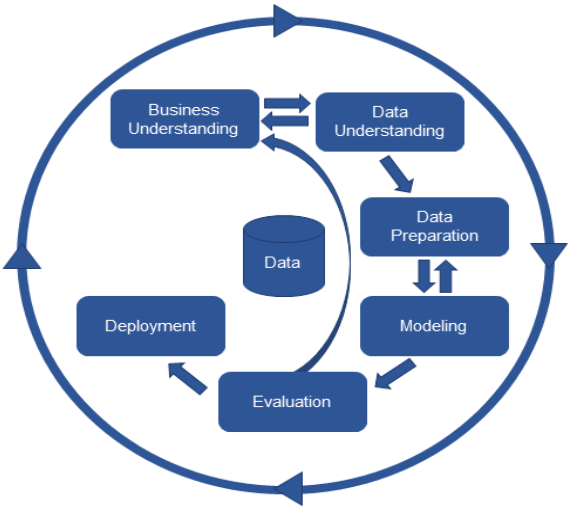
บทที่ 3
วิธีการดำเนินโครงการ

ผู้วิเคราะห์ได้วิเคราะห์ข้อมูลปริมาณน้ำฝนจากกรมอุตุนิยมวิทยา ในปี พ.ศ.2555 – 2557 เพื่อเผยแพร่ข้อมูลบนเว็บไซต์ และการจัดการข้อมูลที่มีอยู่ให้เกิดประโยชน์ และได้สารสนเทศที่สอดคล้องกับข้อมูล ผู้วิเคราะห์ได้วิเคราะห์ และออกแบบข้อมูลใหม่โดยใช้เครื่องมือที่ประกอบไปด้วย tableau กระบวนในการพัฒนาฐานข้อมูล CRISP-DM แบบฟอร์มข้อมูล และปรับโครงสร้างข้อมูลใหม่ เทคนิค Neural network วิเคราะห์ข้อมูลความสัมพันธ์ข้อมูล (Visual Analytics tableau) ในการวิเคราะห์ข้อมูลข้อมูลซึ่งในบทนี้จะแสดงถึงวิธีในการดำเนินงาน ดังนี้

3.1 กระบวนการ CRISP-DM

กระบวนการ CRISP-DM นี้จะประกอบด้วย 6 ขั้นตอนแต่ละขั้นตอนในรูปจะเป็นขั้นตอนที่ต่อเนื่องกันนั่นคือขั้นตอนถัดไปจะรอผลลัพธ์จากขั้นตอนก่อนหน้าซึ่งแสดงด้วยลูกศรที่เชื่อมระหว่างกล่องสี่เหลี่ยมแต่ละกล่อง ตัวอย่างเช่นเมื่อได้ผลลัพธ์จากขั้นตอนการเตรียมข้อมูล (Data Preparation) แล้วจะนำไปสร้างโมเดลจำแนกประเภทข้อมูลในขั้น Modeling และหลังจากนั้นอาจจะย้อนกลับมาเปลี่ยนแปลงข้อมูลให้ถูกต้องมากขึ้นเพื่อหวังว่าจะโมเดลที่ให้ความถูกต้องมากขึ้นก็ได้



ภาพที่ 3.1 แสดงกระบวนการ CRISP-DM

ที่มา : dataminingtrend.com (ม.ป.ป)

โดยกระบวนการ CRISP-DM จะประกอบไปด้วย 6 ขั้นตอน ดังนี้

1.) Business Understanding

เป็นขั้นตอนแรกที่ทำความเข้าใจปัญหาที่ได้ให้อยู่ในรูปโจทย์ของการวิเคราะห์ข้อมูลทางดาต้าไมน์นิ่ง และวางแผนในการดำเนินการโดยการวิเคราะห์ข้อมูลรูปแบบความสัมพันธ์ที่ซ่อนอยู่ในข้อมูลของข้อมูล

2.) Data Understanding

เริ่มทำการเก็บรวบรวมข้อมูล และตรวจสอบข้อมูลที่ได้ ดูความถูกต้อง และความน่าเชื่อถือของข้อมูล นำมาพิจารณาว่าจะใช้ข้อมูลทั้งหมดหรือจำเป็นต้องเลือกข้อมูลบางส่วนมาใช้ในการวิเคราะห์

3.) Data Preparation

ทำการแปลงข้อมูลที่ได้ทำการเก็บรวบรวมมาให้กลายเป็นข้อมูลที่สามารถนำไปวิเคราะห์ในขั้นถัดไปได้ โดยการทำข้อมูลให้ถูกต้องโดยใช้กระบวนการ data cleaning เพื่อให้ได้ข้อมูลอยู่ในรูปแบบและมาตรฐานเดียวกัน และกำจัดข้อมูลเสียออก

4.) Modeling

เป็นการวิเคราะห์ข้อมูลด้วยเทคนิคทาง Data Mining เช่น การจำแนกประเภทข้อมูล หรือแบ่งกลุ่มข้อมูล ซึ่งขั้นตอนนี้หลายเทคนิคจะถูกนำมาใช้เพื่อได้คำตอบที่ดีที่สุด

5.) Evaluation

ผลการวิเคราะห์ข้อมูลด้วยเทคนิคทาง Data Mining จากการสร้างโมเดล time series สามารถวัดประสิทธิภาพของผลลัพธ์ที่ได้ว่าตรงกับวัตถุประสงค์ที่ได้ตั้งไว้

6.) Deployment

การนำผลองค์ความรู้ที่ได้จากการได้ผลลัพธ์ด้วยเทคนิค Data Mining ไปใช้ประโยชน์ต่อในองค์กร หรือบริษัท

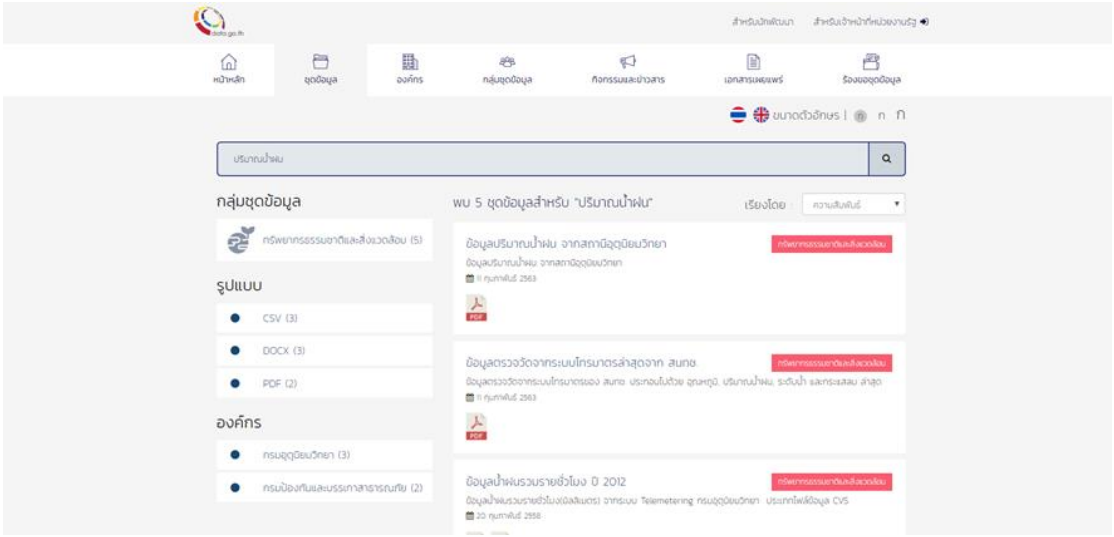
3.1.1 ความเข้าใจในธุรกิจ (Business Understanding)

ผู้วิเคราะห์ทำความเข้าใจกระบวนการทางข้อมูล และรับฟังปัญหารวมถึงการมองหาปัญหาจากเรื่องต่าง ๆ เพื่อที่นำผลการวิเคราะห์ข้อมูลไปใช้ให้เกิดประโยชน์ โดยเรียงลำดับความสำคัญ และกำหนดวัตถุประสงค์ที่จะนำไปสู่รูปแบบการวิเคราะห์ข้อมูล ให้อยู่ในรูปโจทย์ของการวิเคราะห์ฐานข้อมูล และวางแผนในการดำเนินการโดยการวิเคราะห์ข้อมูลรูปแบบความสัมพันธ์ที่ซ่อนอยู่ในข้อมูลของข้อมูลปริมาณน้ำฝนจากกรมอุตุนิยมวิทยาจากช่วงเวลาที่มีความถี่หรืออัตราที่เกิดขึ้นซ้ำ ๆ กัน เพื่อที่จะนำผลการวิเคราะห์ข้อมูลไปใช้ประโยชน์ตามความต้องการ และตามวัตถุประสงค์ที่จะนำไปสู่รูปแบบการวิเคราะห์ข้อมูล

3.1.2 การทำความเข้าใจข้อมูล (Data Understanding)

ผู้วิเคราะห์ข้อมูลทำการจัดเก็บรวบรวมข้อมูล และพิจารณาตรวจสอบข้อมูลที่ได้รับดูความถูกต้อง มีความน่าเชื่อถือของข้อมูล โดยพิจารณาว่าจะใช้ข้อมูลทั้งหมดหรือจำเป็นต้องเลือกข้อมูลบางส่วนมาใช้ในการวิเคราะห์ ให้สอดคล้องกับวัตถุประสงค์ที่กำหนดไว้ ผู้วิเคราะห์ได้สำรวจภาพรวมก่อนนำข้อมูลมาใช้ในการวิเคราะห์ข้อมูล จากเว็บไซต์ https://data.go.th ซึ่งเป็นศูนย์กลางการให้บริการข้อมูลเปิดภาครัฐ (Open Government Data) ภายใต้ชื่อ data.go.th ที่สำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน) ผู้วิเคราะห์ข้อมูลจึงได้ดำเนินการตามขั้นตอนดังนี้

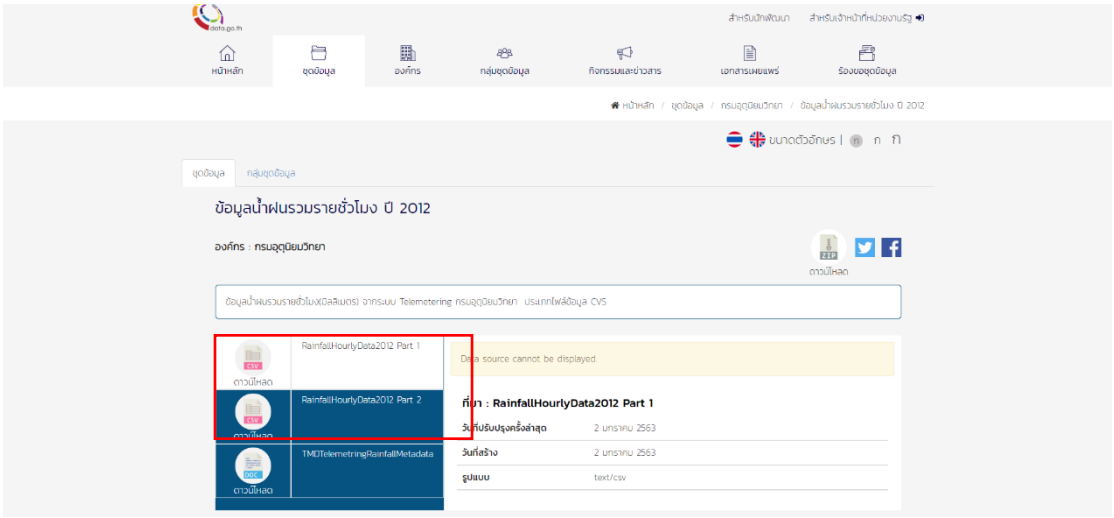
1.) ผู้วิเคราะห์รวบรวมข้อมูลจาก จากเว็บไซต์ https://data.go.th ผู้วิเคราะห์ทำการเลือกกลุ่มชุดข้อมูลทรัพยากรธรรมชาติ และสิ่งแวดล้อม จากนั้นเลือกหัวข้อข้อมูลปริมาณน้ำฝน จากกรมอุตุนิยมวิทยา



ภาพที่ 3.2 แสดงขั้นตอนการหาข้อมูลจากเว็บไซต์

ที่มา : data.go.th (2563)

2.) ผู้วิเคราะห์ทำการจัดเก็บ และรวบรวมข้อมูลทั้งหมด ของข้อมูลปริมาณน้ำฝนในช่วงปี พ.ศ. 2555 – 2557



ภาพที่ 3.3 แสดงขั้นตอนการดาวน์โหลดข้อมูล

ที่มา : data.go.th (2563)

3.) ตรวจสอบความถูกต้อง และความน่าเชื่อถือของข้อมูล โดยในไฟล์ข้อมูลจะบันทึกข้อมูลปริมาณน้ำฝนเป็นรายชั่วโมง โดยมีข้อมูล 7 แอตทริบิวท์คือ Station_id, Station_Name, Location, Latitude, Longitude, Province_Name, Date และแยกเป็นค่าน้ำฝนรายชั่วโมงทั้งหมด 24 แอตทริบิวท์ H00_01, H01_02, H02_03, H03_04, H04_05, H05_06, H06_07, H07_08, H08_09, H09_10, H10_11, H11_12, H12_13, H13_14, H14_15, H15_16, H16_17, H17_18, H18_19, H19_20, H20_21, H21_22, H22_23, H23_24 รวมทั้งหมด 31 แอตทริบิวท์

H00_01	H01_02	H02_03	H03_04	H04_05	H05_06	H06_07	H07_08	H08_09	H09_10	H10_11	H11_12	H12_13	H13_14	H14_15	H15_16	H16_17	H17_18	H18_19	H19_20	H20_21	H21_22	H22_23	H23_24
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	2	4	0	3	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0.5	0	0	6.5	0	0	0	13	0	6.5	1.5	0	0	0	1	0	0.5
0	4.5	4	0	1	0	1	0	0	0	0	0	0	0	2	1.5	0.5	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	1	0.5	0	0	0	5.5	2.5	10	9	0	0.5	1	0	0	1
0.5	0.5	1	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3.5
0	0	0	0	0	0	0	0	0	0	0	0	2	1	2	2.5	0	0.5	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	20.5	11.5	0.5	0	0	0	0	0	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0.5	0	0	0	0	1	0	0	0	0	0	0
0	0	0	4	4.5	3.5	0	0.5	0	0	0	0	0	0	0	0.5	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1.5	28.5	0	10	0	0.5	6.5	1.5	0	0	0	0	0	0	11	4.5	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1.5	0	0	0	0	0	1.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

ภาพที่ 3.4 ไฟล์ข้อมูลการบันทึกปริมาณน้ำฝนรายชั่วโมง

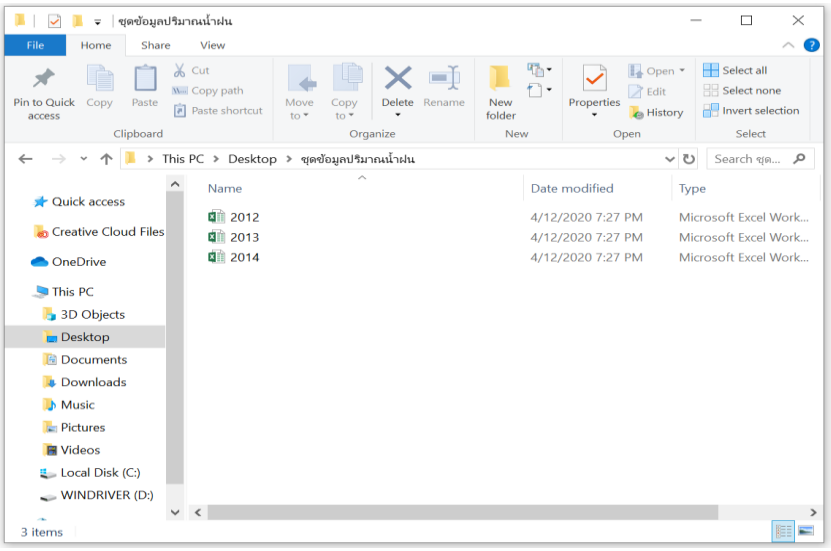
ที่มา : data.go.th (2563)

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

3.1.3 การเตรียมข้อมูล (Data Preparation)

ผู้วิเคราะห์ข้อมูลทำการแปลงข้อมูลที่ได้ทำการเก็บรวบรวม และทำการจัดการข้อมูลที่ได้ให้อยู่ในรูปแบบที่สามารถนำไปวิเคราะห์ในขั้นถัดไปได้ เพื่อลดความผิดพลาดการทำข้อมูลให้ถูกต้อง โดยใช้กระบวนการ data cleaning เพื่อให้ได้ข้อมูลอยู่ในรูปแบบและมาตรฐานเดียวกัน และกำจัดข้อมูลเสียออก ผู้วิเคราะห์ข้อมูลจึงได้ดำเนินการตามขั้นตอนดังนี้

1.) การปรับมาตรฐาน (Standardizing) ผู้วิเคราะห์ข้อมูลทำการแปลงข้อมูลให้อยู่ในรูปแบบเดียวกันพร้อมทั้งกำหนดแอตทริบิวท์เพื่อให้เป็นคีย์หลักของแต่ละไฟล์



ภาพที่ 3.5 ไฟล์ข้อมูลปริมาณน้ำฝนทั้งหมดที่ได้จากเว็บไซต์

ผู้วิเคราะห์ได้ทำการจัดการข้อมูลให้อยู่ในรูปแบบเดียวกัน โดยการรวบรวมข้อมูลจากรายชั่วโมงให้เป็นรายเดือน เพื่อให้ข้อมูลรายเดือนไปพยากรณ์ตามวัตถุประสงค์ที่กำหนดไว้ข้างต้น

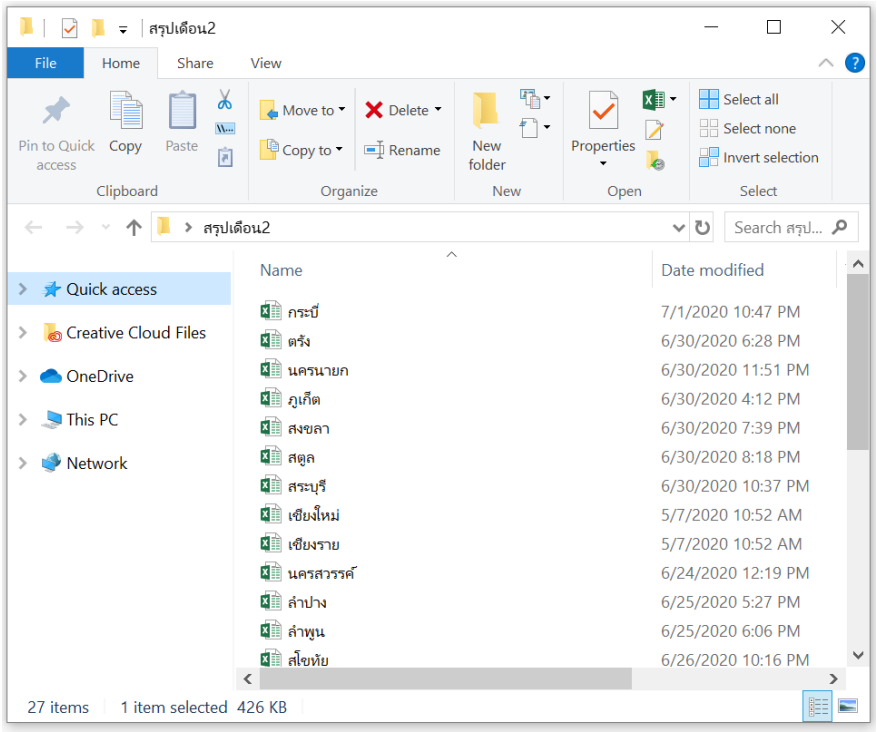
Month	ค่าปริมาณน้ำฝนจริง
1	440.5
2	1045
3	199.5
4	461.5
5	1155
6	795.5
7	1124.5
8	945.5
9	1533
10	1033.5
11	260
12	34

ภาพที่ 3.6 ข้อมูลที่รวบรวมให้เป็นรายเดือนในแต่ละจังหวัด

2.) การจัดหมวดหมู่ (Transfrom) ผู้วิเคราะห์นำ dataset ที่ผ่านการรวบรวมรายเดือนมาจัดหมวดหมู่เพื่อความถูกต้อง ผู้วิเคราะห์ดำเนินการจัดกลุ่มข้อมูลแบ่งตามภูมิภาค โดยการสร้างแอตทริบิวท์เพิ่มขึ้น 1 ตัว คือ แอตทริบิวท์ Provincial (ภูมิภาค) เพื่อแบ่งหมวดหมู่และแอตทริบิวท์ที่กำหนดให้ Province_Name เป็นคีย์หลักเพื่อเป็นตัวที่ใช้ระบุความสัมพันธ์ของ

ข้อมูลที่จะแสดงข้อมูลภาคต่าง ๆ ในแต่ละปี เตรียมความพร้อมการนำข้อมูลไปแสดงบน Tableau Public โดยดำเนินตามขั้นตอนดังนี้

2.1 การจัดกลุ่มตามภูมิภาค



ภาพที่ 3.7 ไฟล์ข้อมูลที่ต้องการจัดกลุ่มให้รวมอยู่ในภาคเดียวกัน

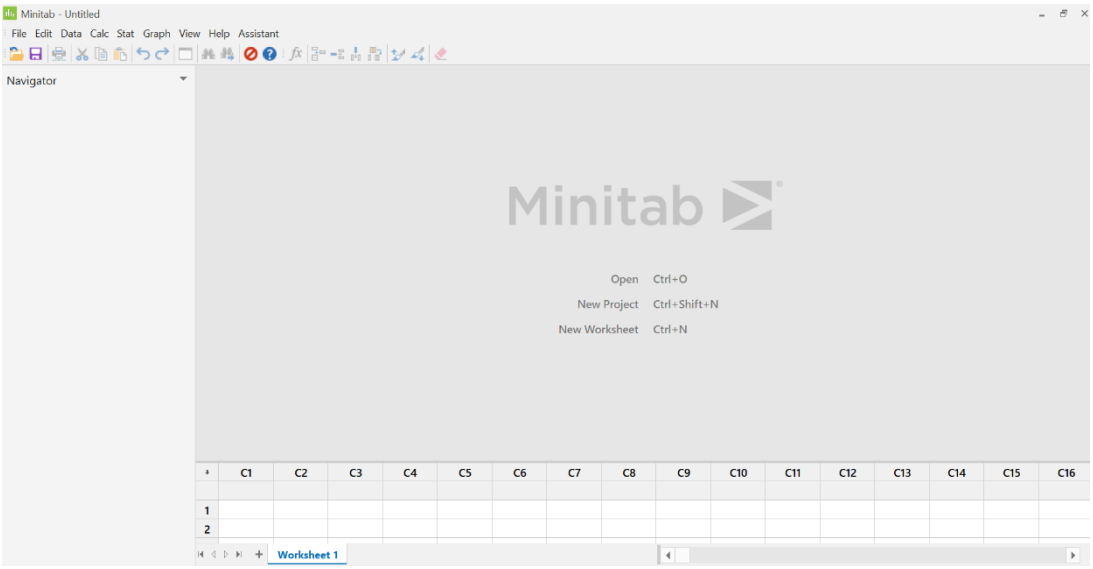
2.2 นำข้อมูลที่ผ่านการรวบรวมเป็นรายเดือนมาไว้ในไฟล์ภูมิภาคเดียวกัน

Province_Name	Provincial	Year	Month	Rainfall
ChiangMai	เหนือ	2555	4	1030
ChiangMai	เหนือ	2555	5	968.5
ChiangMai	เหนือ	2555	6	648.5
ChiangMai	เหนือ	2555	7	714
ChiangMai	เหนือ	2555	8	956.5
ChiangMai	เหนือ	2555	9	1049.5
ChiangMai	เหนือ	2555	10	630.5
ChiangRai	เหนือ	2555	4	995
ChiangRai	เหนือ	2555	5	1920.5
ChiangRai	เหนือ	2555	6	961
ChiangRai	เหนือ	2555	7	1644.5
ChiangRai	เหนือ	2555	8	1716.5
ChiangRai	เหนือ	2555	9	1452.5
ChiangRai	เหนือ	2555	10	1123.5
Phrae	เหนือ	2555	4	707
Phrae	เหนือ	2555	5	1272

ภาพที่ 3.8 แสดงการจัดกลุ่มภาคเหนือ

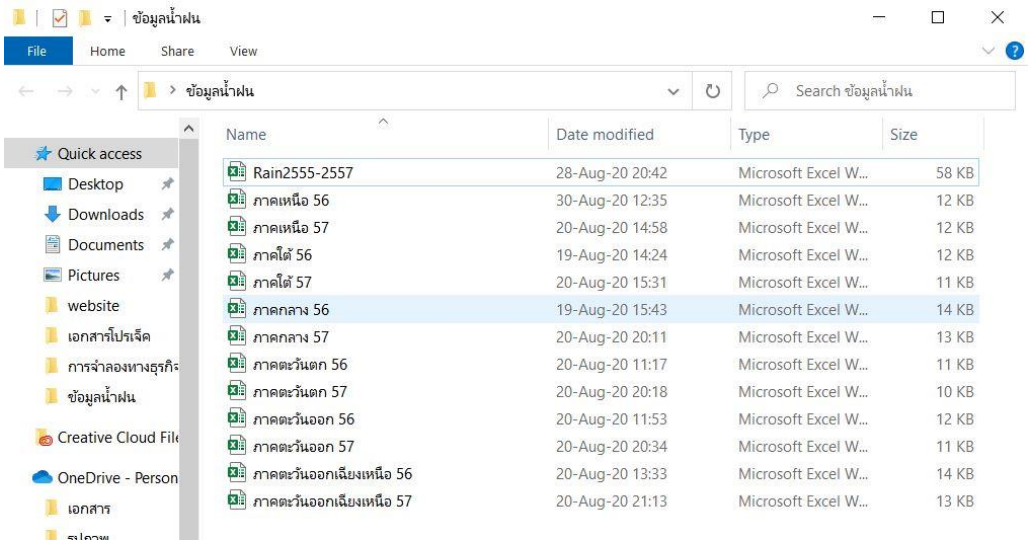
3. การอัปโหลดข้อมูล (Loading) โดยใช้ Minitab ดำเนินกระบวนการ การโหลดข้อมูลเข้าระบบฐานข้อมูลของ Minitab เพื่อเป็นการพยากรณ์ของข้อมูล โดยมีขั้นตอนดำเนินการดังนี้

3.1 โหลดข้อมูลที่ได้เตรียมไว้แล้วลงโปรแกรม Minitab



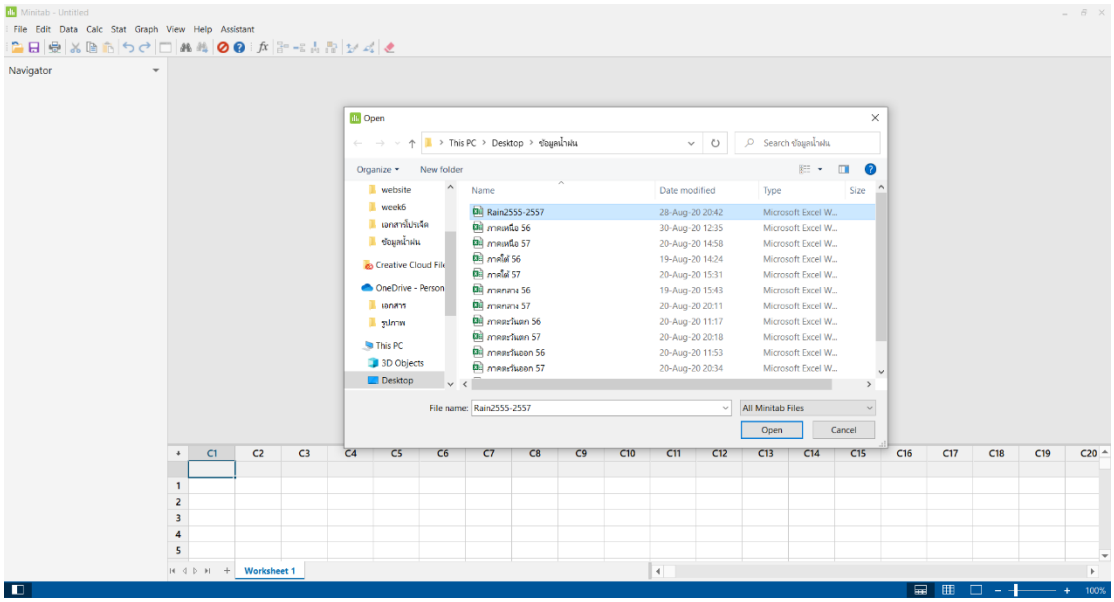
ภาพที่ 3.9 หน้าต่างพื้นที่การทำงานของโปรแกรม Minitab

3.2 จัดเตรียมข้อมูลที่นำมาใช้ในการพยากรณ์



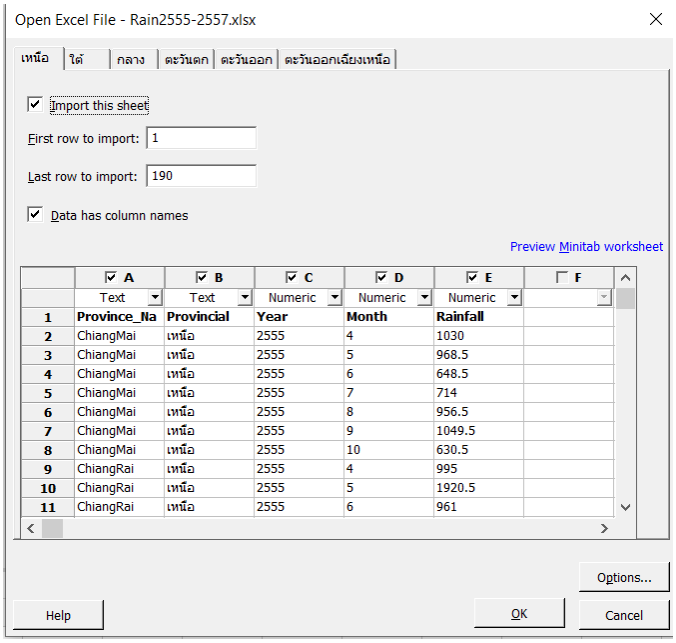
ภาพที่ 3.10 ข้อมูลที่นำมาใช้ในการพยากรณ์

3.3 ใช้ฟังก์ชัน File open เพื่อทำการนำข้อมูลเข้าฐานข้อมูลรูปแบบของ Microsoft Excel และเลือกข้อมูลที่ผ่านการจัดกลุ่มข้อมูลแล้ว และนำข้อมูลเข้าโปรแกรม



ภาพที่ 3.11 เลือกไฟล์ที่ต้องการนำข้อมูลมาพยากรณ์

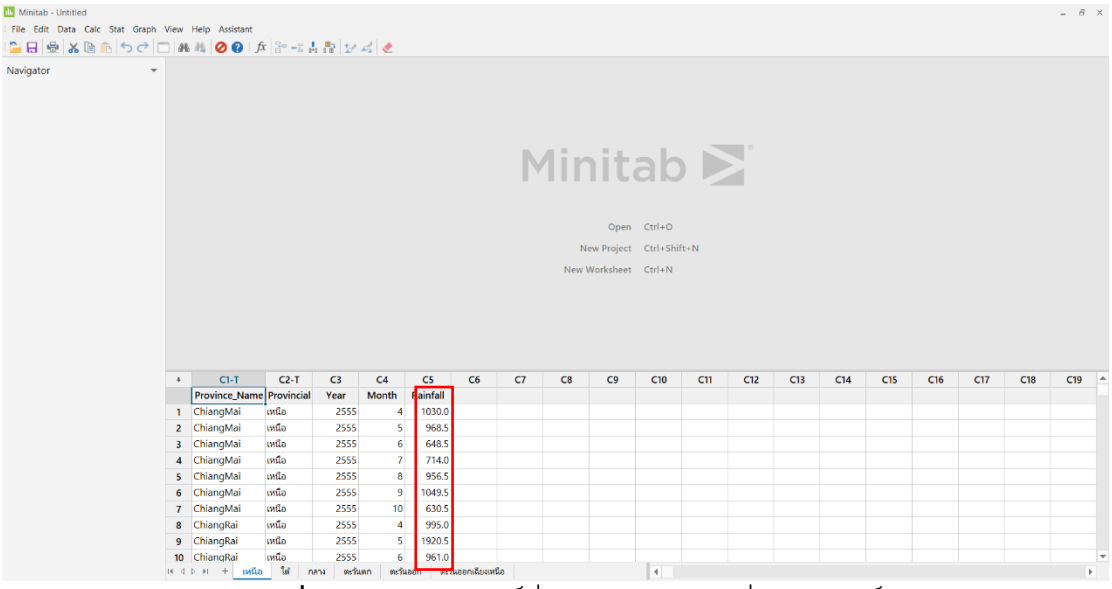
3.4 ทำการ Import ข้อมูลทั้งหมดของแต่ละภาคเข้าโปรแกรมเพื่อทำการหาค่าผลลัพธ์ของการพยากรณ์



	A	B	C	D	E	F
	Text	Text	Numeric	Numeric	Numeric	
1	Province_Na	Provincial	Year	Month	Rainfall	
2	ChiangMai	เหนือ	2555	4	1030	
3	ChiangMai	เหนือ	2555	5	968.5	
4	ChiangMai	เหนือ	2555	6	648.5	
5	ChiangMai	เหนือ	2555	7	714	
6	ChiangMai	เหนือ	2555	8	956.5	
7	ChiangMai	เหนือ	2555	9	1049.5	
8	ChiangMai	เหนือ	2555	10	630.5	
9	ChiangRai	เหนือ	2555	4	995	
10	ChiangRai	เหนือ	2555	5	1920.5	
11	ChiangRai	เหนือ	2555	6	961	

ภาพที่ 3.12 Import ข้อมูลเพื่อหาค่าการพยากรณ์

3.5 ข้อมูลที่ถูก Import ลงโปรแกรมจะมีทั้งหมด 5 แอตทริบิวท์ ได้แก่ Province_Name, Provincial, year, Month, Rainfall และจะนำแอตทริบิวท์ Rainfall มาคำนวณหาค่าพยากรณ์ด้วยวิธีต่าง ๆ



	C1-T	C2-T	C3	C4	C5
	Province_Name	Provincial	Year	Month	Rainfall
1	ChiangMai	เหนือ	2555	4	1030.0
2	ChiangMai	เหนือ	2555	5	968.5
3	ChiangMai	เหนือ	2555	6	648.5
4	ChiangMai	เหนือ	2555	7	714.0
5	ChiangMai	เหนือ	2555	8	956.5
6	ChiangMai	เหนือ	2555	9	1049.5
7	ChiangMai	เหนือ	2555	10	630.5
8	ChiangRai	เหนือ	2555	4	995.0
9	ChiangRai	เหนือ	2555	5	1920.5
10	ChiangRai	เหนือ	2555	6	961.0

ภาพที่ 3.13 แอตทริบิวท์ที่นำมาคำนวณหาค่าพยากรณ์

3.1.4 การสร้างแบบจำลอง (Modeling)

ผู้วิเคราะห์ข้อมูลจะวิเคราะห์ด้วยเทคนิคทางสถิติ ทำการเรียนรู้จากข้อมูลเดิมของปริมาณน้ำฝนเพื่อให้ได้คำตอบที่ดีที่สุด การเปรียบเทียบความคลาดเคลื่อนของวิธีการพยากรณ์ โดยนำโมเดล Time series analysis มาช่วยในการหาค่าความคลาดเคลื่อนของการพยากรณ์ในวิธีต่าง ๆ มาเปรียบเทียบตัวโปรแกรมว่าโปรแกรมไหนให้ความถูกต้องแม่นยำมากที่สุดนำมาใช้ในการพยากรณ์การเกิดฝน การพยากรณ์ในรูปแบบคณิตศาสตร์ใช้รูปแบบสมการทางคณิตศาสตร์ ผู้วิเคราะห์ข้อมูลได้เลือกใช้เทคนิครูปแบบอนุกรมเวลา Time series Models ด้วยชุดข้อมูลปริมาณน้ำฝนรายเดือนของแต่ละภูมิภาค โดยมีขั้นตอนการดำเนินการดังนี้

ในการพยากรณ์โดยใช้รูปแบอนุกรมเวลาจะเป็นการพยากรณ์ข้อมูลในอนาคตจากข้อมูลในอดีตเท่านั้น ตัวแปรอื่น ๆ จะไม่นำมาพิจารณา ผู้วิเคราะห์ข้อมูลได้เลือกข้อมูลในรูปแบบมีแนวโน้ม (Trend)

	A	B	C	D	E
1	Province_Name	Provincial	Year	Month	Rainfall
2	ChiangMai	เหนือ	2555	4	1030
3	ChiangMai	เหนือ	2555	5	968.5
4	ChiangMai	เหนือ	2555	6	648.5
5	ChiangMai	เหนือ	2555	7	714
6	ChiangMai	เหนือ	2555	8	956.5
7	ChiangMai	เหนือ	2555	9	1049.5
8	ChiangMai	เหนือ	2555	10	630.5
9	ChiangRai	เหนือ	2555	4	995
10	ChiangRai	เหนือ	2555	5	1920.5
11	ChiangRai	เหนือ	2555	6	961
12	ChiangRai	เหนือ	2555	7	1644.5
13	ChiangRai	เหนือ	2555	8	1716.5
14	ChiangRai	เหนือ	2555	9	1452.5
15	ChiangRai	เหนือ	2555	10	1123.5

ภาพที่ 3.14 ตารางข้อมูลปริมาณน้ำฝน

จากรูปภาพที่ 3.14 ประกอบไปด้วย 5 แอตทริบิวท์ คือ

1.) Province_Name เก็บข้อมูลชื่อจังหวัดทั้งหมดทั่วประเทศ ทำหน้าที่เป็นคีย์หลักในการอ้างอิงถึงข้อมูลอื่น ๆ

2.) Provincial เก็บข้อมูลภูมิภาคต่าง ๆ

3.) Year ปีของปริมาณน้ำฝน

4.) Month เก็บข้อมูลเดือนทุกเดือนตั้งแต่ เดือน เมษายน ถึง ตุลาคม เป็นตัวเลข

5.) Rainfall เก็บค่าปริมาณน้ำฝนจริงรายเดือนตั้งแต่ 0 ถึง ค่ามากที่สุด

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

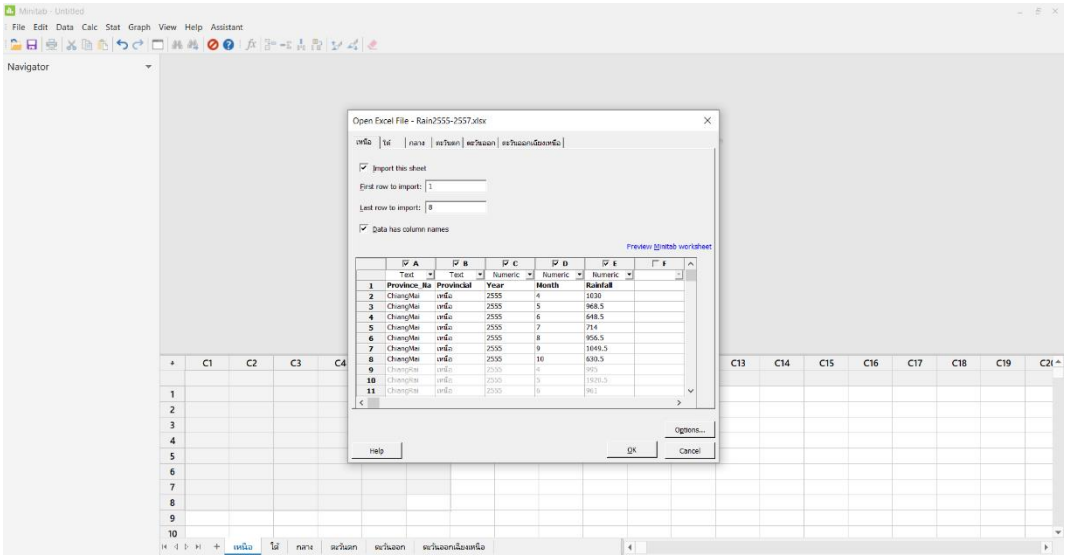
$$EMA_n = aP_n + EMA_{n-1}(1 - a)$$

Province_Nam	Provincial	Year	Month	Rainfall	EMA
ChiangMai	เหนือ	2555	4	1030	894.50
ChiangMai	เหนือ	2555	5	968.5	983.93
ChiangMai	เหนือ	2555	6	648.5	973.75
ChiangMai	เหนือ	2555	7	714	759.08
ChiangMai	เหนือ	2555	8	956.5	729.33
ChiangMai	เหนือ	2555	9	1049.5	879.26
ChiangMai	เหนือ	2555	10	630.5	991.62
ChiangRai	เหนือ	2555	4	995	1448.33
ChiangRai	เหนือ	2555	5	1920.5	1149.13
ChiangRai	เหนือ	2555	6	961	1658.24
ChiangRai	เหนือ	2555	7	1644.5	1198.06
ChiangRai	เหนือ	2555	8	1716.5	1492.71
ChiangRai	เหนือ	2555	9	1452.5	1640.41
ChiangRai	เหนือ	2555	10	1123.5	1516.39
Phrae	เหนือ	2555	4	707	1101.5
Phrae	เหนือ	2555	5	1272	841.13

ภาพที่ 3.15 ตารางการพยากรณ์อนุกรมเวลาด้วยวิธี Single Exponential Moving Average

การพยากรณ์โดยการวิเคราะห์อนุกรมเวลาของการทำเหมืองข้อมูล ผ่านโปรแกรม Minitab โดยผลลัพธ์ที่ได้จากการพยากรณ์ล่วงหน้า ผู้วิเคราะห์ข้อมูลได้นำข้อมูลที่ผ่านการวิเคราะห์ด้วยเทคนิค Data Mining จากการสร้างโมเดล Time Series Models จึงทำข้อมูลดังกล่าวมาทดสอบกับโปรแกรม Minitab ซึ่งมีขั้นตอนการทำงานดังนี้

ขั้นตอนที่ 1 ทำการอัปโหลดข้อมูลปริมาณน้ำฝนทั้งหมด 7 เดือนในช่วงฤดูฝนเข้าโปรแกรม เพื่อเตรียมความพร้อมในการพยากรณ์

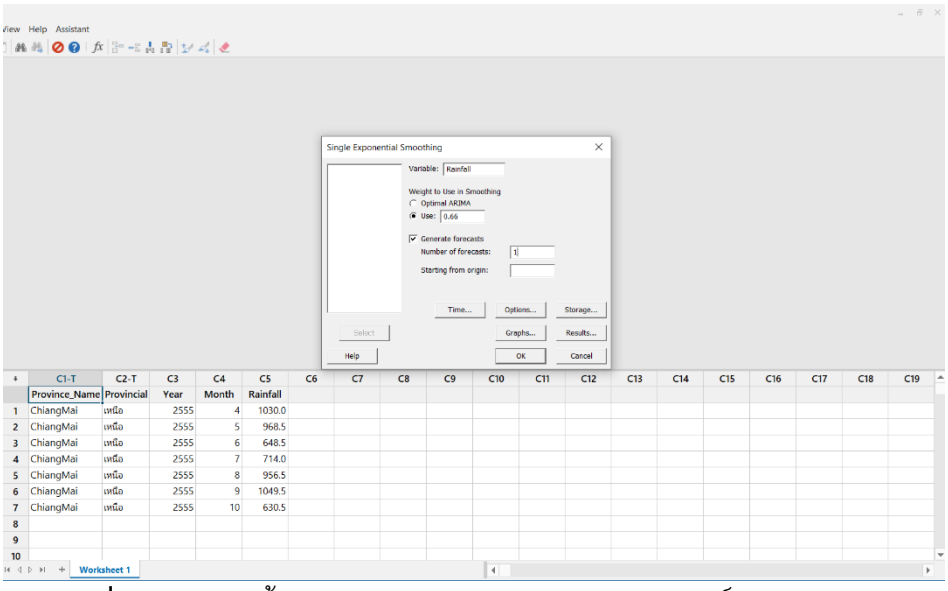


ภาพที่ 3.16 แสดงขั้นตอนการนำข้อมูลเข้าโปรแกรม Minitab

1 Single Exponential Moving Average

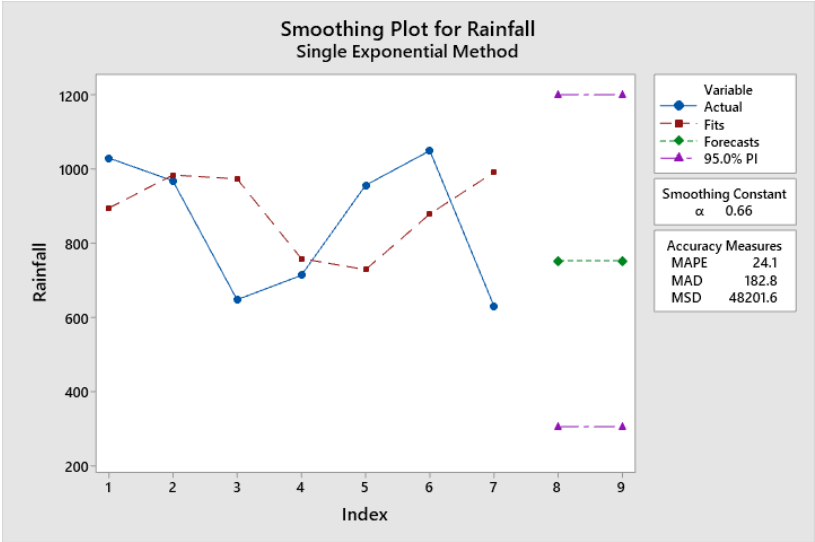
ขั้นตอนที่ 1 เลือกโมเดล Time Series ใช้วิธี Single Exponential Moving Average และกำหนด Variable เป็นข้อมูลที่ต้องการจะทำการพยากรณ์

ขั้นตอนที่ 2 กำหนดค่า α เท่ากับ 0.66 และกำหนดจำนวนเลขที่จะพยากรณ์



	C1-T	C2-T	C3	C4	C5
	Province_Name	Provincial	Year	Month	Rainfall
1	ChiangMai	เหนือ	2555	4	1030.0
2	ChiangMai	เหนือ	2555	5	968.5
3	ChiangMai	เหนือ	2555	6	648.5
4	ChiangMai	เหนือ	2555	7	714.0
5	ChiangMai	เหนือ	2555	8	956.5
6	ChiangMai	เหนือ	2555	9	1049.5
7	ChiangMai	เหนือ	2555	10	630.5

ภาพที่ 3.17 แสดงขั้นตอนการกำหนดรูปแบบการพยากรณ์วิธี Single EMA

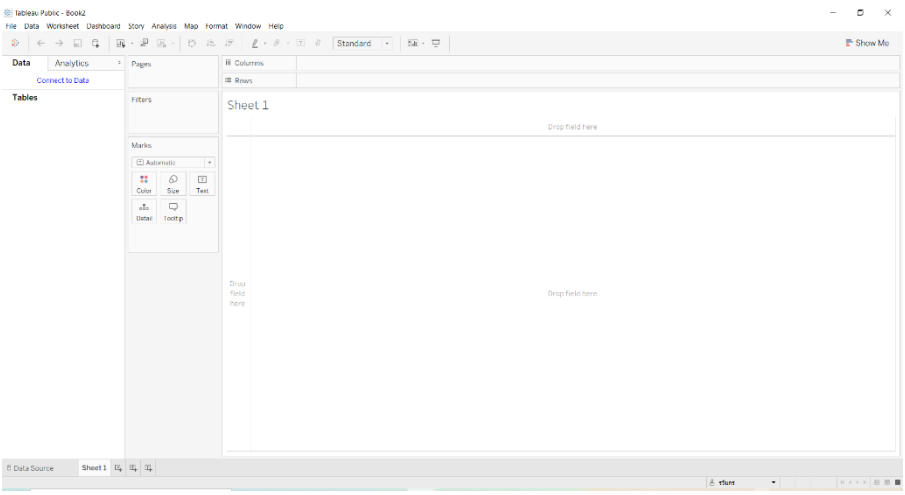


ภาพที่ 3.18 แสดงผลลัพธ์กราฟวิธี Single EMA

จากภาพที่ 3.18 จะเห็นได้ว่าการใช้วิธีการ Single Exponential Moving Average ในการพยากรณ์จะมีความคลาดเคลื่อน MAPE 24.1 %

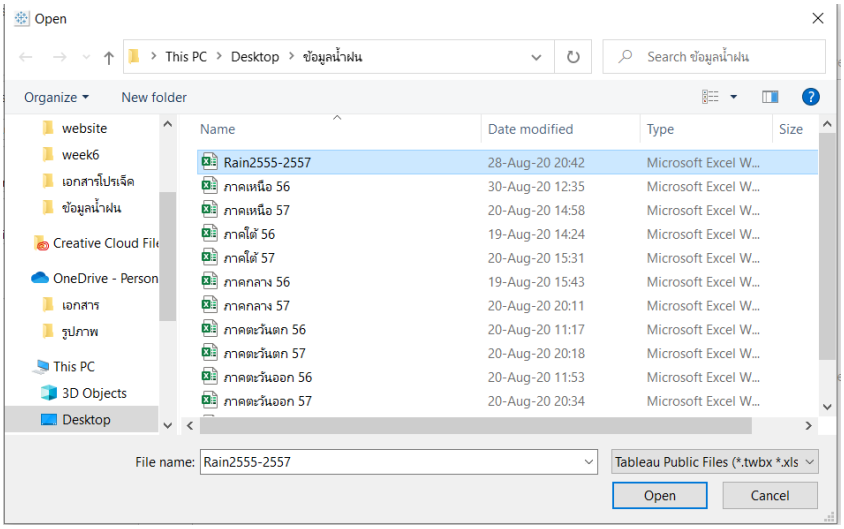
2 การพยากรณ์โดย การวิเคราะห์อนุกรมเวลาของข้อมูล (Data mining time series analysis) ผ่านโปรแกรม Tableau Public ผู้วิเคราะห์ได้ข้อมูลที่ผ่านการวิเคราะห์ข้อมูลด้วยเทคนิค Data Mining จากการสร้างโมเดล Time Series Models จึงนำข้อมูลดังกล่าวมาทดสอบกับโปรแกรม Tableau Public ซึ่งมีขั้นตอนการทำงานดังนี้

ขั้นตอนที่ 1 หน้าต่างพื้นที่ทำงานของโปรแกรม tableau



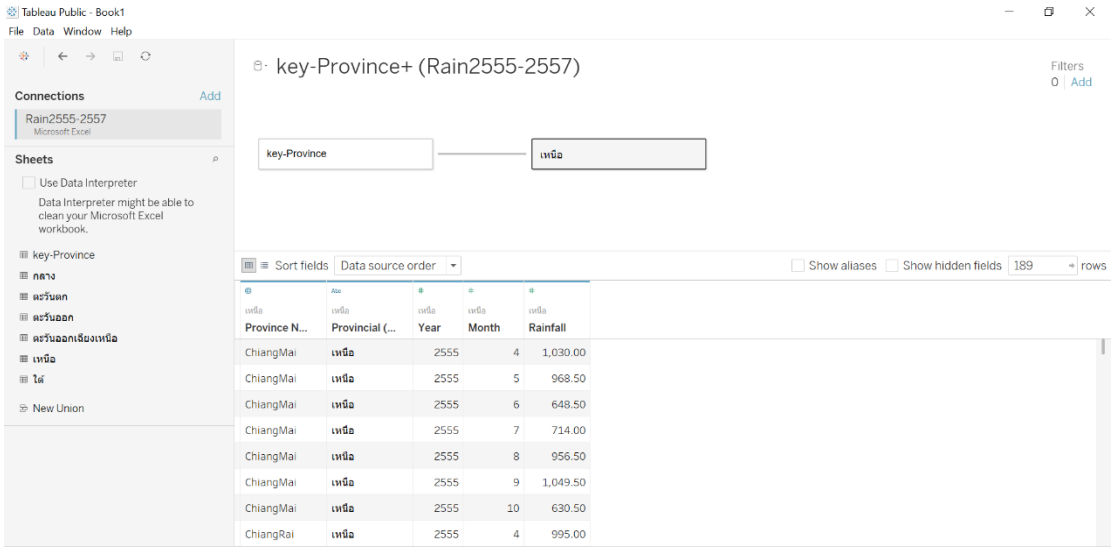
ภาพที่ 3.19 หน้าต่างพื้นที่ทำงานของโปรแกรม tableau

ขั้นตอนที่ 2 จัดเตรียมข้อมูลที่จะนำมาแสดงผลลัพธ์



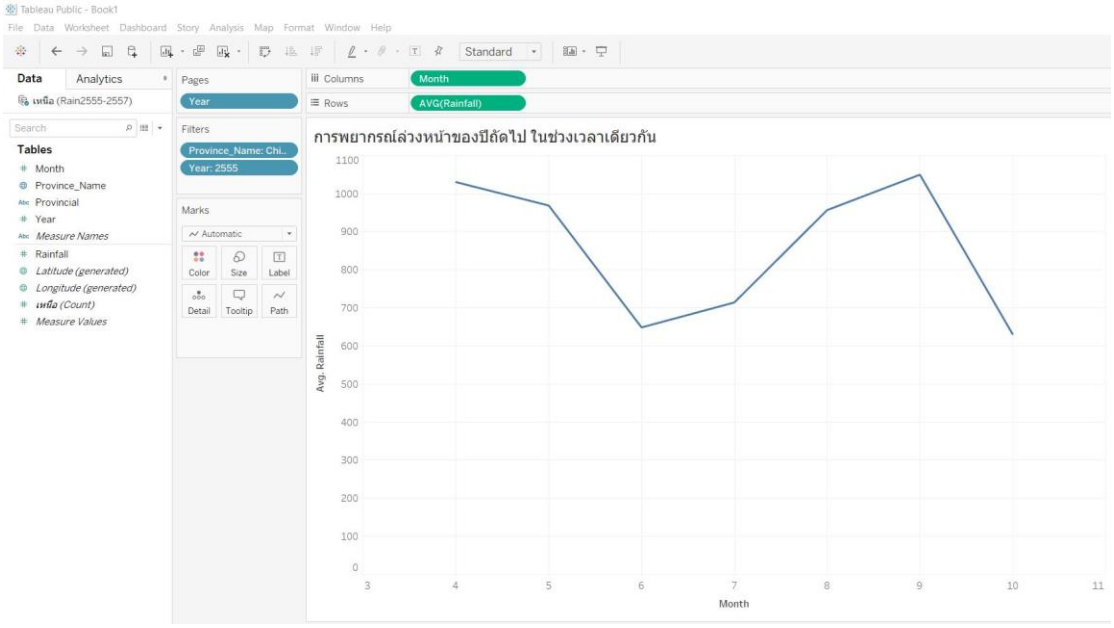
ภาพที่ 3.20 Import ข้อมูลที่ต้องการหาผลลัพธ์

ขั้นตอนที่ 3 ทำการใช้เครื่องมือทำเหมืองข้อมูล (Data Mining) ของโปแกรม Tableau Public ในการพยากรณ์ข้อมูลปริมาณน้ำฝน



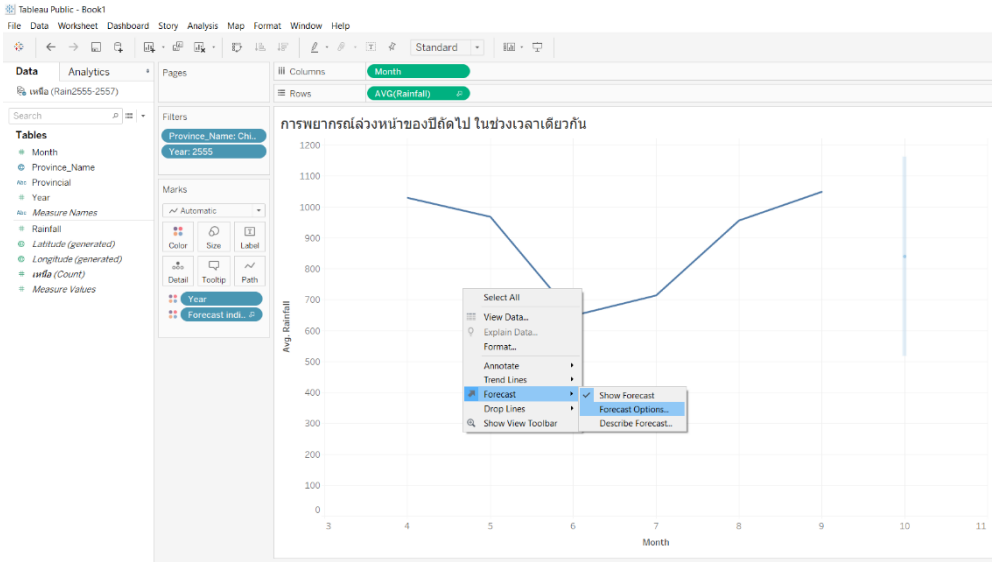
ภาพที่ 3.21 Import แสดงผลการโหลดฐานข้อมูลเข้าสู่โปรแกรม Tableau Public

ขั้นตอนที่ 4 การเตรียมกราฟข้อมูลก่อนการวิเคราะห์ข้อมูล



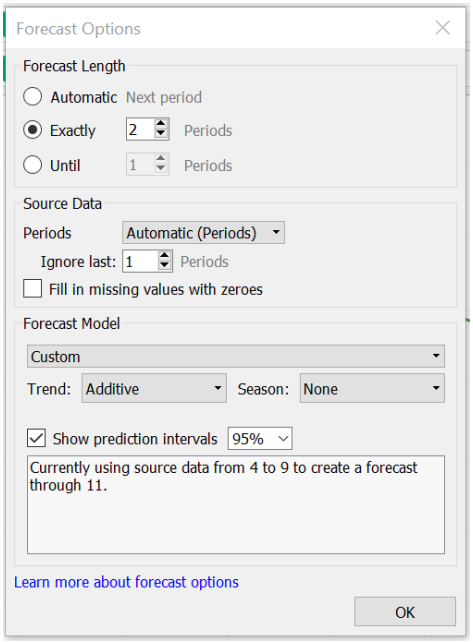
ภาพที่ 3.22 แสดงการเตรียมกราฟข้อมูลก่อนการวิเคราะห์ข้อมูล

ขั้นตอนที่ 5 เครื่องมือในการวิเคราะห์ข้อมูลของโปรแกรม Tableau Public



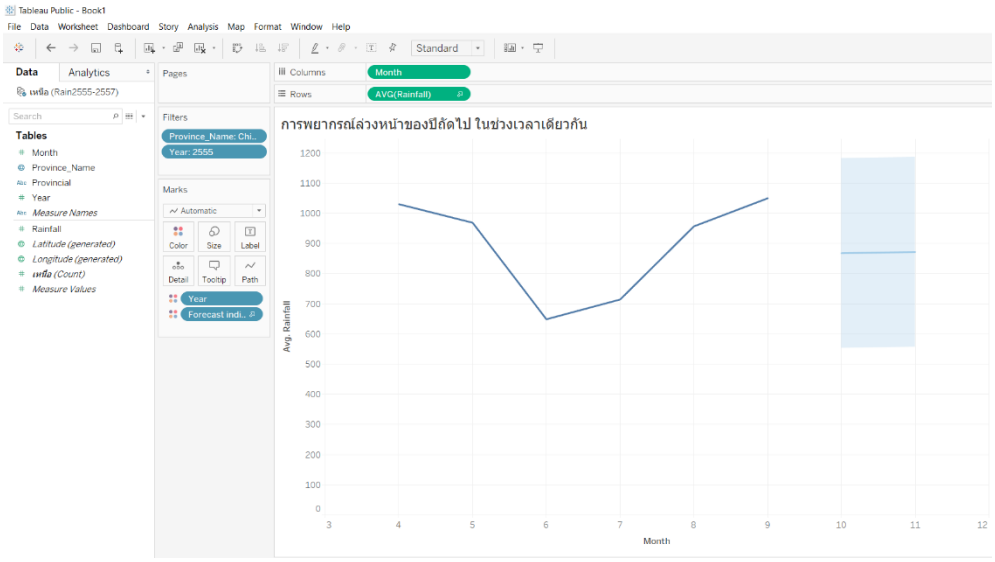
ภาพที่ 3.23 แสดงเครื่องมือในการวิเคราะห์ข้อมูลของโปรแกรม Tableau Public

ขั้นตอนที่ 6 การใช้เครื่องมือทำเหมืองข้อมูล (Data Mining) ของโปแกรม Tableau Public ในการพยากรณ์ข้อมูลปริมาณน้ำฝน กำหนดค่า Exactly เป็น 2 ของปีถัดไปในช่วงเวลาเดียวกัน



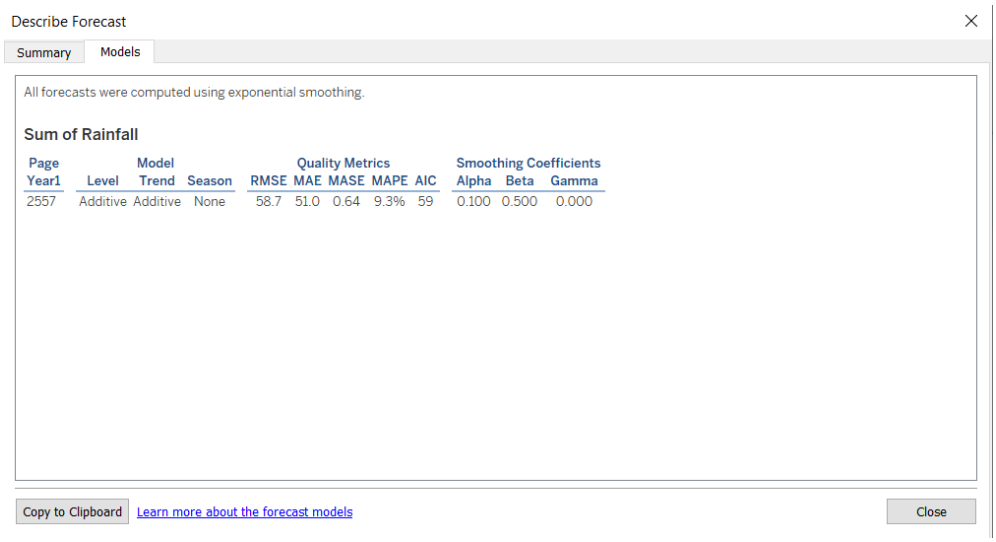
ภาพที่ 3.24 แสดงการใช้เครื่องมือทำเหมืองข้อมูลของโปรแกรม Tableau Public

ขั้นตอนที่ 7 กราฟผลลัพธ์การพยากรณ์โดย Tableau Public



ภาพที่ 3.25 กราฟแสดงผลลัพธ์การพยากรณ์โดย Tableau Public

ขั้นตอนที่ 8 ผลการคำนวณความคลาดเคลื่อน MAPE โดยโปรแกรม Tableau Public



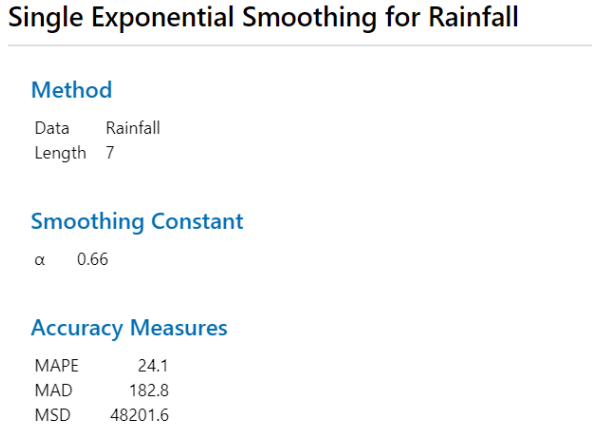
ภาพที่ 3.26 แสดงผลการคำนวณความคลาดเคลื่อน MAPE โดยโปรแกรม Tableau Public

จะเห็นได้ว่าการใช้โปรแกรม Tableau Public ในการพยากรณ์จะมีความคลาดเคลื่อน MAPE 9.3 %

3.1.5 การประเมินประสิทธิภาพ (Evaluation)

เป็นขั้นตอนก่อนนำผลลัพธ์ที่ได้จากขั้นตอนที่ 4 ไปใช้งานผู้วิเคราะห์ข้อมูลวัดประสิทธิภาพของผลลัพธ์ได้กับวัตถุประสงค์ที่ตั้งไว้ในขั้นตอนแรก ว่ามีนัยสำคัญหรือความน่าเชื่อถือมากน้อยเพียงใด โดยการนำไปตรวจสอบค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percentage Error : MAPE) จากการทดลองโมเดลเทคนิค Time Series Models ด้วย การวิเคราะห์อนุกรมเวลาของข้อมูล (Data mining time series analysis) ผ่าน 2 โปรแกรมมีค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ยดังนี้

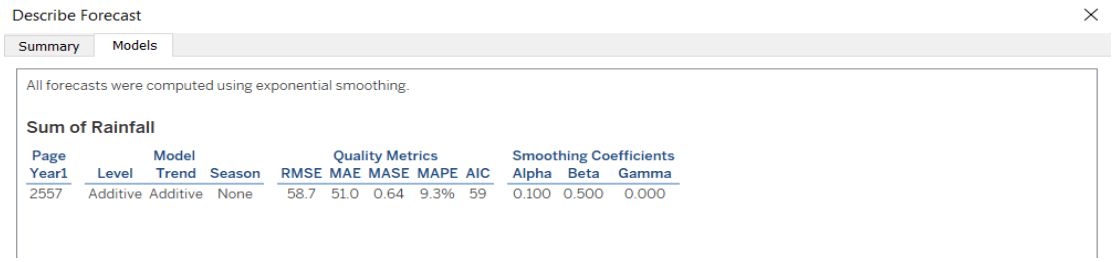
1) การวัดประสิทธิภาพของผลลัพธ์ด้วยโปรแกรม Minitab



ภาพที่ 3.27 ผลลัพธ์การวัดประสิทธิภาพของผลลัพธ์ด้วยโปรแกรม Minitab

ผลการตรวจสอบค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percentage Error : MAPE) จากการทดลองโมเดลเทคนิค Time Series Models ด้วยการวิเคราะห์อนุกรมเวลาของข้อมูล (Data mining time series analysis) ผ่านโปรแกรม Minitab ได้ผลลัพธ์ถึง 24.1%

2) การวัดประสิทธิภาพของผลลัพธ์ด้วยโปรแกรม Tableau Public



ภาพที่ 3.28 ผลลัพธ์การวัดประสิทธิภาพของผลลัพธ์ด้วยโปรแกรม Tableau Public

ผลการตรวจสอบค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percentage Error : MAPE) จากการทดลองโมเดลเทคนิค Time Series Models ด้วยการวิเคราะห์อนุกรมเวลาของข้อมูล (Data mining time series analysis) ผ่านโปรแกรม Tableau Public ได้ผลลัพธ์ต่ำสุดถึง 9.3 %

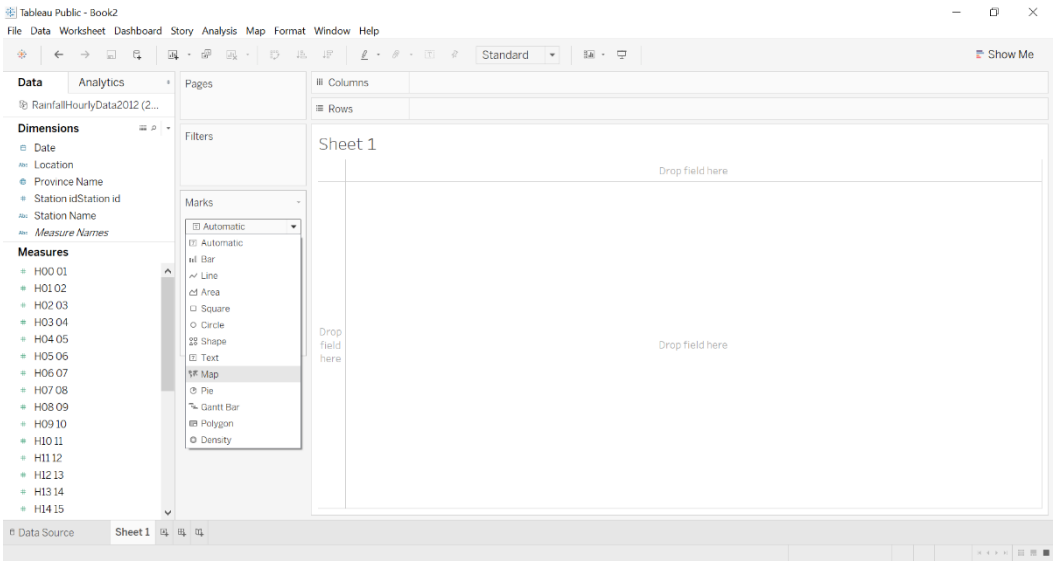
ผลการทดสอบประสิทธิภาพโมเดล และการประเมินค่าความคลาดเคลื่อน MAPE ที่ค่าคงที่เกิดยิ่งน้อยยิ่งเป็นผลดีแสดงว่าค่าพยากรณ์ออกมานั้นมีความคลาดเคลื่อนจากค่าจริงเพียงใด ผู้วิเคราะห์จึงเลือกใช้โมเดลการวิเคราะห์อนุกรมเวลาของข้อมูล (Data mining time series analysis) ผ่านโปรแกรม Tableau Public ให้ผลลัพธ์ที่น่าพอใจอยู่ในระดับค่อนข้างดี

สามารถนำโมเดลไปใช้งานได้ ผู้วิเคราะห์จึงเลือกใช้โมเดลอนุกรมเวลาของ Tableau Public ในการพยากรณ์นำไปสู่การแสดงสารสนเทศสร้างประโยชน์แก่ข้อมูลปริมาณน้ำฝนที่มีอยู่ในการเตรียมพร้อม และวางแผนการปฏิบัติตัวในพื้นที่บุคคลนั้นอาศัยอยู่

3.1.6 การนำผลลัพธ์ไปใช้งาน (Deployment)

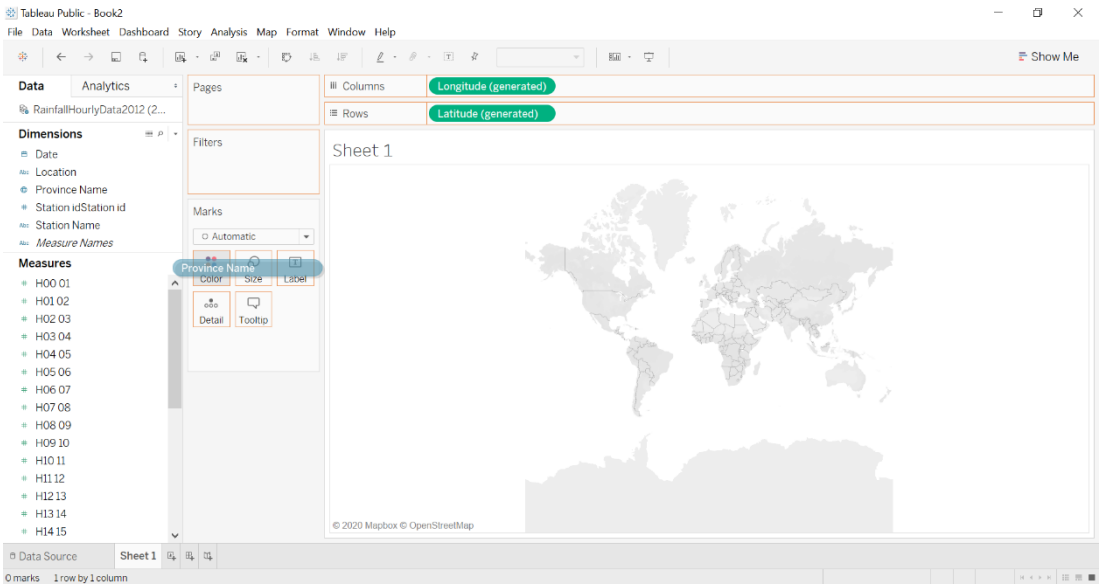
ผู้วิเคราะห์ข้อมูลนำผลองค์ความรู้ที่ได้เหล่านี้ไปนำเสนอข้อมูลแบบ visualization ด้วยการแสดงผลข้อมูลในรูปแบบของภาพโดยใช้โปรแกรม Tableau Public โดยดำเนินการตามขั้นตอนดังนี้

6.1 เข้าหน้า DashBoard การทำงานข้อมูลที่ Import จะอยู่แถบด้านซ้าย



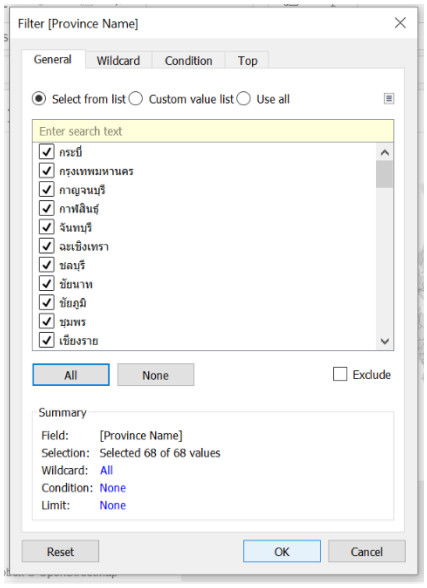
ภาพที่ 3.29 แสดงขั้นตอนการปรับ Mark ให้เป็น Map

6.2 กำหนด Columns เป็น Longitude กำหนด Rows เป็น Latitude และกำหนด Province Name ใส่ใน Color

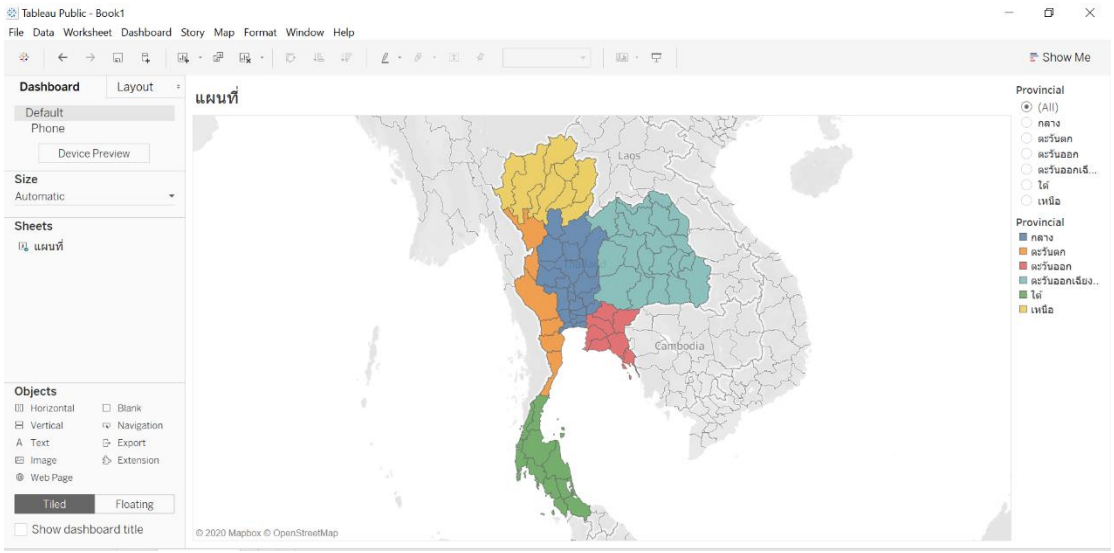


ภาพที่ 3.30 แสดงขั้นตอนการโชว์พื้นที่จังหวัดบน Map

6.3 กำหนด Filter and then add ทำการเลือกชื่อจังหวัดทั้งหมดเพื่อไปแสดงบนแผนที่

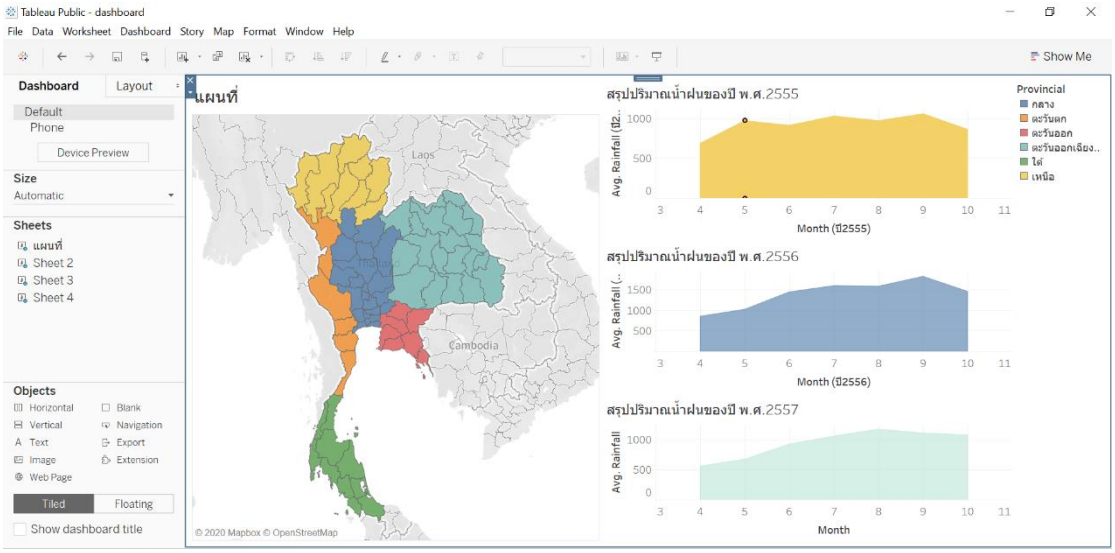


ภาพที่ 3.31 แสดงขั้นตอนการเลือกชื่อจังหวัดทั้งหมด



ภาพที่ 3.32 แสดงผลลัพธ์แผนที่ประเทศไทยเป็นสี

6.4 หน้าแสดงผลข้อมูลปริมาณน้ำฝนทั้ง 74 จังหวัด



ภาพที่ 3.33 แสดงผลข้อมูลปริมาณน้ำฝนทั้ง 74 จังหวัด

3.2 การออกแบบและพัฒนาเว็บไซต์

การวางแผนการจัดลำดับ เนื้อหาสาระของเว็บไซต์ ออกเป็นหมวดหมู่ เพื่อจัดทำเป็นโครงสร้างในการจัดวางหน้าเว็บเพจทั้งหมด ทำให้เห็นโครงสร้างทั้งหมดของเว็บไซต์และการออกแบบโครงสร้างหรือจัดระเบียบของข้อมูลที่ชัดเจน แยกย่อยเนื้อหาออกเป็นหาออกเป็นส่วนต่าง ๆ ที่สัมพันธ์และให้มีอยู่ในมาตรฐานเดียวกัน จะช่วยให้นำมาใช้งานและง่าย ต่อการเข้าอ่านเนื้อหาของผู้ใช้เว็บไซต์

3.2.1 การออกแบบ Wireframe หน้าจอเว็บไซต์

1.) หน้าแรกของเว็บไซต์ แสดงเมนูต่าง ๆ ของหน้าเว็บไซต์

ภาพที่ 3.34 แสดงหน้าแรกของเว็บไซต์แสดงเมนูต่าง ๆ ของหน้าเว็บไซต์

2.) หน้าแสดงข่าวสาร และแหล่งความรู้

ภาพที่ 3.35 หน้าแสดงข่าวสาร และแหล่งความรู้

3.) หน้าแสดงข้อมูลปริมาณน้ำฝนย้อนหลัง

ภาพที่ 3.36 หน้าแสดงข้อมูลปริมาณน้ำฝนย้อนหลัง

4.) หน้าแสดงโมเดลที่ใช้ในการพยากรณ์

ภาพที่ 3.37 หน้าแสดงโมเดลที่ใช้ในการพยากรณ์

5.) หน้าแสดงการพยากรณ์ปริมาณน้ำฝนของภาคเหนือ

ภาพที่ 3.38 หน้าแสดงการพยากรณ์ปริมาณน้ำฝนของภาคเหนือ

6.) หน้าแสดงการพยากรณ์ปริมาณน้ำฝนของภาคตะวันออกเฉียงเหนือ

ภาพที่ 3.39 หน้าแสดงการพยากรณ์ปริมาณน้ำฝนของภาคตะวันออกเฉียงเหนือ

7.) หน้าแสดงการพยากรณ์ปริมาณน้ำฝนของภาคตะวันออก

ภาพที่ 3.40 หน้าแสดงการพยากรณ์ปริมาณน้ำฝนของภาคตะวันออก

8.) หน้าแสดงการพยากรณ์ปริมาณน้ำฝนของภาคตะวันตก

ภาพที่ 3.41 หน้าแสดงการพยากรณ์ปริมาณน้ำฝนของภาคตะวันตก

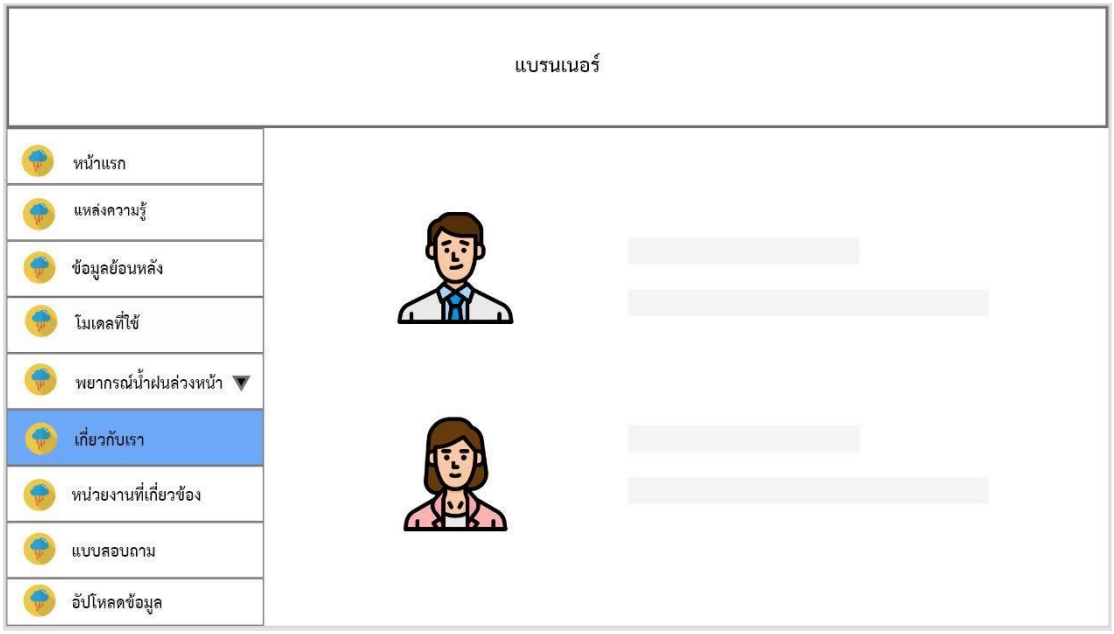
9.) หน้าแสดงการพยากรณ์ปริมาณน้ำฝนของภาคกลาง

ภาพที่ 3.42 หน้าแสดงการพยากรณ์ปริมาณน้ำฝนของภาคกลาง

10.) หน้าแสดงการพยากรณ์ปริมาณน้ำฝนของภาคใต้

ภาพที่ 3.43 หน้าแสดงการพยากรณ์ปริมาณน้ำฝนของภาคใต้

11.) หน้าแสดงเกี่ยวกับเรา

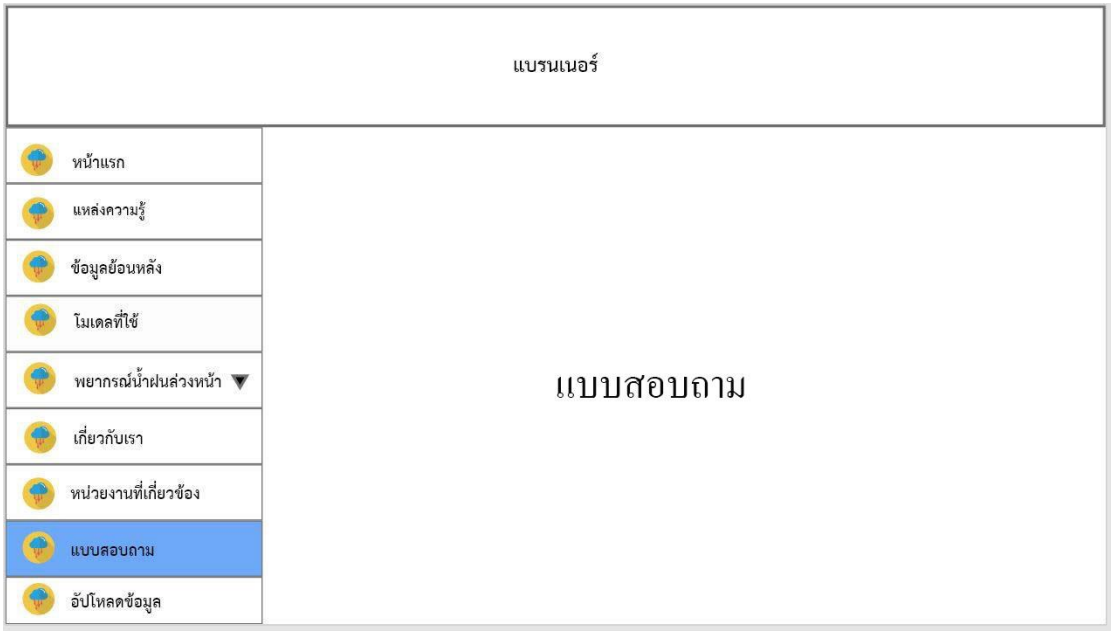


ภาพที่ 3.44 หน้าแสดงเกี่ยวกับเรา

12.) หน้าแสดงหน่วยงานที่เกี่ยวข้อง

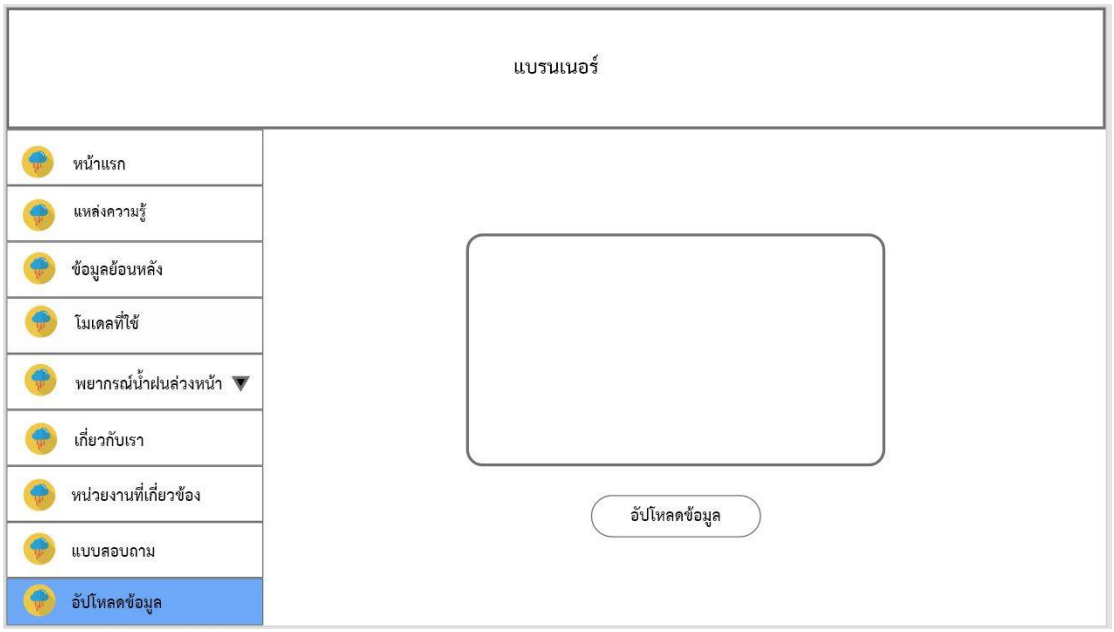
ภาพที่ 3.45 หน้าแสดงหน่วยงานที่เกี่ยวข้อง

11.) หน้าแสดงแบบสอบถาม



ภาพที่ 3.46 หน้าแสดงแบบสอบถาม

12.) หน้าแสดงอัปโหลดข้อมูล



ภาพที่ 3.47 หน้าแสดงอัปโหลดข้อมูล

3.3 บทสรุป

จากขั้นตอนการดำเนินงานผู้วิเคราะห์ได้แสดงกระบวนการวิเคราะห์การพยากรณ์ข้อมูลปริมาณน้ำฝน ด้วยกระบวนการวิเคราะห์ข้อมูล CRISP-DM มาใช้ในการรวบรวมข้อมูล วิเคราะห์ข้อมูลที่เกี่ยวข้องกับงาน เพื่อให้ได้สารสนเทศของงานอย่างเพียงพอในการนำไปใช้ประโยชน์ โดยผู้วิเคราะห์ได้วิเคราะห์ข้อมูลเป็นหลักสำคัญที่จะช่วยให้เข้าใจในงานแต่ละส่วน จากการวิเคราะห์ข้อมูลด้วยกระบวนการวิเคราะห์ข้อมูล CRISP-DM รวมถึงการสร้างโมเดล Time series analysis โดยเลือกใช้โปรแกรม Tableau Public ในการพยากรณ์ค่าปริมาณน้ำฝนล่วงหน้าของปีถัดไปในช่วงเวลาเดียวกัน และนำผลการพยากรณ์ที่ได้ไปประเมินประสิทธิภาพของโมเดล ด้วยวิธีการคำนวณหาค่าเฉลี่ยเปอร์เซ็นต์ความคลาดเคลื่อนสมบูรณ์ (MAPE) ที่ต่ำที่สุด ผู้วิเคราะห์ได้นำข้อมูลสารสนเทศมาทำการแสดงผลแบบ visualization โดยใช้โปรแกรม Tableau Public และออกแบบ Wireframe ของเว็บไซต์ที่จะเผยแพร่บนเว็บไซต์