



รายงานการปฏิบัติงานสหกิจศึกษา

การวิเคราะห์พฤติกรรมของผู้ใช้งานจากร้านค้าออนไลน์ด้วยเทคนิค

การวิเคราะห์ตะกร้าสินค้า

Analysis of User Behavior from Online Stores using
Market Basket Analysis Technique

บริษัท ดาต้า เฟิร์ส จำกัด

Data First Company Limited

โดย

นายนารารธร

ปิ่นสุข

6504800012

นายสุภวินท์

ปิติโชตินันท์

6504800028

รายงานนี้เป็นส่วนหนึ่งของรายวิชา 128-491 สหกิจศึกษาสำหรับนักวิทยาการคอมพิวเตอร์ 1

หลักสูตรวิทยาศาสตรบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์

คณะวิทยาศาสตร์ มหาวิทยาลัยสยาม

ภาคการศึกษาที่ 3 ปีการศึกษา 2567



รายงานการปฏิบัติงานสหกิจศึกษา

การวิเคราะห์พฤติกรรมของผู้ใช้งานจากร้านค้าออนไลน์ด้วยเทคนิค

การวิเคราะห์ตะกร้าสินค้า

Analysis of User Behavior from Online Stores using
Market Basket Analysis Technique

บริษัท ดาต้า เฟิร์ส จำกัด

Data First Company Limited

โดย

นายนารธร

ปิ่นสุข

6504800012

นายสุภวินท์

ปิติโชตินันท์

6504800028

รายงานนี้เป็นส่วนหนึ่งของรายวิชา 128-491 สหกิจศึกษาสำหรับนักวิทยาการคอมพิวเตอร์ 1

หลักสูตรวิทยาศาสตรบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์

คณะวิทยาศาสตร์ มหาวิทยาลัยสยาม

ภาคการศึกษาที่ 3 ปีการศึกษา 2567

หัวข้อโครงการ

การวิเคราะห์พฤติกรรมของผู้ใช้งานจากร้านค้าออนไลน์ด้วยเทคนิค การวิเคราะห์ตะกร้าสินค้า

Analysis of User Behavior from Online Stores using Market Basket Analysis Technique

บริษัท ดาต้า เฟิร์ส จำกัด

รายชื่อคณะผู้จัดทำ

นาย นารัทร ปินสุข

6504800012

นาย สุภวินท์ ปิติโชติพันธ์

6504800028

หลักสูตร

วิทยาศาสตร์บัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์

อาจารย์นิเทศ

อาจารย์จรรยา แหม่มเจริญ

อนุมัติให้โครงการนี้เป็นส่วนหนึ่งของการปฏิบัติสหกิจศึกษาและการศึกษาเชิงบูรณาการกับการทำงาน หลักสูตรวิทยาศาสตร์บัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยสยาม ภาคการศึกษาที่ 3 ปีการศึกษา 2567

คณะกรรมการสอบโครงการ

.....ลพพร แหม่มเจริญ.....อาจารย์นิเทศ

(อาจารย์จรรยา แหม่มเจริญ)

.....นายอนันต์ ตีระบุรณะพงษ์.....ผู้นิเทศ

(นายอนันต์ ตีระบุรณะพงษ์)

.....ดร.อรุณ รอดชีวิต.....กรรมการกลาง

(อาจารย์ธนาภรณ์ รอดชีวิต)

.....ดร.มารุจ ลิ้มปะวัฒน์.....ผู้ช่วยอธิการบดีและผู้อำนวยการสำนักสหกิจศึกษา

(ผู้ช่วยศาสตราจารย์ ดร.มารุจ ลิ้มปะวัฒน์)

จดหมายนำส่งรายงาน

วันที่ 29 เดือน สิงหาคม พ.ศ. 2568

เรื่อง ขอส่งรายงานการปฏิบัติสหกิจศึกษา
เรียน อาจารย์นิเทศ หลักสูตรวิทยาศาสตรบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์
อาจารย์ จรรยา แหยมเจริญ

ตามที่ นายนารธร ปีนสุข และ นายสุภวินท์ ปิติโชตินันท์ นักศึกษาหลักสูตรวิทยาศาสตรบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยสยาม ได้ไปปฏิบัติสหกิจศึกษา และการศึกษาเชิงบูรณาการกับการทำงาน ระหว่างวันที่ 19 พฤษภาคม พ.ศ. 2568 ถึงวันที่ 29 สิงหาคม พ.ศ. 2568 ในตำแหน่ง นักวิทยาศาสตร์ข้อมูล (Data Scientist) ณ บริษัท ดาต้า เฟิร์ส จำกัด และได้รับมอบหมายจากผู้เินเทศให้ศึกษาและทำรายงานเรื่องการวิเคราะห์พฤติกรรมของผู้ใช้งานจากร้านค้าออนไลน์ด้วยการเทคนิคการวิเคราะห์ตะกร้าสินค้า (Market Basket Analysis)

บัดนี้การปฏิบัติสหกิจศึกษาและการศึกษาเชิงบูรณาการกับการทำงานได้สิ้นสุดแล้ว คณะผู้จัดทำ จึงขอส่งรายงานดังกล่าวมาพร้อมกันนี้จำนวน 1 เล่ม เพื่อขอรับคำปรึกษาต่อไป

จึงเรียนมาเพื่อโปรดพิจารณา

ขอแสดงความนับถือ

ลงชื่อนายนารธร ปีนสุข.....
(นายนารธร ปีนสุข)

ลงชื่อนายสุภวินท์ ปิติโชตินันท์.....
(นายสุภวินท์ ปิติโชตินันท์)

คณะผู้จัดทำ

กิตติกรรมประกาศ (Acknowledgement)

การที่คณะผู้จัดทำได้มาปฏิบัติสหกิจศึกษาในตำแหน่งนักวิทยาศาสตร์ข้อมูล (Data Scientist) ณ บริษัท ดาต้า เฟิร์ส จำกัด ตั้งแต่วันที่ 19 พฤษภาคม พ.ศ. 2568 ถึงวันที่ 29 สิงหาคม พ.ศ. 2568 ได้สำเร็จลุล่วงตามวัตถุประสงค์ด้วยดี ส่งผลให้คณะผู้จัดทำ ได้รับความรู้ ประสบการณ์ การทำงานต่างๆ และความเข้าใจในชีวิตการทำงานจริง ที่เป็นประโยชน์ต่อการเรียนและสามารถนำ ความรู้ประสบการณ์ที่ได้ไปใช้ในการประกอบอาชีพในอนาคต ด้วยความอนุเคราะห์อย่างยิ่งจาก บริษัท ดาต้า เฟิร์ส ที่ให้โอกาสคณะผู้จัดทำเข้ามาปฏิบัติสหกิจศึกษา กรุณาเสียสละเวลาอบรม สอน งาน และช่วยเหลือด้านต่างๆ ตลอดระยะเวลาในการปฏิบัติสหกิจศึกษาในครั้งนี้ จึงขอขอบพระคุณ อย่างสูงมา ณ ที่นี้ จากการสนับสนุนหลายฝ่าย ดังนี้

1. คุณอนันท์ ตีระบุรณะพงษ์ (Vice Executive Data & Innovation Director)
2. อาจารย์จรรยา แหยมเจริญ (อาจารย์นิเทศ)
3. อาจารย์ธนาภรณ์ รอดชีวิต (อาจารย์ผู้ร่วมนิเทศ)

และบุคคลที่ไม่ได้กล่าวนามทุกท่าน ที่ได้ให้คำแนะนำในการจัดทำรายงานสหกิจศึกษาฉบับนี้จนเสร็จ สมบูรณ์

คณะผู้จัดทำหวังเป็นอย่างยิ่งว่ารายงานฉบับนี้จะเป็นประโยชน์ต่อ บริษัท ดาต้าเฟิร์ส จำกัด และผู้สนใจปฏิบัติสหกิจศึกษาของบริษัทเพื่อเป็นแนวทางเบื้องต้นในการทำความเข้าใจและพัฒนา โครงการต่อไป รวมทั้งในการค้นคว้าของผู้สนใจทั่วไปด้วย หากรายงานฉบับนี้มีข้อผิดพลาดประการใด ผู้จัดทำ ก็ขออภัยมา ณ ที่นี้

นาราร ปิ่นสุข
สุภาวิณ์ ปิติโชตินันท์

คณะผู้จัดทำ

29 / สิงหาคม / 2568

ชื่อโครงการ : การวิเคราะห์พฤติกรรมของผู้ใช้งานจากร้านค้าออนไลน์ด้วยเทคนิค การวิเคราะห์ตะกร้าสินค้า

หน่วยกิต : 5 หน่วยกิต

คณะผู้จัดทำ : นาย นารารธ ปิ่นสุข 6504800012
นาย สุภวินท์ ปิติโชติพันธ์ 6504800028

อาจารย์ที่ปรึกษา : อาจารย์จรรยา แหยมเจริญ

ระดับการศึกษา : ปริญญาตรี

หลักสูตร : วิทยาศาสตร์บัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์

คณะ : วิทยาศาสตร์

ภาคการศึกษา/ปีการศึกษา : 3/2567

บทคัดย่อ

โครงการนี้มุ่งเน้นการใช้เทคนิคการวิเคราะห์ตะกร้าสินค้า (Market Basket Analysis, MBA) เพื่อศึกษาพฤติกรรมผู้ใช้งานร้านค้าออนไลน์ โดยเฉพาะในมิติเกี่ยวกับแบรนด์สินค้าและช่องทางการขาย เทคนิคนี้ช่วยค้นหารูปแบบความสัมพันธ์ที่เกิดขึ้นร่วมกันในข้อมูลพฤติกรรมการซื้อของผู้ใช้ที่สามารถนำไปประยุกต์ใช้ในการกำหนดกลยุทธ์การตลาดเชิงบูรณาการ เช่น การโปรโมตข้ามแบรนด์ (Cross-Brand Promotion) และการเลือกช่องทางการขายที่เหมาะสม เพื่อเพิ่มประสิทธิภาพในการตลาดและยกระดับความสามารถในการแข่งขันขององค์กร ผลการวิเคราะห์จากการใช้อัลกอริธึม Apriori แสดงให้เห็นถึงความสัมพันธ์ระหว่างแบรนด์และช่องทางการขายที่สามารถนำไปใช้ในการวางแผนและดำเนินกลยุทธ์ทางการตลาดในอนาคต

คำสำคัญ : การวิเคราะห์ตะกร้าสินค้า, พฤติกรรมผู้ใช้งาน, แบรนด์สินค้า, กลยุทธ์ทางการตลาด

Project Title : Analysis of User Behavior from Online Stores using
Market Basket Analysis Technique

Credits : 5 Units

By : Mr. Narathon Pinsuk 6504800012
Mr. Supawin Pitichotenant 6504800028

Advisor : Miss Janya Yaemcharoen

Degree : Bachelor of Science

Major : Computer Science

Faculty : Science

Semester / Academic year : 3 / 2024

Abstract

This project focuses on utilizing the Market Basket Analysis (MBA) technique to study user behavior on online stores, particularly in terms of brand products and sales channels. This technique helps identify co-occurrence patterns in user purchase behavior, which can be applied to develop integrated marketing strategies such as Cross-Brand Promotion and choosing the appropriate sales channels. The results from the use of the Apriori algorithm reveal relationships between brands and sales channels, which can be used for planning and implementing future marketing strategies to enhance efficiency and strengthen an organization's competitive ability.

Keywords: market asket analysis, user behavior, cross-brand promotion,
marketing strategies, apriori algorithm

นางสาว จุฬารัตน์

(Co-op Advisor.)

Approved by

g

สารบัญ

	หน้า
จดหมายนำส่งรายงาน	ก
กิตติกรรมประกาศ	ข
บทคัดย่อ	ค
บทคัดย่อภาษาอังกฤษ (Abstract)	ง
บทที่ 1 บทนำ	
1.1 ที่มาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์ของโครงการ.....	2
1.3 ขอบเขตของโครงการ.....	2
1.4 ประโยชน์ที่ได้รับ.....	2
1.5 ขั้นตอนและวิธีการดำเนินงาน.....	3
1.6 ระยะเวลาในการดำเนินงาน.....	4
1.7 อุปกรณ์และเครื่องมือที่ใช้.....	5
บทที่ 2 การทบทวนเอกสารและวรรณกรรมที่เกี่ยวข้อง	
2.1 โปรแกรม Tableau.....	6
2.2 ภาษาไพธอน (Python).....	7
2.3 Google Colab.....	8
2.4 Data Preparation.....	8
2.5 Microsoft Excel.....	10
2.6 Market Basket Analysis.....	11
2.3 Apriori Algorithm.....	12
บทที่ 3 รายละเอียดการปฏิบัติงาน	
3.1 ชื่อและที่ตั้งสถานประกอบการ.....	15
3.2 ลักษณะการประกอบการ ผลิตภัณฑ์การให้บริการหลักขององค์กร.....	16
3.4 ตำแหน่งงานและลักษณะงานที่ได้รับมอบหมาย.....	16
3.4 ชื่อและตำแหน่งงานของพนักงานที่ปรึกษา.....	22
3.5 ระยะเวลาที่ปฏิบัติงาน.....	22

สารบัญ (ต่อ)

	หน้า
บทที่ 4 ผลการปฏิบัติงานโครงการ	
4.1 รายละเอียดของโครงการ.....	23
4.2 ขั้นตอนในการวิเคราะห์ข้อมูล.....	23
บทที่ 5 สรุปผลและข้อเสนอแนะ	
5.1 สรุปผลโครงการ.....	42
5.2 สรุปผลการปฏิบัติสหกิจศึกษา.....	43
บรรณานุกรม.....	45
ภาคผนวก.....	46
ภาคผนวก ก ภาพขณะปฏิบัติงานสหกิจศึกษา.....	47
ประวัติผู้จัดทำ	
แบบสรุปโครงการสหกิจศึกษาและการศึกษาเชิงบูรณาการกับการทำงาน (CWIE)	

สารบัญตาราง

	หน้า
ตารางที่ 1.1 ขั้นตอนและระยะเวลาในการดำเนินโครงการ.....	4
ตารางที่ 2.1 ตารางแสดงตัวอย่างข้อมูลธุรกรรมจำลอง เพื่อสาธิตการประยุกต์ใช้ อัลกอริทึม Apriori ในการค้นหาความสัมพันธ์.....	13
ตารางที่ 2.2 ตารางการแสดงผลการแปลงข้อมูลธุรกรรมด้วยเทคนิค One-hot Encoding.....	14
ตารางที่ 2.3 ตัวอย่างกฎความสัมพันธ์ที่ได้จากการวิเคราะห์ตะกร้าสินค้า โดยใช้ค่า ตัวชี้วัด Support, Confidence และ Lift.....	14

สารบัญรูปภาพ

	หน้า
รูปที่ 2.1 ตัวอย่างหน้าโปรแกรม Tableau.....	6
รูปที่ 2.2 แสดงการใช้ Library และ Import ไฟล์ Excel โดย Python.....	8
รูปที่ 2.3 ตัวอย่างหน้า Google Colab.....	8
รูปที่ 2.4 ชุดคำสั่ง Python ที่ใช้เลือกและจัดรูปแบบให้เหมาะกับการทำ MBA.....	10
รูปที่ 3.1 แผนที่ตั้ง บริษัท ดาต้า เพิร์ส จำกัด.....	15
รูปที่ 3.2 คอร์ส Meta Data Analyst Professional Certificate.....	17
รูปที่ 3.3 ตัวอย่างข้อมูล Keyword Labeling สุนัขและแมว.....	18
รูปที่ 3.4 ตัวอย่างข้อมูล การรวบรวมรายชื่อผู้เข้าร่วมงาน Data First Campus.....	18
รูปที่ 3.5 ตัวอย่างการการนำเสนอข้อมูล.....	19
รูปที่ 3.6 ตัวอย่างข้อมูล Keyword Labeling ผลิตภัณฑ์ดูแลช่องปาก.....	20
รูปที่ 4.1 ตัวอย่าง Grouped Matrix-based Visualization.....	21
รูปที่ 4.2 ตัวอย่างข้อมูลดิบ (raw data) บางส่วน.....	26
รูปที่ 4.3 ตัวอย่างชุดคำสั่ง Python ที่ใช้ทำ one-hot encoding.....	27
รูปที่ 4.4 ตัวอย่างข้อมูลที่จัดเตรียมให้อยู่ในรูปแบบที่เหมาะสม บางส่วน.....	27
รูปที่ 4.5 ตัวอย่างชุดคำสั่ง Python สำหรับสร้างกฎความสัมพันธ์ด้วยอัลกอริทึม Apriori.....	28
รูปที่ 4.6 ผลลัพธ์กฎความสัมพันธ์ (Association Rules) ที่ได้จากอัลกอริทึม Apriori.....	28
รูปที่ 4.7 ผลลัพธ์การใช้เทคนิค Grouped Matrix-based Visualization กับ ข้อมูลแรนด์.....	30
รูปที่ 4.8 ความสัมพันธ์ระหว่างแบรนต์ผลิตภัณฑ์ดูแลผมร่วง H และแบรนต์ของใช้เด็ก K.....	30
รูปที่ 4.9 ความสัมพันธ์ระหว่างแบรนต์ผลิตภัณฑ์สำหรับสัตว์เลี้ยง P และแบรนต์สินค้าทำ ความสะอาดมือ F.....	32
รูปที่ 4.10 ความสัมพันธ์ระหว่างแบรนต์จำหน่ายยาสีฟันเฉพาะทาง A และแบรนต์ผลิต ผลิตภัณฑ์ดูแลช่องปากครบวงจร B.....	33
รูปที่ 4.11 ผลลัพธ์การใช้เทคนิค Grouped Matrix-based Visualization กับ แบรนต์และ ช่องทางการขาย.....	35
รูปที่ 4.12 ผลลัพธ์การใช้ Filter เพื่อคัดเลือกและแสดงเฉพาะข้อมูลที่เกี่ยวข้อง.....	36
รูปที่ 4.13 ความสัมพันธ์ระหว่างช่องทางการขาย L และแบรนต์ผลิตภัณฑ์ดูแลสุขภาพ G.....	37

สารบัญรูปภาพ (ต่อ)

	หน้า
รูปที่ 4.14 ความสัมพันธ์ระหว่างช่องทางการขาย L และแบรนด์ของใช้เด็ก K.....	38
รูปที่ 4.15 ความสัมพันธ์ระหว่างช่องทางการขาย Z และแบรนด์สกินแคร์ R.....	39
รูปที่ 4.16 ความสัมพันธ์ระหว่างช่องทางการขายทั้งหมด และแบรนด์ยาสระผม Q.....	40
รูปที่ 4.17 ความสัมพันธ์ระหว่างช่องทางการขายทั้งหมด และแบรนด์น้ำยาปรับผ้านุ่ม E.....	41
รูปที่ ก.1 ขณะปฏิบัติงาน.....	47
รูปที่ ก.2 ขณะปฏิบัติงาน.....	47



บทที่ 1

บทนำ

1.1 ที่มาและความสำคัญของปัญหา

บริษัท ดาต้า เฟิร์ส จำกัด ในฐานะผู้นำด้านการให้คำปรึกษาและการประยุกต์ใช้ข้อมูลเชิงธุรกิจแบบครบวงจร มีความเชี่ยวชาญทั้งด้านการตลาดออนไลน์และออฟไลน์ การจัดการข้อมูลเพื่อสร้างคุณค่าทางธุรกิจที่ยั่งยืน ในยุคดิจิทัลที่มีการเปลี่ยนแปลงอย่างรวดเร็ว พฤติกรรมของผู้บริโภคบนแพลตฟอร์มร้านค้าออนไลน์ (E-commerce Platforms) กลายเป็นปัจจัยสำคัญที่องค์กรต่าง ๆ ต้องทำความเข้าใจอย่างลึกซึ้ง เพื่อระบุโอกาสทางการตลาดและพัฒนากลยุทธ์ที่ตอบสนองต่อความต้องการที่เปลี่ยนแปลงไปของผู้ใช้ การเข้าใจความสัมพันธ์ระหว่าง แแบรนด์ สินค้าและช่องทางการขาย ถือเป็นหนึ่งในกุญแจสำคัญของธุรกิจค้าปลีกออนไลน์ การวิเคราะห์ว่าผู้ใช้ที่สนใจแบรนด์หนึ่งมีแนวโน้มที่จะเข้าถึงแบรนด์อื่น ๆ หรือใช้ช่องทางการขายใดร่วมด้วย สามารถช่วยให้องค์กรกำหนดกลยุทธ์การทำ Cross-Brand Promotion เลือกช่องทางการขายที่เหมาะสม และเพิ่มประสิทธิภาพของแคมเปญทางการตลาดได้ดียิ่งขึ้น

ด้วยเหตุนี้ ผู้จัดทำจึงเห็นความสำคัญในการนำ เทคนิคการวิเคราะห์ตะกร้าตลาด (Market Basket Analysis: MBA) มาประยุกต์ใช้ในการศึกษาพฤติกรรมผู้ใช้งานร้านค้าออนไลน์ โดยเฉพาะในมิติของ แแบรนด์และช่องทางการขาย เพื่อค้นหารูปแบบความสัมพันธ์ที่มักเกิดขึ้นร่วมกัน ซึ่งจะช่วยให้สามารถทำความเข้าใจแนวโน้มของผู้ใช้งานในเชิงลึก ตลอดจนระบุโอกาสทางการตลาดใหม่ ๆ ที่สามารถนำไปประยุกต์ใช้ในการเพิ่มยอดขายและเสริมสร้างความสามารถในการแข่งขันขององค์กร การดำเนินการโครงการนี้ครอบคลุมตั้งแต่การศึกษาและทำความเข้าใจข้อมูล (Data Understanding) การกำหนดสมมติฐานการวิเคราะห์ (Hypothesis) การเตรียมข้อมูลให้อยู่ในรูปแบบที่เหมาะสม (Data Preparation) การวิเคราะห์ข้อมูลด้วยอัลกอริทึม Apriori และการสรุปผลลัพธ์เพื่อนำเสนอในรูปแบบที่เข้าใจง่าย (Data Presentation) โดยอาศัยเครื่องมือ Google Sheets, Google Colab และ Tableau Public เป็นเครื่องมือสนับสนุน ผลลัพธ์ที่ได้จะเป็นข้อมูลเชิงลึกที่สามารถนำไปใช้ในการออกแบบกลยุทธ์การตลาดเชิงบูรณาการ และเป็นแนวทางให้กับผู้ประกอบการในการใช้ประโยชน์จากข้อมูลเพื่อการตัดสินใจเชิงกลยุทธ์อย่างมีประสิทธิภาพในอนาคต

1.2 วัตถุประสงค์ของโครงการ

เพื่อวิเคราะห์แนวโน้มและรูปแบบความสัมพันธ์ระหว่าง แบรินด์สินค้าและช่องทางการขาย จากพฤติกรรมการเข้าถึงและการมีส่วนร่วมของผู้ใช้งานบนแพลตฟอร์มร้านค้าออนไลน์ โดยใช้เทคนิค Market Basket Analysis (MBA) เพื่อค้นหาความเชื่อมโยงที่มักเกิดขึ้นร่วมกัน อันจะนำไปสู่ความเข้าใจพฤติกรรมผู้ใช้งานเชิงลึกและการกำหนดกลยุทธ์ทางการตลาดที่มีประสิทธิภาพ

1.3 ขอบเขตของโครงการ

1.3.1 แหล่งข้อมูลที่น่าเชื่อถือมาใช้งานมาจาก Holistica

1.3.2 การวิเคราะห์ข้อมูลแบ่งเป็น 4 ส่วน

1.3.2.1 Data Understanding

ศึกษาและทำความเข้าใจกับโครงสร้างและคุณลักษณะของข้อมูลที่ได้รับ

1.3.2.2 Data Preparation

จัดเตรียมและทำความสะอาดข้อมูลให้อยู่ในรูปแบบที่เหมาะสม เพื่อให้ได้ข้อมูลที่มีคุณภาพมากที่สุด

1.3.2.3 Data Analytics

วิเคราะห์ข้อมูลเพื่อค้นหาความสัมพันธ์ระหว่างแบรินด์และช่องทางการขาย โดยใช้เทคนิค Market Basket Analysis ด้วยอัลกอริทึม Apriori

1.3.2.4 Data Visualization

นำเสนอผลการวิเคราะห์ในรูปแบบของแดชบอร์ด ที่ประกอบด้วยแผนภาพและกราฟ เพื่อให้เข้าใจได้ง่ายและนำไปใช้ประโยชน์ต่อการตัดสินใจได้จริง

1.4 ประโยชน์ที่ได้รับ

1.4.1 ช่วยให้เข้าใจภาพรวมของพฤติกรรมของผู้ใช้งานร้านค้าออนไลน์ ความเชื่อมโยงระหว่างแบรินด์สินค้าและช่องทางการขาย รวมถึงแนวโน้มของการเข้าถึงและการมีส่วนร่วมของผู้ใช้

1.4.2 ช่วยสนับสนุนการวางแผนกลยุทธ์ทางการตลาด เช่น การทำ Cross-Brand Promotion และการเลือกช่องทางการขายที่เหมาะสม เพื่อเพิ่มประสิทธิภาพในการเข้าถึงลูกค้าที่เป็นกลุ่มเป้าหมาย

- 1.4.3 ช่วยในการประเมินโอกาสทางธุรกิจและความเสี่ยงจากพฤติกรรมผู้บริโภค รวมถึงการตัดสินใจเชิงกลยุทธ์ที่เกี่ยวข้องกับการจัดการแบรนด์และช่องทางการขาย
- 1.4.4 ช่วยเพิ่มทักษะการใช้โปรแกรม Tableau ในการสร้าง Visualization เพื่อนำเสนอข้อมูลให้อยู่ในรูปแบบที่เข้าใจง่ายและสื่อสารได้อย่างมีประสิทธิภาพ
- 1.4.5 ช่วยเพิ่มทักษะการใช้ภาษา Python สำหรับการทำความสะอาด จัดเตรียม และวิเคราะห์ข้อมูลด้วยเทคนิค Market Basket Analysis และอัลกอริทึม Apriori

1.5 ขั้นตอนและวิธีการดำเนินงาน

ในการดำเนินงานจัดทำโครงการสหกิจศึกษาวิเคราะห์พฤติกรรมของผู้ใช้งานบนร้านค้าออนไลน์ มีลำดับขั้นตอนการจัดทำ ดังนี้

1.5.1 การศึกษาและรวบรวมข้อมูล (Data Collection)

รวบรวมข้อมูลที่เกี่ยวข้องกับเรื่องที่สนใจ โดยปรึกษากับพนักงานที่ปรึกษาว่าข้อมูลที่ได้รับมาจาก Holistica สามารถนำไปวิเคราะห์ได้มากน้อยเพียงใด และทำความเข้าใจกับข้อมูลที่ได้มา

1.5.2 การตั้งเป้าหมาย (Goal)

ตั้งโจทย์หรือคำถาม เพื่อหาแนวทางวิเคราะห์ข้อมูล และสามารถนำผลการวิเคราะห์ไปต่อยอดด้านใดได้บ้าง โดยคณะผู้จัดทำตั้งโจทย์เกี่ยวกับพฤติกรรมของผู้ใช้งานว่ามีความสนใจในสินค้าแบรนด์ใดร่วมกับแบรนด์หรือช่องทางใดบ้างที่ผู้ใช้งานมักกดเข้ามาดูสินค้าของแบรนด์นั้น ความถี่ (Frequency) ในการลงโพสต์ของสินค้าแบรนด์นั้น การจับคู่สินค้านี้ระหว่างแบรนด์ เช่น ข้อมูลของร้านค้าที่คณะผู้จัดทำได้นำมาเป็นตัวอย่างในโครงการนี้ เปรียบเทียบความสัมพันธ์กับข้อมูลของช่องทางการจำหน่ายและแบรนด์ จากพฤติกรรมการใช้งานของผู้ใช้ที่ได้จาก Holistica กับข้อมูลการโพสต์และขายบนแพลตฟอร์ม E-Commerce ได้แก่ Line, Shopee, Lazada และ Website

1.5.3 การเตรียมข้อมูล (Data Preparation)

ในการเตรียมข้อมูลเริ่มต้นด้วยการตรวจสอบความถูกต้องของข้อมูลที่ต้องใช้ในการวิเคราะห์และเตรียมข้อมูล เช่น จัดการกับค่าที่หายไป (Missing value) แปลงข้อมูลให้อยู่ในรูปแบบที่เหมาะสมกับการทำ Market Basket Analysis และการแปลงให้อยู่ในรูปแบบ One-Hot Encoding ด้วยภาษา Python ผ่าน Google Colab รวมถึงการจัดรูปแบบให้อยู่ในรูปของตารางสำหรับอัลกอริทึม Apriori

1.5.4 การวิเคราะห์ข้อมูล (Data Analytics)

ทำการวิเคราะห์ข้อมูลตามโจทย์ที่ตั้งไว้ โดยคณะผู้จัดทำนำข้อมูลที่ได้จัดเตรียมไว้เข้าสู่กระบวนการวิเคราะห์ โดยใช้เทคนิค Market Basket Analysis (MBA) อัลกอริทึม Apriori เพื่อตรวจหาความสัมพันธ์ของชุดข้อมูล (Association Rules) ตัวชี้วัดที่นำมาใช้ในการวิเคราะห์ประกอบด้วย

ค่า Support คือ ความถี่ของการเกิดร่วมกันของชุดข้อมูล

ค่า Confidence คือ ความน่าจะเป็นที่เกิดการซื้อหรือการเข้าถึงร่วมกัน

ค่า Lift คือ ความสัมพันธ์ที่แท้จริงเทียบกับความบังเอิญ

1.5.5 การจัดทำแผนภาพข้อมูลเพื่อการนำเสนอ (Data Visualization)

สรุปผลและแสดงข้อมูลในรูปของแผนภาพ กราฟ และตาราง โดยเน้นการนำเสนอในรูปแบบที่เข้าใจง่ายผ่านการมองเห็น ในรูปแบบ Dashboard โดยใช้ Tableau Public เพื่อสื่อสารข้อมูลเชิงลึกให้เข้าใจได้ง่าย

1.5.6 จัดทำเอกสาร (Create Document)

จัดทำรูปเล่มประกอบโครงงาน แนวทางในการจัดทำโครงงาน วิธีและขั้นตอนการดำเนินการ เพื่อเสนอต่ออาจารย์ที่ปรึกษา และคู่มือการใช้งานสำหรับสถานประกอบการใช้อ้างอิงในอนาคต

1.6 ระยะเวลาในการดำเนินงาน

ตารางที่ 1.1 ขั้นตอนและระยะเวลาในการดำเนินโครงงาน

ขั้นตอนการดำเนินงาน	พ.ค. 68	มิ.ย. 68	ก.ค. 68	ส.ค. 68	ก.ย. 68
1. ศึกษาข้อมูล	←→				
2. ตั้งเป้าหมาย			←→		
3. เตรียมข้อมูล			←→		
4. วิเคราะห์ข้อมูล					
5. จัดทำแผนภาพเพื่อการนำเสนอ				←→	←→
6. จัดทำเอกสาร					

1.7 อุปกรณ์และเครื่องมือที่ใช้

1.7.1 ฮาร์ดแวร์

1.7.1.1 เครื่องคอมพิวเตอร์โน้ตบุค Acer Nitro V 15

- 13th Gen Intel Core i5
- Ram 16 GB
- SSD 477 GB
- NVIDIA GeForce RTX 3050 6GB
- Windows 11 Home Single Language

1.7.1.1 เครื่องคอมพิวเตอร์โน้ตบุค Acer Nitro 5

- 8th Gen Intel Core i5
- Ram 16 GB
- SSD 1200 GB
- NVIDIA GeForce GTX 1050 4GB
- Windows 11 Home Single Language

1.7.2 ซอฟต์แวร์

1.7.2.1 ระบบปฏิบัติการ Window 11

1.7.2.2 เครื่องมือ Google Colab

1.7.2.3 โปรแกรม Microsoft Excel

1.7.2.4 โปรแกรม Tableau

บทที่ 2

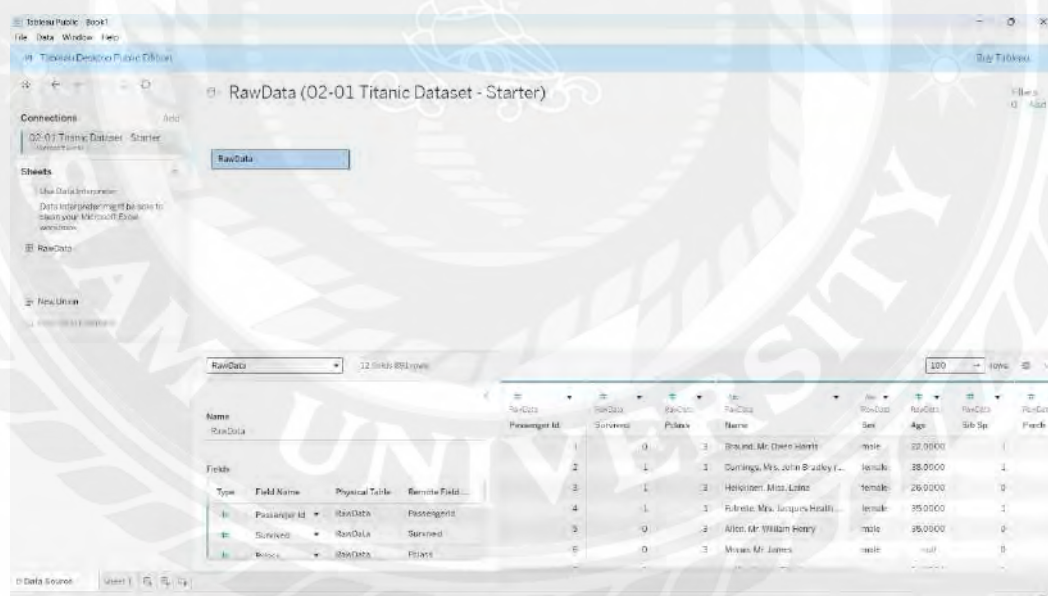
การทบทวนเอกสารและวรรณกรรมที่เกี่ยวข้อง

โครงการสหกิจศึกษาเรื่อง การวิเคราะห์พฤติกรรมของผู้ใช้งานจากร้านค้าออนไลน์ด้วยเทคนิคการวิเคราะห์ตะกร้าสินค้า มีแนวคิดทฤษฎีและงานวิจัยที่เกี่ยวข้องดังต่อไปนี้

2.1 โปรแกรม Tableau¹

Tableau เป็นโปรแกรมสำหรับการวิเคราะห์ข้อมูล และหาคำตอบในเชิงธุรกิจอย่างรวดเร็ว โดยการใช้ Tableau จะเป็นการช่วยให้ผู้ใช้งานมีความเข้าใจในข้อมูลมากขึ้น รวมถึงผู้ใช้งานสามารถนำข้อมูลที่มีอยู่มาวิเคราะห์และแสดงผลอย่างมีประสิทธิภาพ

โดยคณะผู้จัดทำได้เลือกใช้โปรแกรม Tableau มาใช้ในการจัดทำแผนภาพข้อมูล หรือ Visualization เพื่อให้ง่ายต่อการวิเคราะห์ข้อมูลเชิงลึก (Insight) และการนำเสนอผลลัพธ์ (Presentation) โดยนำมาประยุกต์ใช้ในการวิเคราะห์ข้อมูลเชิงลึกทั้งหมด ทั้งการศึกษาการใช้งาน การสร้างแผนภาพหรือกราฟ เพื่อนำเสนอข้อมูลให้ผู้ใช้และผู้อ่านผลเข้าใจง่าย



รูปที่ 2.1 ตัวอย่างหน้าโปรแกรม Tableau

¹ <https://www.bac.co.th/web/products-and-services/tableau-software/>

2.2 ภาษาไพธอน (Python)²

ภาษาไพธอน (Python) เป็นภาษาคอมพิวเตอร์ที่ได้รับความนิยมอย่างแพร่หลายในสำหรับการเขียนโปรแกรมด้านวิทยาศาสตร์ข้อมูล (Data Science) และการวิเคราะห์ข้อมูล (Data Analytics) เนื่องจากเป็นภาษาที่เข้าใจง่าย เรียนรู้ได้รวดเร็ว มีชุมชนนักพัฒนาขนาดใหญ่ และมีไลบรารี (Libraries) จำนวนมากที่ช่วยสนับสนุนการทำงานทั้งด้านการประมวลผลข้อมูล การวิเคราะห์ และการแสดงผลเชิงภาพ

สำหรับโครงการนี้ ซึ่งมุ่งเน้นการวิเคราะห์พฤติกรรมผู้ใช้งานร้านค้าออนไลน์ด้วยเทคนิค Market Basket Analysis (MBA) คณะผู้จัดทำได้เลือกใช้ Python ร่วมกับไลบรารีที่เกี่ยวข้อง ดังนี้

2.2.1 Libraries ที่ใช้ในการเตรียมและจัดการข้อมูล (Data Preparation)

2.2.1.1 Pandas

ใช้สำหรับอ่านไฟล์ข้อมูล (CSV/Excel) และจัดการข้อมูลในรูปแบบ DataFrame เช่น การ groupby, crosstab และการแปลงข้อมูลให้อยู่ในรูปแบบตารางธุรกรรม (Transaction Table) ซึ่งจำเป็นต่อการทำ Market Basket Analysis

2.2.1.2 NumPy

ใช้ร่วมกับ Pandas และ mlxtend เพื่อช่วยในการคำนวณเชิงตัวเลข และการจัดการข้อมูลแบบอาร์เรย์

2.2.2 Libraries ที่ใช้ในการวิเคราะห์ (Data Analytics)

2.2.2.1 mlxtend (Machine Learning Extensions)

ใช้สำหรับการทำ Apriori Algorithm และ Association Rules ซึ่งเป็นหัวใจหลักของ Market Basket Analysis ในการค้นหาความสัมพันธ์ของชุดข้อมูล เช่น การคำนวณค่า Support, Confidence และ Lift

2.2.2.2 Scikit-learn

ใช้เป็นพื้นฐานสนับสนุนการทำงานของ mlxtend และการเตรียมข้อมูลเบื้องต้น เช่น การจัดการค่าข้อมูล และการแปลงข้อมูลให้อยู่ในรูปแบบ Boolean

² <https://getcodecamp.com/2020/08/26/data-science-ทำไมต้องใช้-python-และ-ทำเจ/>

```
import pandas as pd
import numpy as np
df2 = data=pd.read_excel('exported_customers_1558699178 (1).xlsx')
df2
```

รูปที่ 2.2 แสดงการใช้ Library และ Import ไฟล์ Excel โดย Python

2.3 Google Colab³

Google Colaboratory หรือ Google Colab เป็นแพลตฟอร์มออนไลน์ที่ให้บริการสมุดบันทึกเชิงโต้ตอบ (Interactive Notebook) สำหรับการเขียนและรันชุดคำสั่งภาษา Python ผ่านเว็บเบราว์เซอร์ โดยไม่จำเป็นต้องติดตั้งซอฟต์แวร์เพิ่มเติม ผู้ใช้สามารถเข้าถึงและใช้งานได้โดยไม่มีค่าใช้จ่ายผ่านบัญชีของ Google และมีจุดเด่นที่รองรับการประมวลผลด้วย GPU/TPU ซึ่งเหมาะสำหรับงานด้านวิทยาศาสตร์ข้อมูล (Data Science) และการเรียนรู้ของเครื่อง (Machine Learning)



รูปที่ 2.3 ตัวอย่างหน้าจอโปรแกรม Google Colab

2.4 Data Preparation⁴

การเตรียมข้อมูล (Data Preparation) หมายถึงกระบวนการปรับปรุงและแปลงข้อมูลดิบ (Raw Data) ให้อยู่ในรูปแบบที่พร้อมนำไปใช้งาน เช่น การวิเคราะห์ทางสถิติหรือการประมวลผลด้วยอัลกอริทึม การทำ Data Preparation จึงเป็นขั้นตอนสำคัญที่จะช่วยให้ข้อมูลมี

³ <https://www.novelbiz.co.th/google-colab/>

⁴ <https://bzinsight.wordpress.com/2014/06/11/การทำ-data-preparation-อย่างมีอาชีพ/>

คุณภาพ ถูกต้อง และครบถ้วน โดยทั่วไปจะรวมถึงการทำความสะอาดข้อมูล (Data Cleaning) และการจัดโครงสร้างข้อมูลให้อยู่ในรูปแบบที่เหมาะสม (Data Conversion)

คุณลักษณะสำคัญของการเตรียมข้อมูลที่มีประสิทธิภาพ ได้แก่

- ข้อมูลครบถ้วนสมบูรณ์และปราศจากความผิดพลาด
- มีการกำหนดนิยามข้อมูลที่ชัดเจนและสอดคล้องกัน
- มีการบันทึกขั้นตอนการจัดการข้อมูลอย่างเป็นระบบเพื่อให้ง่ายต่อการตรวจสอบ
- มีการปรับกระบวนการให้เป็นกึ่งอัตโนมัติหรือตั้งค่าเป็น Script เพื่อช่วยลดข้อผิดพลาดจากการทำงานด้วยมือ

ในการดำเนินโครงการนี้ ผู้จัดทำได้นำแนวคิดการเตรียมข้อมูลมาประยุกต์ใช้กับข้อมูลพฤติกรรมผู้ใช้งานของร้านค้าหนึ่งบนแพลตฟอร์ม LINE โดยมีขั้นตอนหลักดังนี้

- 2.4.1 การลบข้อมูลที่ไม่จำเป็น เช่น คอลัมน์ที่ไม่เกี่ยวข้องกับการวิเคราะห์
- 2.4.2 การเลือกข้อมูลที่เราสนใจ เช่น เลือกคอลัมน์ ที่เราสนใจจะนำมาใช้วิเคราะห์เพียงบางส่วน
- 2.4.3 การปรับรูปแบบข้อมูล เช่น การเปลี่ยนประเภทข้อมูล (Data Types) ให้ถูกต้อง
- 2.4.4 การจัดรูปแบบวันที่ให้อยู่ในมาตรฐานเดียวกัน
- 2.4.5 การแปลงข้อมูลให้อยู่ในรูปแบบที่เหมาะสมกับ Market Basket Analysis เช่น การใช้ group by และ cross tab เพื่อสร้างตารางธุรกรรม (Transaction Table) และ การทำ One-Hot Encoding ให้ได้ข้อมูลในรูปแบบ 0 และ 1
- 2.4.6 การบันทึกขั้นตอนการทำงาน ไว้อย่างละเอียดด้วยโค้ด Python ใน Google Colab เพื่อความโปร่งใสและสามารถทำซ้ำได้ในอนาคต

การเตรียมข้อมูลที่มีคุณภาพทำให้สามารถนำชุดข้อมูลไปใช้กับอัลกอริทึม Apriori ได้อย่างมีประสิทธิภาพ และช่วยให้ผลลัพธ์การวิเคราะห์พฤติกรรมผู้ใช้งานมีความน่าเชื่อถือมากยิ่งขึ้น



```

from mlxtend.frequent_patterns import apriori
from mlxtend.frequent_patterns import association_rules

columns_of_interest = [...]

existing_columns = [col for col in columns_of_interest if col in df2.columns]

# Pass the DataFrame with selected columns to apriori
frequent_itemsets = apriori(df2[existing_columns], astype=bool, min_support=0.005, use_colnames=True)

# frequent_itemsets_freq = frequent_itemsets.sort_values("support", axis=0, ascending=False).head(50)

rules = association_rules(frequent_itemsets, metric="lift")

# Display rules
rules = rules[
    (rules["antecedents"].apply(lambda x: len(x) == 1)) &
    (rules["consequents"].apply(lambda x: len(x) == 1))
]

display(rules)

```

รูปที่ 2.4 ชุดคำสั่ง Python ที่ใช้เลือกและจัดรูปแบบให้เหมาะสมกับการวิเคราะห์ MBA

2.5 Microsoft Excel⁵

Microsoft Excel เป็นโปรแกรมตารางคำนวณ (Spreadsheet) ที่นิยมใช้กันอย่างแพร่หลาย สามารถทำงานด้านการจัดการข้อมูล การคำนวณ และการนำเสนอข้อมูลได้อย่างมีประสิทธิภาพ

ในการจัดทำโครงงานนี้ ผู้จัดทำได้ใช้ Excel ในการ จัดการข้อมูลเบื้องต้น (Data Preparation) โดยเน้นการใช้งานฟังก์ชันพื้นฐาน ได้แก่

- การกรองข้อมูล (Filter): เพื่อเลือกเฉพาะข้อมูลที่สนใจ
- การลบหรือซ่อนข้อมูลที่ไม่ต้องการ (Remove/Hide): เพื่อลดความซับซ้อนของชุดข้อมูล
- การเลือกคอลัมน์ที่เกี่ยวข้อง (Select Columns): ทำให้ข้อมูลที่เหลืออยู่ตรงตามที่ต้องการนำไปวิเคราะห์
- การคำนวณเบื้องต้นด้วยสูตร (Formula): เช่น SUM, AVERAGE, IF เพื่อช่วยตรวจสอบหรือสรุปผล

แม้ว่า Excel จะมีเครื่องมือขั้นสูง เช่น Power Query หรือ Power Pivot แต่โครงงานสหกิจศึกษา เน้นการใช้งานมุ่งเน้นไปที่การจัดการข้อมูลอย่างง่าย ด้วยฟังก์ชันพื้นฐานของ Excel เพื่อให้ได้ข้อมูลที่สะอาดและพร้อมนำไปใช้ในการวิเคราะห์ต่อไป

⁵ <https://www.9experttraining.com/articles/มีอะไรใน-microsoft-excel-365>

2.6 Market Basket Analysis⁶

Market Basket Analysis (MBA) หรือ การวิเคราะห์ตะกร้าตลาด เป็นเทคนิคที่ใช้ในการค้นหาความสัมพันธ์ของชุดรายการสินค้าที่ถูกซื้อร่วมกันโดยลูกค้าในการทำธุรกรรมหนึ่งๆ โดยอาศัยแนวคิดของ กฎความสัมพันธ์ (Association Rules) MBA ได้รับความนิยมอย่างแพร่หลายในภาคธุรกิจโดยเฉพาะในอุตสาหกรรมค้าปลีก ร้านสะดวกซื้อ และแพลตฟอร์ม E-commerce ซึ่งมักนำผลการวิเคราะห์ไปใช้ในการวางกลยุทธ์ทางการตลาด เช่น การจัดวางสินค้า การออกแบบโปรโมชั่น และระบบแนะนำสินค้า (Recommender System) การวิเคราะห์ตะกร้าสินค้าจะเริ่มจากการตรวจสอบชุดรายการสินค้าที่มักถูกซื้อร่วมกันเป็นประจำ ซึ่งในทางเทคนิคเรียกว่า Frequent Itemsets จากนั้นจึงนำไปสร้างเป็น กฎความสัมพันธ์ (Association Rules) ที่สามารถนำไปประยุกต์ใช้ได้จริง กฎที่ได้จะอยู่ในรูปแบบ ถ้าซื้อ A $\rightarrow$ มีแนวโน้มจะซื้อ B ด้วย โดยจะใช้ค่าตัวชี้วัด เช่น Support, Confidence และ Lift เพื่อประเมินความน่าเชื่อถือและความสัมพันธ์ของกฎนั้นๆ

2.6.1 การคำนวณค่ากฎความสัมพันธ์ (Association Rules)

ในการวิเคราะห์ตะกร้าสินค้า (Market Basket Analysis) การประเมินความน่าเชื่อถือของกฎความสัมพันธ์ (Association Rules) จำเป็นต้องใช้ตัวชี้วัดทางสถิติที่สำคัญ 3 ค่า ได้แก่ Support, Confidence และ Lift โดยแต่ละค่ามีความหมายและวิธีการคำนวณดังต่อไปนี้

2.6.1.1 ค่า Support (ซัพพอร์ต) คือ ค่าความถี่สัมพัทธ์ของการเกิดขึ้นร่วมกันของรายการสินค้าทั้งฝั่งซ้าย (Antecedent) และฝั่งขวา (Consequent) ของกฎภายในชุดข้อมูลธุรกรรมทั้งหมด

สูตรการคำนวณ

$$\text{Support}(A \rightarrow B) = \frac{\text{จำนวนธุรกรรมที่มีทั้ง } A \text{ และ } B}{\text{จำนวนธุรกรรมทั้งหมด}}$$

2.6.1.2 ค่า Confidence (คอนฟิเดนซ์) คือ ค่าความน่าจะเป็นที่รายการ B จะถูกซื้อพร้อม เมื่อมีการซื้อรายการ A แล้ว กล่าวคือเป็นการวัดความเชื่อมั่นในความสัมพันธ์ระหว่าง A และ B

⁶ <https://lengyi.medium.com/inside-purchase-behaviour-with-market-basket-analysis-307ccfc980b5>

สูตรการคำนวณ

$$Confidence(A \rightarrow B) = \frac{\text{จำนวนธุรกรรมที่มีทั้ง } A \text{ และ } B}{\text{จำนวนธุรกรรมที่มี } A}$$

2.6.1 ค่า Lift (ลิฟต์) คือ ค่าที่ใช้วัดความแข็งแกร่งของความสัมพันธ์ระหว่าง A และ B โดยเปรียบเทียบกับความน่าจะเป็นที่ B จะเกิดขึ้นโดยไม่ขึ้นอยู่กับ A หากค่า Lift มากกว่า 1 แสดงว่าการเกิดขึ้นของ A เพิ่มโอกาสที่ B จะเกิดขึ้น (มีความสัมพันธ์ในเชิงบวก)

สูตรการคำนวณ

$$Lift(A \rightarrow B) = \frac{Confidence(A \rightarrow B)}{Support(B)}$$

2.7 Apriori Algorithm⁷

Apriori Algorithm เป็นอัลกอริทึมพื้นฐานที่นิยมใช้ในกระบวนการวิเคราะห์กฎความสัมพันธ์ (Association Rule Mining) โดยเฉพาะในบริบทของการวิเคราะห์ตะกร้าสินค้า (Market Basket Analysis) มีวัตถุประสงค์เพื่อค้นหากลุ่มสินค้าที่มักปรากฏร่วมกันในธุรกรรมของลูกค้า โดยอาศัยหลักการ “หากชุดรายการสินค้าหนึ่งไม่มีความถี่ในการปรากฏมากเพียงพอ ชุดที่มีรายการเหล่านั้นเป็นส่วนประกอบย่อมไม่มีความถี่มากเพียงพอเช่นกัน” หลักการดังกล่าวเรียกว่า Apriori Principle

อัลกอริทึมนี้เหมาะสำหรับข้อมูลประเภทไม่ระบุผลลัพธ์ (Unsupervised Data) และเป็นรากฐานสำคัญของการวิเคราะห์ข้อมูลในระบบแนะนำสินค้า การจัดวางสินค้าร่วมกัน และการออกแบบโปรโมชั่นที่มีประสิทธิภาพ

2.7.1 หลักการทำงานของ Apriori Algorithm

กระบวนการทำงานของ Apriori Algorithm ประกอบด้วยขั้นตอนหลักดังนี้

2.7.1.1 การค้นหา Frequent 1-itemsets ตรวจสอบชุดรายการสินค้าทั้งหมดในแต่ละธุรกรรม และคำนวณค่าความถี่ในการปรากฏ (Support) เพื่อ

⁷ <https://bdi.or.th/big-data-101/frequent-itemsets-apriori-algorithm/>

คัดกรองเฉพาะชุดรายการสินค้าที่มีความถี่มากกว่าค่าขั้นต่ำ
(Minimum Support Threshold)

2.7.1.2 การสร้าง Candidate Itemsets ขนาดใหญ่ขึ้น ผสาน Frequent Itemsets ที่พบในขั้นตอนก่อนหน้านี้เพื่อสร้าง กลุ่มชุดรายการสินค้า ขนาด 2 รายการ (2-itemsets), 3 รายการ (3-itemsets) เป็นต้น แล้วทำการตรวจสอบค่า Support ซ้ำอีกครั้ง

2.7.1.3 การทำซ้ำจนไม่สามารถสร้าง Frequent Itemsets ได้
เพิ่มเติม กระบวนการจะหยุดเมื่อไม่สามารถค้นพบ Frequent Itemsets ใหม่ที่ผ่านเกณฑ์การคัดกรองได้อีก

2.7.1.4 การสร้างกฎความสัมพันธ์ (Association Rules) จาก Frequent Itemsets ที่ได้ จะถูกนำมาแปลงเป็นกฎใน รูปแบบ $A \rightarrow B$ โดยมีการ คำนวณค่าตัวชี้วัด คือ Support, Confidence และ Lift

2.7.2 ตัวอย่างการประยุกต์ใช้ Apriori Algorithm

ตารางที่ 2.1 ตารางแสดงตัวอย่างข้อมูลธุรกรรมจำลอง เพื่อสาธิตการประยุกต์ใช้อัลกอริทึม Apriori ในการค้นหาความสัมพันธ์

ลำดับธุรกรรม	ชุดรายการสินค้า
1	ขนมปัง, ไข่, เนย
2	ขนมปัง, น้ำเปล่า
3	ขนมปัง, ไข่, เนย
4	ขนมปัง, น้ำเปล่า

2.7.3 การเตรียมข้อมูลด้วยเทคนิค One-hot Encoding

ในการนำข้อมูลธุรกรรมไปวิเคราะห์ด้วยอัลกอริทึม Apriori จำเป็นต้องแปลงข้อมูลให้อยู่ในรูปแบบที่เครื่องสามารถประมวลผลได้ โดยเฉพาะการแสดงผลข้อมูลในรูปแบบ Binary Matrix ซึ่งแต่ละคอลัมน์แทนการมีหรือไม่มีของสินค้านั้นรายการใดรายการหนึ่ง เทคนิค One-hot Encoding เป็นวิธีที่นิยมใช้ โดยจะแปลงรายการสินค้าให้กลายเป็นค่าทางตรรกะ (1 หมายถึงมีสินค้า, 0 หมายถึงไม่มี) ตามตัวอย่างในตารางที่ 2.2

ตารางที่ 2.2 ตารางการแสดงผลการแปลงข้อมูลธุรกรรมด้วยเทคนิค One-hot Encoding

ลำดับธุรกรรม	ขนมปัง	ไข่	เนย	น้ำเปล่า
1	1	1	1	0
2	1	0	0	1
3	1	1	1	0
4	1	0	0	1

ตารางที่ 2.3 ตัวอย่างกฎความสัมพันธ์ที่ได้จากการวิเคราะห์ตะกร้าสินค้า โดยใช้ค่าตัวชี้วัด

Support, Confidence และ Lift

กฎความสัมพันธ์	Support	Confidence	Lift
{ขนมปัง} $\rightarrow$ {เนย}	0.50	0.67	1.34
{ขนมปัง} $\rightarrow$ {น้ำเปล่า}	0.50	0.50	0.90
{ไข่, ขนมปัง} $\rightarrow$ {เนย}	0.50	1.00	2.00

จากตารางข้างต้น พบว่า กฎ {ขนมปัง} $\rightarrow$ {เนย} มีค่า Confidence และ Lift ในระดับที่น่าสนใจ แสดงให้เห็นถึงแนวโน้มที่ลูกค้าจะซื้อสินค้าทั้งสองรายการร่วมกัน

บทที่ 3

รายละเอียดการปฏิบัติงาน

3.1 ชื่อและที่ตั้งสถานประกอบการ

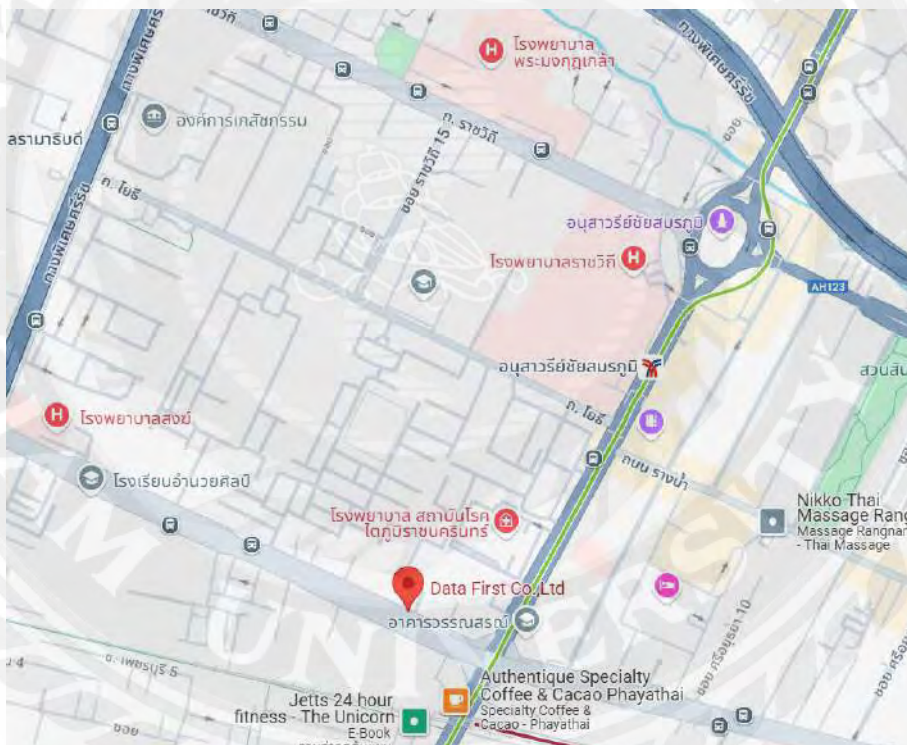
ชื่อสถานประกอบการ : บริษัท ดาต้า เฟิร์ส จำกัด สำนักงานใหญ่ (Data First Co.,Ltd.)

ที่ตั้ง : 465/1 ถนนศรีอยุธยา แขวงทุ่งพญาไท เขตราชเทวี กรุงเทพฯ
10400

โทรศัพท์ : 66(0) 2354 3570

E-mail : support@datafirst.id

เว็บไซต์ : <https://datafirst.co.th>



รูปที่ 3.1 แผนที่ตั้ง บริษัท ดาต้า เฟิร์ส จำกัด

3.2 ลักษณะการประกอบการ ผลิตภัณฑ์การให้บริการหลักขององค์กร

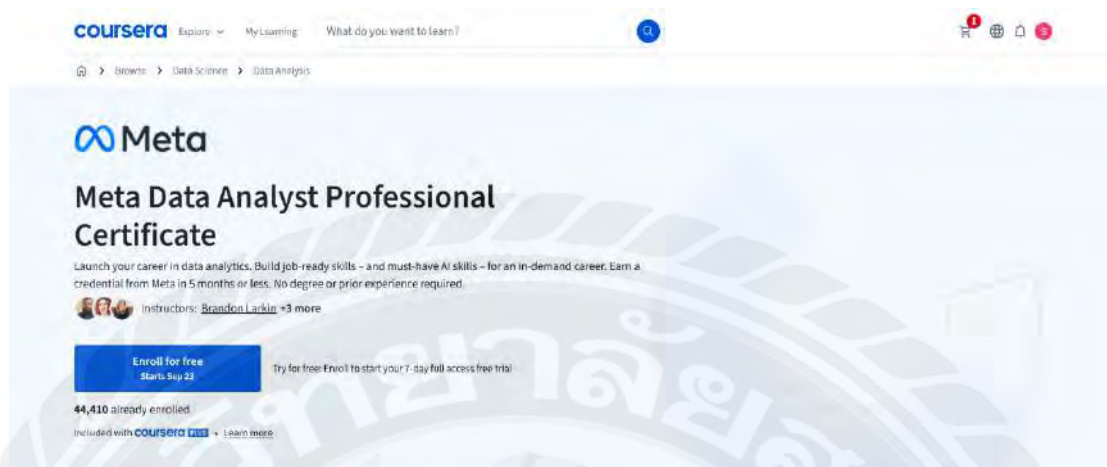
บริษัท ดาต้า เพิร์ส จำกัด ผู้นำด้านให้คำปรึกษาด้านการประยุกต์ใช้ข้อมูลครบวงจร โดดเด่นด้วยประสบการณ์ด้านโฆษณา ทั้งออนไลน์ ออฟไลน์และการประยุกต์ใช้ข้อมูลได้อย่างมีประสิทธิภาพ ทำให้บริษัท ดาต้า เพิร์ส จำกัด สามารถผสานความรู้ด้านการเข้าใจโซเชียลมีเดียในปัจจุบันได้เป็นอย่างดี เป็นผู้ให้บริการโซลูชันด้านข้อมูลและการวิเคราะห์เชิงลึก โดยมุ่งให้ลูกค้าสามารถ “ก้าวกระโดด” ทางธุรกิจได้ผ่านการระบุปัญหาทางธุรกิจที่แท้จริง และจัดเก็บ – เลือกใช้ชุดข้อมูลที่เหมาะสม เพื่อเปิดโอกาสเชิงธุรกิจที่อาจถูกมองข้าม พร้อมทั้งปลดล็อกศักยภาพเต็มที่ขององค์กรด้วยคำว่า “Actionable Data Solution” ในแง่ของบริการเฉพาะ บริษัทให้บริการครอบคลุมตั้งแต่ การให้คำปรึกษาด้านการเปลี่ยนผ่านองค์กรสู่การเป็น Data-Driven Organization (“Data Transformation Consulting”) รวมถึงโซลูชันด้านพาณิชย์อิเล็กทรอนิกส์, ประสบการณ์ลูกค้า (Experience/CRM), สื่อแบบผสมผสาน (Hybrid Media) ที่เชื่อมโยงตั้งแต่สื่อดั้งเดิมไปจนถึงดิจิทัล พร้อมด้วยฐานข้อมูลผู้ใช้ “Single-Source Data Solutions” ซึ่งครอบคลุมหลายหมวดหมู่และแบรนด์มากมาย

3.3 ตำแหน่งงานและลักษณะงานที่ได้รับมอบหมาย

ตำแหน่งงานที่ผู้จัดทำได้รับหมาย คือ Data Analyst โดยระหว่างปฏิบัติงานสหกิจศึกษาที่สถานประกอบการได้รับมอบหมายให้ปฏิบัติงาน ดังต่อไปนี้

สัปดาห์ที่ 1-2

- ศึกษาหน้าที่และความรับผิดชอบของตำแหน่ง Data Analyst ผ่านการเรียนรู้คอร์ส Meta Data Analyst Professional Certificate บนแพลตฟอร์ม Coursera ซึ่งเป็นคอร์สที่พัฒนาโดย Meta โดยมีจุดประสงค์เพื่อเสริมสร้างความรู้ความเข้าใจที่จำเป็นสำหรับการทำงานจริงในสายอาชีพ Data Analytics นอกจากนี้ ผู้จัดทำยังได้รับมอบหมายให้จัดทำสไลด์สรุปสิ่งที่ได้เรียนรู้ในแต่ละสัปดาห์ เพื่อใช้ในการนำเสนอแก่ผู้ดูแลการปฏิบัติงาน



รูปที่ 3.2 คอร์ส Meta Data Analyst Professional Certificate

เนื้อหาที่ศึกษาครอบคลุมถึง:

- ความสำคัญและบทบาทของ Data Analyst ในองค์กร
- ขั้นตอนของการทำงานวิเคราะห์ข้อมูลตามกระบวนการมาตรฐาน (OSEMN) เช่น การรวบรวม (Obtaining), การทำความสะอาด (Scrubbing), การสำรวจข้อมูล (Exploring), การสร้างแบบจำลอง (Modeling) และการตีความผลลัพธ์ (Interpreting)
- การใช้เครื่องมือพื้นฐานสำหรับการวิเคราะห์ข้อมูล เช่น สเปรดชีต (Spreadsheet), SQL และแนวทางการใช้เครื่องมือ AI เพื่อสนับสนุนการวิเคราะห์
- การฝึกทักษะด้านการคิดเชิงวิเคราะห์ (Analytical Thinking) และการสื่อสารผลการวิเคราะห์ให้นำไปใช้ในการตัดสินใจเชิงธุรกิจ

การเรียนรู้เนื้อหาดังกล่าวมีส่วนช่วยให้ผู้จัดทำเข้าใจบทบาทของ Data Analyst ได้ชัดเจนยิ่งขึ้น และเตรียมความพร้อมในการนำทักษะไปประยุกต์ใช้กับงานจริงระหว่างการปฏิบัติงานสหกิจศึกษา

- ได้รับมอบหมายให้ปฏิบัติงานด้าน Keyword Labeling สำหรับข้อมูลการค้นหาที่เกี่ยวข้องกับสุนัขและแมวจาก Google

- ได้รับมอบหมายให้ดำเนินการ รวบรวมรายชื่อผู้เข้าร่วมงาน Data First Campus เพื่อนำไปใช้ในการจัดทำ Certificate ให้กับผู้เข้าร่วมงาน

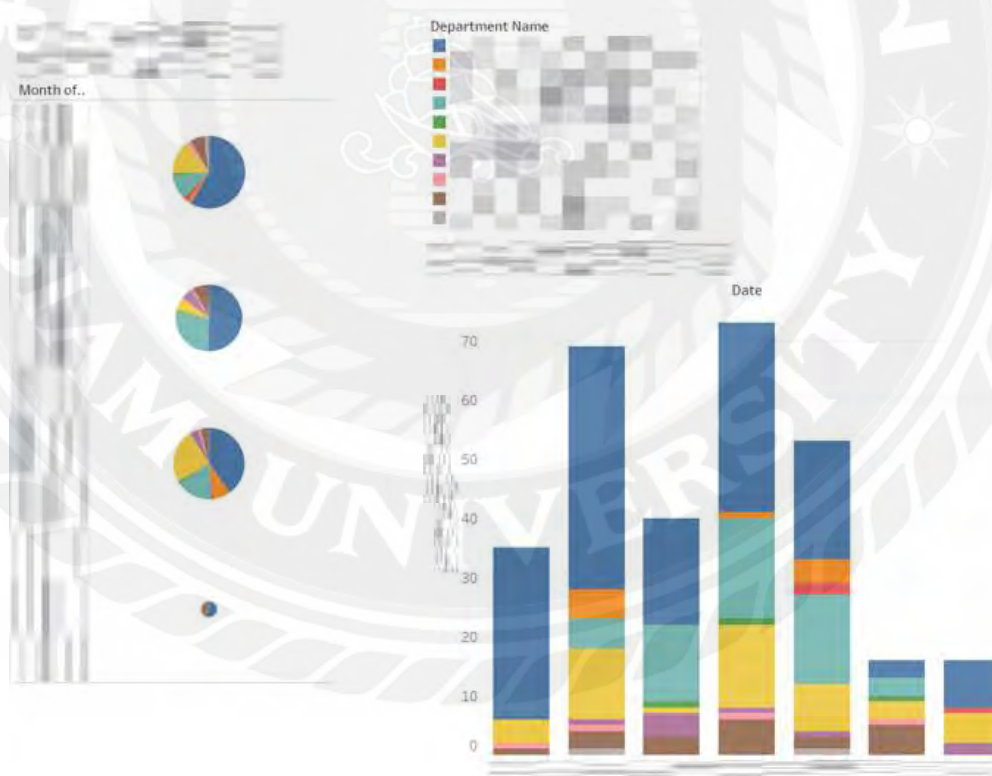
[illegible]

รูปที่ 3.4 ตัวอย่างข้อมูล การรวบรวมรายชื่อผู้เข้าร่วมงาน Data First Campus

จากรูปที่ 3.4 แสดงผลการรวบรวมรายชื่อผู้เข้าร่วมงาน Data First Campus โดยมุ่งวิเคราะห์จำนวนครั้งในการเข้าร่วมของผู้เข้าร่วมแต่ละคน เพื่อนำข้อมูลไปใช้ในการจัดทำและมอบใบ Certificate แก่ผู้เข้าร่วมงาน

สไลด์ที่ 3-4

- ศึกษาบทความ Use Case Personalization จากเว็บไซต์ <https://www.warc.com> โดยเลือกหัวข้อที่น่าสนใจมาทำการวิเคราะห์ เพื่อทำความเข้าใจแนวทางการประยุกต์ใช้ Personalization ในเชิงธุรกิจ
- ศึกษาวิธีการใช้งานโปรแกรม Tableau Public เพื่อพัฒนาทักษะการนำเสนอข้อมูลให้อยู่ในรูปแบบที่เข้าใจง่ายและสื่อความหมายได้ชัดเจนยิ่งขึ้น พร้อมทั้งได้รับมอบหมายให้นำข้อมูลภายใน (Confidential) ของใบเบิกงบประมาณบริษัท มาสร้างเป็น Visualization เพื่อการฝึกปฏิบัติ

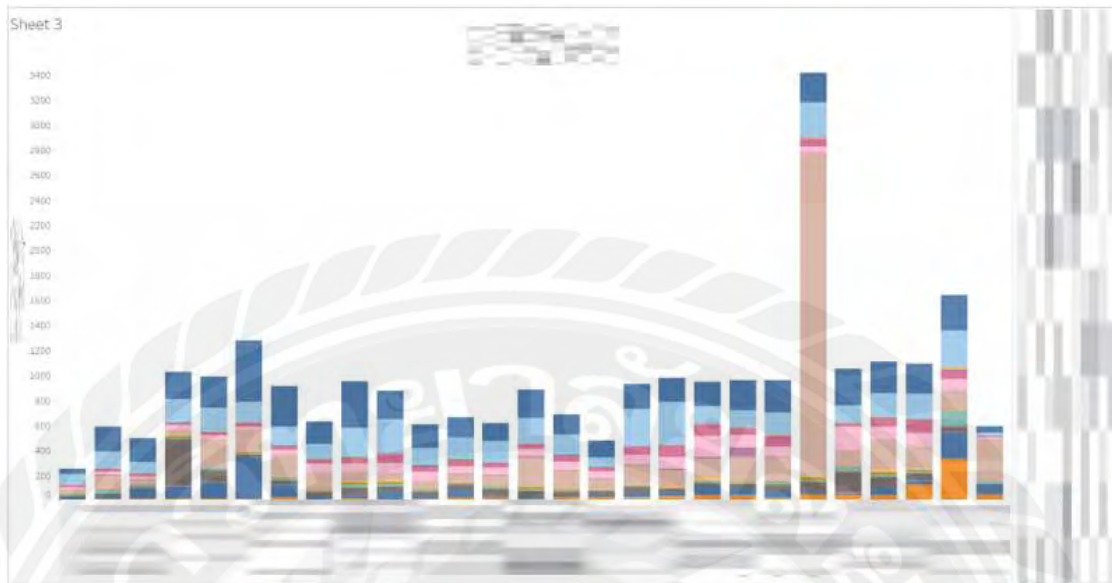


รูปที่ 3.5 ตัวอย่างการนำเสนอข้อมูล

- ศึกษาวิธีการใช้งานโปรแกรม Tableau Prep เพื่อเตรียมและจัดการข้อมูลให้มีความพร้อมสำหรับการนำเสนอได้อย่างถูกต้องและรวดเร็วยิ่งขึ้น
- ได้รับมอบหมายให้ปฏิบัติงานด้าน Keyword Labeling โดยใช้ข้อมูล Social Listening ที่ดึงมาจากแพลตฟอร์มโซเชียลมีเดีย ซึ่งมีการพูดถึงและแสดงความคิดเห็นเกี่ยวกับผลิตภัณฑ์ดูแลช่องปาก



- ได้รับมอบหมายให้ใช้โปรแกรม Tableau Public และ Tableau Prep ในการจัดการและแสดงผลข้อมูลที่ผ่านมาการทำ Keyword Labeling แล้ว เพื่อตรวจสอบความถูกต้องและความสมบูรณ์ของข้อมูลที่ถูกจัดหมวดหมู่



รูปที่ 3.7 ตัวอย่างข้อมูล Keyword Labeling ที่ถูกนำมาแสดงผลผ่าน Tableau

สัปดาห์ที่ 9-10

- ได้รับมอบหมายให้ทบทวนการเขียนโปรแกรมภาษา Python โดยเฉพาะการใช้งานไลบรารี Pandas ผ่านแพลตฟอร์ม Google Colab

```

# Import pandas as pd
import pandas as pd

# Load the dataset
data = pd.read_csv('data.csv')

# Display the first few rows of the dataset
data.head()

```

	age	job	marital	education	default	balance	housing	loan	contact	day	month	duration	campaign	previous	poutcome	y	
0	51	manager	married	tertiary	no	2142	yes	no	telephone	5	may	261	1	1	0	success	no
1	44	technician	single	secondary	no	29	yes	no	telephone	5	may	151	1	1	0	success	no
2	33	technician	married	secondary	no	2	yes	yes	telephone	5	may	36	1	1	0	success	no
3	47	blue-collar	married	secondary	no	1566	yes	no	telephone	5	may	52	1	1	0	success	no
4	33	unemployed	single	university	no	1	no	no	telephone	5	may	168	1	1	0	success	no

รูปที่ 3.8 ตัวอย่างทบทวนการเขียนโปรแกรมภาษา Python

- ศึกษา การวิเคราะห์ตะกร้าสินค้า (Market Basket Analysis: MBA)

สัปดาห์ที่ 10-11

- ได้รับมอบหมายให้ศึกษาและวิเคราะห์ข้อมูลพฤติกรรมความสนใจของลูกค้า โดยประยุกต์ใช้เทคนิค Market Basket Analysis (MBA) เพื่อวิเคราะห์เปรียบเทียบรูปแบบ ความสัมพันธ์ และพฤติกรรมของผู้บริโภคในมิติต่างๆ

สัปดาห์ที่ 12-16

- ดำเนินการจัดทำโครงการในหัวข้อ “การวิเคราะห์พฤติกรรมของผู้ใช้งานจากร้านค้าออนไลน์ ด้วยการทำ Market Basket Analysis”
- ทำการวิเคราะห์ข้อมูลด้วยภาษา Python บนแพลตฟอร์ม Google Colab เพื่อค้นหารูปแบบและความสัมพันธ์ระหว่างแบรนด์ที่ลูกค้ามักให้ความสนใจร่วมกัน รวมถึงความสัมพันธ์ระหว่างแบรนด์กับช่องทางการซื้อ
- ทำการสรุปผลการวิเคราะห์ออกมาในรูปแบบการนำเสนอข้อมูลด้วยภาพ (Data Visualization) โดยใช้โปรแกรม Tableau Public

3.4 ชื่อและตำแหน่งงานของพนักงานที่ปรึกษา

ชื่อ – นามสกุล : คุณอนันต์ ตีระบูรณะพงษ์
 ตำแหน่ง : Vice Executive Data & Innovation Director
 เบอร์ติดต่อ : 086-666-7070
 อีเมล : anan@datafirst.co.th

3.5 ระยะเวลาที่ปฏิบัติงาน

ได้เข้าปฏิบัติงานสหกิจศึกษา ณ บริษัท ดาต้า เฟิร์ส จำกัด ตั้งแต่วันที่ 19 พฤษภาคม พ.ศ. 2568 ถึงวันที่ 29 สิงหาคม พ.ศ. 2568

บทที่ 4

ผลการปฏิบัติงานโครงการ

4.1 รายละเอียดของโครงการ

โครงการนี้มุ่งเน้นการนำข้อมูลพฤติกรรมการใช้งาน (User Behavior Data) ของลูกค้าบนแพลตฟอร์มขายของออนไลน์มาวิเคราะห์และประมวลผล เพื่อค้นหารูปแบบการใช้งาน ความสนใจ และแนวโน้มพฤติกรรมของลูกค้าอย่างเป็นระบบ

เครื่องมือหลักที่ใช้ ได้แก่ Google Colab สำหรับการเขียนและประมวลผลชุดคำสั่งวิเคราะห์ข้อมูล ครอบคลุมตั้งแต่ขั้นตอนการเตรียมและทำความสะอาดข้อมูล การทำความเข้าใจเชิงลึกตลอดจนการวิเคราะห์ความสัมพันธ์ระหว่างสินค้าและโปรโมชั่นด้วยเทคนิค Market Basket Analysis (MBA) นอกจากนี้ยังใช้ Tableau Public เพื่อสร้างแผนภาพข้อมูล (Data Visualization) และจัดทำ Dashboard ในรูปแบบที่เข้าใจง่าย สนับสนุนการวางแผนการตลาดอย่างมีประสิทธิภาพ

4.2 ขั้นตอนในการวิเคราะห์ข้อมูล

4.2.1 ศึกษา Market Basket Analysis (MBA)

คณะผู้จัดทำได้ศึกษาหลักการและแนวคิดของ Market Basket Analysis (MBA) ซึ่งเป็นเทคนิคที่ใช้วิเคราะห์ความสัมพันธ์ของสินค้าและบริการ โดยอาศัยข้อมูลการเลือกซื้อหรือการมีปฏิสัมพันธ์ของลูกค้า เพื่อนำไปค้นหารูปแบบ (Patterns) หรือกฎความสัมพันธ์ (Association Rules) ระหว่างแบรนด์สินค้าและช่องทางการขายที่มักปรากฏร่วมกัน การศึกษาดังกล่าวครอบคลุมทั้งในด้านทฤษฎี เช่น หลักการของ Association Rule Mining การคำนวณค่า Support, Confidence และ Lift ตลอดจนตัวอย่างงานวิจัยที่เกี่ยวข้อง เพื่อสร้างความเข้าใจที่ถูกต้องและสามารถนำไปปรับใช้กับข้อมูลจริงของบริษัทได้อย่างเหมาะสม

4.2.2 ศึกษาบทความ Visualizing Association Rules in Hierarchical Groups” (Hahsler & Chelluboina, 2011)

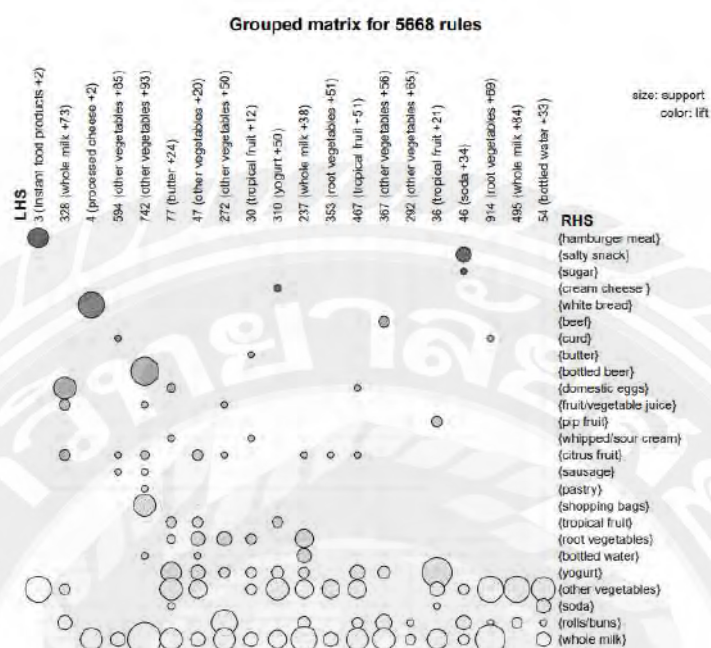
การทำ Association Rule Mining มักให้ผลลัพธ์เป็นกฎจำนวนมากไม่ถ้วน ทำให้การตีความและเลือกเฉพาะกฎที่น่าสนใจเป็นเรื่องยาก Hahsler และ Chelluboina (2011) จึงได้เสนอแนวทาง Grouped Matrix-based Visualization เพื่อแก้ปัญหานี้

โดยมี แนวคิดหลักดังนี้:

- ใช้ Matrix-based Visualization วาง Antecedent (LHS) บนแกน x และ Consequent (RHS) บนแกน y พร้อมแสดงค่า Interest Measure (เช่น Lift) ใน Cell
- ปัญหาของ Matrix แบบเดิมคือ ขนาดใหญ่และ Label อ่านยากเมื่อกลุ่มมีจำนวนมาก
- เพิ่มขั้นตอน Clustering ด้วย k-means เพื่อจัดกลุ่ม Antecedents ที่มีความสัมพันธ์คล้ายกัน
- ผลลัพธ์ถูกนำเสนอในรูปแบบ Balloon Plot
 - o ขนาดของบอลูนแสดงค่า Support
 - o สีของบอลูนแสดงค่า Lift (เข้ม = ความสัมพันธ์น่าสนใจมากขึ้น)

จากการศึกษาบทความดังกล่าว ผู้จัดทำได้เรียนรู้ว่า การทำ Visualization ที่มีประสิทธิภาพสามารถช่วยลดภาระในการวิเคราะห์ข้อมูล และทำให้มองเห็นความสัมพันธ์ที่สำคัญได้ชัดเจนยิ่งขึ้น สำหรับโครงการสหกิจที่เกี่ยวข้องกับการทำ Market Basket Analysis และการวิเคราะห์ความสัมพันธ์ของข้อมูลลูกค้า การประยุกต์ใช้แนวคิด Grouped Matrix-based Visualization ถือเป็นแนวทางที่มีคุณค่า เนื่องจากช่วยให้สามารถโฟกัสที่กฎที่มีค่า Lift สูงและมีความน่าสนใจ โดยไม่จำเป็นต้องเสียเวลาในการตรวจสอบกฎจำนวนมากที่ละรายการ

นอกจากนี้ การจัดกลุ่ม antecedent ยังเปิดโอกาสให้ค้นพบรูปแบบเชิงลึก เช่น การซื้อสินค้าทดแทน หรือสินค้าที่มักปรากฏร่วมกันในธุรกรรม ซึ่งสะท้อนถึงความสัมพันธ์ทางสถิติที่มีความหมายต่อการวิเคราะห์ ดังนั้น งานวิจัยของ Hahsler และ Chelluboina (2011) จึงถือเป็นแนวทางสำคัญที่สามารถต่อยอดการประยุกต์ใช้ Association Rules ให้สอดคล้องกับการใช้งานจริงทั้งในเชิงธุรกิจและในโครงการวิเคราะห์ข้อมูล



รูปที่ 4.1 ตัวอย่าง Grouped Matrix-based Visualization¹

4.2.3 กำหนดเป้าหมายในการวิเคราะห์ข้อมูล (Research Understanding)

เป้าหมายหลัก คือทำความเข้าใจพฤติกรรมลูกค้าบนแพลตฟอร์มขายของออนไลน์ โดยวิเคราะห์ความสัมพันธ์ระหว่างแบรนด์สินค้าและช่องทางการขาย ผ่านเทคนิค Market Basket Analysis (MBA) และการทำ Visualization เพื่อให้ได้ข้อมูลเชิงลึกที่สามารถนำไปประยุกต์ใช้จริงได้

4.2.4 ทำความเข้าใจข้อมูล (Data Understanding)

ผู้จัดทำได้ทำการศึกษาข้อมูลที่ได้รับ และตีความว่าเป็น ข้อมูลพฤติกรรมการใช้งานของลูกค้าบนแพลตฟอร์มขายของออนไลน์ โดยข้อมูลดังกล่าวถูกเก็บรวบรวมจากการที่ลูกค้ามีการกดเข้าไปดูหรือมีปฏิสัมพันธ์ (interaction) กับฟีเจอร์ต่าง ๆ ภายในแอปพลิเคชัน ไม่ว่าจะเป็นการเปิดดูรายละเอียดโปรโมชั่น หรือการเข้าร่วมกิจกรรมที่จัดขึ้น

¹ Hahsler, M., & Chelluboina, S. (2011). *Visualizing association rules in hierarchical groups*. 42nd Symposium on the Interface: Statistical, Machine Learning, and Visualization Algorithms (Interface 2011).

ทุกครั้งที่ผู้ใช้งานมีการกระทำใด ๆ ระบบจะทำการบันทึกแท็ก (tags) กำกับไว้ เพื่อระบุว่าลูกค้ารายนั้นเคยมีพฤติกรรมดังกล่าวมาก่อน ข้อมูลที่ได้มีจำนวนทั้งหมด 287,088 แถว และ 5 คอลัมน์ ซึ่งเพียงพอสำหรับการนำมาวิเคราะห์เพื่อค้นหารูปแบบและแนวโน้มพฤติกรรมของผู้ใช้งาน

อย่างไรก็ตาม ผู้จัดทำยังจำเป็นต้องทำความเข้าใจที่มาของแต่ละแท็ก (tags) ว่าเกิดจากการปฏิสัมพันธ์ประเภทใดบ้าง แต่ในส่วนนี้ไม่สามารถเปิดเผยรายละเอียดได้เนื่องจากข้อจำกัดด้านความลับของข้อมูล

4.2.5 การเตรียมข้อมูล (Data Preparation)

จากข้อมูลที่ได้รับมานั้นยังไม่อยู่ในรูปแบบที่พร้อมต่อการวิเคราะห์ ผู้จัดทำจึงได้ทำการปรับแต่งและจัดเตรียมข้อมูลให้อยู่ในลักษณะที่เหมาะสม โดยใช้ภาษา Python เป็นเครื่องมือหลักในการประมวลผล ในขั้นตอนนี้ได้ทำการคัดเลือกเฉพาะข้อมูลที่เป็นสำหรับการวิเคราะห์ ได้แก่ uid และ tags

ข้อมูลในคอลัมน์ tags ถูกจัดเก็บในลักษณะข้อความที่รวมแท็กหลายรายการไว้ในคอลัมน์เดียว โดยคั่นด้วยเครื่องหมาย “,” เพื่อให้สามารถนำไปประยุกต์ใช้กับกระบวนการ Market Basket Analysis (MBA) ได้ จึงมีความจำเป็นต้องปรับให้อยู่ในรูปแบบ one-hot encoding



uid	tags	...
1	1,2,3,4,5	...
2	6,7,8,9,10	...
3	11,12,13,14,15	...
4	16,17,18,19,20	...
5	21,22,23,24,25	...

รูปที่ 4.2 ตัวอย่างข้อมูลดิบ (Raw Data)

```
df['tags'] = df['tags'].str.replace(' ', '', regex=False)
df_tags_encoded = df['tags'].str.get_dummies(sep=',')
df_tags_encoded
```

รูปที่ 4.3 ตัวอย่างชุดคำสั่ง Python ที่ใช้ทำ One-hot encoding



รูปที่ 4.4 ตัวอย่างข้อมูลที่จัดเตรียมให้อยู่ในรูปแบบที่เหมาะสมบางส่วน

จากรูปที่ 4.2 และ 4.3 แสดงถึงการปรับข้อมูลดังกล่าว ส่งผลให้คอลัมน์ tags ซึ่งเดิมถูกรวมอยู่ในคอลัมน์เดียว ถูกแยกออกมาเป็น 192 คอลัมน์ (แท็ก) ในรูปแบบ one-hot encoding โดยแต่ละคอลัมน์จะแสดงการมีอยู่ (ค่า 1) หรือไม่มีอยู่ (ค่า 0) ของแท็กนั้น ๆ สำหรับผู้ใช้งานแต่ละราย

4.2.6 การวิเคราะห์ความสัมพันธ์ของข้อมูล (Association Rule Mining: Apriori Algorithm)

หลังจากที่ข้อมูลได้ถูกจัดให้อยู่ในรูปแบบ one-hot encoding แล้ว ผู้จัดทำได้นำข้อมูลดังกล่าวมาวิเคราะห์เพื่อค้นหากฎความสัมพันธ์ (Association Rules) โดยใช้อัลกอริทึม Apriori ซึ่งเป็นหนึ่งในเทคนิคพื้นฐานของการทำ Market Basket Analysis (MBA) ในกระบวนการนี้ ได้ใช้ไลบรารี mlxtend.frequent_patterns ของภาษา Python โดยมีขั้นตอนสำคัญดังนี้

1. ค้นหาชุดรายการที่พบบ่อย (Frequent Itemsets):

ใช้ฟังก์ชัน apriori() เพื่อหาชุดแท็ก (itemsets) ที่ปรากฏร่วมกันบ่อย โดยกำหนดค่า min_support = 0.004 เพื่อกรองเฉพาะชุดข้อมูลที่มีอัตราการเกิดซ้ำ (support) ไม่น้อยกว่าร้อยละ 0.4

2. สร้างกฎความสัมพันธ์ (Association Rules):

ใช้ฟังก์ชัน `association_rules()` เพื่อสร้างกฎความสัมพันธ์จาก frequent itemsets โดยเลือกใช้ `metric = "lift"` เพื่อวัดความน่าสนใจของความสัมพันธ์ระหว่าง antecedent และ consequent

3. กรองกฎ (Rule Filtering):

เลือกเฉพาะกฎที่มี antecedent และ consequent อย่างละ 1 รายการ (single $\rightarrow$ single) เพื่อให้ง่ายต่อการตีความ โดยใช้เงื่อนไข `len(x) == 1` ทั้งฝั่ง antecedents และ consequents

```
from mlxtend.frequent_patterns import apriori
from mlxtend.frequent_patterns import association_rules

frequent_itemsets = apriori(df_tags_encoded, min_support=0.004, use_colnames=True)

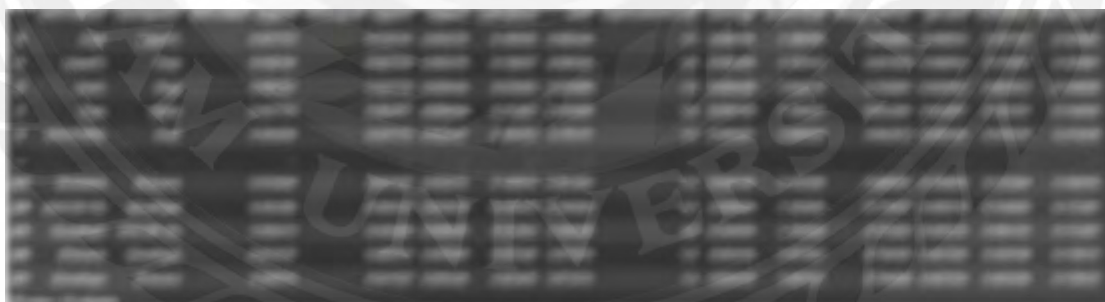
#frequent_itemsets_final=frequent_itemsets.sort_values("support",axis = 0, ascending = False).head(50)

rules = association_rules(frequent_itemsets, metric="lift")

#single
rules = rules[
    (rules['antecedents'].apply(lambda x: len(x) == 1)) &
    (rules['consequents'].apply(lambda x: len(x) == 1))
]

rules
```

รูปที่ 4.5 ตัวอย่างชุดคำสั่ง Python สำหรับสร้างกฎความสัมพันธ์ด้วยอัลกอริทึม Apriori



รูปที่ 4.6 ผลลัพธ์กฎความสัมพันธ์ (Association Rules) ที่ได้จากอัลกอริทึม Apriori

จากรูปที่ 4.6 แสดงผลลัพธ์ของการสร้างกฎความสัมพันธ์ โดยในตารางประกอบด้วยตัวชี้วัดสำคัญ ได้แก่ Antecedents, Consequents, Support, Confidence และ Lift

4.2.7 การนำเสนอแผนภาพของข้อมูล (Data Visualization)

หลังจากการสร้างกฎความสัมพันธ์ (Association Rules) เสร็จสิ้น ผลลัพธ์ที่ได้ถูกนำมาสร้างเป็นแผนภาพ (Data Visualization) โดยใช้เทคนิค Grouped Matrix-based Visualization ที่ศึกษาและอ้างอิงจากงานวิจัย เพื่อช่วยให้สามารถมองเห็นภาพรวมของพฤติกรรมการใช้งานของลูกค้าได้อย่างชัดเจนยิ่งขึ้น แผนภาพดังกล่าวมีบทบาทสำคัญในการสะท้อนประเด็นและความสนใจหลักของลูกค้า ซึ่งถือเป็นข้อมูลเชิงลึกที่สามารถนำไปต่อยอดในการวิเคราะห์เชิงลึก รวมถึงการกำหนดกลยุทธ์ทางธุรกิจอย่างมีประสิทธิภาพ การนำเสนอผลลัพธ์ดังกล่าวแบ่งออกเป็น 2 มิติหลัก ได้แก่ ความสัมพันธ์ระหว่างแบรนด์ด้วยกัน และความสัมพันธ์ระหว่างแบรนด์กับช่องทางการขาย

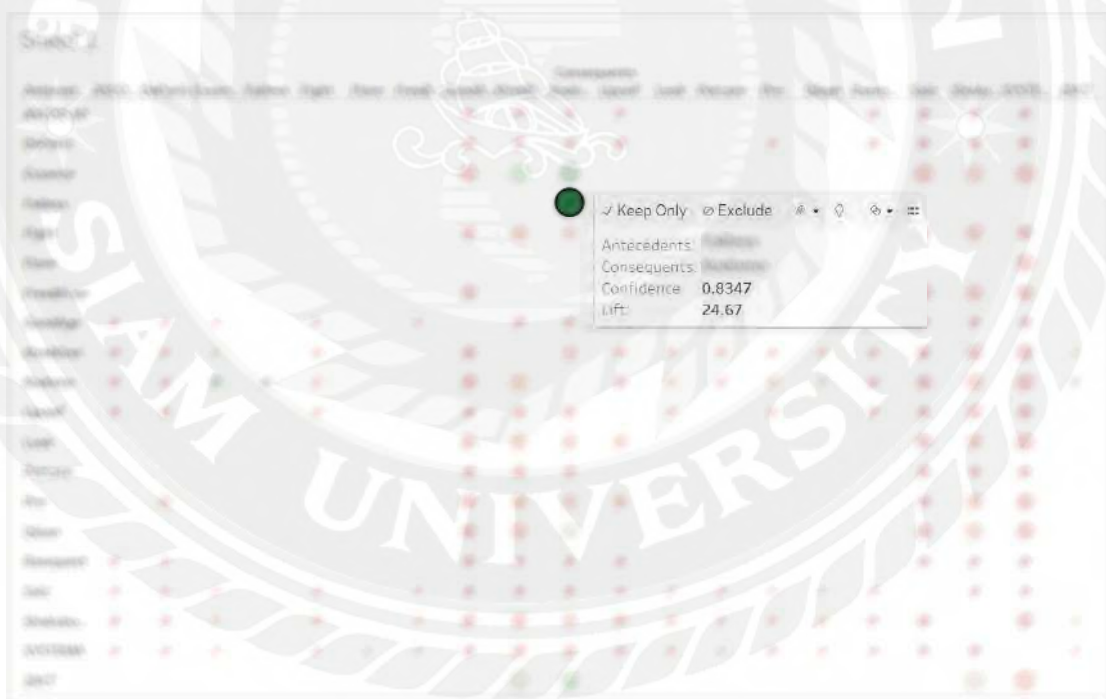
4.2.6.1 ความสัมพันธ์ระหว่างแบรนด์ด้วยกัน

ผู้จัดทำได้คัดเลือกเฉพาะ แท็กที่ระบุชื่อแบรนด์ เพื่อนำมาวิเคราะห์ความสัมพันธ์ระหว่างแบรนด์ต่าง ๆ โดยมุ่งศึกษาว่าลูกค้ามีแนวโน้มที่จะให้ความสนใจในแบรนด์ใดบ้างพร้อมกัน ผลการวิเคราะห์ดังกล่าวสะท้อนให้เห็นถึงรูปแบบการเชื่อมโยงของความสนใจระหว่างแบรนด์ ซึ่งสามารถนำไปใช้ในการแบ่งกลุ่มลูกค้าได้อย่างเหมาะสม

การเข้าใจความสัมพันธ์นี้มีประโยชน์ต่อการวางกลยุทธ์ทางการตลาด เช่น การออกแบบแคมเปญร่วม (Co-branding) หรือการจัดโปรโมชั่นที่นำเสนอสินค้าของสองแบรนด์ควบคู่กัน เพื่อเพิ่มโอกาสในการดึงดูดและตอบสนองต่อความต้องการของกลุ่มลูกค้าที่มีความสนใจร่วมกัน



รูปที่ 4.7 ผลลัพธ์การใช้เทคนิค Grouped Matrix-based Visualization กับ ข้อมูลแบรนด์



รูปที่ 4.8 ความสัมพันธ์ระหว่างแบรนด์ผลิตภัณฑ์แลมร่ว่ง H และแบรนด์ของใช้เด็ก K

จาก รูปที่ 4.8 พบความสัมพันธ์ที่น่าสนใจระหว่าง ปัจจัยที่มาก่อน (Antecedents) คือ แบรินด์ H ซึ่งจำหน่ายผลิตภัณฑ์ประเภทแชมพูลดอาการผดผื่น และ ปัจจัยผลลัพธ์ (Consequents) คือ แบรินด์ K ซึ่งเป็นผู้ผลิตสินค้าของใช้เด็กแบบครบวงจร เช่น น้ำยาซักผ้าเด็ก แป้งเด็ก แชมพูเด็ก สบู่ อุปกรณ์ดูแลช่องปาก รวมถึงผลิตภัณฑ์ล้างขูดนม ความสัมพันธ์นี้ดูเหมือนจะเป็นสิ่งที่ไม่น่าจะเกี่ยวข้องกันโดยตรง เนื่องจากแบรินด์ K เองก็มีผลิตภัณฑ์แชมพูสำหรับเด็กอยู่แล้ว อย่างไรก็ตาม การที่ลูกค้าเลือกซื้อแชมพูลดผดผื่นจากแบรินด์ H ควบคู่กับการเลือกซื้อสินค้าของใช้เด็กจากแบรินด์ K สะท้อนให้เห็นถึงรูปแบบพฤติกรรมที่เฉพาะเจาะจง ซึ่งมีความเป็นไปได้สูงว่าจะเป็นกลุ่ม คุณแม่่มือใหม่ ที่เพิ่งคลอดบุตร เนื่องจากหลังคลอด ฮอร์โมนบางชนิดในร่างกายจะลดลง ส่งผลให้เกิดปัญหาผดผื่น จึงมีความต้องการใช้แชมพูลดผดผื่นควบคู่กับการซื้อของใช้จำเป็นสำหรับทารก จากการคำนวณกฎความสัมพันธ์ พบว่า ความน่าจะเป็นแบบมีเงื่อนไข (Confidence) มีค่าสูงถึง 0.8347 และค่าตัวชี้วัดความแข็งแรงของความสัมพันธ์ (Lift) เท่ากับ 24.67 ซึ่งสูงกว่าค่าเฉลี่ยอย่างมาก แสดงให้เห็นว่าความสัมพันธ์ดังกล่าวไม่ได้เกิดขึ้นเพียงโดยบังเอิญ แต่สะท้อนพฤติกรรมจริงของลูกค้ากลุ่มนี้ อย่างไรก็ตาม การตีความยังต้องอาศัยการพิจารณาในมิติอื่นร่วมด้วย เนื่องจากอาจมีบางกรณีที่ผู้ซื้อไม่ใช่คุณแม่่มือใหม่โดยตรง เช่น ผู้ที่มีปัญหาผดผื่นและซื้อของใช้เด็กเป็นของฝากให้หลาน เป็นต้น อีกหนึ่งประเด็นที่น่าสนใจคือ แบรินด์ H ไม่ปรากฏการจับคู่กับแบรินด์อื่น ๆ เลยนอกจากแบรินด์ K ซึ่งบ่งชี้ว่ากลุ่มลูกค้าของแบรินด์ H ส่วนใหญ่มีพฤติกรรมการเลือกซื้อที่เฉพาะเจาะจง และความสัมพันธ์กับแบรินด์ K น่าจะเป็นความเชื่อมโยงหลักเพียงช่องทางเดียว



รูปที่ 4.9 ความสัมพันธ์ระหว่างแบรนด์ผลิตภัณฑ์สำหรับสัตว์เลี้ยง P และแบรนด์สินค้าทำความสะอาดมือ F

จาก รูปที่ 4.9 พบความสัมพันธ์ที่น่าสนใจระหว่าง ปัจจัยที่มาก่อน (Antecedents) คือ แบรินด์ P ซึ่งจำหน่ายผลิตภัณฑ์สำหรับสัตว์เลี้ยง เช่น แชมพูอาบน้ำแห้ง กระดาษทิชชูเปียกสำหรับสัตว์ รวมถึงยาสีฟันและแปรงสีฟันสำหรับสัตว์ และ ปัจจัยผลลัพธ์ (Consequents) คือ แบรินด์ F ผู้ผลิตสินค้าทำความสะอาดมือที่อ่อนโยนต่อผิว อาทิ โฟมล้างมือ เจลล้างมือ และสเปรย์ล้างมือ การที่แบรินด์ P ปรากฏในฐานะ Antecedent สื่อได้ว่ากลุ่มลูกค้ากลุ่มนี้คือ ผู้เลี้ยงสัตว์ ซึ่งโดยพฤติกรรมแล้ว มักจะต้องสัมผัสสัตว์เลี้ยงและสิ่งแวดล้อมที่เกี่ยวข้องอยู่เสมอ ความสนใจที่เชื่อมโยงไปยังแบรินด์ F สามารถอธิบายได้ว่า กลุ่มผู้เลี้ยงสัตว์มีความจำเป็นในการล้างมือบ่อยกว่าคนทั่วไป เพื่อรักษาความสะอาดและลดความกังวลของคนรอบข้างที่อาจไม่ชอบหรือแพ้อื่นจากสัตว์เลี้ยง ดังนั้น การเลือกซื้อผลิตภัณฑ์สำหรับสัตว์เลี้ยงจึงสัมพันธ์กับการเลือกซื้อผลิตภัณฑ์ทำความสะอาดมือไปพร้อมกัน ผลการวิเคราะห์กฎความสัมพันธ์ พบว่า ค่าความน่าจะเป็นแบบมีเงื่อนไข (Confidence) อยู่ที่ 0.3133 แม้จะไม่สูงนัก แต่ก็ถือว่าสูงกว่าความสัมพันธ์กับแบรินด์อื่น ๆ ที่ปรากฏ และยังสะท้อนให้เห็นว่าลูกค้าที่ซื้อสินค้าของแบรินด์ P มีโอกาสซื้อสินค้าของแบรินด์ F ร่วมด้วยมากที่สุด อีกทั้งค่าตัวชี้วัดความแข็งแกร่งของความสัมพันธ์ (Lift) อยู่ที่ 6.57 ซึ่งสูงกว่าค่าเฉลี่ยอย่างมีนัยสำคัญ แสดงให้เห็นว่า

ความสัมพันธ์นี้ไม่ได้เกิดขึ้นโดยบังเอิญ แต่มีพื้นฐานจากพฤติกรรมจริงของลูกค้ากลุ่มผู้เลี้ยงสัตว์ที่มีความต้องการดูแลสุขภาพอนามัยของตนเองควบคู่ไปด้วย



รูปที่ 4.10 ความสัมพันธ์ระหว่างแบรนด์จำหน่ายยาสีฟันเฉพาะทาง A และแบรนด์ผลิตภัณฑ์ดูแลช่องปากครบวงจร B

จาก รูปที่ 4.10 พบความสัมพันธ์ที่น่าสนใจระหว่าง ปัจจัยที่มาก่อน (Antecedents) คือ แบนด์ A ซึ่งจำหน่ายยาสีฟันเฉพาะทางสำหรับลดคราบเหลืองจากชา กาแฟ และบุหรี่ และ ปัจจัยผลลัพธ์ (Consequents) คือ แบนด์ B ผู้ผลิตผลิตภัณฑ์ดูแลช่องปากครบวงจร เช่น แปรงสีฟัน ยาสีฟัน น้ำยาล้างปาก และไหมขัดฟัน การที่แบนด์ A ปรากฏในฐานะ Antecedent สะท้อนว่าลูกค้าที่เลือกสนใจสินค้านี้มักเป็นผู้ที่ประสบปัญหาเฉพาะ เช่น คราบเหลืองจากชา กาแฟ หรือการสูบบุหรี่ ซึ่งกลุ่มลูกค้าดังกล่าวมีแนวโน้มที่จะให้ความสำคัญกับสุขภาพช่องปากโดยรวม และต้องการการดูแลที่ครอบคลุมมากกว่าการใช้ยาสีฟันเฉพาะทางเพียงอย่างเดียว ความสนใจที่เชื่อมโยงไปยังแบนด์ B จึงสามารถอธิบายได้ว่า ลูกค้ากลุ่มนี้มักมองหาผลิตภัณฑ์เสริมสำหรับการดูแลสุขภาพช่องปากแบบครบวงจร ซึ่งสอดคล้องกับภาพลักษณ์ของแบนด์ B ในฐานะแบนด์ผู้เชี่ยวชาญด้านการดูแลช่องปาก ผลการวิเคราะห์กฎความสัมพันธ์ พบว่า ค่าความน่าจะเป็นแบบมีเงื่อนไข (Confidence) อยู่ที่ 0.7339 หมายความว่าลูกค้าที่ซื้อสินค้าของแบนด์ A มีโอกาสมากถึง 73.39% ที่จะซื้อผลิตภัณฑ์ของแบนด์ B ร่วมด้วย อีกทั้งยังมี ค่าตัวชี้วัดความแข็งแกร่งของความสัมพันธ์ (Lift) เท่ากับ 8.69 ซึ่งสูงกว่าค่าเฉลี่ย

อย่างมาก แสดงให้เห็นว่าความสัมพันธ์นี้ไม่ได้เกิดขึ้นโดยบังเอิญ แต่เป็นผลจากพฤติกรรมจริงของกลุ่มผู้บริโภคที่ต้องการการดูแลสุขภาพช่องปากอย่างครบถ้วน กล่าวได้ว่า ความสัมพันธ์นี้สะท้อนถึงพฤติกรรมการบริโภคที่เชื่อมโยงกันระหว่างสินค้าเฉพาะทาง (Specialized Product) และสินค้าในหมวดดูแลสุขภาพช่องปากครบวงจร (Comprehensive Oral Care) ซึ่งสามารถนำไปใช้กำหนดกลยุทธ์ทางการตลาดได้ เช่น การจัดทำ โปรโมชันร่วม (Bundle Promotion) หรือ Cross-selling ระหว่างยาสีฟันของแบรนด์ A และผลิตภัณฑ์ของแบรนด์ B เพื่อเข้าถึงกลุ่มผู้บริโภคที่มีความกังวลเรื่องสุขภาพช่องปากและต้องการการดูแลสุขภาพที่ครอบคลุม

การเปรียบเทียบกับแบรนด์ C

ถึงแม้ว่า แบรนด์ C จะมีผลิตภัณฑ์ที่คล้ายกับแบรนด์ B ได้แก่ ยาสีฟัน น้ำยาบ้วนปาก และแปรงสีฟัน แต่จากผลการวิเคราะห์ ไม่ปรากฏถึงความสัมพันธ์ระหว่างแบรนด์ A และแบรนด์ C ซึ่งสามารถอธิบายได้ดังนี้

1. ความแตกต่างด้านภาพลักษณ์ของแบรนด์ (Brand Positioning)

- o **แบรนด์ B** วางตำแหน่งตนเองเป็นผู้เชี่ยวชาญด้านการดูแลสุขภาพช่องปากเชิงลึก และครบวงจร
- o **แบรนด์ C** มีจุดเด่นแตกต่าง โดยเน้นภาพลักษณ์ของ “ยาสีฟันสูตรเกลือ” ซึ่งเน้นการป้องกันกลิ่นปากและทำความสะอาดเป็นหลัก จึงไม่ถูกมองว่าเป็นสินค้าที่ต่อยอดโดยตรงจากการใช้แบรนด์ A

2. การจัดจำหน่ายและการทำโปรโมชันร่วม (Cross-promotion)

- o มีความเป็นไปได้ว่าแบรนด์ A และแบรนด์ B มีการจัดวางขายร่วมกันในช่องทางจำหน่าย หรือมีการทำโปรโมชัน/แคมเปญการตลาดร่วม ทำให้ลูกค้ามีแนวโน้มที่จะเลือกซื้อทั้งสองแบรนด์ในธุรกรรมเดียวกัน
- o ในขณะที่แบรนด์ C อาจไม่มีการเชื่อมโยงเชิงการตลาดกับแบรนด์ A อย่างเด่นชัด

3. กลุ่มลูกค้าเป้าหมาย (Customer Segment)

- o ลูกค้าที่ซื้อแบรนด์ A มักเป็นผู้ที่มีปัญหาเฉพาะ และต้องการผลิตภัณฑ์เสริมที่ช่วยแก้ปัญหาเชิงลึก → ทำให้แบรนด์ B ตอบโจทย์ได้ตรงกว่า

- o ส่วนแบรนด์ C แม้จะมีสินค้าในหมวดเดียวกัน แต่ถูกมองว่าเป็น “ผลิตภัณฑ์ดูแลช่องปากพื้นฐาน” มากกว่าที่จะเป็นแบรนด์เสริมสำหรับผู้มีปัญหาเฉพาะทาง

4.2.6.1 ความสัมพันธ์ระหว่างแบรนด์กับช่องทางการขาย

คณะผู้จัดทำได้ทำการวิเคราะห์ความสัมพันธ์ระหว่าง แท็กที่ระบุชื่อแบรนด์ และ แท็กที่ระบุช่องทางการขาย เพื่อตรวจสอบว่าลูกค้ามีแนวโน้มให้ความสนใจสินค้าแบรนด์ใดในช่องทางการขายแบบใดบ้าง ผลการวิเคราะห์ดังกล่าวสะท้อนให้เห็นว่า ความนิยมของแบรนด์ไม่ได้ขึ้นอยู่กับตัวสินค้าเพียงอย่างเดียว แต่ยังสัมพันธ์กับช่องทางการขายที่ลูกค้าเลือกใช้

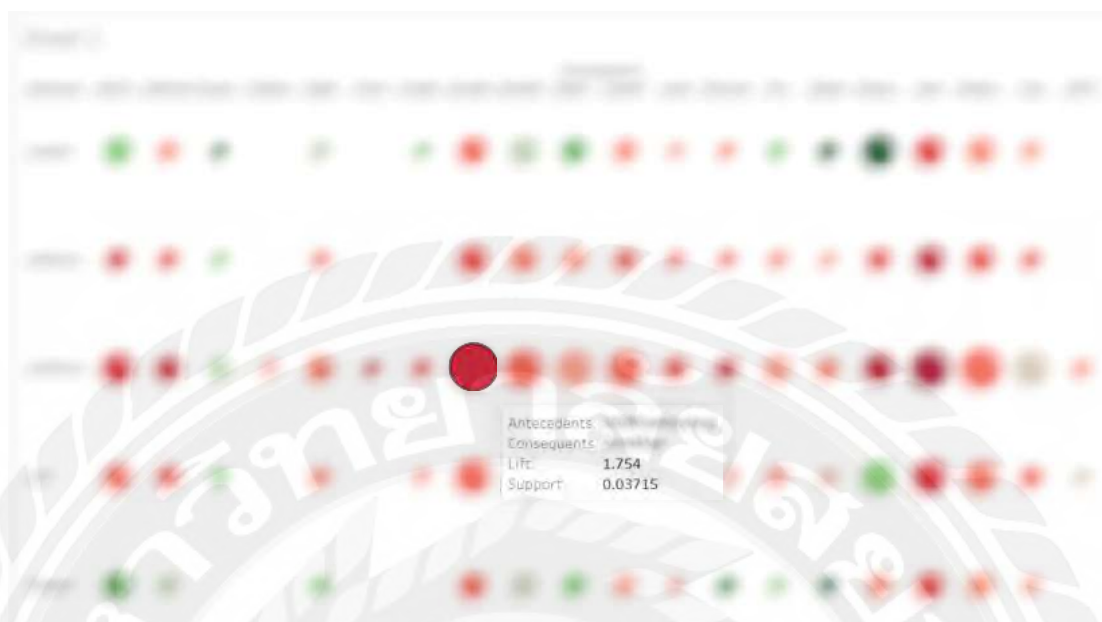
รูปที่ 4.11 ผลลัพธ์การใช้เทคนิค Grouped Matrix-based Visualization กับ แบรนด์และช่องทางการขาย

จาก รูปที่ 4.11 ข้อมูลที่ได้รับมาในเบื้องต้น ยังอยู่ในลักษณะที่ไม่เหมาะสมต่อการนำไปวิเคราะห์โดยตรง ดังนั้นจึงจำเป็นต้องมีการจัดการและปรับแต่งข้อมูล โดยเฉพาะการใช้ Filter เพื่อคัดเลือกและแสดงเฉพาะข้อมูลที่เกี่ยวข้องและอยู่ในรูปแบบที่ต้องการสำหรับการวิเคราะห์ต่อไป



รูปที่ 4.12 ผลลัพธ์การใช้ Filter เพื่อคัดเลือกและแสดงเฉพาะข้อมูลที่เกี่ยวข้อง

จาก รูปที่ 4.12 จะเห็นได้ว่าเมื่อได้ทำการคัดกรองข้อมูลตามที่ต้องการแล้ว จะสามารถมองเห็นรูปแบบของข้อมูลได้อย่างชัดเจนมากขึ้น โดยในการแสดงผล ข้อมูลฝั่งซ้ายเป็น ช่องทางการขาย ส่วนด้านบนเป็น แบนด์ของสินค้าที่ถูกนำมาลงขายในแต่ละแพลตฟอร์ม การวิเคราะห์มุ่งเน้นไปที่ค่า Support ซึ่งสะท้อนถึงจำนวนผู้ใช้งานที่กดเข้ามาดูสินค้าผ่านช่องทางการขายนั้น ๆ และค่า Lift ซึ่งช่วยอธิบายถึงความเป็นไปได้ว่าการที่ผู้ใช้งานกดเข้ามานั้นเกิดจากความตั้งใจจริง (พฤติกรรมเชิงสนใจ) หรือเป็นเพียงการกดเข้ามาโดยบังเอิญ



รูปที่ 4.13 ความสัมพันธ์ระหว่างช่องทางการขาย L และแบรนด์ผลิตภัณฑ์ดูแลสุขภาพ G

จาก รูปที่ 4.13 สังเกตพบว่า แบรนด์ผลิตภัณฑ์ดูแลสุขภาพนี้มีค่า Support ที่สูง แต่ค่า Lift กลับอยู่ในระดับต่ำ สะท้อนว่าผู้ใช้งานมักกดเข้าชมลิงก์สินค้าของแบรนด์นี้บ่อยครั้งบนช่องทางการขาย L ทว่าไม่ได้หมายความว่าผู้ใช้งานมีความตั้งใจจริงที่จะเลือกซื้อเสมอไป แต่เป็นเพียงการกดเข้ามาด้วยความสนใจทั่วไปเท่านั้น ทั้งนี้ สาเหตุหนึ่งอาจเกิดจากการที่สินค้าและช่องทางการขายดังกล่าวถูกนำไปโพสต์ซ้ำหรือปรากฏบ่อยกว่าแบรนด์และช่องทางอื่น ทำให้มีโอกาสถูกกดเข้าชมมากขึ้นโดยอัตโนมัติ

Figure 4.14 Association between the path L and brand K



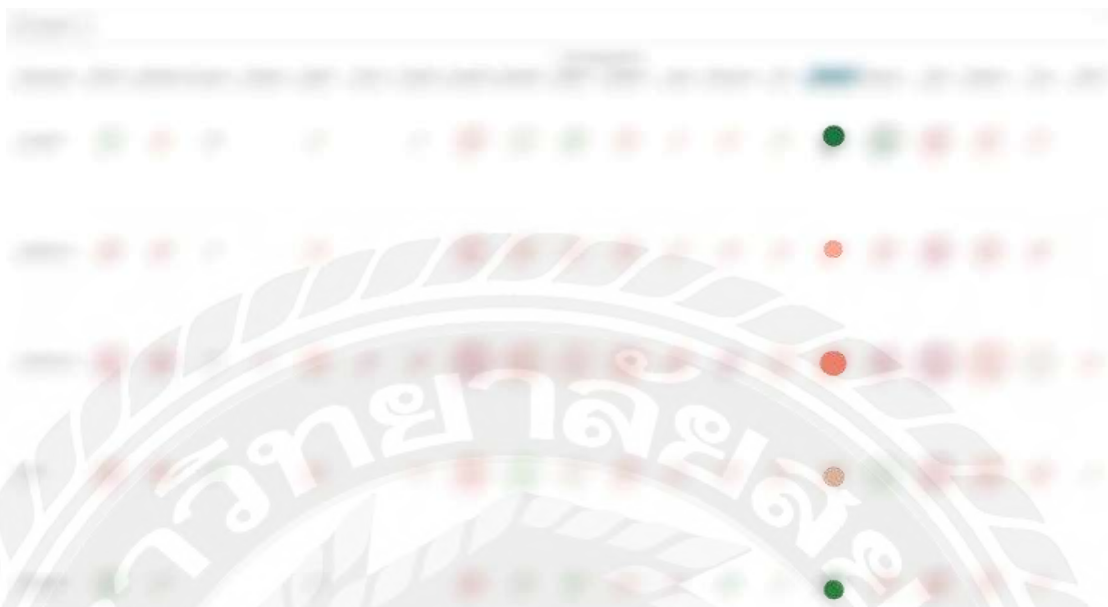
รูปที่ 4.14 ความสัมพันธ์ระหว่างช่องทางการขาย L และแบรนด์ของใช้เด็ก K

จาก รูปที่ 4.14 ในส่วนของ ช่องทางการขาย L ซึ่งถือเป็นช่องทางหลักของร้านค้า พบว่ามี การปรากฏของทุกแบรนด์สินค้าในช่องทางนี้ อย่างไรก็ตาม จากการสังเกตพบว่า แบรนด์ K มีค่า Support และ Lift อยู่ในระดับค่อนข้างสูง ซึ่งสามารถตีความได้ว่าผู้ใช้งานส่วนใหญ่ในช่องทางการ ขาย L มีความสนใจต่อผลิตภัณฑ์ของแบรนด์ K เป็นพิเศษ เมื่อเทียบกับแบรนด์อื่น ๆ ที่ปรากฏใน ช่องทางเดียวกัน



รูปที่ 4.15 ความสัมพันธ์ระหว่างช่องทางการขาย Z และแบรนด์สกินแคร์ R

จาก รูปที่ 4.15 วิเคราะห์พบว่า แบนด์ R บน ช่องทางการขาย Z มีค่า Support และ Lift อยู่ในระดับสูง ซึ่งสะท้อนว่าผู้ใช้งานที่สนใจสินค้าแบรนด์นี้มักเลือกกดเข้าชมผ่านช่องทางการขาย Z มากกว่าช่องทางอื่น ทั้งนี้อาจเป็นผลจากการที่ช่องทางดังกล่าวมีการนำเสนอโปรโมชั่นหรือข้อเสนอพิเศษที่ดึงดูดความสนใจของผู้บริโภคให้เข้ามาชมสินค้าในสัดส่วนที่มากขึ้น



รูปที่ 4.16 ความสัมพันธ์ระหว่างช่องทางการขายทั้งหมด และแบรนด์ยาสระผม Q

จาก รูปที่ 4.16 แบรนด์ Q จะพบว่าค่า Support ในภาพรวมค่อนข้างต่ำเมื่อเทียบกับแบรนด์อื่น อย่างไรก็ตาม ใน ช่องทางการขาย Z และ S กลับมีค่า Lift อยู่ในระดับสูง ซึ่งสามารถตีความได้ว่าผู้ใช้งานมีความสนใจในสินค้าของแบรนด์ Q ผ่านช่องทาง Z และ S มากกว่าช่องทางการขายอื่นๆ ปัจจัยที่อาจส่งผลให้เกิดความแตกต่างนี้ ได้แก่ ความสะดวกในการเข้าถึงและการสั่งซื้อ รวมถึงการมีโปรโมชั่นที่ดึงดูดใจในช่องทางดังกล่าว



รูปที่ 4.17 ความสัมพันธ์ระหว่างช่องทางการขายทั้งหมด และแบรนด์น้ำยาปรับผ้านุ่ม E

จาก รูปที่ 4.17 สำหรับสินค้าจาก แบรินด์ E แม้ว่าจะมีจำนวนผู้ใช้งานที่ทำรายการไม่มากนัก แต่กลับมีค่า Lift อยู่ในระดับสูง ซึ่งสะท้อนว่าสินค้าของแบรนด์นี้สามารถเจาะกลุ่มลูกค้าเฉพาะได้อย่างชัดเจน และได้รับความสนใจอย่างมากจากกลุ่มเป้าหมายนั้นโดยตรง นอกจากนี้ยังพบว่าผู้ใช้งานมีแนวโน้มให้ความสนใจผ่าน ช่องทางการขาย Z มากกว่าช่องทางอื่น ๆ อีกด้วย

บทที่ 5

สรุปผลและข้อเสนอแนะ

5.1 สรุปผลโครงการ

จากการดำเนินงานโครงการสหกิจศึกษา “การวิเคราะห์พฤติกรรมของผู้ใช้งานจากร้านค้าออนไลน์ด้วยเทคนิคการวิเคราะห์ตะกร้าสินค้า” พบว่าการประยุกต์ใช้เทคนิค Market Basket Analysis (MBA) ร่วมกับการนำเสนอผลด้วย Data Visualization สามารถช่วยเปิดเผยรูปแบบความสัมพันธ์ระหว่างแบรนด์สินค้าและช่องทางการขายได้อย่างมีนัยสำคัญ ผลการวิเคราะห์แสดงให้เห็นถึงพฤติกรรมผู้บริโภคที่เชื่อมโยงกันในหลายมิติ เช่น การซื้อพร้อมระหว่างผลิตภัณฑ์ดูแลผมร่วงกับของใช้เด็ก ผลิตภัณฑ์สำหรับสัตว์เลี้ยงกับผลิตภัณฑ์ทำความสะอาดมือ และยาสีฟันเฉพาะทางกับผลิตภัณฑ์ดูแลช่องปากครบวงจร รวมทั้งยังสะท้อนให้เห็นถึงบทบาทสำคัญของช่องทางการขายบางประเภทที่มีอิทธิพลต่อการตัดสินใจซื้อของลูกค้า โดยเฉพาะในกรณีที่มีโปรโมชั่นหรือการเข้าถึงที่สะดวกยิ่งขึ้น นอกจากนี้ การใช้ตัวชี้วัด Support, Confidence และ Lift ยังช่วยยืนยันได้ว่าหลายความสัมพันธ์ที่ค้นพบนั้นเกิดขึ้นจริง มิใช่เพียงความบังเอิญ อันเป็นข้อมูลเชิงลึกที่สามารถนำไปใช้กำหนดกลยุทธ์ทางการตลาดในเชิงรุก ไม่ว่าจะเป็นการทำ Cross-selling, Bundle Promotion, Co-branding หรือการกำหนด Channel Strategy เพื่อเพิ่มประสิทธิภาพในการเข้าถึงกลุ่มลูกค้าเป้าหมายอย่างตรงจุด

5.1.1 ข้อจำกัดหรือปัญหาของโครงการ

แม้ว่าผลการวิเคราะห์จะสามารถสะท้อนพฤติกรรมและแนวโน้มของลูกค้าได้อย่างมีประสิทธิภาพ แต่โครงการยังมีข้อจำกัดบางประการ ได้แก่

5.1.1.1 ความสมบูรณ์ของแท็ก (Tags) ข้อมูลบางส่วนมีการกระจายตัวไม่สม่ำเสมอ ทำให้แบรนด์หรือช่องทางการขายบางประเภทปรากฏน้อย และอาจส่งผลกระทบต่อความแม่นยำในการวิเคราะห์

5.1.1.2 การตีความผลลัพธ์ แม้ว่าตัวชี้วัด Support, Confidence และ Lift จะช่วยอธิบายความสัมพันธ์ได้ แต่การตีความเชิงพฤติกรรมยังขึ้นอยู่กับสมมติฐานของผู้วิเคราะห์ ซึ่งอาจเกิดความคลาดเคลื่อนหากไม่มีข้อมูลเชิงคุณภาพสนับสนุน

5.1.1.3 ข้อจำกัดของเทคนิคที่ใช้ การใช้ Apriori Algorithm มีความเหมาะสมในการค้นหาความสัมพันธ์เบื้องต้น แต่เมื่อข้อมูลมีจำนวนมาก อาจใช้เวลาในการประมวลผลสูง และไม่สามารถสะท้อนความสัมพันธ์ที่ซับซ้อนกว่าการจับคู่แบบ 1 ต่อ 1 ได้อย่างครบถ้วน

5.1.2 ข้อเสนอแนะ

เพื่อพัฒนาผลลัพธ์ของการวิเคราะห์ให้มีความสมบูรณ์และสามารถนำไปใช้เชิงธุรกิจได้อย่างมีประสิทธิภาพยิ่งขึ้น ผู้จัดทำขอเสนอแนวทางดังต่อไปนี้

5.1.2.1 ผสานข้อมูลจากหลายแหล่ง (Multi-source Data Integration) เช่น ข้อมูลจาก Social Media, พฤติกรรมการซื้อซ้ำ (Repeat Purchase) หรือข้อมูลเชิงประชากร (Demographics) เพื่อสร้างภาพรวมที่สมบูรณ์ยิ่งขึ้น

5.1.2.2 พิจารณาเทคนิคการวิเคราะห์ที่หลากหลาย เช่น การใช้ FP-Growth หรือ Machine Learning Model เพื่อค้นหาความสัมพันธ์เชิงลึกและลดข้อจำกัดของ Apriori Algorithm

5.2 สรุปผลการปฏิบัติงานสหกิจ

5.2.1 ข้อดีของการปฏิบัติงานสหกิจ

จากการปฏิบัติสหกิจศึกษา ผู้จัดทำได้รับทั้งความรู้และประสบการณ์ที่มีคุณค่าในด้าน Data Analytics และ Data Science โดยได้มีโอกาสสัมผัสและจัดการกับข้อมูลดิบ (Raw Data) ในการทำงานจริง ศึกษาการใช้เครื่องมือวิเคราะห์ข้อมูล และเรียนรู้วิธีการแปลงข้อมูลให้กลายเป็น Insight ที่สามารถตอบโจทย์เชิงธุรกิจได้อย่างตรงจุด อีกทั้งยังได้เข้าใจมุมมองการทำงานจากผู้บริหาร ว่าข้อมูลควรถูกนำเสนอในรูปแบบใดจึงจะสามารถสนับสนุนการตัดสินใจเชิงกลยุทธ์ได้อย่างมีประสิทธิภาพ นอกจากนี้ ผู้จัดทำยังได้ฝึกฝนทักษะด้าน Data Labeling จากข้อมูลจริง รวมถึงพัฒนาทักษะการทำงานร่วมกับทีม การประสานงานในสภาพแวดล้อมการทำงานจริง และการนำเสนอรายงาน (Report) ให้มีความน่าสนใจและเกิดประสิทธิผลมากยิ่งขึ้น

5.2.2 ปัญหาที่พบของการปฏิบัติงานสหกิจ

ในการปฏิบัติงานสหกิจศึกษา คณะผู้จัดทำพบปัญหาหลักอยู่ 2 ประการ ได้แก่ ความไม่เข้าใจในบาง Syntax ของภาษา Python ทำให้การเขียนโปรแกรมบางครั้งไม่ราบรื่น และความรู้สึก ตื่นเต้นและเกร็งในการนำเสนอผลงานต่อหน้าผู้บริหาร ซึ่งเป็นประสบการณ์ที่ต้องอาศัยการฝึกฝนและความมั่นใจมากขึ้นในการสื่อสารเชิงวิชาชีพ

5.2.3 ข้อเสนอแนะ

สำหรับนักศึกษารุ่นต่อไปที่จะต้องไปปฏิบัติงานสหกิจศึกษา ควรมีการเตรียมในด้านต่อไปนี้

5.2.3.1 ทักษะการเขียนโปรแกรม Python ควรทำความเข้าใจพื้นฐานการเขียนโปรแกรมในระดับที่สามารถนำไปประยุกต์ใช้ได้จริงในการวิเคราะห์ข้อมูล

5.2.3.2 การใช้ Google Sheets เพื่อช่วยในกระบวนการทำความสะอาด (Data Cleaning) และการสำรวจข้อมูลเบื้องต้นอย่างมีประสิทธิภาพ

5.2.3.3 การใช้ Tableau ในการนำเสนอข้อมูลเชิงภาพ (Data Visualization) เพื่อสื่อสารข้อมูลและผลการวิเคราะห์ที่น่าสนใจและเข้าใจง่าย

5.2.3.4 ความรู้ด้านการตลาด (Marketing Analytics) ควรศึกษาเพิ่มเติมเพื่อให้สามารถเชื่อมโยงผลการวิเคราะห์ข้อมูลกับการประยุกต์ใช้ทางธุรกิจได้ชัดเจนยิ่งขึ้น

5.2.3.5 การจดบันทึกงานและสิ่งที่เรียนรู้ ควรมีสุดหรือบันทึกออนไลน์สำหรับจดยละเอียดงานที่ได้รับมอบหมายและประเด็นสำคัญที่ได้เรียนรู้ เพื่อใช้ในการทบทวนและอ้างอิงภายหลัง

บรรณานุกรม

- เก็ตโค้ดแคมป์. (2563, 26 สิงหาคม). *Data Science: ทำไมต้องใช้ Python และทำเงินได้อย่างไร*.
Getcodecamp. <https://getcodecamp.com/2020/08/26/data-science-ทำไมต้องใช้-python-และ-ทำ-เงิน/>
- เก้าเอกเพิท. (2566, 16 ตุลาคม). *มีอะไรใน Microsoft Excel 365*. 9Expert Training.
<https://www.9experttraining.com/articles/มีอะไรใน-microsoft-excel-365>
- โนเวลบิส. (2566, 15 พฤษภาคม). *Google Colab*. [เว็บไซต์].
<https://www.novelbiz.co.th/google-colab/>
- บิสซิเนส แอปพลิเคชัน. (2568, 10 กรกฎาคม). *โปรแกรม Tableau คือโปรแกรมที่ช่วยให้ผู้คน เห็น และทำความเข้าใจกับ Data ได้ดียิ่งขึ้น*. <https://www.bac.co.th/web/products-and-services/tableau-software/>
- บีแซดอินไซด์. (2557, 11 มิถุนายน). *การทำ Data Preparation อย่างมืออาชีพ*. [เว็บไซต์].
<https://bzinsight.wordpress.com/2014/06/11/การทำ-data-preparation-อย่างมืออาชีพ/>
- ศศิวิทย์ ชัยเดชา. (2563, 11 มีนาคม). *Inside purchase behaviour with Market Basket Analysis*. Medium. <https://lengyi.medium.com/inside-purchase-behaviour-with-market-basket-analysis-307ccfc980b5>
- สถาบันบิกดาต้า. (2565, 20 กรกฎาคม). *Frequent Itemsets: Apriori Algorithm*. BDI.
<https://bdi.or.th/big-data-101/frequent-itemsets-apriori-algorithm/>

ภาคผนวก

ภาคผนวก ก.
รูปภาพขณะปฏิบัติงานสหกิจ



รูปที่ ก.1 ขณะปฏิบัติงาน



รูปที่ ก.2 ขณะปฏิบัติงาน

ประวัติคณะผู้จัดทำ



รหัสนักศึกษา : 6504800012

ชื่อ - นามสกุล : นายนารินทร์ ปิ่นสุข

คณะ : วิทยาศาสตร์

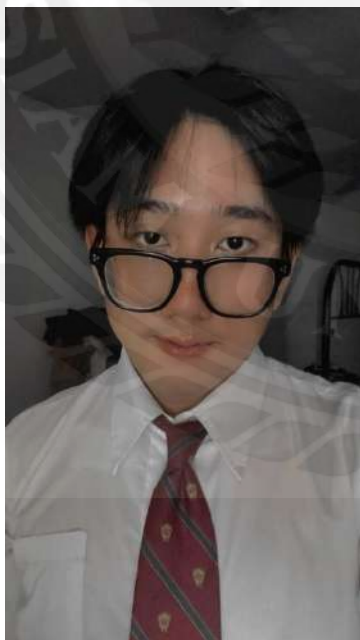
สาขาวิชา : วิทยาการคอมพิวเตอร์

ที่อยู่ : 54 ซ.เทอดไท 57

เขตภาษีเจริญ

แขวงบางหว้า

ผลงาน : การวิเคราะห์พฤติกรรมของ
ผู้ใช้งานจากร้านค้าออนไลน์
ด้วยการทำ Market Basket
Analysis



รหัสนักศึกษา : 6504800028

ชื่อ - นามสกุล : นายสุภวินท์ ปิติโชตินันท์

คณะ : วิทยาศาสตร์

สาขาวิชา : วิทยาการคอมพิวเตอร์

ที่อยู่ : 144 เพชรเกษม 86 หมู่บ้าน

พัชรวิลล์ แขวงบางแคเหนือ

เขตบางแค กรุงเทพมหานคร

ผลงาน : การวิเคราะห์พฤติกรรมของ
ผู้ใช้งานจากร้านค้าออนไลน์
ด้วยการทำ Market Basket
Analysis



แบบสรุปโครงการสหกิจศึกษาและการศึกษาเชิงบูรณาการกับการทำงาน (CWIE)
มหาวิทยาลัยสยาม

ข้อมูลของนักศึกษา

1. ชื่อ-สกุล : นาย นารธร ปีนสุ
นาย สุภาวินท์ ปิติโชติพันธ์
2. สาขาวิชา/คณะ : สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์
3. E-mail นักศึกษา : narathon.pin@siam.edu
supawin.pit@siam.edu
4. ชื่อโครงการ/ผลงาน : การวิเคราะห์พฤติกรรมของผู้ใช้งานจากร้านค้าออนไลน์ด้วยการทำ
Market Basket Analysis
5. ชื่อสถานประกอบการ : บริษัท ดาต้า เฟิร์ส จำกัด
6. ที่อยู่สถานประกอบการ : 465/1 ถนนศรีอยุธยา แขวงทุ่งพญาไท เขตราชเทวี กรุงเทพฯ
10400
7. ระยะเวลาปฏิบัติงาน : 19 พฤษภาคม พ.ศ. 2568 ถึง 29 สิงหาคม พ.ศ. 2568
8. ผู้ให้เตสงานในสถานประกอบการ (พนักงานพี่เลี้ยง)
ชื่อ – สกุล : นาย อนันท์ ตีระบูรณะพงษ์
ตำแหน่ง : Vice Executive Data & Innovation Director

ข้อมูลโครงการ/ผลงาน

1. โครงการ/ผลงาน/งานประจำ ได้รับการจัดระบบการทำงานที่เหมาะสมจากสถานประกอบการ ทั้งลักษณะงานและระยะเวลา มีการจัดระบบพี่เลี้ยงสอนงาน

โครงการ “การวิเคราะห์ความสัมพันธ์ของพฤติกรรมลูกค้าบนแพลตฟอร์มขายของออนไลน์ด้วยเทคนิค **Market Basket Analysis (MBA)**” ได้รับการสนับสนุนจากสถานประกอบการ เป็นอย่างดี ทั้งในด้านการจัดระบบการทำงานที่เหมาะสมกับลักษณะงานและระยะเวลา ตลอดจนการมีพี่เลี้ยงคอยให้คำแนะนำอย่างใกล้ชิด ผู้จัดทำสามารถสอบถามและแลกเปลี่ยนความรู้กับพี่เลี้ยงและทีมงานที่เกี่ยวข้องได้ตลอดเวลา ซึ่งช่วยให้การปฏิบัติงานดำเนินไปอย่างราบรื่นและมีประสิทธิภาพ

โครงการนี้มุ่งเน้นการนำ **ข้อมูลพฤติกรรมการใช้งาน (User Behavior Data)** ของลูกค้าที่มีปฏิสัมพันธ์กับโปรโมชั่นและกิจกรรมต่าง ๆ ภายในแพลตฟอร์ม มาวิเคราะห์ด้วยเทคนิค **Market Basket Analysis** เพื่อค้นหาความสัมพันธ์ระหว่างแบรนด์สินค้าและช่องทางการขายที่มีปรากฏร่วมกัน การสนับสนุนที่ได้รับจากพี่เลี้ยงและทีมงานไม่เพียงช่วยเสริมความเข้าใจเชิงเทคนิค แต่ยังทำให้ผู้จัดทำได้เรียนรู้การทำงานจริงในสภาพแวดล้อมองค์กรอย่างรอบด้าน

2. การดำเนินงานมีความถูกต้อง มีระเบียบแบบแผนและทำให้ผู้จัดทำมีโอกาสประยุกต์ใช้วิชาความรู้/ทักษะตามที่ได้เรียนมา โดยใช้ความรู้ทักษะในการศึกษากระบวนการ การวิเคราะห์ และการแก้ปัญหา หรือสร้างแนวทางใหม่

โครงการนี้ดำเนินการอย่างเป็นระบบและมีระเบียบแบบแผน ครอบคลุมตั้งแต่ขั้นตอนการทำความเข้าใจข้อมูล (Data Understanding), การเตรียมข้อมูล (Data Preparation), การวิเคราะห์ความสัมพันธ์ของข้อมูลด้วยเทคนิค **Association Rule Mining (Apriori Algorithm)** จนถึงการนำเสนอผลลัพธ์ผ่าน **Grouped Matrix-based Visualization** กระบวนการดังกล่าวเปิดโอกาสให้ผู้จัดทำได้นำองค์ความรู้จากรายวิชาที่เกี่ยวข้องมาประยุกต์ใช้กับการทำงานจริงอย่างเป็นรูปธรรม

การใช้ความรู้ในกระบวนการศึกษาและวิเคราะห์ :

- ผู้จัดทำได้นำความรู้ด้าน **Data Analytics** และ **Marketing Analytics** มาประยุกต์ใช้ในการทำความเข้าใจพฤติกรรมของผู้บริโภคบนแพลตฟอร์มออนไลน์
- ใช้เครื่องมือ **Python** (เช่น ไลบรารี **mlxtend** และ **pandas**) สำหรับการประมวลผล การสร้างกฎความสัมพันธ์ (Association Rules) และการจัดการข้อมูล
- ใช้เทคนิค **Data Visualization** เพื่อแปลงข้อมูลเชิงสถิติให้อยู่ในรูปแบบที่เข้าใจง่ายและสามารถสนับสนุนการตัดสินใจเชิงกลยุทธ์ของธุรกิจได้อย่างมีประสิทธิภาพ

การแก้ปัญหาและสร้างแนวทางใหม่ :

- นำแนวคิดจากงานวิจัย (Grouped Matrix-based Visualization) มาประยุกต์ใช้กับข้อมูลจริง เพื่อมองเห็นความสัมพันธ์จำนวนมากได้ชัดเจนและตีความได้ง่ายขึ้น
- ได้เรียนรู้การสร้าง รายงานเชิงวิเคราะห์ ที่สื่อสารผลลัพธ์ทั้งเชิงตัวเลข (Support, Confidence, Lift) และเชิงพฤติกรรมผู้บริโภค เช่น การเชื่อมโยงสินค้ากับกลุ่มคุณแม่มือใหม่ หรือผู้เลี้ยงสัตว์

3. เป็นโครงงาน/ผลงานที่นำไปใช้ประโยชน์ได้อย่างเป็นรูปธรรมในสถานประกอบการ

จากการพัฒนาโครงงาน “การวิเคราะห์ความสัมพันธ์ของพฤติกรรมลูกค้าบนแพลตฟอร์มขายของออนไลน์” สามารถนำไปใช้ประโยชน์ในสถานประกอบการได้อย่างเป็นรูปธรรม ดังนี้

การพัฒนาองค์กรและลดเวลาในการทำงาน

- โครงงานนี้ช่วยให้ทีมการตลาดสามารถเห็นความเชื่อมโยงระหว่าง แบนด์สินค้าและช่องทางการขาย ได้อย่างรวดเร็ว โดยไม่ต้องตรวจสอบกฎจำนวนมากที่สละรายการ
- ทำให้บริษัทเข้าใจพฤติกรรมของลูกค้าเชิงลึก สามารถวางแผนกลยุทธ์ได้อย่างมีข้อมูลสนับสนุน ลดเวลาในการวิเคราะห์ด้วยตนเอง และเพิ่มประสิทธิภาพการตัดสินใจทางธุรกิจ

การลดต้นทุนค่าใช้จ่าย

- การใช้เทคนิค Market Basket Analysis ทำให้บริษัทสามารถระบุได้ว่าสินค้าใดควรถูกทำ **Cross-selling** หรือ **Bundle Promotion** เพื่อลดต้นทุนการทำแคมเปญที่ไม่ตรงกลุ่มเป้าหมาย
- การระบุช่องทางการขายที่มีความสัมพันธ์สูงกับแบรนด์ สามารถช่วยให้บริษัทจัดระบบประมาณด้านการตลาดได้คุ้มค่ามากขึ้น

ความคิดสร้างสรรค์และการพัฒนากระบวนการทำงาน

- โครงงานนี้นำเสนอการใช้ **Grouped Matrix-based Visualization** ซึ่งเป็นแนวทางใหม่ที่ทำให้การวิเคราะห์ข้อมูลขนาดใหญ่เป็นเรื่องง่ายและสื่อสารผลลัพธ์ได้อย่างชัดเจน
- ผลการวิเคราะห์ยังสามารถใช้เป็นแรงบันดาลใจในการออกแบบแคมเปญการตลาดที่ตอบโจทย์กลุ่มลูกค้าเฉพาะ เช่น การออกแบบโปรโมชั่นสำหรับคุณแม่มือใหม่ ผู้เลี้ยงสัตว์ หรือกลุ่มคนที่ใส่ใจสุขภาพช่องปาก

การจดคุ้มครองทรัพย์สินทางปัญญา

ในปัจจุบัน โครงการยังไม่ได้ยื่นจดคุ้มครองทรัพย์สินทางปัญญา แต่แนวทางการวิเคราะห์และรูปแบบ Visualization ที่นำมาประยุกต์ใช้อาจต่อยอดไปสู่การพัฒนาเป็น ซอฟต์แวร์วิเคราะห์พฤติกรรมลูกค้า (**Customer Analytics Tool**) ที่สามารถนำไปจดสิทธิบัตรหรือใช้เป็นผลิตภัณฑ์เชิงพาณิชย์ในอนาคตได้





(<https://bit.ly/3XdFU8E>)

การวิเคราะห์พฤติกรรมของผู้ใช้งานจากร้านค้าออนไลน์ด้วยเทคนิค การวิเคราะห์ตะกร้าสินค้า
Analysis of User Behavior from Online Stores using Market Basket Analysis Technique

บริษัท ดาต้า เฟิร์ส จำกัด

Data First Company Limited

โดย

ชื่อ - นามสกุล นารินทร์ ปิ่นสุข รหัส 6504800012

ชื่อ - นามสกุล สุภวินท์ ปิติโชตินันท์ รหัส 6504800028

รายงานนี้เป็นส่วนหนึ่งของรายวิชา 128-491 สหกิจศึกษาสำหรับนักวิทยาการคอมพิวเตอร์ 1

หลักสูตรวิทยาศาสตรบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์

คณะวิทยาศาสตร์ มหาวิทยาลัยสยาม

ภาคการศึกษาที่ 3 ปีการศึกษา 2567