

บทที่ 3

วิธีการดำเนินงานโครงการ

โครงการเรื่อง การวิเคราะห์ข้อมูลเพื่อสนับสนุนการตัดสินใจในการวางแผนการขาย ร้านกาแฟสวัสดิการแขวงทางหลวงที่ 1 ในบทนี้จะเป็นการวิเคราะห์ข้อมูลด้วยเทคนิคทางดาต้ามาายนิ่ง ซึ่งมีกระบวนการวิเคราะห์ที่สำคัญหลายขั้นตอน เมื่อเสร็จสิ้นจากกระบวนการวิเคราะห์ข้อมูลแล้วจะเป็นการออกแบบเว็บไซต์ และออกแบบรูปแบบการแสดงผลและบทสรุป จากวิธีการดำเนินงาน

3.1 การวิเคราะห์ข้อมูลด้วย CRISP-DM

3.2 การออกแบบเว็บไซต์

3.3 บทสรุป

3.1 การวิเคราะห์ข้อมูลด้วย CRISP-DM

กระบวนการวิเคราะห์ข้อมูลด้วย CRISP-DM หรือ Cross Industry Standard Process for Data Mining พัฒนาขึ้นในปี ค.ศ. 1996 โดยความร่วมมือของ 3 บริษัทคือ Daimler Chrysler, SPSS และ NCR ที่มีการพัฒนาเป็น Workflow มาตรฐานสำหรับการทำเหมืองข้อมูล

ประกอบด้วย 6 ขั้นตอนหลัก ดังนี้

3.1.1 รู้จักและเข้าใจในธุรกิจ (Business Understanding) เป็นขั้นตอนแรกของกระบวนการ ที่มุ่งเน้นไปที่การทำความเข้าใจกระบวนการทางธุรกิจโดยรวม ผู้วิเคราะห์ข้อมูลทำความเข้าใจกับปัญหาให้อยู่ในรูปของการวิเคราะห์ข้อมูลทาง Data Mining โดยการวิเคราะห์ข้อมูลในประเด็นนี้ คือ ต้องการทราบพฤติกรรมการซื้อเครื่องดื่มต่างๆ โดยการหาความสัมพันธ์

3.1.2 จัดเก็บและรวบรวมข้อมูลให้ครบ (Data Understanding) ขั้นตอนการจัดเก็บ และรวบรวมข้อมูล ตลอดจนการพิจารณาตรวจสอบความถูกต้องของข้อมูลที่ได้รับ โดยเลือกว่าจะใช้ ข้อมูลทั้งหมดหรือบางส่วนในการวิเคราะห์ให้สอดคล้องกับวัตถุประสงค์ที่กำหนดไว้ ผู้วิเคราะห์ข้อมูลทำการรวบรวมข้อมูล เพื่อตรวจสอบรายละเอียด และปริมาณของข้อมูล รายงานรายละเอียดยอดขายในแต่ละเดือน ที่ได้จากร้านกาแฟสวัสดิการ แขวงทางหลวงที่1ทางร้านได้มอบข้อมูลให้ผู้วิเคราะห์ เพื่อนำไปวิเคราะห์ให้เกิดประโยชน์อย่างสูงสุดซึ่งข้อมูลตัวอย่างเบื้องต้น ใน

เดือนสิงหาคมของปี 2562 มีจำนวนข้อมูล 566 รายการ ประกอบด้วย 8 แอดทริบิวต์ ประกอบด้วย วันที่ หมวดหมู่อินค้ำ ชื่อสินค้า ประเภทสินค้า ราคาต่อหน่วย จำนวน ยอดสุทธิ การชำระ ยอดการชำระ ซึ่งทั้งหมดนี้ ผู้วิเคราะห์ข้อมูลไม่ได้นำทุกแอดทริบิวต์มาวิเคราะห์ใช้เพียงบางแอดทริบิวต์ที่จำเป็นเท่านั้น เพื่อการวิเคราะห์ข้อมูลที่มีประสิทธิภาพ

#	A	B	C	D	E	F	G	H
1	วันที่	หมวดหมู่อินค้ำ	ชื่อสินค้า	ประเภทสินค้า	ราคาต่อหน่วย	จำนวน	ยอดสุทธิ	การชำระ
2	1/8/2562	COFFEE	AMERICANO	Hot	30	1	30	Cash
3	1/8/2562	COFFEE	AMERICANO	Hot	30	1	30	Cash
4	1/8/2562	COFFEE	AMERICANO	Ice	30	1	30	Cash
5	1/8/2562	COFFEE	AMERICANO	Ice	30	1	30	Cash
6	1/8/2562	COFFEE	AMERICANO	Ice	30	1	30	Cash
7	1/8/2562	COFFEE	AMERICANO	Ice	30	1	30	Cash
8	1/8/2562	COFFEE	ESPRESSO	Hot	35	1	35	Cash
9	1/8/2562	COFFEE	ESPRESSO	Frappe	50	1	50	Cash
10	1/8/2562	COFFEE	ESPRESSO	Frappe	50	1	50	Cash
11	1/8/2562	COFFEE	ESPRESSO	Frappe	50	1	50	Cash
12	1/8/2562	COFFEE	LATTE	Hot	35	1	35	Cash
13	1/8/2562	COFFEE	LATTE	Hot	35	1	35	Cash
14	1/8/2562	COFFEE	MOCHA	Ice	40	1	40	Cash
15	1/8/2562	NON-COFFEE	MILK	Hot	30	1	30	Cash
16	1/8/2562	NON-COFFEE	MILK	Ice	30	1	30	Cash
17	1/8/2562	NON-COFFEE	PINK MILK	Ice	30	1	30	Cash
18	1/8/2562	NON-COFFEE	PINK MILK	Ice	30	1	30	Cash
19	1/8/2562	NON-COFFEE	PINK MILK	Ice	30	1	30	Cash
20	1/8/2562	NON-COFFEE	PINK MILK	Ice	30	1	30	โอนเงิน
21	1/8/2562	NON-COFFEE	PINK MILK	Ice	30	1	30	โอนเงิน
22	1/8/2562	NON-COFFEE	COCO	Ice	30	1	30	Cash
23	1/8/2562	NON-COFFEE	COCO	Ice	30	1	30	Cash
24	1/9/2562	OTHER	M100	Ice	10	1	10	Cash
25	1/10/2562	OTHER	M101	Ice	10	1	10	Cash

ภาพที่ 3.1 ข้อมูลรายงานรายละเอียดยอดขายในแต่ละเดือน

3.1.3 เตรียมข้อมูลให้พร้อมใช้งาน (Data Preparation) ขั้นตอนการแปลงข้อมูลที่ได้รวบรวมมาและเลือกไว้ ให้อยู่ในรูปแบบที่พร้อมสำหรับนำไปวิเคราะห์ในขั้นตอนต่อไปได้ โดยการทำให้เป็นข้อมูลที่ถูกต้อง (Data cleaning) มักใช้เวลาค่อนข้างมาก โดยมีขั้นตอนดังนี้

3.1.3.1 ทำการคัดเลือกข้อมูล (Data Selection) คือการคัดเลือกข้อมูลที่เหมาะสมเพื่อนำมาใช้ในการวิเคราะห์ข้อมูล ผู้วิเคราะห์ข้อมูลทำการคัดเลือกข้อมูล และทำการ Data Cleaning ข้อมูลรายงานรายละเอียดยอดขาย โดยตัดส่วนที่ไม่จำเป็นออก ให้เหลือเฉพาะข้อมูลที่จำเป็นในการวิเคราะห์ในภาพรวม จำนวน 3 แอดทริบิวต์ ได้แก่ วันที่ หมวดหมู่อินค้ำ และ ชื่อสินค้า ซึ่งเป็นข้อมูลที่จำเป็นในการนำไปวิเคราะห์ข้อมูล

#	A	B	C	D	E	F	G	H
1	วันที่	หมวดหมู่อินค้ำ	ชื่อสินค้า	ประเภทสินค้า	ราคาต่อหน่วย	จำนวน	ยอดสุทธิ	การชำระ
2	1/8/2562	COFFEE	AMERICANO	Hot	30	1	30	Cash
3	1/8/2562	COFFEE	AMERICANO	Hot	30	1	30	Cash
4	1/8/2562	COFFEE	AMERICANO	Ice	30	1	30	Cash
5	1/8/2562	COFFEE	AMERICANO	Ice	30	1	30	Cash
6	1/8/2562	COFFEE	AMERICANO	Ice	30	1	30	Cash
7	1/8/2562	COFFEE	AMERICANO	Ice	30	1	30	Cash
8	1/8/2562	COFFEE	ESPRESSO	Hot	35	1	35	Cash
9	1/8/2562	COFFEE	ESPRESSO	Frappe	50	1	50	Cash
10	1/8/2562	COFFEE	ESPRESSO	Frappe	50	1	50	Cash
11	1/8/2562	COFFEE	ESPRESSO	Frappe	50	1	50	Cash
12	1/8/2562	COFFEE	LATTE	Hot	35	1	35	Cash
13	1/8/2562	COFFEE	LATTE	Hot	35	1	35	Cash
14	1/8/2562	COFFEE	MOCHA	Ice	40	1	40	Cash
15	1/8/2562	NON-COFFEE	MILK	Hot	30	1	30	Cash
16	1/8/2562	NON-COFFEE	MILK	Ice	30	1	30	Cash
17	1/8/2562	NON-COFFEE	PINK MILK	Ice	30	1	30	Cash
18	1/8/2562	NON-COFFEE	PINK MILK	Ice	30	1	30	Cash
19	1/8/2562	NON-COFFEE	PINK MILK	Ice	30	1	30	Cash
20	1/8/2562	NON-COFFEE	PINK MILK	Ice	30	1	30	โอนเงิน
21	1/8/2562	NON-COFFEE	PINK MILK	Ice	30	1	30	โอนเงิน
22	1/8/2562	NON-COFFEE	COCO	Ice	30	1	30	Cash
23	1/8/2562	NON-COFFEE	COCO	Ice	30	1	30	Cash
24	1/9/2562	OTHER	M100	Ice	10	1	10	Cash
25	1/10/2562	OTHER	M101	Ice	10	1	10	Cash
26	1/11/2562	OTHER	SPONSER	Ice	10	1	10	Cash
27	1/12/2562	OTHER	SPONSER	Ice	10	1	10	Cash
28	2/8/2562	COFFEE	AMERICANO	Hot	30	1	30	Cash
29	2/8/2562	COFFEE	AMERICANO	Hot	30	1	30	Cash
30	2/8/2562	COFFEE	AMERICANO	Ice	35	1	35	โอนเงิน

ภาพที่ 3.2 ข้อมูลที่ได้ทำการคัดเลือก

3.1.3.2 ทำการกลั่นกรองข้อมูล (Data Cleaning) คือการทำความสะอาดข้อมูล เป็นกระบวนการตรวจสอบและการแก้ไข (หรือลบ) รายการข้อมูลที่ไม่ถูกต้องออกไปจากชุดข้อมูล ตารางหรือฐานข้อมูล ซึ่งเป็นหลักสำคัญของฐานข้อมูล ทางผู้วิเคราะห์ข้อมูลได้ดำเนินการดังนี้

1.1 ทำการแก้ไขข้อมูล ชื่อเมนู หมวดสินค้า ให้เป็นรูปแบบเดียวกัน เพื่อให้สามารถดำเนินการในขั้นตอนถัดไป สรุปได้ว่ามีหมวดสินค้าทั้งหมด 5 หมวด คือ

ตารางที่ 3.1 หมวดหมู่ของสินค้าหลังจากแปลงข้อมูล

หมวดหมู่ของเครื่องดื่ม
1. กาแฟ (COFFEE)
2. ไม่ใช่กาแฟ (NON-COFFEE)
3. ชา (TEA)
4. อิตเลียนโซดา (ITALIAN SODA)
5. เครื่องดื่มอื่นๆ (OTHER)
6. บะหมี่กึ่งสำเร็จรูป (Instant Noodles)

1.2 หลังจากที่ถูกวิเคราะห์ข้อมูลทำการแปลงรูปแบบ หมวดหมู่ของเครื่องดื่มแล้ว ขั้นตอนถัดไปคือทำการแปลงรูปแบบของ ชื่อเมนู ให้เป็นรูปแบบเดียวกัน จากที่ผู้วิเคราะห์ข้อมูลได้ทำการวิเคราะห์ข้อมูลสามารถสรุปรูปแบบจัดเก็บข้อมูลได้ดังนี้

ตารางที่ 3.2 ตารางแสดงข้อมูลหมวดหมู่ COFFEE

COFFEE / กาแฟ	
1. AMERICANO	Hot Ice Frappe
2. ESPRESSO	Hot Ice Frappe

COFFEE / กาแฟ	
3. LATTE	Hot Ice Frappe
4. CAPPUCCINO	Hot Ice Frappe
5. MOCHA	Hot Ice Frappe

ตารางที่ 3.3 ตารางแสดงข้อมูลหมวดหมู่ NON-COFFEE

NON-COFFEE / ไม่ใช่กาแฟ	
1.MILK	Hot Ice Frappe
2. CARAMEL MILK	Hot Ice Frappe
3. PINK MILK	Hot Ice Frappe
4. COCO	Hot Ice Frappe

ตารางที่ 3.4 ตารางแสดงข้อมูลหมวดหมู่ TEA

TEA / ชา	
1. THAI MILK TEA	Hot Ice Frappe
2. THAI BLACK TEA	Hot Ice Frappe
3. NESTEA	Hot Ice Frappe
4. GREEN TEA	Hot Ice Frappe
5. LEMON TEA	Hot Ice Frappe
6. TEA	Hot Ice

ตารางที่ 3.5 ตารางแสดงข้อมูลหมวดหมู่ ITALIAN SODA

ITALIAN SODA / อิตาเลียนโซดา	
1. LYCHEE SODA	Ice
2. GREEN APPLE SODA	Ice
3. CANTALOOPE SODA	Ice
4. KIWI SODA	Ice
5. BULE HAWAII SODA	Ice
6. SATRAWBERRY SODA	Ice
7. BULEBERRY SODA	Ice
ITALIAN SODA / อิตาเลียนโซดา	
8. RASPBERRY SODA	Ice
9. PUNCH SODA	Ice
10. PINEAPPLE SODA	Ice

ตารางที่ 3.6 ตารางแสดงข้อมูลหมวดหมู่ OTHER

OTHER / อื่น ๆ	
1. COKE	Ice
2. COKE Z	Ice
3. COKE L	Ice
4. M100	Ice

OTHER / อื่น ๆ	
5.SPONSER	Ice

ตารางที่ 3.7 ตารางแสดงข้อมูลหมวดหมู่ Instant Noodles

Instant Noodles / บะหมี่กึ่งสำเร็จรูป	
1. หมูสับ	-
2. ต้มยำกุ้ง	-
3. ต้มยำน้ำข้น	-
4. ต้มโคล้ง	-

หลังจากการจัดการรูปแบบให้อยู่ในรูปแบบเดียวกันแล้วโดยการดำเนินการที่ละหมวดสินค้า สรุปได้ว่า

หมวดหมู่ COFFEE	มีอยู่	5 รายการ
หมวดหมู่ NON-COFFEE	มีอยู่	4 รายการ
หมวดหมู่ TEA	มีอยู่	6 รายการ
หมวดหมู่ ITALIAN SODA	มีอยู่	10 รายการ
หมวดหมู่ OTHER	มีอยู่	5 รายการ
หมวดหมู่ INSTANT NOODLE	มีอยู่	4 รายการ

1.3 แปลงข้อมูลให้อยู่ในรูปแบบข้อมูลเพื่อพร้อมนำไปใช้ในการทำโมเดลโดยใช้โปรแกรม

RapidMiner

1	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
2	Date	AMERICAN	ESPRESSO	LATTE	CAPPUCCI	MOCHA	MILK	CARAMEL	PINK MILK	COCO	THAI MILK	THAI BLAC	NESTEA	GREEN TE	LEMON TE	TEA	LYCHEE SC	GREEN API	CANTALOCK	KIWI SOD	BULE HAW	SATRAWBI
3	1	6	4	2	0	1	2	0	5	2	0	0	0	0	0	0	0	0	0	0	0	0
4	2	5	2	1	2	1	1	0	0	0	0	0	3	4	0	0	1	2	0	0	0	0
5	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	5	5	3	2	0	3	6	0	3	0	2	0	0	0	2	0	0	0	0	0	0	0
8	6	6	4	1	0	0	2	1	0	4	1	0	0	0	2	0	0	0	0	0	0	0
9	7	6	2	4	4	3	0	1	0	0	0	0	0	0	1	0	0	0	1	0	1	0
10	8	4	4	3	1	4	0	0	0	0	0	0	1	3	0	0	0	0	0	0	0	0
11	9	1	2	3	2	1	2	0	2	4	2	2	3	2	1	0	0	0	0	0	0	0
12	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	12	4	2	4	4	3	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	13	1	3	0	2	1	3	0	3	1	2	4	1	0	0	0	1	0	1	1	1	0
16	14	3	4	2	4	0	0	1	0	0	0	0	2	2	1	0	1	1	0	0	0	0
17	15	3	0	2	0	0	5	1	2	4	3	1	1	0	0	0	0	0	0	0	0	0
18	16	2	5	0	1	4	0	1	0	0	1	1	2	2	0	0	0	0	0	0	0	0
19	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
21	19	2	2	2	1	0	2	0	2	3	0	0	0	0	0	1	0	0	0	0	0	0
22	20	4	6	1	0	2	0	1	0	4	0	1	2	0	0	0	0	0	0	0	0	0
23	21	5	4	1	3	1	7	1	2	0	1	2	3	0	0	0	0	0	0	2	0	0
24	22	1	5	2	1	4	2	0	1	0	1	3	0	4	0	1	0	0	1	0	1	0
25	23	3	1	0	1	0	0	0	0	0	3	1	0	2	0	0	0	0	0	0	0	0
26	24	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
27	25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28	26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
29	27	4	8	1	0	2	0	1	0	4	0	0	2	0	3	0	0	0	0	0	0	0
30	28	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
31	29	5	0	5	3	2	5	1	3	0	0	0	0	0	0	1	0	1	1	1	0	0

ภาพที่ 3.3 ภาพตัวอย่างของข้อมูลที่น่าไปทำโมเดล

3.1.4 สร้างแบบจำลอง (Modeling) ขั้นตอนการสร้างตัวแบบทางคณิตศาสตร์ และสถิติ เพื่อการวิเคราะห์ข้อมูล โดยสามารถใช้เทคนิควิธีการต่าง ๆ อาทิ การจำแนก (Classification) การแบ่งกลุ่ม (Clustering) การสร้างความสัมพันธ์ (Association rule) และการเลือกแอตทริบิวต์ หรือ คอลัมน์ที่จะนำไปทำการวิเคราะห์ โดยการเลือกแอตทริบิวต์หรือคอลัมน์ (Future Selection) ในการหาความสัมพันธ์การซื้อเมนูรายการอาหารที่นิยมซื้อพร้อมกัน ได้ใช้เทคนิคทำเหมืองข้อมูลการสร้างความสัมพันธ์ (Association rule) คือกระบวนการกระทำกับข้อมูลเพื่อค้นหารูปแบบ กฎ และความสัมพันธ์ของข้อมูลที่อยู่ในชุดข้อมูลนั้นโดย อาศัยหลักสถิติ การรู้จากการเรียนรู้ของเครื่อง และหลักคณิตศาสตร์ เป็นต้น ผลลัพธ์ที่ได้จะทำให้เกิดสารสนเทศที่เป็นประโยชน์ การทำเหมืองข้อมูลสามารถใช้เทคนิคสร้างแบบแผนได้หลายแบบแผน ในที่นี้จะกล่าวเฉพาะในรูปแบบกฎความสัมพันธ์ (Association Rule) ซึ่งเป็นการแสดงความสัมพันธ์ของ เหตุการณ์หรือวัตถุที่เกิดขึ้น จะอยู่ในรูปแบบ $A \Rightarrow B$ การหาความสัมพันธ์มีเทคนิคอยู่หลายวิธีด้วยกัน วิธีที่ ยอมรับและนิยมอย่างหนึ่งคือวิธีอัลกอริทึมออฟริโอริ เป็นอัลกอริทึมใน การสืบค้นกฎความสัมพันธ์ที่เกิดขึ้นในฐานข้อมูล ซึ่งมีรายละเอียด ดังนี้

3.1.4.1 นิยามศัพท์ที่เกี่ยวข้องกับกฎความสัมพันธ์

- ไอเทมเซต (Itemset) คือ ความสัมพันธ์ของข้อมูลที่ได้พื้นฐานข้อมูล โดยไอเทมเซตประกอบด้วย ไอเทมที่ k-itemsets เช่น 4-itemsets ได้แก่ { A, B, C, D }

- ฟรีคว้นไอเทมเซต (Frequent Itemset) คือ ไอเทมเซตที่ผ่านค่าสนับสนุนขั้นต่ำ (Minimum Support) การสร้างกฎความสัมพันธ์จากฟรีคว้นไอเทมเซตที่หาได้ จะพิจารณาจากฟรีคว้นไอเทมเซตที่มีความยาวมากกว่า 2 รายการขึ้นไป เช่น { A, B } จะสร้างเป็นกฎความสัมพันธ์ได้เป็น $A \Rightarrow B$ โดยกฎความสัมพันธ์ทั่วไปเขียนได้ดังนี้ LHS \Rightarrow RHS LHS (Left Hand Side) แสดงรูปแบบของไอเทมเซตด้านซ้ายของกฎความสัมพันธ์และ RHS (Right Hand Side) แสดงรูปแบบของไอเทมเซตด้านขวาของกฎความสัมพันธ์จากตัวอย่างกฎความสัมพันธ์ $A \Rightarrow B$ LHS คือ A และ RHS คือ B

- ค่าสนับสนุน (Support) คือ ค่าความน่าจะเป็นของจำนวนไอเทมเซตที่พบในฐานข้อมูลต่อจำนวนรายการทั้งหมด เช่น คำนวณหาค่าสนับสนุนของไอเทมเซต LHS หาได้จาก

$$\text{Support (LSH)} = \frac{\text{Number of Transactions that contain LHS}}{\text{Number of all Transactions}} \text{ Support (LHS)}$$

คือ ค่าสนับสนุนของไอเทมเซต LHS Number of Transactions that contain LHS คือ จำนวนรายการข้อมูลของไอเทมเซต LHS Number of all Transactions คือ จำนวนรายการข้อมูลทั้งหมด 68

3.1.4.2 การประเมินผล (Evaluation) การพิจารณาว่ากฎความสัมพันธ์ที่สร้างได้ดีหรือไม่จำเป็นจะต้องมีตัววัดประสิทธิภาพของกฎ ที่นิยม คือ

- ค่าความเชื่อมั่น (Confidence) คือ การแสดงค่าความเชื่อมั่นของกฎความสัมพันธ์ เมื่อรูปแบบ LHS ที่อยู่ทางด้านซ้ายของกฎเกิดขึ้นแล้วมีโอกาสเกิดรูปแบบ RHS ที่อยู่ทางด้านขวามากน้อยเท่าใด ซึ่งจะมีค่าอยู่ระหว่าง 0-1 ถ้าใกล้เคียง 1 หมายถึงมีความเชื่อมั่นในการหาความสัมพันธ์มาก อาจคำนวณค่าที่จะ เกิดขึ้นในรูปแบบเปอร์เซ็นต์ก็ได้ วิธีการคำนวณค่าความเชื่อมั่นได้จาก

$$\text{Confidence (LHS} \Rightarrow \text{RHS)} = \frac{\text{Support(LHS, RHS)}}{\text{Support(LHS)}}$$

Support(LHS) คือค่าสนับสนุนที่รูปแบบ LHS และ RHS ของกฎความสัมพันธ์เกิดขึ้นพร้อม ๆ กัน ส่วน Support (LHS) คือค่าสนับสนุนรูปแบบที่อยู่ด้านซ้ายของกฎความสัมพันธ์

- ค่าสหสัมพันธ์หรือเรียกว่าค่าลิฟต์ (Lift) คือค่าที่บ่งบอกว่าการเกิดรูปแบบ LHS และ RHS มีความสัมพันธ์กันมากหรือไม่ โดยถ้าค่าลิฟต์ มีค่าเท่ากับ 1 แสดงว่ารูปแบบ LHS และ RHS ไม่ขึ้นต่อกัน(Independent) แต่ถ้ามีค่ามากกว่า 1 มาก ๆ แสดงว่ากฎทั้งสองมีความสัมพันธ์กันมากด้วยเช่นกัน ค่าลิฟต์คำนวณได้จาก

$$\text{Lift (LSH} \Rightarrow \text{RSH)} = \frac{\text{Support(LHS, RHS)}}{\text{Support(LHS)} \times \text{Support(RHS)}}$$

Support (LHS, RHS) คือค่าสนับสนุนที่รูปแบบ LHS และ RHS ของกฎความสัมพันธ์เกิดขึ้นพร้อม ๆ กัน Support (LHS) คือค่าสนับสนุนรูปแบบที่อยู่ด้านซ้ายของกฎความสัมพันธ์ และ Support (RHS) คือค่าสนับสนุนรูปแบบที่อยู่ด้านขวาของกฎความสัมพันธ์

3.1.4.2 การหากฎความสัมพันธ์ด้วยวิธีอัลกอริทึมอปรไอริขั้นตอนการหากฎความสัมพันธ์ด้วยวิธีอัลกอริทึมอปรไอริ ผู้วิเคราะห์ขอยกตัวอย่างจากข้อมูลของรายงานยอดขายชายของทางร้านกาแฟสวัสดิการแขวงทางหลวงที่ 1 ของเดือนสิงหาคม 2562 จำนวน 566 มาอธิบายการหากฎความสัมพันธ์ ดังนี้

- ตัวอย่างการหาค่าความสัมพันธ์(Confidence) ค่าสนับสนุน(Support) และมีค่าสหสัมพันธ์ (Lift) ค่าความสัมพันธ์ขั้นต่ำที่ 0.10 กำหนดค่าสนับสนุนขั้นต่ำที่ 0.5 ได้ดังนี้

No.	Premises	Conclusion	Support ↓	Confidence	LaPlace	Gain	p-s	Lift	Conviction
731	ESPRESSO	AMERICANO	0.581	1	1	-0.581	0.206	1.550	∞
732	LATTE	AMERICANO	0.548	1	1	-0.548	0.195	1.550	∞
610	ESPRESSO	AMERICANO, LATTE	0.484	0.833	0.939	-0.677	0.165	1.520	2.710
670	LATTE	AMERICANO, ESPRESSO	0.484	0.882	0.958	-0.613	0.165	1.520	3.565
733	MOCHA	AMERICANO	0.484	1	1	-0.484	0.172	1.550	∞
746	ESPRESSO, LATTE	AMERICANO	0.484	1	1	-0.484	0.172	1.550	∞
519	ESPRESSO	AMERICANO, MOCHA	0.452	0.778	0.918	-0.710	0.171	1.607	2.323
729	MOCHA	AMERICANO, ESPRESSO	0.452	0.933	0.978	-0.516	0.171	1.607	6.290
734	CAPPUCCINO	AMERICANO	0.452	1	1	-0.452	0.160	1.550	∞
747	ESPRESSO, MOCHA	AMERICANO	0.452	1	1	-0.452	0.160	1.550	∞
475	ESPRESSO	AMERICANO, CAPPUCCINO	0.419	0.722	0.898	-0.742	0.157	1.599	1.974
506	LATTE	AMERICANO, MOCHA	0.419	0.765	0.917	-0.677	0.154	1.580	2.194
619	MOCHA	AMERICANO, LATTE	0.419	0.867	0.957	-0.548	0.154	1.580	3.387
726	CAPPUCCINO	AMERICANO, ESPRESSO	0.419	0.929	0.978	-0.484	0.157	1.599	5.871
748	ESPRESSO, CAPPUCCINO	AMERICANO	0.419	1	1	-0.419	0.149	1.550	∞
762	LATTE, MOCHA	AMERICANO	0.419	1	1	-0.419	0.149	1.550	∞
404	ESPRESSO	AMERICANO, LATTE, MOCHA	0.387	0.667	0.878	-0.774	0.144	1.590	1.742
463	LATTE	AMERICANO, ESPRESSO, MOCHA	0.387	0.706	0.896	-0.710	0.139	1.563	1.865
584	ESPRESSO, LATTE	AMERICANO, MOCHA	0.387	0.800	0.935	-0.581	0.153	1.653	2.581

ภาพที่ 3.4 ตัวอย่างการหาค่าความเชื่อมั่น(Confidence) ค่าสนับสนุน(Support) และมีค่าสหสัมพันธ์ (Lift)

AssociationRules

```

Association Rules
[AMERICANO] --> [GREEN TEA] (confidence: 0.350)
[AMERICANO] --> [ESPRESSO, CARAMEL MILK] (confidence: 0.350)
[AMERICANO] --> [ESPRESSO, COCO] (confidence: 0.350)
[AMERICANO] --> [ESPRESSO, THAI BLACK TEA] (confidence: 0.350)
[AMERICANO] --> [ESPRESSO, GREEN TEA] (confidence: 0.350)
[AMERICANO] --> [LATTE, THAI MILK TEA] (confidence: 0.350)
[AMERICANO] --> [LATTE, NESTEA] (confidence: 0.350)
[AMERICANO] --> [LATTE, COCO] (confidence: 0.350)
[AMERICANO] --> [MOCHA, THAI MILK TEA] (confidence: 0.350)
[AMERICANO] --> [MOCHA, LEMON TEA] (confidence: 0.350)
[AMERICANO] --> [MOCHA, NESTEA] (confidence: 0.350)
[AMERICANO] --> [CAPPUCCINO, THAI MILK TEA] (confidence: 0.350)
[AMERICANO] --> [CAPPUCCINO, PINK MILK] (confidence: 0.350)
[AMERICANO] --> [CAPPUCCINO, GREEN TEA] (confidence: 0.350)
[AMERICANO] --> [MILK, THAI MILK TEA] (confidence: 0.350)
[AMERICANO] --> [THAI MILK TEA, PINK MILK] (confidence: 0.350)
[AMERICANO] --> [THAI MILK TEA, THAI BLACK TEA] (confidence: 0.350)
[AMERICANO] --> [ESPRESSO, LATTE, PINK MILK] (confidence: 0.350)
[AMERICANO] --> [ESPRESSO, MOCHA, THAI MILK TEA] (confidence: 0.350)
[AMERICANO] --> [ESPRESSO, MOCHA, PINK MILK] (confidence: 0.350)
[AMERICANO] --> [ESPRESSO, MOCHA, LEMON TEA] (confidence: 0.350)
[AMERICANO] --> [ESPRESSO, MOCHA, NESTEA] (confidence: 0.350)
[AMERICANO] --> [ESPRESSO, CAPPUCCINO, MILK] (confidence: 0.350)
[AMERICANO] --> [ESPRESSO, CAPPUCCINO, THAI MILK TEA] (confidence: 0.350)
[AMERICANO] --> [ESPRESSO, CAPPUCCINO, GREEN TEA] (confidence: 0.350)
[AMERICANO] --> [ESPRESSO, MILK, PINK MILK] (confidence: 0.350)
[AMERICANO] --> [LATTE, MOCHA, PINK MILK] (confidence: 0.350)
[AMERICANO] --> [LATTE, CAPPUCCINO, MILK] (confidence: 0.350)
[AMERICANO] --> [MOCHA, CAPPUCCINO, MILK] (confidence: 0.350)

```

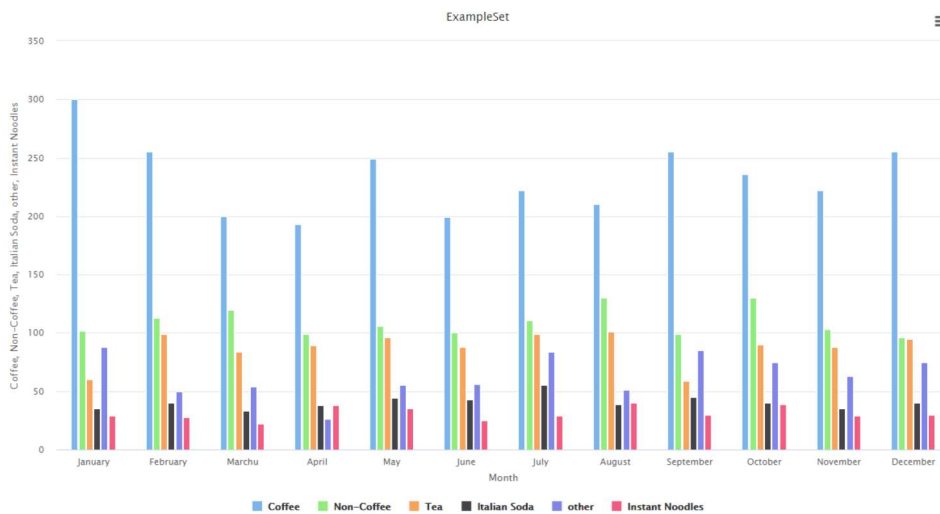
ภาพที่ 3.5 แสดงกฎความสัมพันธ์

3.1.5 เผยแพร่ผลวิเคราะห์ (Deployment) ขั้นตอนการนำผลลัพธ์ที่ได้ไปใช้งาน เป็นการทั่วไป อาจจัดทำเป็นรูปแบบของรายงาน (Report) หรือแผนภาพ (Dashboard)

ที่สามารถนำไปใช้ประโยชน์ในการวางแผน กำหนดกลยุทธ์ และดำเนินการต่าง ๆ ในทางธุรกิจ ผู้วิเคราะห์ข้อมูลนำผลข้อมูลที่ทำกรวิเคราะห์แล้วแสดงผลข้อมูลบน Web Browser โดยใช้ชุดคำสั่ง HTML CSS และ PHP ร่วมกับการนำเสนอข้อมูลแบบ Visualization ด้วยการแสดงผลข้อมูลในรูปแบบของภาพโดยใช้โปรแกรม Tableau Public โดยมีการออกแบบรายงานแสดงผลข้อมูลดังนี้

3.1.5.1 แสดงจำนวนรายการจำนวนปริมาณของยอดขายที่เกิดขึ้นในแต่ละเดือนตั้งแต่เดือนมกราคม 2563 – ธันวาคม 2563 โดยแสดงจำนวนที่เกิดขึ้นตามลำดับโดยทำการจำแนกเป็นหมวดหมู่ดังนี้

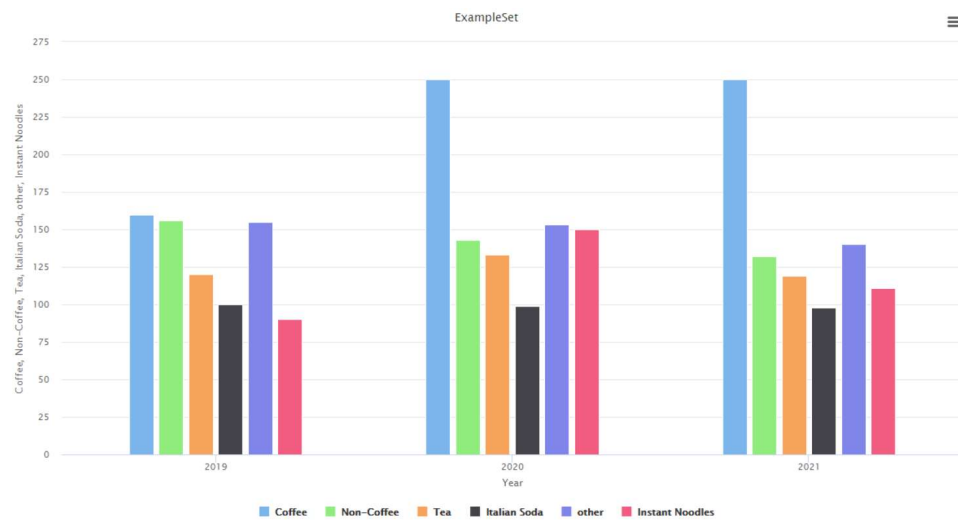
หมวดหมู่ COFFEE	เป็นสีน้ำเงิน
หมวดหมู่ NON-COFFEE	เป็นสีเขียว
หมวดหมู่ TEA	เป็นสีส้ม
หมวดหมู่ ITALIAN SODA	เป็นสีดำ
หมวดหมู่ OTHER	เป็นสีม่วง
หมวดหมู่ INSTANT NOODLE	เป็นสีชมพู



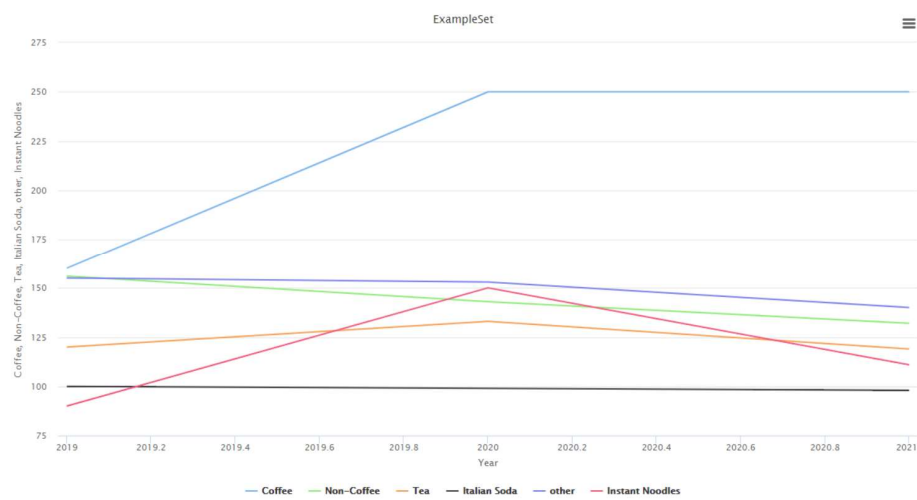
ภาพที่ 3.6 แสดงปริมาณของสินค้าที่ขายได้ในแต่ละเดือน

3.1.5.2 แสดงจำนวนรายการจำนวนปริมาณของยอดขายที่เกิดขึ้นในแต่ละเดือนตั้งแต่ปี 2562-2564 โดยแสดงจำนวนที่เกิดขึ้นตามลำดับโดยทำการจำแนกเป็นหมวดหมู่ดังนี้

หมวดหมู่ COFFEE	เป็นสีน้ำเงิน
หมวดหมู่ NON-COFFEE	เป็นสีเขียว
หมวดหมู่ TEA	เป็นสีส้ม
หมวดหมู่ ITALIAN SODA	เป็นสีดำ
หมวดหมู่ OTHER	เป็นสีม่วง
หมวดหมู่ INSTANT NOODLE	เป็นสีชมพู

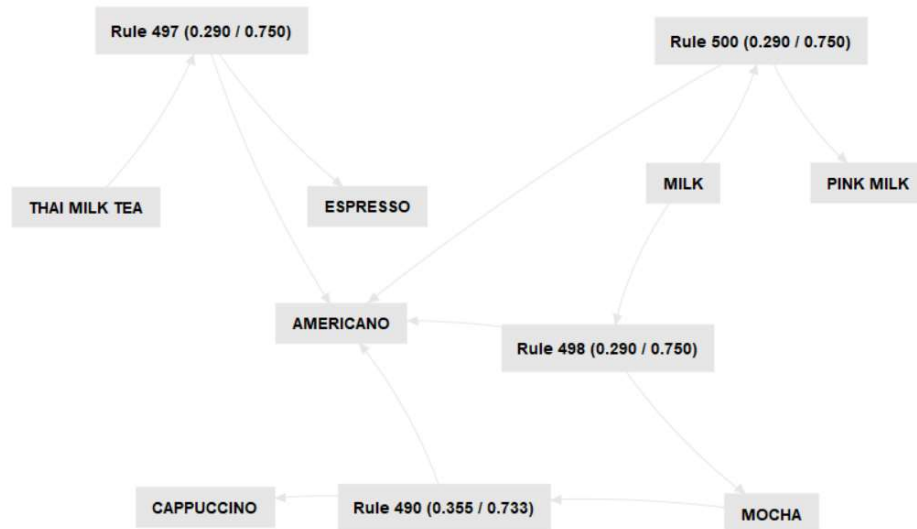


ภาพที่ 3.7 แสดงปริมาณของสินค้าที่ขายได้ในแต่ละปี



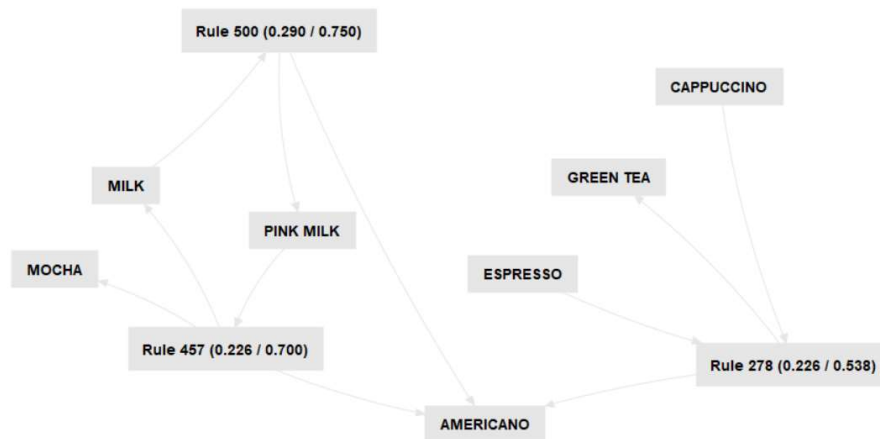
ภาพที่ 3.8 แสดงปริมาณของสินค้าที่ขายได้ปี 2019-2021

3.1.5.3 แสดงจำนวนกฎความสัมพันธ์ระดับค่าความเชื่อมั่น (Confidence) ที่ได้จากการวิเคราะห์



ภาพที่ 3.9 แสดงจำนวนค่าความเชื่อมั่น (Confidence)

3.1.5.4 แสดงจำนวนกฎความสัมพันธ์ระดับค่าสหสัมพันธ์ (Lift) ที่ได้จากการวิเคราะห์



ภาพที่ 3.10 แสดงจำนวนค่าสหสัมพันธ์ (Lift)

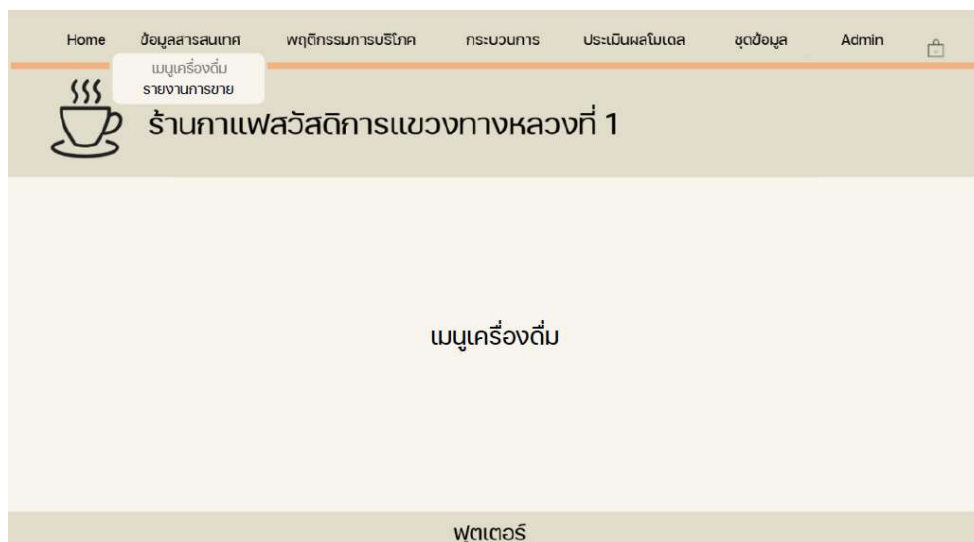
3.2 การออกแบบเว็บไซต์

1. หน้าโฮมเพจของเว็บไซต์แสดงรายละเอียดของร้าน



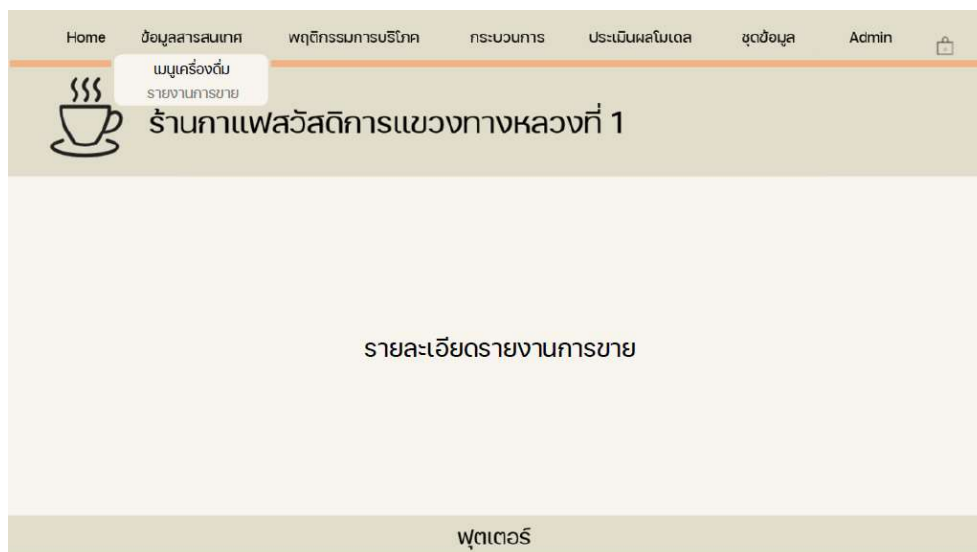
ภาพที่ 3.11 แสดงหน้าโฮมเพจของเว็บไซต์แสดงรายละเอียดของร้าน

2. หน้าเว็บเพจที่ 2 แสดงรายการเครื่องดื่มในแถบเมนูข้อมูลสารสนเทศ



ภาพที่ 3.12 แสดงรายการเครื่องดื่ม

3. หน้าเว็บเพจที่ 3 แสดงรายงานการขายในแถบเมนูข้อมูลสารสนเทศ



ภาพที่ 3.13 แสดงรายงานการขาย

4. หน้าเว็บเพจที่ 4 แสดงพฤติกรรมบริการ



ภาพที่ 3.14 แสดงพฤติกรรมบริการในรูปแบบกราฟฟิก

5. หน้าเว็บเพจที่ 5 แสดงกระบวนการ



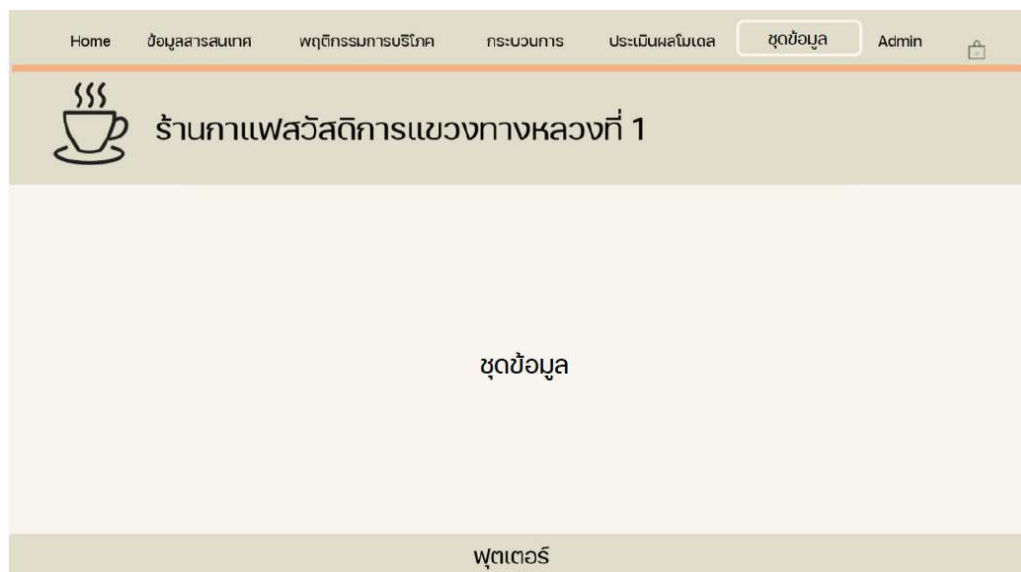
ภาพที่ 3.15 แสดงกระบวนการ

6. หน้าเว็บเพจที่ 6 แสดงผลประเมินโมเดล



ภาพที่ 3.16 แสดงผลประเมินโมเดล

7. หน้าเว็บเพจที่ 7 แสดงชุดข้อมูล



ภาพที่ 3.17 แสดงชุดข้อมูล

3.3 บทสรุป

จากวิธีการดำเนินงานโครงการในข้างต้นทั้งหมดนี้ ผู้วิเคราะห์ข้อมูลได้หาความสัมพันธ์การนิยมซื้ออาหารพร้อมกัน ด้วยเทคนิคการหาความสัมพันธ์ด้วยอัลกอริทึมเอพริออริ ด้วยขั้นตอน กระบวนการวิเคราะห์ข้อมูลด้วย CRISP-DM อย่างละเอียดรวมถึงการสร้างโมเดลจากโปรแกรมที่ใช้ทำเหมืองข้อมูล คือ โปรแกรม Rapid Miner Studio ในการสร้างโมเดล Association Rules ในรูปแบบอัลกอริทึมเอพริออริ ซึ่งผลลัพธ์ที่ได้จะนำข้อมูลสารสนเทศและออกแบบ Wireframe ของเว็บไซต์ที่จะเผยแพร่บน Web Browser ด้วย โปรแกรม Adobe xd