต่ขั้นที่ 1-3 นำมาวาดแบบจำลองแผนภาพ E-R Diagram กำหนดความสัมพันธ์ระหว่างเอนทิตี้ ดังต่อไปนี้



ภาพที่ 3.14 ความสัมพันธ์ของฐานข้อมูลแบบ Crow's Foot Model

3.3.4 พจนานุกรมข้อมูล (Data Dictionary)

จากการออกแบบฐานข้อมูลซึ่งมีการจัดการระบบฐานข้อมูลให้กับระบบที่ประกอบไปด้วยตารางข้อมูลต่าง ๆ ดังตารางที่ 3.18

ตารางที่ 3.18 แสดงชื่อตารางทั้งหมดของระบบฐานข้อมูล

ลำดับ	**ชื่อตาราง**	**ประเภท**	**รายละเอียด**
1	rice_export	Transaction	เก็บข้อมูลการส่งออกข้าว
2	continent	Reference	เก็บข้อมูลทวีป
3	country	Reference	เก็บข้อมูลประเทศ
4	category	Reference	เก็บข้อมูลประเภทข้าว

คำอธิบาย ประเภทของตาราง ได้แก่

Transaction หมายถึง ตารางที่มีการเปลี่ยนแปลงของข้อมูล

Reference หมายถึง ตารางที่มีการอ้างอิงถึงข้อมูล

จากตารางที่ 3.18 สามารถแสดงรายละเอียดของแต่ละตาราง rice_export, continent, country, category ได้ดังต่อไปนี้

ตารางที่ 3.19 แสดงรายละเอียดของตาราง rice_export

ชื่อตาราง : rice_export ประเภทตาราง : Transaction คำอธิบาย : เก็บข้อมูลการส่งออกข้าว คีย์หลัก : rice_id คีย์รอง : category_id, continent_id, country_id			
เขตข้อมูล	**ชนิดและขนาด**	**ความหมาย**	**ตัวอย่าง**
rice_id	INT(6)	รหัสการส่งออกข้าว	exp001
category_id	Char(6)	ประเภทข้าว	cate01
continent_id	Char(6)	ทวีป	con01
country_id	Char(6)	ประเทศ	cou81
year	year(4)	ปี	2013
volume	double	ปริมาณการส่งออก	0.013
value	double	มูลค่าการส่งออก	527346.56
Ton_Baht	double	ตันต่อบาท	26,324

ตารางที่ 3.20 แสดงรายละเอียดของตาราง continent

ชื่อตาราง : continent ประเภทตาราง : Reference คำอธิบาย : เก็บข้อมูลทวีป คีย์หลัก : continent_id คีย์รอง : –			
เขตข้อมูล	**ชนิดและขนาด**	**ความหมาย**	**ตัวอย่าง**
continent_id	INT(5)	รหัสทวีป	con01
continent_name	VARCHAR(25)	ชื่อทวีป	เอเชีย

ตารางที่ 3.21 แสดงรายละเอียดของตาราง country

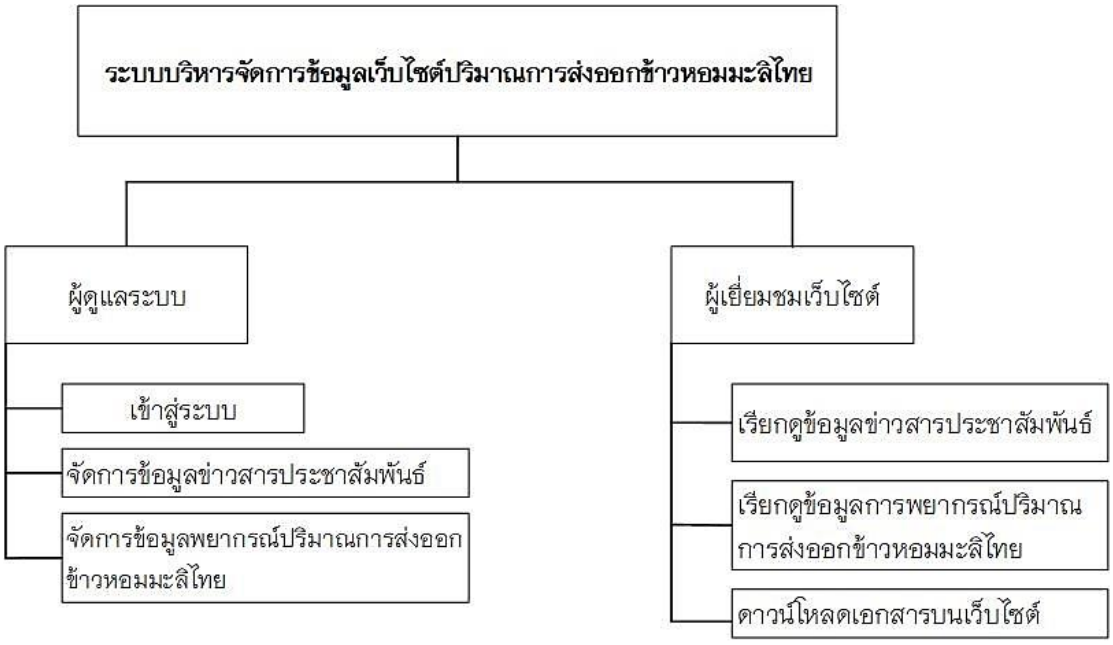
ชื่อตาราง : country ประเภทตาราง : Reference คำอธิบาย : เก็บข้อมูลประเทศ คีย์หลัก : country_id คีย์รอง : –			
เขตข้อมูล	**ชนิดและขนาด**	**ความหมาย**	**ตัวอย่าง**
country_id	INT(5)	รหัสประเทศ	cou81
country_name	VARCHAR(30)	ชื่อประเทศ	ญี่ปุ่น

ตารางที่ 3.22 แสดงรายละเอียดของตาราง category

ชื่อตาราง : category ประเภทตาราง : Reference คำอธิบาย : เก็บข้อมูลประเทศ คีย์หลัก : category_id คีย์รอง : –			
เขตข้อมูล	**ชนิดและขนาด**	**ความหมาย**	**ตัวอย่าง**
category_id	INT(5)	รหัสประเภทข้าว	cate01
category_name	VARCHAR(50)	ชื่อประเภทข้าว	ข้าวขาว100%ชั้น 1

3.4 โครงสร้างระบบ (System Structure)

ลักษณะภายในของโครงสร้างระบบแสดงถึงโครงสร้างการใช้งานในระบบฐานข้อมูลการส่งออกข้าวหอมมะลิไทย ส่วนที่เป็นส่วนสำคัญโดยตรง ได้แก่ ผู้ดูแลระบบ และผู้เยี่ยมชมเว็บไซต์ ดังภาพที่ 3.15

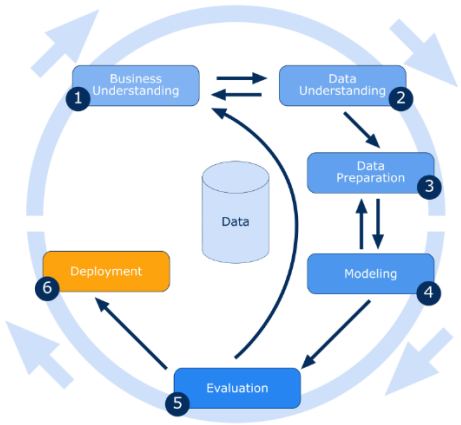


ภาพที่ 3.15 แผนผังโครงสร้างการใช้งานในระบบฐานข้อมูลการส่งออกข้าวหอมมะลิไทย

จากภาพที่ 3.15 เพื่อความสะดวกในการออกแบบฐานข้อมูลผู้วิเคราะห์จึงจัดแบ่งระบบออกเป็นส่วนย่อย ๆ มีทั้งหมด 2 ส่วนด้วยกัน และให้แต่ละส่วนมีหน้าที่รับผิดชอบการทำงานในแต่ละด้าน โดยไม่คาบเกี่ยวกันแต่สัมพันธ์กัน ส่วนแรกคือส่วนของผู้ดูแลระบบ จะเป็นส่วนที่สามารถจัดการข้อมูลทั้งหมดในระบบได้ ส่วนที่สองคือส่วนผู้เยี่ยมชมเว็บไซต์ เป็นส่วนที่บุคคลภายนอกสามารถเข้ามาเยี่ยมชมข้อมูลบนเว็บไซต์ได้และสามารถดาวน์โหลดข้อมูลบนเว็บไซต์ได้

3.5 กระบวนการ CRISP-DM

กระบวนการ CRISP-DM จะประกอบด้วย 6 ขั้นตอนแต่ละขั้นตอนในรูปจะเป็นขั้นตอนที่ต่อเนื่องกันนั่นคือขั้นตอนถัดไปจะรอผลลัพธ์จากขั้นตอนก่อนหน้าซึ่งแสดงด้วยลูกศรที่เชื่อมระหว่างกล่องสี่เหลี่ยมแต่ละกล่อง ตัวอย่างเช่น เมื่อได้ผลลัพธ์จากขั้นตอนการเตรียมข้อมูล (Data Preparation) แล้วจะนำไปสร้างโมเดลจำแนกประเภทข้อมูลในขั้น Modeling และหลังจากนั้นอาจจะย้อนกลับมาเปลี่ยนแปลงข้อมูลให้ถูกต้องมากขึ้นเพื่อหวังว่าจะได้โมเดลที่ให้ความถูกต้องมากขึ้น



ภาพที่ 3.16 แสดงกระบวนการ CRISP-DM

ที่มา : barnraisersllc.com (2561) , 12 ธันวาคม 2563.

โดยกระบวนการ CRISP-DM จะประกอบไปด้วย 6 ขั้นตอน ดังนี้

1.) Business Understanding

เป็นขั้นตอนแรกที่ทำความเข้าใจปัญหาที่ได้ให้อยู่ในรูปโจทย์ของการวิเคราะห์ข้อมูลทางดาต้า ไมน์นิ่ง และวางแผนในการดำเนินการโดยการวิเคราะห์ข้อมูลรูปแบบความสัมพันธ์ที่ซ่อนอยู่ในข้อมูลของข้อมูล ขั้นตอนการทำความเข้าใจธุรกิจต้องเข้าใจวัตถุประสงค์ทางธุรกิจอย่างชัดเจนและค้นหาว่าอะไรคือความต้องการของธุรกิจจากนั้นให้ประเมินสถานการณ์ปัจจุบันโดยการค้นหาทรัพยากรสมมติฐานข้อ จำกัด และปัจจัยสำคัญอื่น ๆ ที่ควรพิจารณาจากวัตถุประสงค์ทางธุรกิจและสถานการณ์ปัจจุบันให้สร้างเป้าหมายการขุดข้อมูลเพื่อให้บรรลุวัตถุประสงค์ทางธุรกิจภายในสถานการณ์ปัจจุบัน สุดท้ายต้องมีการกำหนดแผนการขุดข้อมูลที่ดีเพื่อให้บรรลุเป้าหมายทั้งทางธุรกิจและการขุดข้อมูล แผนควรมีรายละเอียดมากที่สุด

2.) Data Understanding

เริ่มทำการเก็บรวบรวมข้อมูล และตรวจสอบข้อมูลที่ได้ ดูความถูกต้อง และความน่าเชื่อถือของข้อมูล นำมาพิจารณาว่าจะใช้ข้อมูลทั้งหมดหรือจำเป็นต้องเลือกข้อมูลบางส่วนมาใช้ในการวิเคราะห์ ขั้นตอนการทำความเข้าใจข้อมูลเริ่มต้นด้วยการรวบรวมข้อมูลเบื้องต้นซึ่งรวบรวมจากแหล่งข้อมูลที่มีอยู่เพื่อช่วยให้คุ้นเคยกับข้อมูล กิจกรรมที่สำคัญบางอย่างต้องดำเนินการรวมถึงการโหลดข้อมูลและการรวมข้อมูลเพื่อให้การรวบรวมข้อมูลประสบความสำเร็จถัดไปคุณสมบัติ “ขั้นต้น” หรือ "พื้นผิว" ของข้อมูลที่ได้มาจะต้องได้รับการตรวจสอบอย่างรอบคอบจากนั้นข้อมูลจะต้องได้รับการสำรวจโดยการจัดการกับคำถามเกี่ยวกับการขุดข้อมูลซึ่งสามารถแก้ไขได้โดยใช้การสอบถามการรายงานและการแสดงภาสุดท้ายนี้ต้องตรวจสอบคุณภาพของข้อมูลโดยตอบคำถามสำคัญบางอย่างเช่น ข้อมูลที่ได้มานั้นสมบูรณ์หรือไม่มีค่าที่ขาดหายไปในข้อมูลที่ได้มาหรือไม่

3.) Data Preparation

ทำการแปลงข้อมูลที่ได้ทำการเก็บรวบรวมมาให้กลายเป็นข้อมูลที่สามารถนำไปวิเคราะห์ในขั้นถัดไปได้ โดยการการทำข้อมูลให้ถูกต้องโดยใช้กระบวนการ data cleaning เพื่อให้ได้ข้อมูลอยู่ในรูปแบบและมาตรฐานเดียวกัน และกำจัดข้อมูลเสียออก โดยทั่วไปการเตรียมข้อมูลจะใช้เวลาประมาณ 90% ของเวลาของโครงการ ผลลัพธ์ของขั้นตอนการเตรียมข้อมูลคือชุดข้อมูลสุดท้าย เมื่อระบุแหล่งข้อมูลที่พร้อมใช้งานแล้วจะต้องมีการเลือกทำความสะอาดสร้างและจัดรูปแบบให้อยู่ในรูปแบบที่ต้องการ งานสำรวจข้อมูลที่มีความลึกมากขึ้นอาจดำเนินการในช่วงนี้เพื่อสังเกตเห็นรูปแบบตามความเข้าใจทางธุรกิจ

4.) Modeling

ขั้นตอนนี้ต้องเลือกเทคนิคการสร้างแบบจำลองเพื่อใช้กับชุดข้อมูลที่เตรียมไว้ จากนั้นต้องสร้างสถานการณ์การทดสอบเพื่อตรวจสอบคุณภาพและความถูกต้องของแบบจำลองโมเดลอย่างน้อยหนึ่งแบบจะถูกสร้างขึ้นบนชุดข้อมูลที่เตรียมไว้ สุดท้ายโมเดลต้องได้รับการประเมินอย่างรอบคอบเกี่ยวกับผู้มีส่วนได้ส่วนเสียเพื่อให้แน่ใจว่าแบบจำลองที่สร้างขึ้นนั้นตรงตามความคิดริเริ่มทางธุรกิจเป็นการวิเคราะห์ข้อมูลด้วยเทคนิคทาง Data Mining

เช่น การจำแนกประเภทข้อมูล หรือแบ่งกลุ่มข้อมูล ซึ่งขั้นตอนนี้หลานเทคนิคจะถูกนำมาใช้เพื่อได้คำตอบที่ดีที่สุด

5.) Evaluation

ขั้นตอนการประเมินผลของแบบจำลองจะต้องได้รับการประเมินในบริบทของวัตถุประสงค์ทางธุรกิจในระยะแรก ในระยะนี้ความต้องการทางธุรกิจใหม่อาจเพิ่มขึ้นเนื่องจากรูปแบบใหม่ที่ถูกค้นพบในผลลัพธ์ของแบบจำลองหรือจากปัจจัยอื่น ๆ การทำความเข้าใจธุรกิจเป็นกระบวนการซ้ำ ๆ ในการตัดสินใจทำหรือไม่ทำในขั้นตอนนี้เพื่อย้ายไปยังขั้นตอนการปรับใช้ผลการวิเคราะห์ข้อมูลด้วยเทคนิคทาง Data Mining จากการสร้างโมเดล time series สามารถวัดประสิทธิภาพของผลลัพธ์ที่ได้ว่าตรงกับวัตถุประสงค์ที่ได้ตั้งไว้

6.) Deployment

ความรู้หรือข้อมูลที่ได้รับจากกระบวนการวิเคราะห์จำเป็นต้องนำเสนอในลักษณะที่ผู้มีส่วนได้ส่วนเสียสามารถนำไปใช้ได้เมื่อพวกเขาต้องการ ในขั้นตอนการปรับใช้ต้องมีการสร้างแผนสำหรับการปรับใช้การบำรุงรักษาและการตรวจสอบสำหรับการนำไปใช้งานและการสนับสนุนในอนาคตด้วย จากมุมมองของโครงการรายงานขั้นสุดท้ายของโครงการจำเป็นต้องสรุปประสบการณ์ของโครงการและทบทวนโครงการเพื่อดูสิ่งที่ต้องปรับปรุงบทเรียนที่สร้างขึ้นการนำผลองค์ความรู้ที่ได้จากการหาผลลัพธ์ด้วยเทคนิค Data Mining ไปใช้ ประโยชน์ต่อในองค์กร หรือบริษัท

ขั้นตอนในการดำเนินการวิเคราะห์ข้อมูล ผู้วิเคราะห์ได้เลือกวิธีการวิเคราะห์ข้อมูลโดยใช้โมเดล Time Series เพื่อต้องการทราบข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทยของปีถัดไป ที่จะมีแนวโน้มเพิ่มขึ้นหรือลดลง ใช้ในการตัดสินใจในการผลิตหรือสต๊อกข้าว และข้อมูลที่นำมาวิเคราะห์เป็นข้อมูลที่เปลี่ยนแปลงไปตามเวลาที่มีการส่งออกข้าว มีการจัดเก็บข้อมูลเป็นช่วงเวลามีการเพิ่มขึ้นหรือลดลงอย่างต่อเนื่อง สามารถนำข้อมูลมาพยากรณ์หรือทำนายข้อมูลที่จะเกิดขึ้นในอนาคตได้ จากการที่ผู้วิเคราะห์ได้ศึกษาวรรณกรรมที่เกี่ยวข้องงานวิจัยที่พบส่วนใหญ่จะใช้โมเดล Time Series ในการวิเคราะห์ข้อมูล

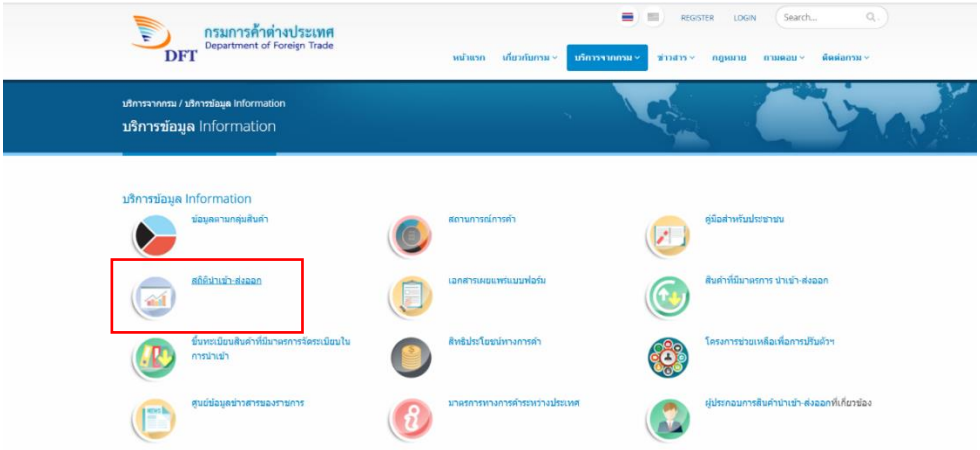
3.5.1 ความเข้าใจในธุรกิจ (Business Understanding)

ผู้วิเคราะห์ทำความเข้าใจกระบวนการทางข้อมูล และรับรู้ปัญหารวมถึงการมองหาปัญหาจากเรื่องต่าง ๆ เพื่อที่นำผลการวิเคราะห์ข้อมูลไปใช้ให้เกิดประโยชน์ โดยเรียงลำดับความสำคัญ และกำหนดวัตถุประสงค์ที่จะนำไปสู่รูปแบบการวิเคราะห์ข้อมูล ให้อยู่ในรูปโจทย์ของการวิเคราะห์ฐานข้อมูล และวางแผนในการดำเนินการโดยการวิเคราะห์ข้อมูลรูปแบบความสัมพันธ์ที่ซ่อนอยู่ในข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย เพื่อที่จะนำผลการวิเคราะห์ข้อมูลไปใช้ประโยชน์ตามความต้องการ และตามวัตถุประสงค์ที่จะนำไปสู่รูปแบบการวิเคราะห์ข้อมูล

3.5.2 การทำความเข้าใจข้อมูล (Data Understanding)

ผู้วิเคราะห์ข้อมูลทำการจัดเก็บรวบรวมข้อมูล และพิจารณาตรวจสอบข้อมูลที่ได้รับดูความถูกต้อง มีความน่าเชื่อถือของข้อมูล โดยพิจารณาว่าจะใช้ข้อมูลทั้งหมดหรือจำเป็นต้องเลือกข้อมูลบางส่วนมาใช้ในการวิเคราะห์ ให้สอดคล้องกับวัตถุประสงค์ที่กำหนดไว้ ผู้วิเคราะห์ได้สำรวจภาพรวมของข้อมูลก่อนนำข้อมูลนั้นมาใช้ในการวิเคราะห์ จากเว็บไซต์ https://www.dft.go.th/th-th/index ซึ่งเป็นศูนย์กลางการให้บริการของกรมการค้าต่างประเทศ เป็นข้อมูลเปิดของภาครัฐ (Open Data) ผู้วิเคราะห์ข้อมูลจึงได้ดำเนินการตามขั้นตอนดังนี้

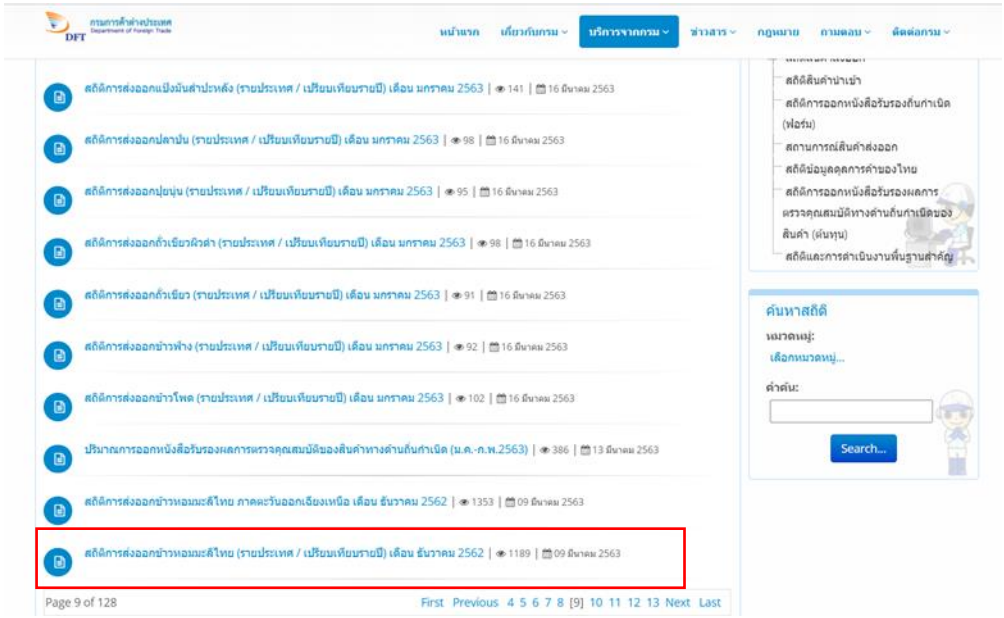
1) ผู้วิเคราะห์รวบรวมข้อมูลจากเว็บไซต์ https://www.dft.go.th/th-th/index ผู้วิเคราะห์ทำการ เลือกเมนูบริการข้อมูล Information กลุ่มชุดข้อมูลสถิตินำเข้า-ส่งออก



ภาพที่ 3.17 แสดงขั้นตอนการหาข้อมูลจากเว็บไซต์

ที่มา : dft.go.th (2563) , 12 ธันวาคม 2563.

2) ผู้วิเคราะห์เลือกหาเอกสารที่เกี่ยวข้องกับข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทยรายประเภทตามปีที่มีการส่งออก ทำการจัดเก็บ และรวบรวมข้อมูลทั้งหมด ของข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทยในช่วงปี 2013–2019



ภาพที่ 3.18 แสดงขั้นตอนการดาวน์โหลดข้อมูล

ที่มา : dft.go.th (2563) , 12 ธันวาคม 2563.

3) การนำข้อมูลมาตรวจสอบความถูกต้อง และความน่าเชื่อถือ โดยไฟล์ข้อมูลจะบันทึกข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทยรายปี โดยมีข้อมูล 5 แอตทริบิวท์ คือ ทวีป, ประเทศ, ปริมาณการส่งออก, มูลค่าการส่งออก, ราคาตัน/บาท และแยกเป็นประเภทของข้าว 4 ประเภท คือ ข้าวขาว100%ชั้น 1, ข้าวขาวหักเอวันเลิศพิเศษ, ข้าวขาวหักเอวันเลิศ, ข้าวกล้อง 100%ชั้น 1 ดังนี้

สถิติการส่งออกสินค้ามาตรฐานข้าวหอมมะลิไทยตามใบรับรองมาตรฐานสินค้า ปี 2013

ปริมาณ : พันเมตริกตัน

มูลค่า : ล้านบาท

ข้าวขาว 100% ชั้น 1

ประเทศ		ปี 2013		
		ปริมาณ	มูลค่า	ตัน/บาท
เอเชีย	ญี่ปุ่น	0.011	408.474	37,134
	มาเลเซีย	0.022	751.718	34,169
	สหรัฐอาหรับเอมิเรตส์	0.024	856.160	35,673
	สาธารณรัฐประชาชนจีน	0.028	1,173.431	41,318
	สาธารณรัฐสังคมนิยมเวียดนาม	0.014	672.311	48,022
	ฮ่องกง	0.033	1,808.169	54,793
	รวม	0.132	5,670.263	251,110
ยุโรป	ตุรกี	0.019	726.997	38,263
	นอร์เวย์	0.040	1,504.080	37,602
	เนเธอร์แลนด์	0.021	894.159	42,579
	ฝรั่งเศส	0.027	1,400.517	51,871
	สวิตเซอร์แลนด์	0.045	1,625.490	36,122
	สวีเดน	0.033	1,256.409	38,073
	สหพันธ์สาธารณรัฐเยอรมัน	0.041	1,573.990	38,390
	รวม	0.226	8,981.642	282,900

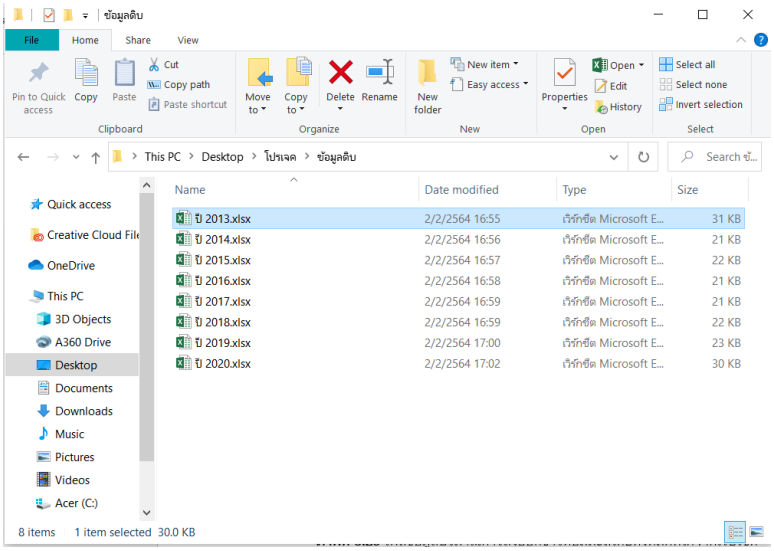
ภาพที่ 3.19 ไฟล์ข้อมูลการบันทึกปริมาณการส่งออกข้าวหอมมะลิไทย

ขั้นตอนการแปลงข้อมูลที่ได้รวบรวม ให้อยู่ในรูปแบบที่พร้อมนำไปวิเคราะห์ในขั้นตอนต่อไป โดยการทำเป็นข้อมูลที่ถูกต้อง (Data cleaning) เป็นกระบวนการตรวจสอบ และแก้ไขรายการข้อมูลที่ไม่ถูกต้องออกไปจากชุดข้อมูล ตารางหรือฐานข้อมูลซึ่งเป็นหลักสำคัญของฐานข้อมูล เพราะหมายถึงความไม่สมบูรณ์ ความไม่ถูกต้อง ความไม่สัมพันธ์กับข้อมูลอื่น ๆ เป็นต้น จึงต้องมีการแทนที่ การปรับปรุง การลบข้อมูลที่ไม่ถูกต้องเหล่านี้ออกไป เพื่อให้ข้อมูลมีคุณภาพ และระบบการรับข้อมูลเข้าสู่ระบบที่ทันสมัยในปัจจุบันจะลดการคีย์ข้อมูลจากคนให้น้อยที่สุด แต่จะใช้วิธีการสแกน การเลือก เพื่อลดความผิดพลาดให้น้อยที่สุด การลดข้อผิดพลาดของข้อมูลได้มากเท่าใดก็จะยิ่งมีประสิทธิภาพมากขึ้น

3.5.3 การเตรียมข้อมูล (Data Preparation)

ผู้วิเคราะห์ข้อมูลทำการแปลงข้อมูลที่ได้ทำการเก็บรวบรวม และทำการจัดการข้อมูลที่ได้ให้อยู่ในรูปแบบที่สามารถนำไปวิเคราะห์ในขั้นถัดไปได้ เพื่อลดความผิดพลาดการทำข้อมูลให้ถูกต้อง โดยใช้กระบวนการ data cleaning เพื่อให้ได้ข้อมูลอยู่ในรูปแบบและมาตรฐานเดียวกัน และกำจัดข้อมูลเสียออก ผู้วิเคราะห์ข้อมูลจึงได้ดำเนินการตามขั้นตอนดังนี้

1) การปรับมาตรฐาน (Standardizing) ผู้วิเคราะห์ข้อมูลทำการแปลงข้อมูลให้อยู่ในรูปแบบเดียวกันพร้อมทั้งกำหนดแอตทริบิวท์เพื่อให้เป็นคีย์หลักของแต่ละไฟล์



ภาพที่ 3.20 ไฟล์ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทยทั้งหมดที่ได้จากเว็บไซต์

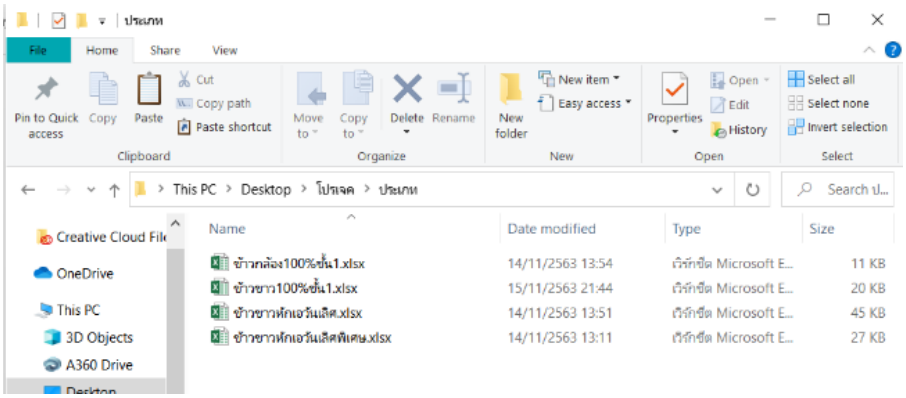
ผู้วิเคราะห์ได้ทำการจัดการข้อมูลให้อยู่ในรูปแบบเดียวกัน โดยการรวบรวมข้อมูลจากรายปีให้เป็นประเภทของข้าว เพื่อให้ข้อมูลสามารถนำไปพยากรณ์ตามวัตถุประสงค์ที่กำหนดไว้ข้างต้น

category_name	year	volume
ข้าวกล่อง 100% ชั้น 1	2013	0.005
ข้าวกล่อง 100% ชั้น 1	2013	0.006
ข้าวกล่อง 100% ชั้น 1	2013	0.006
ข้าวกล่อง 100% ชั้น 1	2013	0.007
ข้าวกล่อง 100% ชั้น 1	2013	0.005
ข้าวกล่อง 100% ชั้น 1	2014	0.007
ข้าวกล่อง 100% ชั้น 1	2014	0.011
ข้าวกล่อง 100% ชั้น 1	2014	0.008
ข้าวกล่อง 100% ชั้น 1	2014	0.005
ข้าวกล่อง 100% ชั้น 1	2014	0.006
ข้าวกล่อง 100% ชั้น 1	2015	0.006
ข้าวกล่อง 100% ชั้น 1	2015	0.007
ข้าวกล่อง 100% ชั้น 1	2015	0.007
ข้าวกล่อง 100% ชั้น 1	2015	0.006
ข้าวกล่อง 100% ชั้น 1	2015	0.005
ข้าวกล่อง 100% ชั้น 1	2016	0.008
ข้าวกล่อง 100% ชั้น 1	2016	0.010
ข้าวกล่อง 100% ชั้น 1	2016	0.007
ข้าวกล่อง 100% ชั้น 1	2016	0.012
ข้าวกล่อง 100% ชั้น 1	2016	0.011

ภาพที่ 3.21 ข้อมูลที่รวบรวมให้เป็นประเภทในแต่ละปี

2.) การจัดหมวดหมู่ (Transform) ผู้วิเคราะห์นำ Dataset มาจัดหมวดหมู่เพื่อความถูกต้อง จัดกลุ่มข้อมูลแบ่งตามประเภทข้าว กำหนดให้ category_name เป็นคีย์หลักเป็นตัวที่ใช้ระบุความสัมพันธ์ของข้อมูลที่จะแสดงปริมาณการส่งออกข้าวในแต่ละประเทศ ตามรายปี เพื่อเตรียมความพร้อมการนำข้อมูลไปวิเคราะห์แบบอนุกรมเวลาบนโปรแกรม Tableau Public โปรแกรม Minitab และโปรแกรม IBM SPSS โดยดำเนินตามขั้นตอนดังนี้

2.1 การจัดกลุ่มตามประเภทข้าว



ภาพที่ 3.22 ไฟล์ข้อมูลที่ต้องการแยกกลุ่มให้อยู่ในไฟล์ประเภท

2.2 นำข้อมูลที่ผ่านการรวบรวมเป็นรายเดือนมาไว้ในไฟล์ภูมิภาคเดียวกัน

category_name	year	continent_name	country_name	volume	value	Ton / Baht
ข้าวขาว 100% ชั้น 1	2013	เอเชีย	ญี่ปุ่น	0.011	0.000408474	37,134
ข้าวขาว 100% ชั้น 1	2013	เอเชีย	มาเลเซีย	0.022	0.000751718	34,169
ข้าวขาว 100% ชั้น 1	2013	เอเชีย	สหรัฐอาหรับเอมิเรตส์	0.024	0.000847440	35,310
ข้าวขาว 100% ชั้น 1	2013	เอเชีย	สาธารณรัฐประชาชนจีน	0.028	0.000900223	31698
ข้าวขาว 100% ชั้น 1	2013	เอเชีย	สาธารณรัฐสังคมนิยมเวียดนาม	0.014	0.000578452	41,318
ข้าวขาว 100% ชั้น 1	2013	เอเชีย	ฮ่องกง	0.033	0.001808169	54,793
ข้าวขาว 100% ชั้น 1	2013	ยุโรป	ตุรกี	0.019	0.000729410	38,390
ข้าวขาว 100% ชั้น 1	2013	ยุโรป	นอร์เวย์	0.040	0.001504080	37,602
ข้าวขาว 100% ชั้น 1	2013	ยุโรป	เนเธอร์แลนด์	0.021	0.000894159	42,579
ข้าวขาว 100% ชั้น 1	2013	ยุโรป	ฝรั่งเศส	0.027	0.000979830	36,290
ข้าวขาว 100% ชั้น 1	2013	ยุโรป	สวิตเซอร์แลนด์	0.045	0.002334195	51,871
ข้าวขาว 100% ชั้น 1	2013	ยุโรป	สวีเดน	0.033	0.001266870	38,390
ข้าวขาว 100% ชั้น 1	2013	ยุโรป	สหพันธ์สาธารณรัฐเยอรมัน	0.041	0.001560993	38,073
ข้าวขาว 100% ชั้น 1	2013	อเมริกาเหนือ	แคนาดา	0.042	0.001403598	33,419
ข้าวขาว 100% ชั้น 1	2013	แอฟริกา	กินี	0.040	0.001518400	37,960
ข้าวขาว 100% ชั้น 1	2013	แอฟริกา	คอโมโรส	0.035	0.001080380	30,868
ข้าวขาว 100% ชั้น 1	2013	แอฟริกา	สาธารณรัฐคองโก	0.057	0.001936062	33,966
ข้าวขาว 100% ชั้น 1	2013	แอฟริกา	แองโกลา	0.042	0.001443456	34,368
ข้าวขาว 100% ชั้น 1	2013	ออสเตรเลียและโอเชียนเนีย	นิวซีแลนด์	0.078	0.002818608	36,136
ข้าวขาว 100% ชั้น 1	2013	ออสเตรเลียและโอเชียนเนีย	ออสเตรเลีย	0.076	0.002913688	38,338
ข้าวขาว 100% ชั้น 1	2014	เอเชีย	ญี่ปุ่น	0.017	0.000689607	40,565
ข้าวขาว 100% ชั้น 1	2014	เอเชีย	มาเลเซีย	0.024	0.000780246	32,510
ข้าวขาว 100% ชั้น 1	2014	เอเชีย	สหรัฐอาหรับเอมิเรตส์	0.022	0.000816948	37,134
ข้าวขาว 100% ชั้น 1	2014	เอเชีย	สาธารณรัฐประชาชนจีน	0.027	0.000915732	33,916
ข้าวขาว 100% ชั้น 1	2014	เอเชีย	สาธารณรัฐสังคมนิยมเวียดนาม	0.025	0.001368475	54,739
ข้าวขาว 100% ชั้น 1	2014	เอเชีย	ฮ่องกง	0.031	0.001796021	57,936
ข้าวขาว 100% ชั้น 1	2014	ยุโรป	ตุรกี	0.031	0.001186153	38,263
ข้าวขาว 100% ชั้น 1	2014	ยุโรป	นอร์เวย์	0.036	0.001271962	35,332

ภาพที่ 3.23 แสดงการจัดกลุ่มประเภทข้าว

3.5.4 การสร้างแบบจำลอง (Modeling)

ผู้วิเคราะห์ข้อมูลได้วิเคราะห์ด้วยเทคนิคทางสถิติ ทำการเรียนรู้จากข้อมูลเดิมของการส่งออกข้าวหอมมะลิเพื่อให้ได้คำตอบที่ดีที่สุด การเปรียบเทียบความคลาดเคลื่อนของวิธีการพยากรณ์ โดยนำโมเดล Time series analysis มาช่วยในการหาค่าความคลาดเคลื่อนของการพยากรณ์ในวิธีต่าง Single Exponential Moving Average โดยนำมาเปรียบเทียบกับโปรแกรมว่าโปรแกรมใดให้ค่าความถูกต้องแม่นยำมากที่สุดนำมาใช้ในการพยากรณ์ข้อมูลการส่งออกข้าวหอมมะลิไทย การพยากรณ์ในรูปแบบคณิตศาสตร์ใช้รูปแบบสมการทางคณิตศาสตร์ ผู้วิเคราะห์ข้อมูลได้เลือกใช้เทคนิครูปแบบอนุกรมเวลา Time series Models ด้วยชุดข้อมูลปริมาณการส่งออกของแต่ละประเภทข้าว โดยมีขั้นตอนการดำเนินการดังนี้

ในการพยากรณ์โดยใช้รูปแบอนุกรมเวลาจะเป็นการพยากรณ์ข้อมูลในอนาคตจากข้อมูลในอดีตเท่านั้น ตัวแปรอื่น ๆ จะไม่นำมาพิจารณา ผู้วิเคราะห์ข้อมูลได้เลือกข้อมูลในรูปแบบมีแนวโน้ม (Trend)

category_name	year	continent_name	country_name	volume	value	Ton / Baht
ข้าวขาว 100% ชั้น 1	2013	เอเชีย	ญี่ปุ่น	0.011	0.000408474	37,134
ข้าวขาว 100% ชั้น 1	2013	เอเชีย	มาเลเซีย	0.022	0.000751718	34,169
ข้าวขาว 100% ชั้น 1	2013	เอเชีย	สหรัฐอาหรับเอมิเรตส์	0.024	0.000847440	35,310
ข้าวขาว 100% ชั้น 1	2013	เอเชีย	สาธารณรัฐประชาชนจีน	0.028	0.000900223	31698
ข้าวขาว 100% ชั้น 1	2013	เอเชีย	สาธารณรัฐสังคมนิยมเวียดนาม	0.014	0.000578452	41,318
ข้าวขาว 100% ชั้น 1	2013	เอเชีย	ฮ่องกง	0.033	0.001808169	54,793
ข้าวขาว 100% ชั้น 1	2013	ยุโรป	ตุรกี	0.019	0.000729410	38,390
ข้าวขาว 100% ชั้น 1	2013	ยุโรป	นอร์เวย์	0.040	0.001504080	37,602
ข้าวขาว 100% ชั้น 1	2013	ยุโรป	เนเธอร์แลนด์	0.021	0.000894159	42,579
ข้าวขาว 100% ชั้น 1	2013	ยุโรป	ฝรั่งเศส	0.027	0.000979830	36,290
ข้าวขาว 100% ชั้น 1	2013	ยุโรป	สวิตเซอร์แลนด์	0.045	0.002334195	51,871
ข้าวขาว 100% ชั้น 1	2013	ยุโรป	สวีเดน	0.033	0.001266870	38,390
ข้าวขาว 100% ชั้น 1	2013	ยุโรป	สหพันธ์สาธารณรัฐเยอรมัน	0.041	0.001560993	38,073
ข้าวขาว 100% ชั้น 1	2013	อเมริกาเหนือ	แคนาดา	0.042	0.001403598	33,419
ข้าวขาว 100% ชั้น 1	2013	แอฟริกา	กินี	0.040	0.001518400	37,960
ข้าวขาว 100% ชั้น 1	2013	แอฟริกา	คอโมโรส	0.035	0.001080380	30,868
ข้าวขาว 100% ชั้น 1	2013	แอฟริกา	สาธารณรัฐคองโก	0.057	0.001936062	33,966
ข้าวขาว 100% ชั้น 1	2013	แอฟริกา	แองโกลา	0.042	0.001443456	34,368
ข้าวขาว 100% ชั้น 1	2013	ออสเตรเลียและโอเชียนเนีย	นิวซีแลนด์	0.078	0.002818608	36,136
ข้าวขาว 100% ชั้น 1	2013	ออสเตรเลียและโอเชียนเนีย	ออสเตรเลีย	0.076	0.002913688	38,338
ข้าวขาว 100% ชั้น 1	2014	เอเชีย	ญี่ปุ่น	0.017	0.000689607	40,565
ข้าวขาว 100% ชั้น 1	2014	เอเชีย	มาเลเซีย	0.024	0.000780246	32,510
ข้าวขาว 100% ชั้น 1	2014	เอเชีย	สหรัฐอาหรับเอมิเรตส์	0.022	0.000816948	37,134
ข้าวขาว 100% ชั้น 1	2014	เอเชีย	สาธารณรัฐประชาชนจีน	0.027	0.000915732	33,916
ข้าวขาว 100% ชั้น 1	2014	เอเชีย	สาธารณรัฐสังคมนิยมเวียดนาม	0.025	0.001368475	54,739
ข้าวขาว 100% ชั้น 1	2014	เอเชีย	ฮ่องกง	0.031	0.001796021	57,936
ข้าวขาว 100% ชั้น 1	2014	ยุโรป	ตุรกี	0.031	0.001186153	38,263
ข้าวขาว 100% ชั้น 1	2014	ยุโรป	นอร์เวย์	0.036	0.001271962	35,332

ภาพที่ 3.24 ตารางข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย

จากรูปภาพที่ 3.32 ประกอบไปด้วย 5 แอตทริบิวท์ คือ

1.) category_name เก็บข้อมูลประเภทข้าวทั้งหมดที่ส่งออกในแต่ละปี ทำหน้าที่เป็นคีย์หลักในการอ้างอิงถึงข้อมูลอื่น ๆ

2.) Year ปีของปริมาณการส่งออกข้าวหอมมะลิไทยตั้งแต่ ปี 2013 – 2019

3.) continent เก็บข้อมูลทวีปทั้งหมด 6 ทวีป

4.) country เก็บข้อมูลประเทศทั้งหมด 54 ประเทศ

5.) volume เก็บค่าปริมาณการส่งออกข้าวหอมมะลิไทย

6) Ton/Baht เก็บค่าราคาข้าวต่อตันหน่วยเป็นบาท

การสร้างโมเดล Time series จะทำการคัดเลือกแอตทริบิวท์ข้อมูลเชิงปริมาณที่จัดเก็บในช่วงเวลาหนึ่ง หลังจากนั้นก็คำนวณหาค่าความเป็นไปได้เชิงพยากรณ์ในการหาค่าคำนวณได้จากสมการทางคณิตศาสตร์ ดังนี้

1) โดยการใช้เทคนิค Single Exponential Moving Average

เป็นเส้นค่าเฉลี่ยเคลื่อนที่ เส้นค่าเฉลี่ยเคลื่อนที่ EMA เป็นการคำนวณในลักษณะถ่วงน้ำหนักที่ให้ความสำคัญกับตัวแปรที่ทำให้มีผลต่อการเปลี่ยนแปลงของข้อมูลค่อนข้างเร็ว และการถ่วงน้ำหนักจะให้ค่าสุดท้ายมีความสำคัญเพิ่มขึ้น

ขั้นตอนที่ 1 คำนวณหาค่า Smoothing Factor จากสมการ

$$2 / (n + 1)$$

ขั้นตอนที่ 2 คำนวณ EMA

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

category_name	year	continent	country	volume	EMA
ข้าวขาว 100% ชั้น 1	2013	เอเชีย	ญี่ปุ่น	0.011	0.0140
ข้าวขาว 100% ชั้น 1	2014	เอเชีย	ญี่ปุ่น	0.013	0.0120
ข้าวขาว 100% ชั้น 1	2015	เอเชีย	ญี่ปุ่น	0.016	0.0127
ข้าวขาว 100% ชั้น 1	2016	เอเชีย	ญี่ปุ่น	0.012	0.0149
ข้าวขาว 100% ชั้น 1	2017	เอเชีย	ญี่ปุ่น	0.014	0.0130
ข้าวขาว 100% ชั้น 1	2018	เอเชีย	ญี่ปุ่น	0.018	0.0137
ข้าวขาว 100% ชั้น 1	2019	เอเชีย	ญี่ปุ่น	0.019	0.0165
ข้าวขาว 100% ชั้น 1	2013	เอเชีย	มาเลเซีย	0.022	0.0182
ข้าวขาว 100% ชั้น 1	2014	เอเชีย	มาเลเซีย	0.024	0.0207
ข้าวขาว 100% ชั้น 1	2015	เอเชีย	มาเลเซีย	0.023	0.0229
ข้าวขาว 100% ชั้น 1	2016	เอเชีย	มาเลเซีย	0.055	0.0230
ข้าวขาว 100% ชั้น 1	2017	เอเชีย	มาเลเซีย	0.019	0.0441
ข้าวขาว 100% ชั้น 1	2018	เอเชีย	มาเลเซีย	0.019	0.0275
ข้าวขาว 100% ชั้น 1	2019	เอเชีย	มาเลเซีย	0.027	0.0219
ข้าวขาว 100% ชั้น 1	2013	เอเชีย	สหรัฐอาหรับเอมิเรตส์	0.024	0.0253
ข้าวขาว 100% ชั้น 1	2014	เอเชีย	สหรัฐอาหรับเอมิเรตส์	0.022	0.0244
ข้าวขาว 100% ชั้น 1	2015	เอเชีย	สหรัฐอาหรับเอมิเรตส์	0.022	0.0228
ข้าวขาว 100% ชั้น 1	2016	เอเชีย	สหรัฐอาหรับเอมิเรตส์	0.052	0.0223
ข้าวขาว 100% ชั้น 1	2017	เอเชีย	สหรัฐอาหรับเอมิเรตส์	0.020	0.0419
ข้าวขาว 100% ชั้น 1	2018	เอเชีย	สหรัฐอาหรับเอมิเรตส์	0.023	0.0274
ข้าวขาว 100% ชั้น 1	2019	เอเชีย	สหรัฐอาหรับเอมิเรตส์	0.030	0.0245
ข้าวขาว 100% ชั้น 1	2013	เอเชีย	สาธารณรัฐประชาชนจีน	0.028	0.0281
ข้าวขาว 100% ชั้น 1	2014	เอเชีย	สาธารณรัฐประชาชนจีน	0.027	0.0283
ข้าวขาว 100% ชั้น 1	2015	เอเชีย	สาธารณรัฐประชาชนจีน	0.024	0.0274
ข้าวขาว 100% ชั้น 1	2016	เอเชีย	สาธารณรัฐประชาชนจีน	0.057	0.0252
ข้าวขาว 100% ชั้น 1	2017	เอเชีย	สาธารณรัฐประชาชนจีน	0.018	0.0458
ข้าวขาว 100% ชั้น 1	2018	เอเชีย	สาธารณรัฐประชาชนจีน	0.024	0.0275
ข้าวขาว 100% ชั้น 1	2019	เอเชีย	สาธารณรัฐประชาชนจีน	0.034	0.0252
ข้าวขาว 100% ชั้น 1	2013	เอเชีย	สาธารณรัฐสังคมนิยมเวียดนาม	0.014	0.0310
ข้าวขาว 100% ชั้น 1	2014	เอเชีย	สาธารณรัฐสังคมนิยมเวียดนาม	0.025	0.0198
ข้าวขาว 100% ชั้น 1	2015	เอเชีย	สาธารณรัฐสังคมนิยมเวียดนาม	0.020	0.0232
ข้าวขาว 100% ชั้น 1	2016	เอเชีย	สาธารณรัฐสังคมนิยมเวียดนาม	0.039	0.0211

ภาพที่ 3.25 ตารางการพยากรณ์อนุกรมเวลาด้วยวิธี Single Exponential Moving Average

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

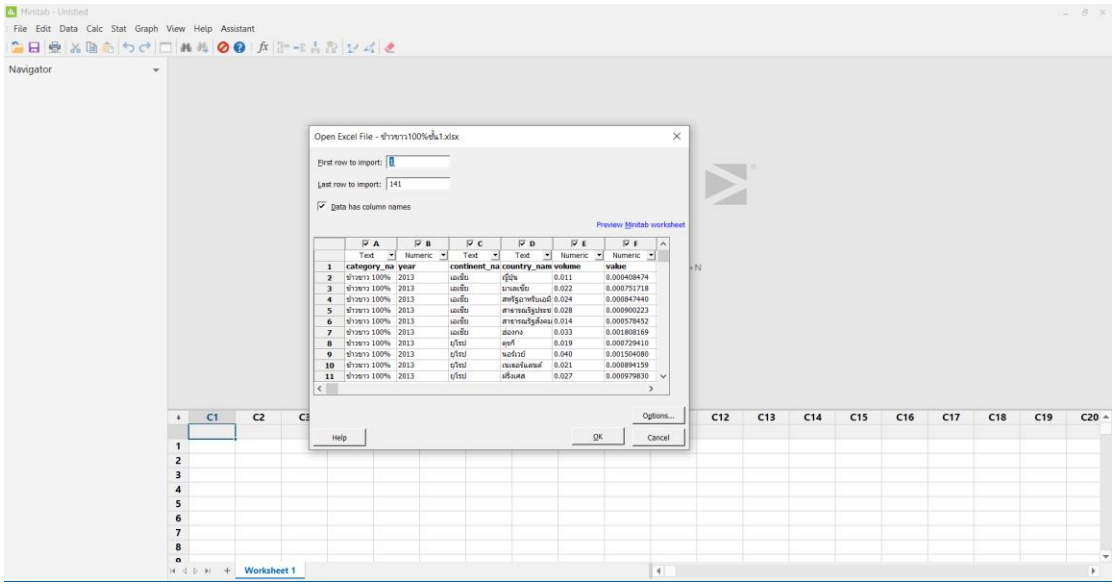
ตารางที่ 3.23 ตารางเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE)

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.

ผู้วิเคราะห์ข้อมูลได้นำข้อมูลที่ผ่านการวิเคราะห์ด้วยเทคนิค Single Exponential จากการสร้างโมเดล Time Series Models โดยใช้โปรแกรม Minitab ซึ่งมีขั้นตอนการทำงานดังนี้

ขั้นตอนที่ 1 ทำการอัปโหลดข้อมูลปริมาณการส่ออกข้าวหอมมะลิไทยทั้งหมด 7 ปี ตามประเภทข้าวเข้าโปรแกรม เพื่อเตรียมความพร้อมในการพยากรณ์

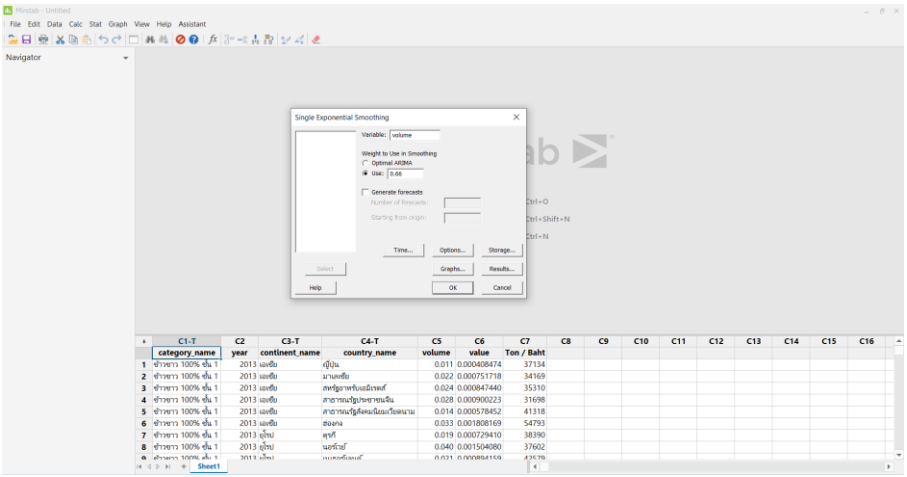


ภาพที่ 3.26 แสดงขั้นตอนการนำข้อมูลเข้าโปรแกรม Minitab

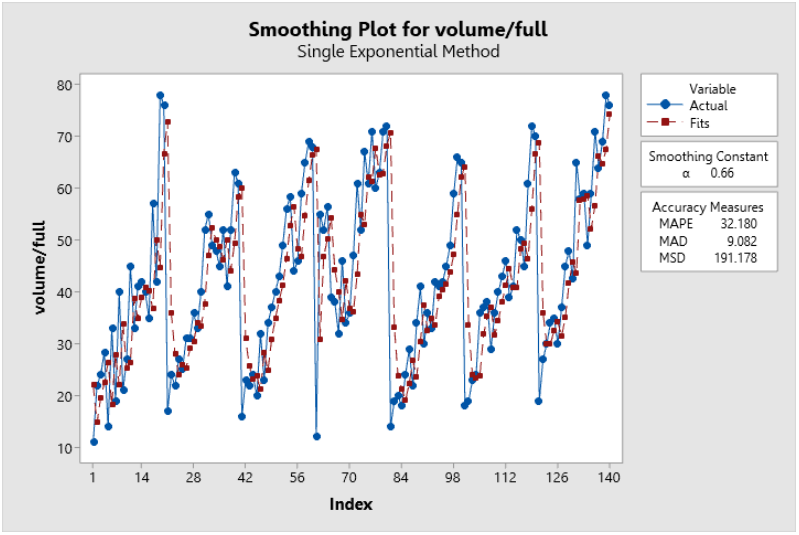
1. Single Exponential

ขั้นตอนที่ 1 เลือกโมเดล Time Series ใช้วิธี Single Exponential และกำหนด Variable เป็นข้อมูลที่ต้องการจะทำการพยากรณ์

ขั้นตอนที่ 2 กำหนดค่า α เท่ากับ 0.66 และกำหนดจำนวนเลขที่จะพยากรณ์



ภาพที่ 3.27 แสดงขั้นตอนการกำหนดรูปแบบการพยากรณ์วิธี Single Exponential

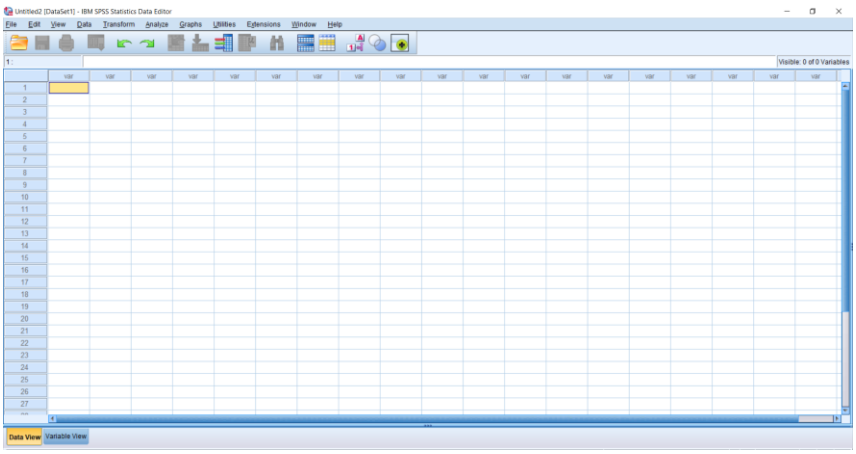


ภาพที่ 3.28 แสดงผลลัพธ์ของการพยากรณ์ด้วยโปรแกรม Minitab

จากภาพที่ 3.28 จะเห็นได้ว่าการใช้วิธีการ Single Exponential จากโปรแกรม Minitab ในการพยากรณ์จะมีค่าความคลาดเคลื่อน MAPE เท่ากับ 32.180%

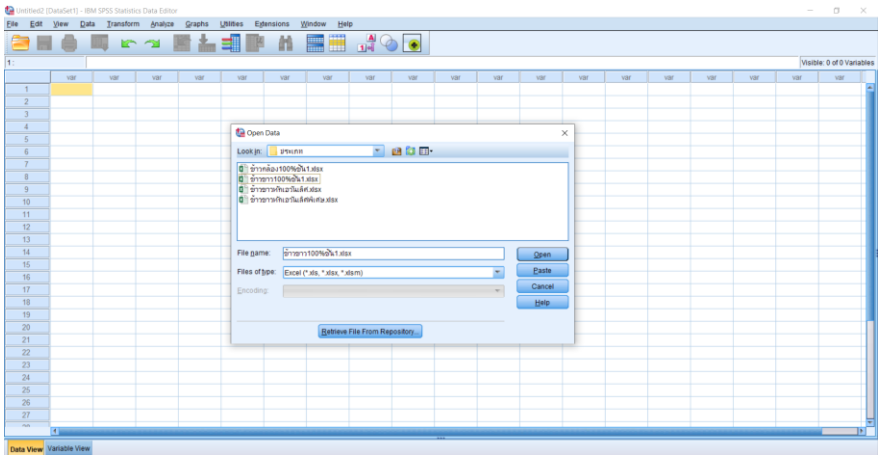
2. การพยากรณ์โดย การวิเคราะห์อนุกรมเวลาของข้อมูล (Data mining time series analysis) ผ่านโปรแกรม IBM SPSS ผู้วิเคราะห์ได้ข้อมูลที่ผ่านการวิเคราะห์ข้อมูลด้วยเทคนิค Data Mining จากการสร้างโมเดล Time Series Models จึงนำข้อมูลดังกล่าว มาทดสอบกับโปรแกรม IBM SPSS ซึ่งมีขั้นตอนการทำงานดังนี้

ขั้นตอนที่ 1 หน้าต่างพื้นที่ทำงานของโปรแกรม IBM SPSS



ภาพที่ 3.29 หน้าต่างพื้นที่ทำงานของโปรแกรม IBM SPSS

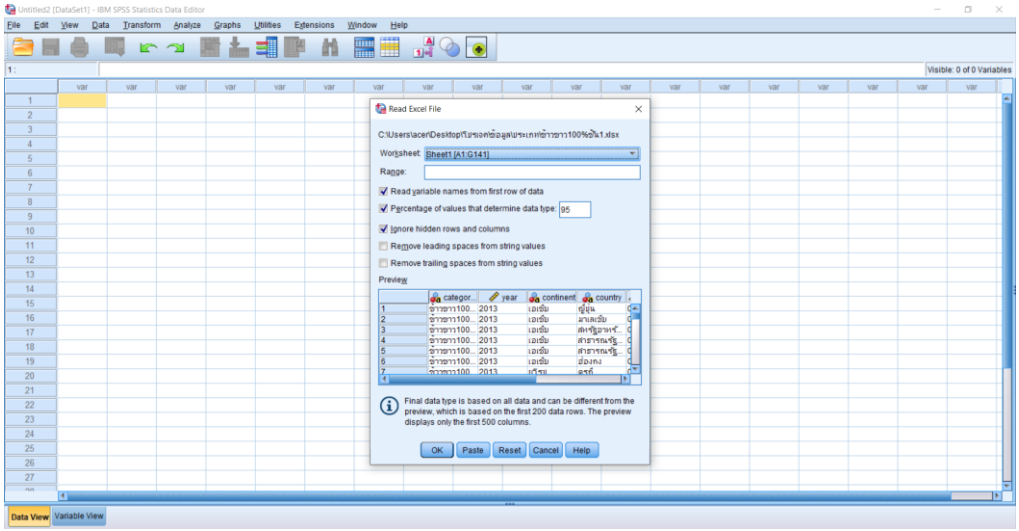
ขั้นตอนที่ 2 ทำการ Import ข้อมูลทั้งหมดของแต่ละภาคเข้าโปรแกรมเพื่อทำการหาค่าผลลัพธ์ของการพยากรณ์



ภาพที่ 3.30 Import ข้อมูลที่ต้องการหาผลลัพธ์

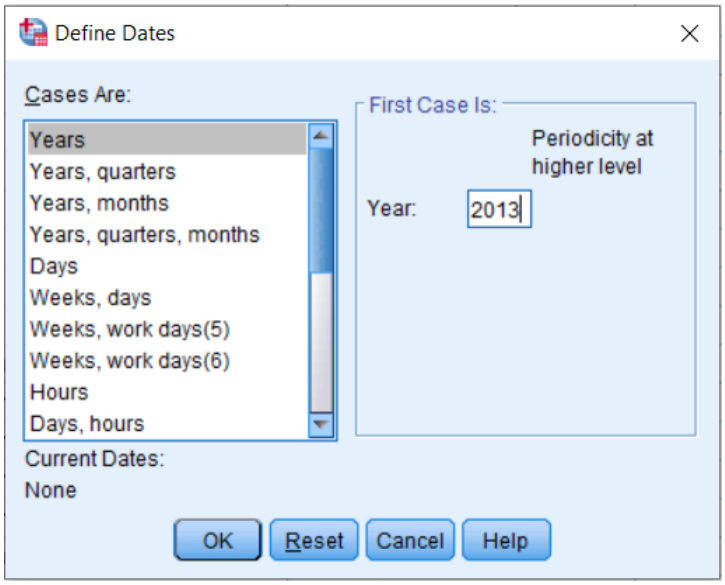
ขั้นตอนที่ 3 ทำการอัปโหลดข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทยทั้งหมด 7 ปี เพื่อเตรียมความพร้อมในการพยากรณ์

เนื่องจากโปรแกรม IBM SPSS รองรับชุดข้อมูลที่เป็นภาษาอังกฤษ ผู้วิเคราะห์จึงได้ทำชุดข้อมูลภาษอังกฤษขึ้นมาเพื่อทำให้โปรแกรมสามารถหาค่าพยากรณ์ได้



ภาพที่ 3.31 Import ข้อมูลเพื่อหาค่าการพยากรณ์

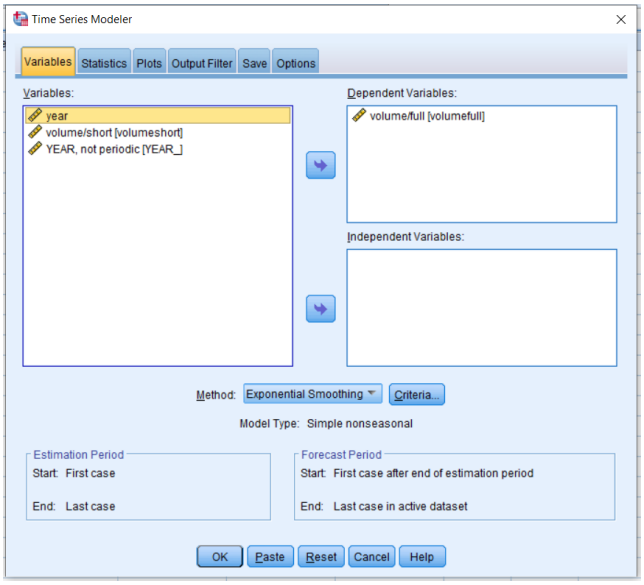
ขั้นตอนที่ 4 กำหนดระยะเวลาในการพยากรณ์ Define Date and Time โดยเลือก Year ในการพยากรณ์เริ่มจากปี 2013



ภาพที่ 3.32 กำหนดระยะเวลาในการพยากรณ์

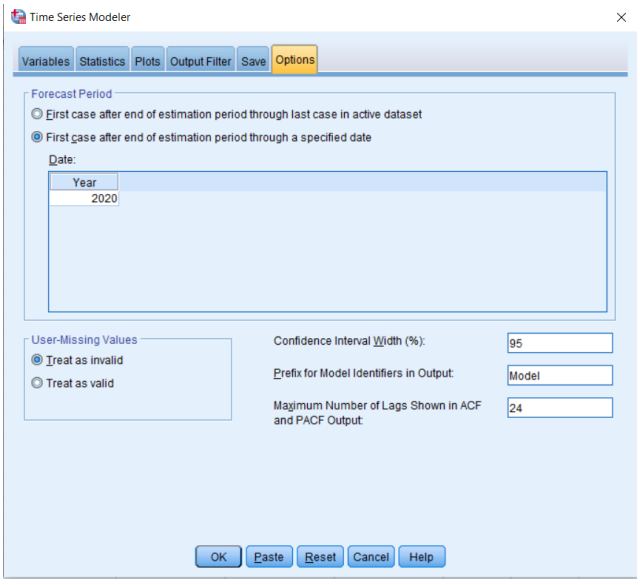
ขั้นตอนที่ 5 ขั้นตอนการพยากรณ์ด้วยวิธี Single Exponential

ขั้นตอนที่ 1 เลือกโมเดล Forecasting ใช้วิธี Single Exponential และกำหนด Volume เป็นข้อมูลที่ต้องการจะทำการพยากรณ์



ภาพที่ 3.33 แสดงขั้นตอนการกำหนดรูปแบบการพยากรณ์วิธี Single Exponential

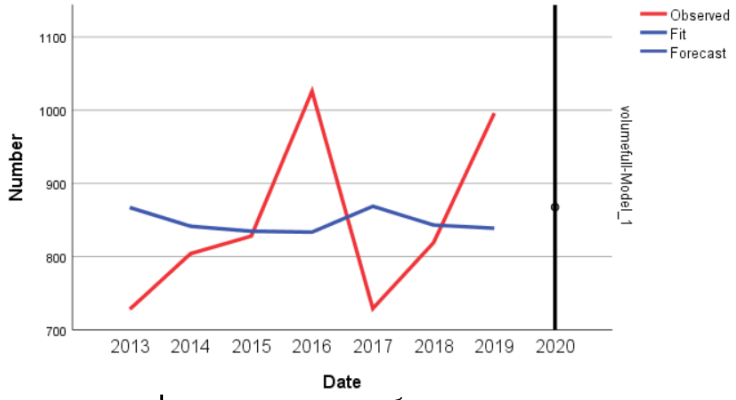
ขั้นตอนที่ 2 กำหนดระยะเวลาในการพยากรณ์ในปีถัดไปคือ 2020



ภาพที่ 3.34 กำหนดระยะเวลาในการพยากรณ์ในปีถัดไปคือ 2020

Model Statistics

Model	Number of Predictors	Model Fit statistics: Stationary R-squared	MAPE	MAE	Ljung-Box Q(18): Statistics	DF	Sig.	Number of Outliers
volume/full-Model_1	0	.483	11.604	99.543	.	0	.	0

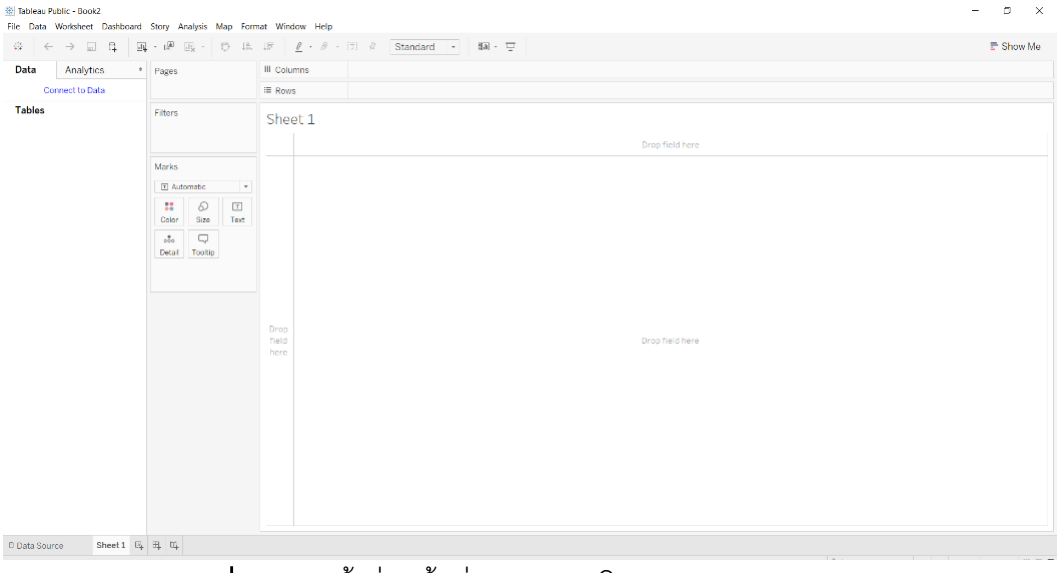


ภาพที่ 3.35 แสดงผลลัพธ์กราฟวิธี Single Exponential

จากภาพที่ 3.35 จะเห็นได้ว่าการใช้วิธีการ Single Exponential จากโปรแกรม IBM SPSS ในการพยากรณ์จะมีความคลาดเคลื่อน MAPE 11.604%

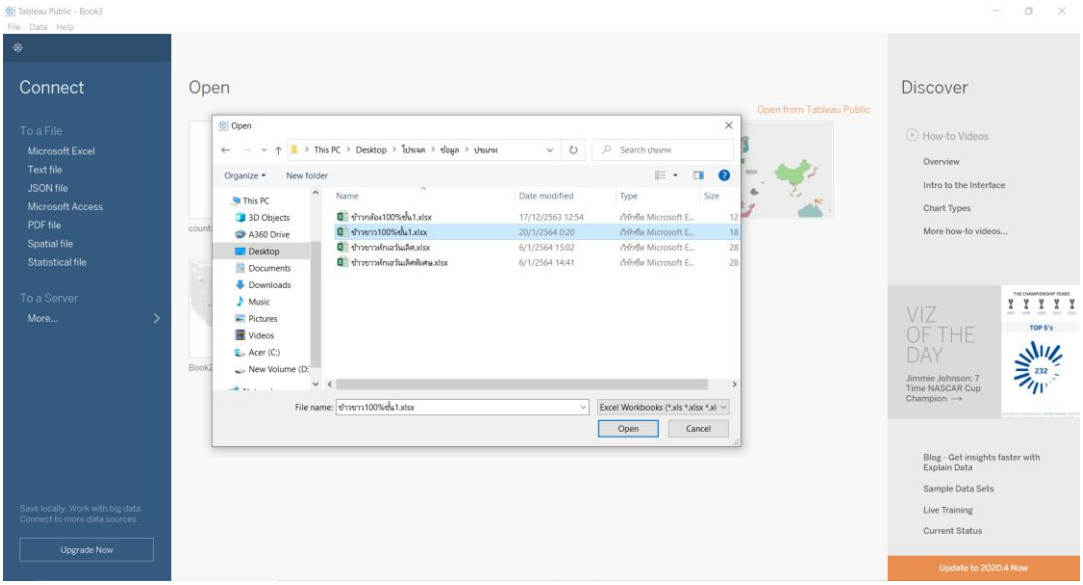
3) การพยากรณ์โดยการวิเคราะห์อนุกรมเวลาของข้อมูล (Data mining time series analysis) ผ่านโปรแกรม Tableau Public ผู้วิเคราะห์ได้ข้อมูลที่ผ่านการวิเคราะห์ข้อมูลด้วยเทคนิค Data Mining จากการสร้างโมเดล Time Series Models จึงนำข้อมูลดังกล่าว มาทดสอบกับโปรแกรม Tableau Public ซึ่งมีขั้นตอนการทำงานดังนี้

ขั้นตอนที่ 1 หน้าต่างพื้นที่ทำงานของโปรแกรม tableau



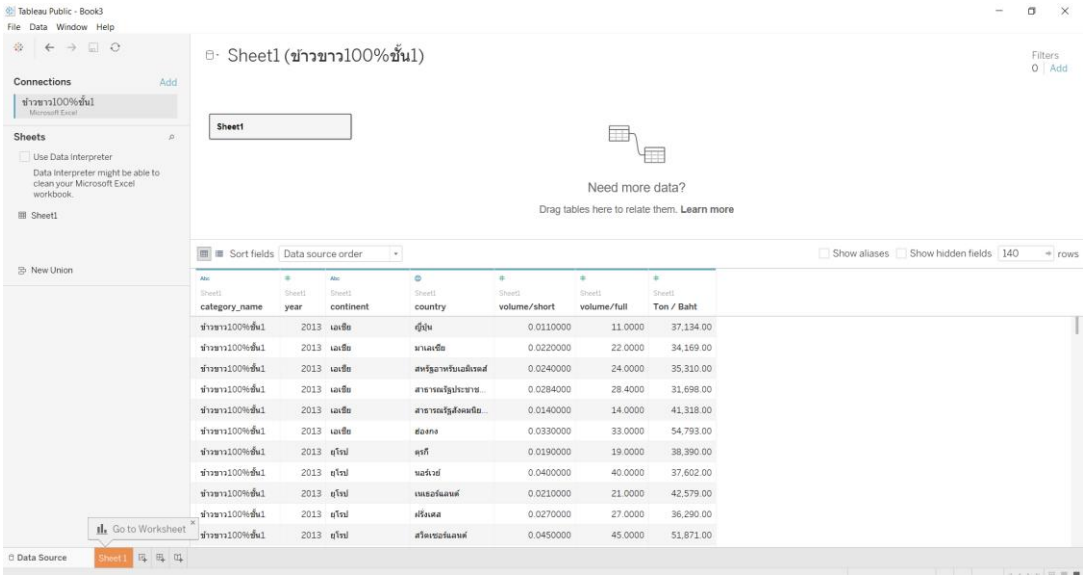
ภาพที่ 3.36 หน้าต่างพื้นที่ทำงานของโปรแกรม Tableau Public

ขั้นตอนที่ 2 จัดเตรียมข้อมูลที่จะนำมาแสดงผลลัพธ์



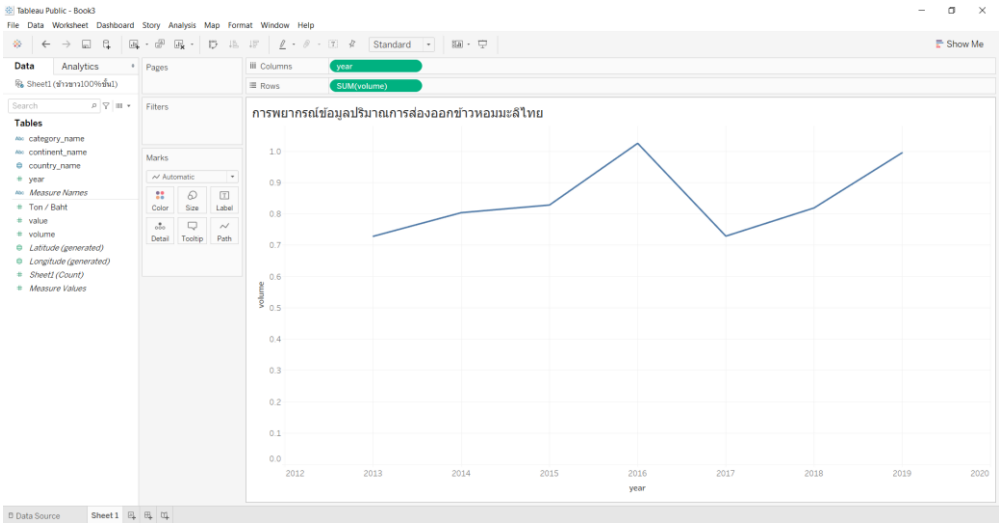
ภาพที่ 3.37 Import ข้อมูลที่ต้องการหาผลลัพธ์

ขั้นตอนที่ 3 ทำการใช้เครื่องมือทำเหมืองข้อมูล (Data Mining) ของโปแกรม Tableau Public ในการพยากรณ์ข้อมูล



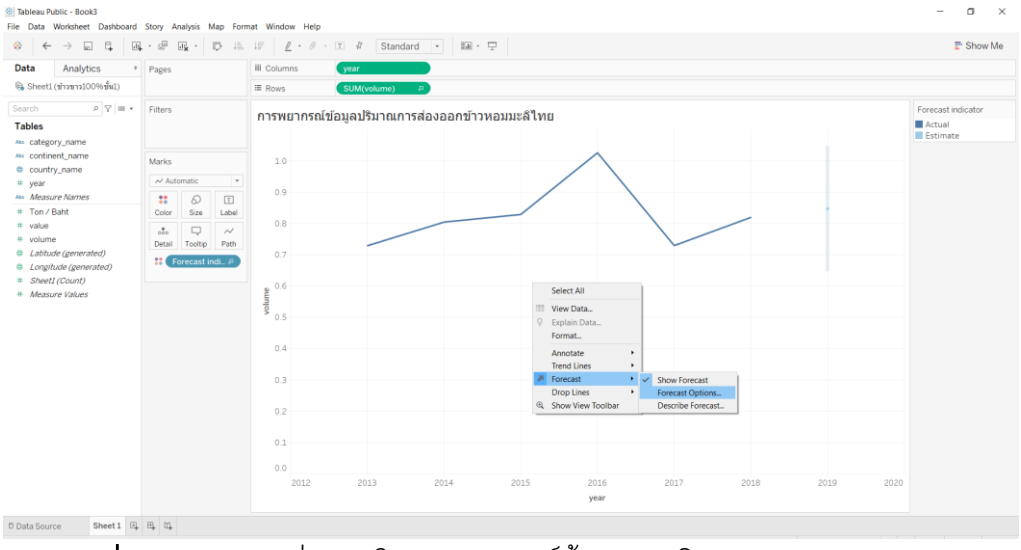
ภาพที่ 3.38 Import แสดงผลการโหลดฐานข้อมูลเข้าสู่โปรแกรม Tableau Public

ขั้นตอนที่ 4 การเตรียมกราฟข้อมูลก่อนการวิเคราะห์ข้อมูล



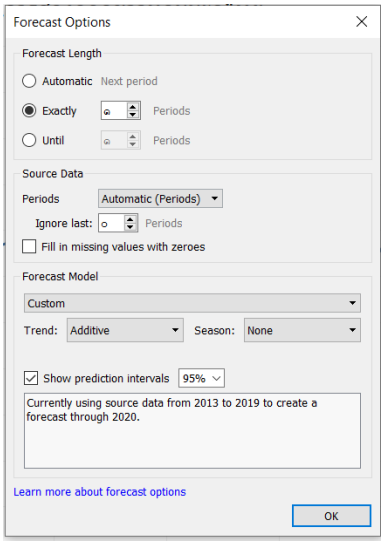
ภาพที่ 3.39 แสดงการเตรียมกราฟข้อมูลก่อนการวิเคราะห์ข้อมูล

ขั้นตอนที่ 5 เครื่องมือในการวิเคราะห์ข้อมูลของโปรแกรม Tableau Public



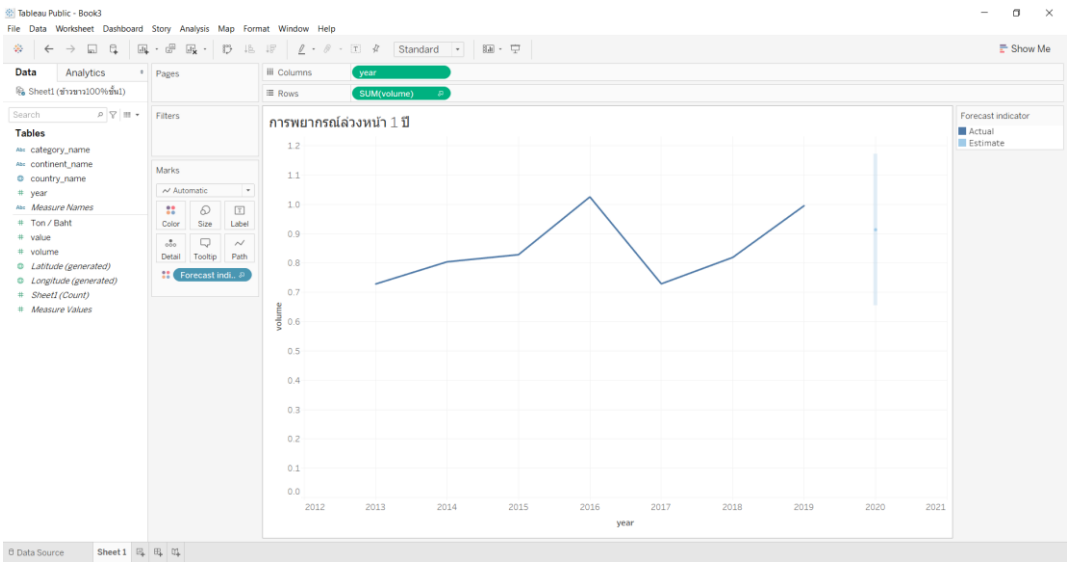
ภาพที่ 3.40 แสดงเครื่องมือในการวิเคราะห์ข้อมูลของโปรแกรม Tableau Public

ขั้นตอนที่ 6 การใช้เครื่องมือทำเหมืองข้อมูล (Data Mining) ของโปแกรม Tableau Public ในการพยากรณ์ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทย กำหนดค่า Exactly เป็น 1 ของปีถัดไปในช่วงเวลาเดียวกัน



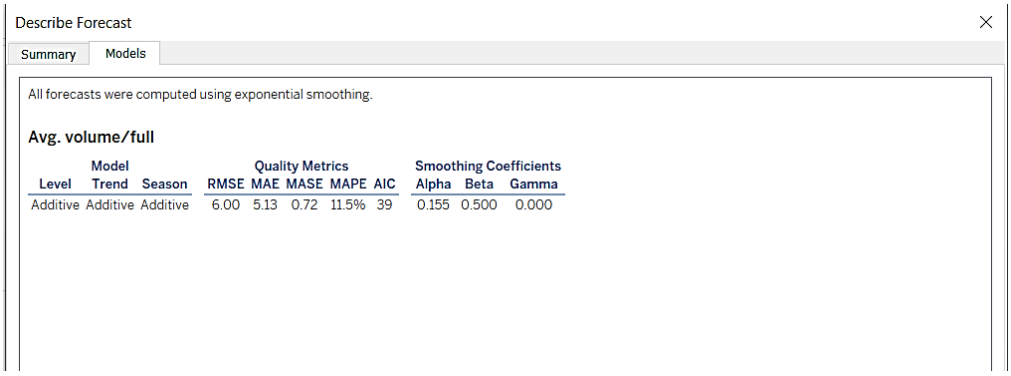
ภาพที่ 3.41 แสดงการใช้เครื่องมือทำเหมืองข้อมูลของโปรแกรม Tableau Public

ขั้นตอนที่ 7 กราฟผลลัพธ์การพยากรณ์โดย Tableau Public



ภาพที่ 3.42 กราฟแสดงผลลัพธ์การพยากรณ์โดย Tableau Public

ขั้นตอนที่ 8 ผลลัพธ์การคำนวณค่าความคลาดเคลื่อน MAPE โดยโปรแกรม Tableau Public


Describe Forecast

Summary | Models

All forecasts were computed using exponential smoothing.

Avg. volume/full

Model			Quality Metrics					Smoothing Coefficients		
Level	Trend	Season	RMSE	MAE	MASE	MAPE	AIC	Alpha	Beta	Gamma
Additive	Additive	Additive	6.00	5.13	0.72	11.5%	39	0.155	0.500	0.000

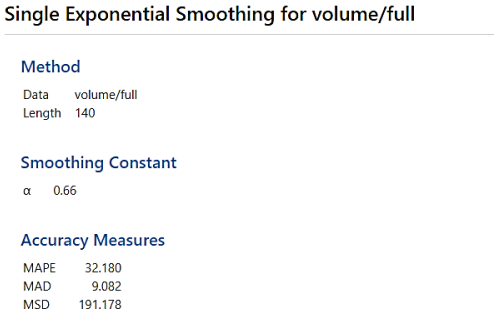
ภาพที่ 3.43 แสดงผลการคำนวณความคลาดเคลื่อน MAPE โดยโปรแกรม Tableau Public

จากภาพที่ 3.43 จะเห็นได้ว่าการพยากรณ์จากโปรแกรม Tableau Public การพยากรณ์จะมีค่าความคลาดเคลื่อน MAPE 11.5 %

3.5.5 การประเมินประสิทธิภาพ (Evaluation)

เป็นขั้นตอนก่อนนำผลลัพธ์ที่ได้จากขั้นตอนที่ 4 ไปใช้งาน ผู้วิเคราะห์ข้อมูลวัดประสิทธิภาพของผลลัพธ์ได้กับวัตถุประสงค์ที่ตั้งไว้ในขั้นตอนแรก ว่ามีนัยสำคัญหรือความน่าเชื่อถือมากน้อยเพียงใด โดยการนำไปตรวจสอบค่าเปอร์เซ็นต์ความคลาดเคลื่อนสมบูรณ์เฉลี่ย (Mean Absolute Percentage Error : MAPE) จากการทดลองโมเดลเทคนิค Time Series Models ด้วยการวิเคราะห์อนุกรมเวลาของข้อมูล (Data mining time series analysis) ผ่าน 3 โปรแกรมมีค่าเปอร์เซ็นต์ความคลาดเคลื่อนสมบูรณ์เฉลี่ยดังนี้

1) การวัดประสิทธิภาพของผลลัพธ์ด้วยโปรแกรม Minitab


Single Exponential Smoothing for volume/full

Method

Data volume/full
Length 140

Smoothing Constant

α 0.66

Accuracy Measures

MAPE 32.180
MAD 9.082
MSD 191.178

ภาพที่ 3.44 ผลลัพธ์การวัดประสิทธิภาพของผลลัพธ์ด้วยโปรแกรม Minitab

ผลการตรวจสอบค่าเปอร์เซ็นต์ความคลาดเคลื่อนสมบูรณ์เฉลี่ย (Mean Absolute Percentage Error : MAPE) จากการทดลองโมเดลเทคนิค Time Series Models ด้วยการวิเคราะห์อนุกรมเวลาของข้อมูล (Data mining time series analysis) ผ่านโปรแกรม Minitab ได้ผลลัพธ์ถึง 32.180%

2) การวัดประสิทธิภาพของผลลัพธ์ด้วยโปรแกรม IBM SPSS

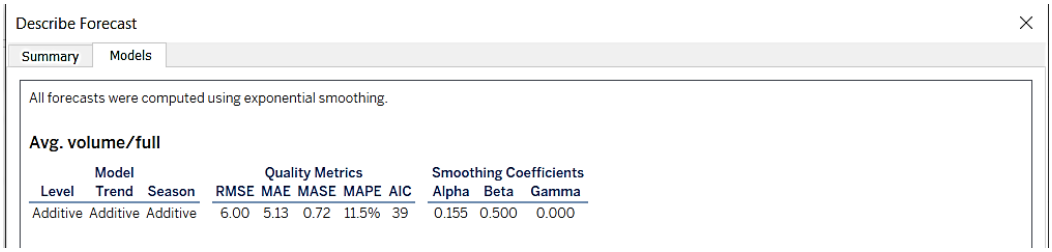
Model Statistics

Model	Number of Predictors	Model Fit statistics: Stationary R-squared	Model Fit statistics: MAPE	Model Fit statistics: MAE	Ljung-Box Q(18): Statistics	Ljung-Box Q(18): DF	Ljung-Box Q(18): Sig.	Number of Outliers
volume/full-Model_1	0	.483	11.604	99.543	.	0	.	0

ภาพที่ 3.45 ผลลัพธ์การวัดประสิทธิภาพของผลลัพธ์ด้วยโปรแกรม IBM SPSS

ผลการตรวจสอบค่าเปอร์เซ็นต์ความคลาดเคลื่อนสมบูรณ์เฉลี่ย (Mean Absolute Percentage Error : MAPE) จากการทดลองโมเดลเทคนิค Time Series Models ด้วยการวิเคราะห์อนุกรมเวลาของข้อมูล (Data mining time series analysis) ผ่านโปรแกรม IBM SPSS ได้ผลลัพธ์ถึง 11.604%

3) การวัดประสิทธิภาพของผลลัพธ์ด้วยโปรแกรม Tableau Public



ภาพที่ 3.46 ผลลัพธ์การวัดประสิทธิภาพของผลลัพธ์ด้วยโปรแกรม Tableau Public

ผลการตรวจสอบค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percentage Error : MAPE) จากการทดลองโมเดลเทคนิค Time Series Models ด้วยการวิเคราะห์อนุกรมเวลาของข้อมูล (Data mining time series analysis) ผ่านโปรแกรม Tableau Public ได้ผลลัพธ์ต่ำสุดถึง 11.5%

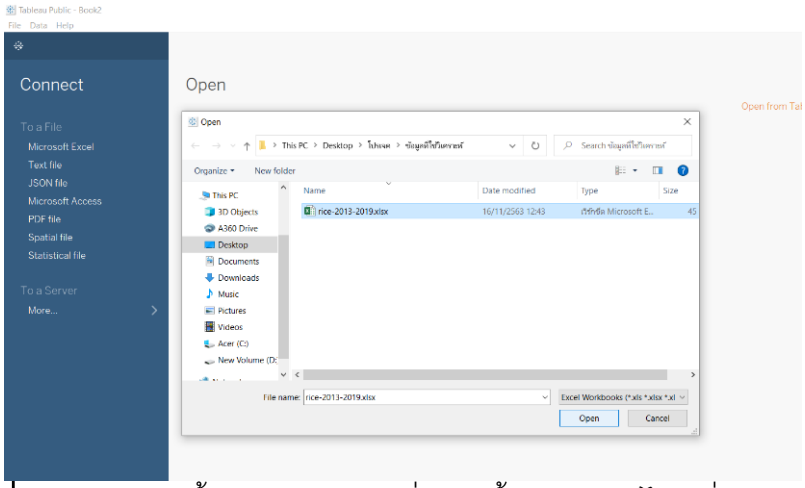
จากการทดสอบประสิทธิภาพโมเดล และการประเมินค่าความคลาดเคลื่อน MAPE ที่ค่าคงที่เกิดยิ่งน้อยยิ่งเป็นผลดีแสดงว่าค่าพยากรณ์ออกมานั้นมีความคลาดเคลื่อนจากค่าจริงเพียงใด ผู้วิเคราะห์จึงเลือกใช้โมเดลการวิเคราะห์อนุกรมเวลาของข้อมูล (Data mining time series analysis) ผ่านโปรแกรม Tableau Public ให้ผลลัพธ์ที่น่าพึงพอใจอยู่ในระดับใช้ได้ดี สามารถนำโมเดลไปใช้งานได้ ผู้วิเคราะห์จึงเลือกใช้โมเดลอนุกรมเวลาของ Tableau Public ในการพยากรณ์ปริมาณการส่งออกข้าวหอมมะลิไทยนำไปสู่การแสดงสารสนเทศสร้างประโยชน์ในการตัดสินใจสำหรับเกษตรกรผู้เพาะปลูกข้าวผู้ประกอบการส่งออกข้าว และหน่วยงานที่เกี่ยวข้องในการวางแผนการผลิตและการส่งออกให้สอดคล้องกับความต้องการของตลาดต่อไป

3.5.6 การนำผลลัพธ์ไปใช้งาน (Deployment)

ผู้วิเคราะห์ข้อมูลนำผลองค์ความรู้ที่ได้เหล่านี้ไปนำเสนอข้อมูลแบบ visualization ด้วยการแสดงผลข้อมูลในรูปแบบของภาพโดยใช้โปรแกรม Tableau Public แสดงข้อมูลการพยากรณ์ และใช้ highcharts.js แสดงข้อมูลแบบแดชบอร์ด โดยดำเนินการตามขั้นตอนดังนี้

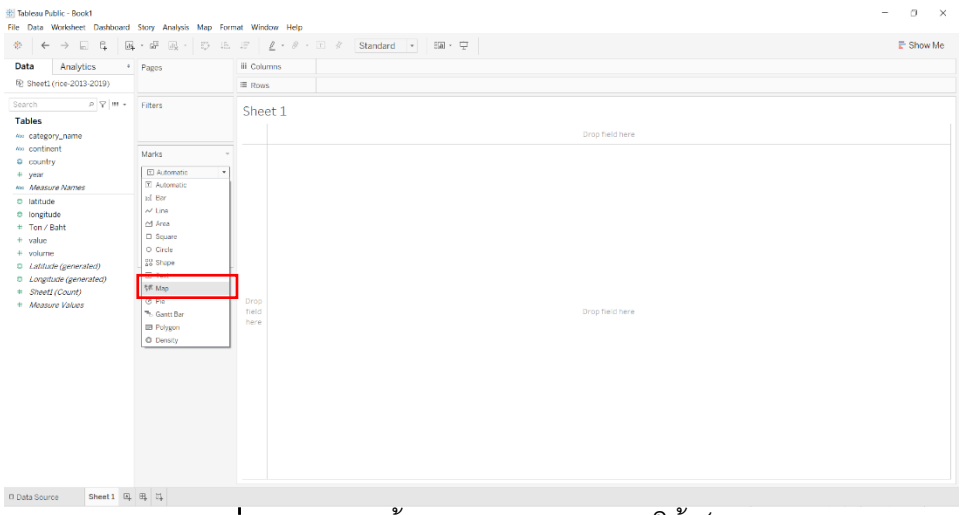
1. การนำเสนอข้อมูลแบบ visualization ด้วยการแสดงผลข้อมูลในรูปแบบของภาพโดยใช้โปรแกรม Tableau Public

1.1) ทำการ Import ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทยเข้าโปรแกรม Tableau Public



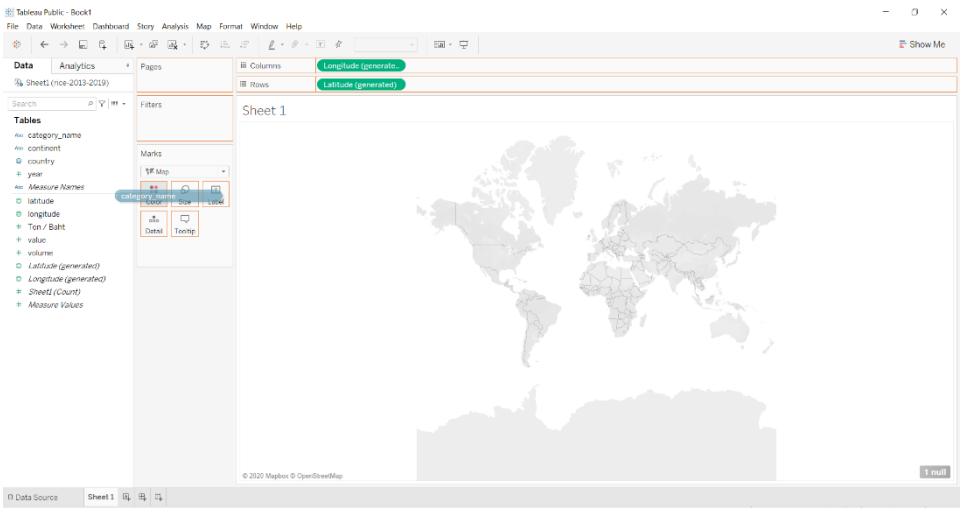
ภาพที่ 3.47 Import ข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทยเพื่อหาผลลัพธ์

1.2) เข้าหน้า DashBoard การทำงานข้อมูลที่ Import จะอยู่แถบด้านซ้าย



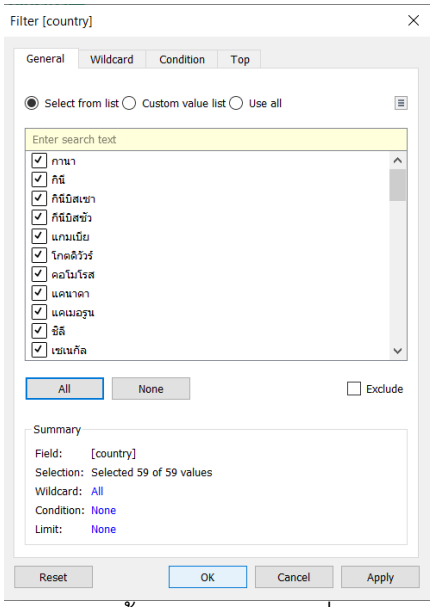
ภาพที่ 3.48 แสดงขั้นตอนการปรับ Mark ให้เป็น Map

1.3) กำหนด Columns เป็น Longitude กำหนด Rows เป็น Latitude และกำหนด country_name ใส่ใน Color



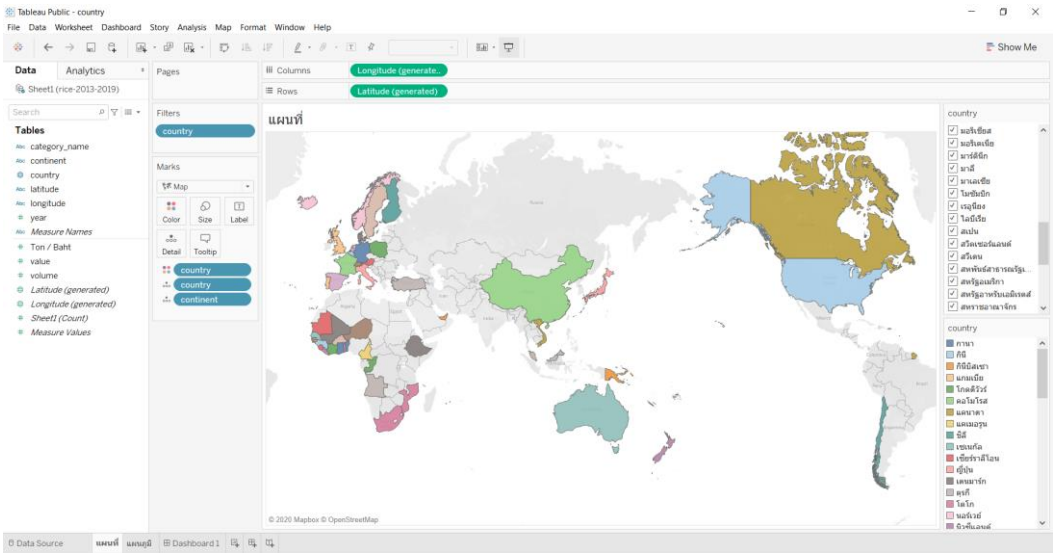
ภาพที่ 3.49 แสดงขั้นตอนการโชว์ประเภทข้าวบน Map

1.4) กำหนด Filter and then add ทำการเลือกชื่อประเทศทั้งหมดเพื่อไปแสดงบนแผนที่



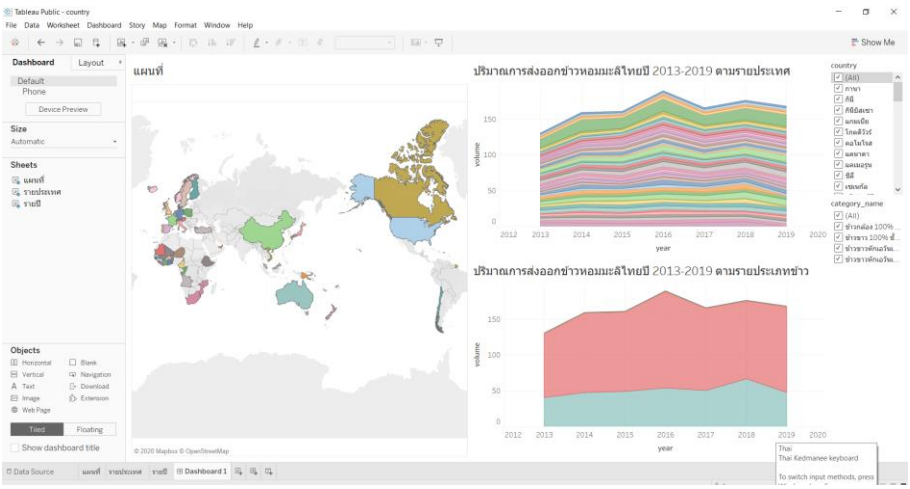
ภาพที่ 3.50 แสดงขั้นตอนการเลือกชื่อประเทศทั้งหมด

1.5) การแสดงบนแผนที่โลก



ภาพที่ 3.51 แสดงผลลัพธ์แผนที่ประเทศเป็นสี

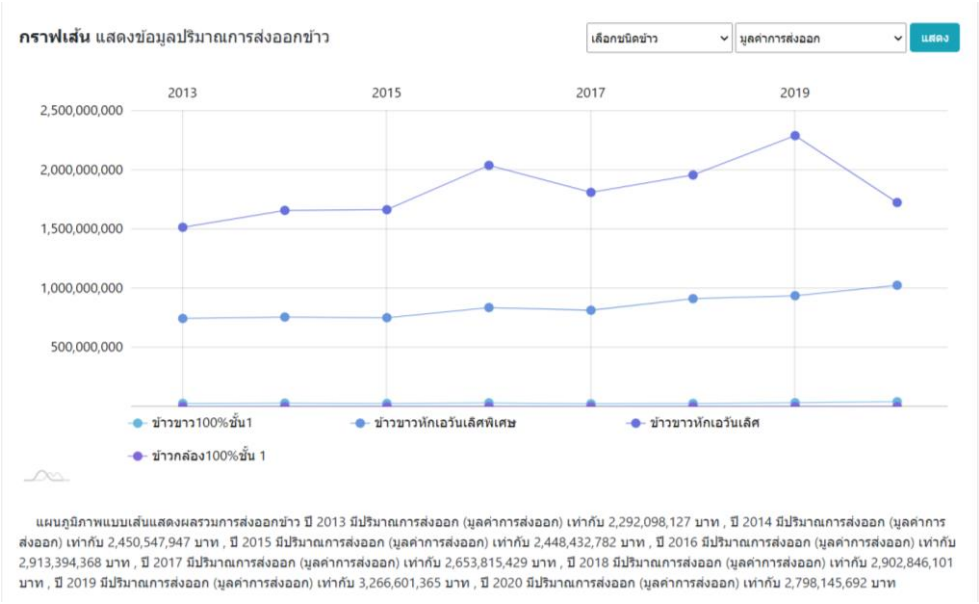
1.6) การแสดงหน้าแดชบอร์ดบนโปรแกรม Tableau Public



ภาพที่ 3.52 แสดงผลข้อมูลปริมาณการส่งออกข้าวหอมมะลิไทยทั้ง 54 ประเทศ

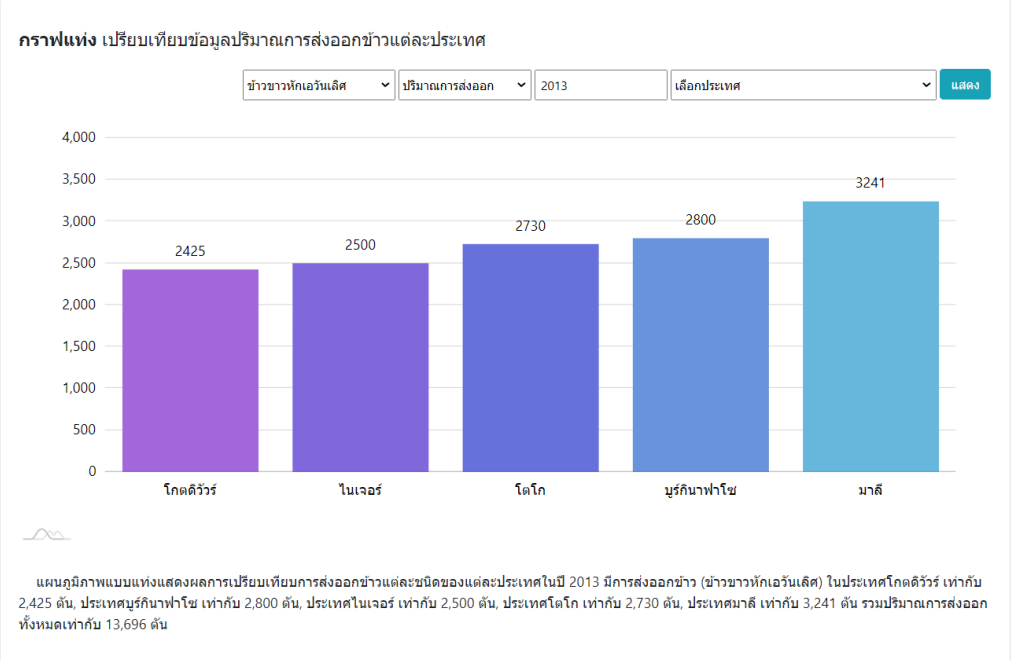
2. นำเสนอข้อมูลแบบ visualization ด้วยการแสดงผลข้อมูลในรูปแบบของภาพโดยใช้การแสดงข้อมูลโดยใช้ highcharts.js แสดงข้อมูลแบบแดชบอร์ด โดยดำเนินการตามขั้นตอนดังนี้

2.1) กราฟเส้น แสดงข้อมูลปริมาณการส่งออกข้าว



ภาพที่ 3.53 แสดงกราฟเส้น ข้อมูลปริมาณการส่งออกข้าว

2.2) กราฟเส้น เปรียบเทียบข้อมูลปริมาณการส่งออกข้าวแต่ละประเทศ

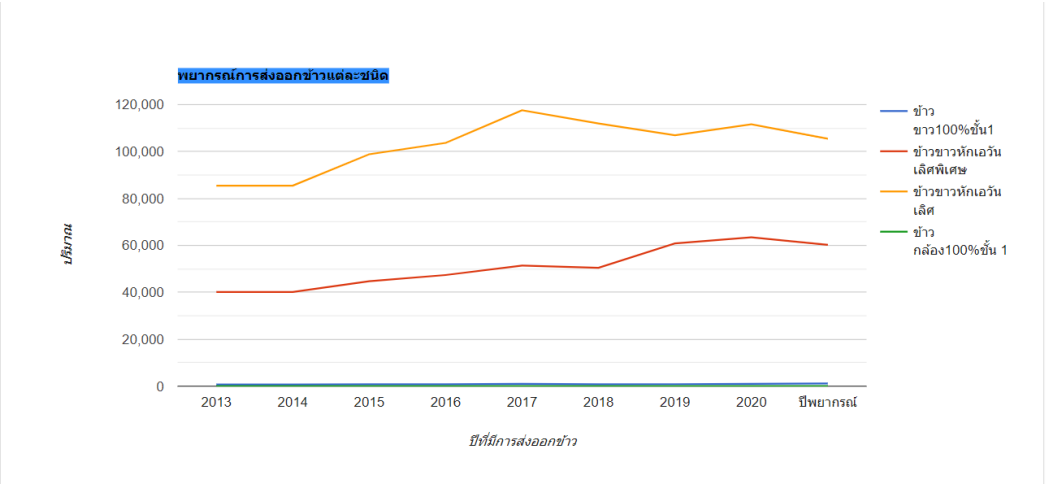


ภาพที่ 3.54 แสดงกราฟเส้น เปรียบเทียบข้อมูลปริมาณการส่งออกข้าวแต่ละประเทศ

2.3) แผนที่กราฟ แสดงข้อมูลปริมาณการส่งออกข้าวแต่ละทวีป

ภาพที่ 3.55 แสดงแผนที่กราฟ แสดงข้อมูลปริมาณการส่งออกข้าวแต่ละทวีป

2.4) กราฟพยากรณ์การส่งออกข้าวแต่ละชนิด



ภาพที่ 3.56 แสดงกราฟพยากรณ์การส่งออกข้าวแต่ละชนิด

3.6 การออกแบบและพัฒนาเว็บไซต์

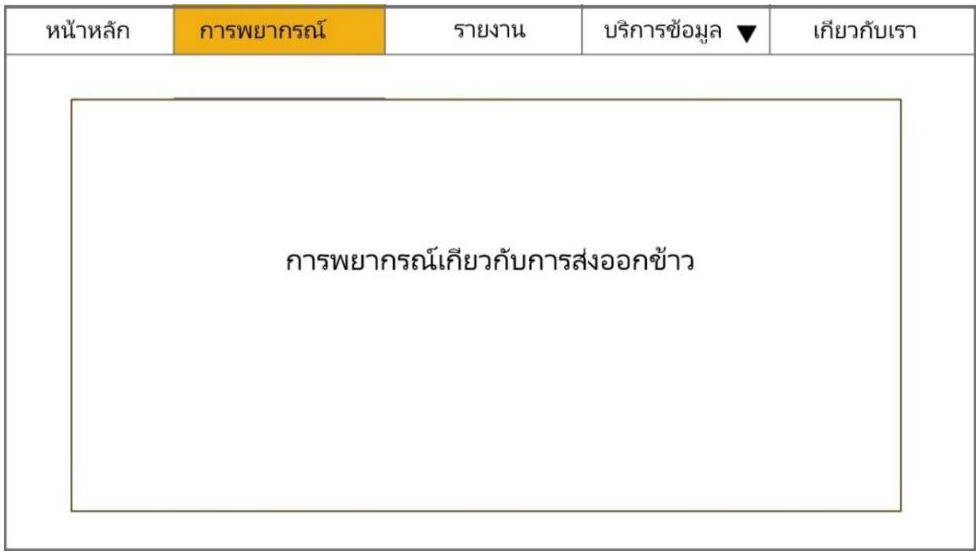
การวางแผนการจัดลำดับ เนื้อหาสาระของเว็บไซต์ ออกเป็นหมวดหมู่ เพื่อจัดทำเป็นโครงสร้างในการจัดวางหน้าเว็บเพจทั้งหมด ทำให้เห็นโครงสร้างทั้งหมดของเว็บไซต์และการออกแบบโครงสร้างหรือจัดระเบียบของข้อมูลที่ชัดเจน แยกย่อยเนื้อหาออกเป็นหาออกเป็นส่วนต่าง ๆ ที่สัมพันธ์และให้มีอยู่ในมาตรฐานเดียวกัน จะช่วยให้นำมาใช้งานและง่าย ต่อการเข้าอ่านเนื้อหาของผู้ใช้เว็บไซต์

3.6.1 การออกแบบ Wireframe หน้าจอเว็บไซต์

1.) หน้าแรกของเว็บไซต์ แสดงเมนูต่าง ๆ ของหน้าเว็บไซต์

ภาพที่ 3.57 หน้าแรกของเว็บไซต์ แสดงเมนูต่าง ๆ ของหน้าเว็บไซต์

2) หน้าแสดงข้อมูลพยากรณ์เกี่ยวกับการส่งออกข้าว



ภาพที่ 3.58 หน้าแสดงข้อมูลพยากรณ์เกี่ยวกับการส่งออกข้าว

3) หน้าแสดงรายงาน Dashboard ข้อมูลปริมาณการส่งออกข้าว

ภาพที่ 3.59 หน้าแสดงรายงาน Dashboard ข้อมูลปริมาณการส่งออกข้าว

4) หน้าแสดงบริการข้อมูลต่างๆ เกี่ยวกับการพยากรณ์ข้อมูล

ภาพที่ 3.60 หน้าแสดงบริการข้อมูลต่างๆ เกี่ยวกับการพยากรณ์ข้อมูล

5) หน้าแสดงบริการข้อมูลต่างๆ เกี่ยวกับหน่วยงานที่เกี่ยวข้อง

ภาพที่ 3.61 หน้าแสดงบริการข้อมูลต่างๆ เกี่ยวกับหน่วยงานที่เกี่ยวข้อง

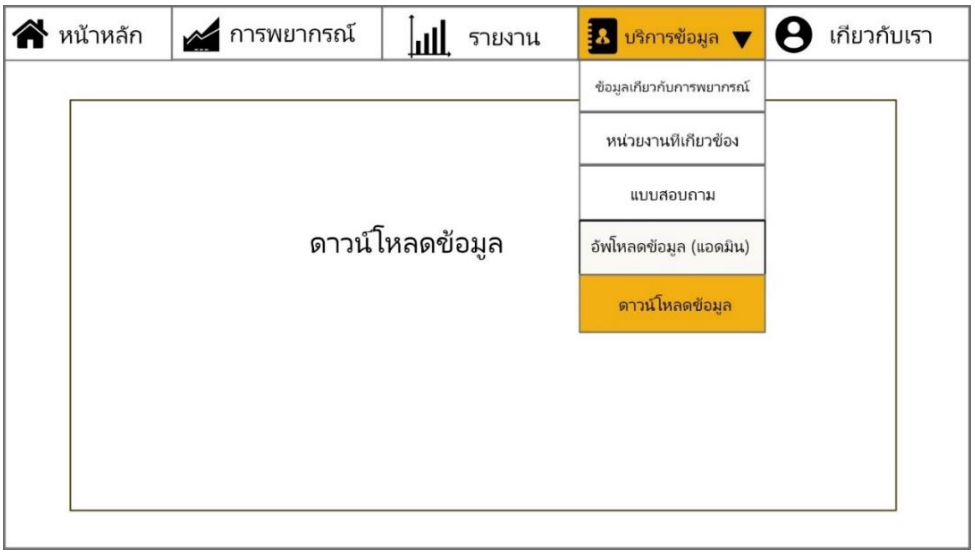
6) หน้าแสดงบริการข้อมูลต่างๆ แบบสอบถาม

ภาพที่ 3.62 หน้าแสดงบริการข้อมูลต่างๆ แบบสอบถาม

7) หน้าแสดงบริการข้อมูลต่างๆ เกี่ยวกับการอัพโหลดข้อมูลเฉพาะแอดมิน

ภาพที่ 3.63 หน้าแสดงบริการข้อมูลต่างๆ เกี่ยวกับการอัพโหลดข้อมูลเฉพาะแอดมิน

8) หน้าแสดงบริการข้อมูลต่างๆ เกี่ยวกับการดาวน์โหลดข้อมูล



ภาพที่ 3.64 หน้าแสดงบริการข้อมูลต่างๆ เกี่ยวกับการดาวน์โหลดข้อมูล

9) หน้าแสดงข้อมูลเกี่ยวกับผู้จัดทำ

ภาพที่ 3.65 หน้าแสดงข้อมูลเกี่ยวกับผู้จัดทำ

3.7 บทสรุป

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