

วิธีอ้างอิงบทความนี้: ศิวรักษ์ กิจวัฒน์โกสิน และลลันลลิต เหลาโชติ. (2569). การแบ่งกลุ่มลูกค้าด้วยวิธีผสมระหว่างการจัดกลุ่มแบบเคมีนส์และอาร์เอฟเอ็ม: กรณีศึกษาในธุรกิจค้าปลีกออนไลน์. วารสารปฏิบัติการโลจิสติกส์และซัพพลายเชน, 12(2), 145-160. <https://doi.org/10.53848/jlscov12i2.299053>

Received:	March	05, 2026
Revised:	April	22, 2026
Accepted:	Augusst	28, 2026

การแบ่งกลุ่มลูกค้าด้วยวิธีผสมระหว่างการจัดกลุ่มแบบเคมีนส์และอาร์เอฟเอ็ม: กรณีศึกษาในธุรกิจค้าปลีกออนไลน์

ศิวรักษ์ กิจวัฒน์โกสิน¹ และ ลลันลลิต เหลาโชติ^{2*}

บทคัดย่อ

ในยุคที่ธุรกิจค้าปลีกออนไลน์ต้องเผชิญกับข้อมูลลูกค้าจำนวนมากมหาศาล การแบ่งกลุ่มลูกค้าแบบดั้งเดิมยังมีข้อจำกัดในการสะท้อนพฤติกรรมที่ซับซ้อน งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาการแบ่งกลุ่มลูกค้าด้วยวิธีผสมระหว่างการใช้คะแนน RFM และเทคนิค K-Means clustering โดยใช้ข้อมูลธุรกรรมค้าปลีกออนไลน์จำนวน 406,573 รายการเป็นกรณีศึกษา กระบวนการวิจัยเริ่มจากการทำความสะอาดข้อมูลและคัดเลือกข้อมูลที่มีคุณภาพ จากนั้นคำนวณตัวแปร Recency, Frequency และ Monetary และทำการปรับสเกลข้อมูลก่อนนำไปจัดกลุ่มด้วยวิธี RFM scoring ซึ่งแบ่งลูกค้าออกเป็น 5 กลุ่มหลัก และ K-Means clustering โดยกำหนดจำนวนคลัสเตอร์ที่เหมาะสมด้วย Elbow method และเลือกใช้ 4 กลุ่ม ผลการศึกษาพบว่าวิธี RFM ให้ผลลัพธ์ที่เข้าใจง่ายและเหมาะสำหรับการนำไปใช้งานเชิงปฏิบัติ ขณะที่ K-Means สามารถสะท้อนความแตกต่างเชิงโครงสร้างของลูกค้าได้ละเอียดกว่า โดยสามารถแยกกลุ่มลูกค้ามูลค่าสูงมากออกจากกลุ่มอื่นได้อย่างชัดเจนแม้มีจำนวนลูกค้าเพียงส่วนน้อย การบูรณาการทั้งสองวิธีช่วยเพิ่มความแม่นยำและความครอบคลุมของการแบ่งกลุ่มลูกค้า และสามารถสนับสนุนการกำหนดกลยุทธ์ทางการตลาด การรักษาลูกค้า และการเพิ่มมูลค่าลูกค้าในธุรกิจค้าปลีกออนไลน์ได้อย่างมีประสิทธิภาพ

คำสำคัญ: การแบ่งกลุ่มลูกค้า, การเรียนรู้ของเครื่อง, K-Means clustering, RFM, ค้าปลีกออนไลน์

ประเภทบทความ: บทความวิจัย

*ผู้รับผิดชอบงานหลัก

¹บริษัท นิวเจน อีเนอจี้ โซลูชั่น จำกัด, อีเมล: sivarak@ngesth.com

²คณะศิลปศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลรัตนโกสินทร์, อีเมล: lanlalit.l@mutr.ac.th

Customer Segmentation Using a Hybrid Approach of K-Means Clustering and RFM: An Online Retail Case Study

Sivarak Kijwattanaphokin¹ and Lanlalit Lhaochot^{2*}

Abstract

Traditional customer segmentation methods are inadequate for understanding complex consumer behaviors, as online shopping requires extensive data analysis. This research aims to formulate a hybrid customer segmentation methodology by amalgamating RFM scoring with K-Means clustering, employing a dataset from a genuine online retail transaction consisting of 406,573 records. The research methodology encompasses data cleansing and preprocessing, the computation of the Recency, Frequency, and Monetary (RFM) variables, and data normalization. We use RFM scoring to put customers into five groups and K-Means clustering with the Elbow method to find the best number of clusters to put customers into four groups. The findings demonstrate that RFM produces distinct and intelligible groups relevant to practical situations. K-Means, on the other hand, wants to find deeper structural differences between customers. This is especially helpful for finding customer groups that are worth a lot. Using both methods together makes customer segmentation more accurate and complete. Online stores can use this information to plan their marketing, keep customers, and increase the value of their business.

Keywords: Customer segmentation, Machine learning, K-Means clustering, RFM, Online retail.

Type of Article: Research Article

* Corresponding author

¹Newgen Energy Solution Co., Ltd., E-mail: sivarak@ngesth.com

²Faculty of Liberal Arts, Rajamangala University of Technology Rattanakosin, E-mail: Lanlalit.l@rmutr.ac.th

ผ่านการรับรองคุณภาพจากศูนย์ดัชนีการอ้างอิงวารสารไทย (TCI.) อยู่ในกลุ่ม 2 สาขามนุษยศาสตร์และสังคมศาสตร์

1. บทนำ

การเติบโตของพาณิชย์อิเล็กทรอนิกส์ทำให้ธุรกิจค้าปลีกออนไลน์สามารถเข้าถึงข้อมูลพฤติกรรมลูกค้าในระดับธุรกรรมจำนวนมาก ส่งผลให้แนวคิดการแบ่งกลุ่มลูกค้า (customer segmentation) เปลี่ยนจากการใช้ข้อมูลประชากรศาสตร์ไปสู่การวิเคราะห์เชิงพฤติกรรมที่ขับเคลื่อนด้วยข้อมูล (Cooil et al., 2007; Kajornkasirat et al., 2024) อย่างไรก็ตาม วิธีการแบบดั้งเดิม เช่น การวิเคราะห์ RFM แม้จะได้รับความนิยม แต่ยังมีข้อจำกัดด้านความยืดหยุ่นและไม่สามารถสะท้อนรูปแบบพฤติกรรมที่ซับซ้อนในสภาพแวดล้อมดิจิทัลได้อย่างครบถ้วน (Blattberg et al., 2008; Mahboob et al., 2025)

ในขณะเดียวกัน เทคนิคการเรียนรู้ของเครื่อง (Machine Learning: ML) โดยเฉพาะการจัดกลุ่มแบบไม่มีผู้สอน (Unsupervised clustering) ได้ถูกนำมาใช้เพื่อค้นหารูปแบบพฤติกรรมที่ซ่อนอยู่ในข้อมูลลูกค้า และสามารถเพิ่มความละเอียดของการแบ่งกลุ่มได้อย่างไรก็ตาม จากการทบทวนวรรณกรรมพบว่างานวิจัยส่วนใหญ่ยังมุ่งเน้นการเปรียบเทียบประสิทธิภาพของอัลกอริทึมเป็นหลัก ขณะที่ยังขาดการพัฒนา “กรอบการแบ่งกลุ่มแบบบูรณาการ” ที่สามารถเชื่อมโยงความเข้าใจเชิงธุรกิจจาก RFM กับความสามารถในการค้นหารูปแบบเชิงลึกของ ML ได้อย่างเป็นระบบ นอกจากนี้ ยังมีข้อจำกัดด้านการประเมินความเสถียรของกลุ่ม (Cluster stability) และการเชื่อมโยงผลลัพธ์ไปสู่การใช้งานเชิงกลยุทธ์อย่างเป็นรูปธรรม

โดยเฉพาะในบริบทประเทศไทย แม้ว่าธุรกิจค้าปลีกออนไลน์จะเติบโตอย่างรวดเร็ว แต่การประยุกต์ใช้ ML กับข้อมูลธุรกรรมจริงเพื่อพัฒนาการแบ่งกลุ่มลูกค้ายังมีจำกัด งานวิจัยในประเทศจำนวนมากยังคงอาศัยข้อมูลแบบสอบถามหรือการวิเคราะห์เชิงพรรณนา ซึ่งยังไม่สามารถสะท้อนพฤติกรรมลูกค้าจริงในระดับธุรกรรมได้อย่างครบถ้วน ช่องว่างดังกล่าว

สะท้อนถึงความจำเป็นในการพัฒนาวิธีการแบ่งกลุ่มลูกค้าที่มีความเข้มแข็งเชิงวิธีวิทยาและสามารถนำไปใช้เชิงธุรกิจได้จริง

ดังนั้น การศึกษานี้จึงมุ่งพัฒนารอบการแบ่งกลุ่มลูกค้าแบบผสม (Hybrid segmentation framework) โดยบูรณาการวิธีการ RFM ซึ่งให้ความเข้าใจเชิงการตลาดที่ชัดเจน เข้ากับ K-Means clustering ซึ่งสามารถสะท้อนโครงสร้างข้อมูลเชิงลึกได้ดีกว่า งานวิจัยนี้มีความแตกต่างจากงานวิจัยก่อนหน้าใน 3 ประเด็นสำคัญ ได้แก่ (1) การผสานแนวทางเชิง Heuristic (RFM scoring) และ Data-driven clustering อย่างเป็นระบบ (2) การใช้ข้อมูลธุรกรรมจริงในบริบทค้าปลีกออนไลน์ของประเทศไทย และ (3) การเชื่อมโยงผลการแบ่งกลุ่มไปสู่ข้อเสนอเชิงกลยุทธ์สำหรับการบริหารลูกค้า ผลลัพธ์ของงานวิจัยนี้จึงมีศักยภาพในการเติมเต็มช่องว่างทางวิชาการ และสนับสนุนการตัดสินใจเชิงการตลาดในธุรกิจค้าปลีกออนไลน์ได้อย่างมีประสิทธิภาพ

2. วัตถุประสงค์ของการวิจัย

1. เพื่อพัฒนารอบการแบ่งกลุ่มลูกค้าแบบผสมโดยบูรณาการวิธี RFM และเทคนิค K-Means clustering เพื่อเพิ่มความแม่นยำและความครอบคลุมในการวิเคราะห์พฤติกรรมลูกค้าในธุรกิจค้าปลีกออนไลน์

2. เพื่อประเมินประสิทธิภาพและเปรียบเทียบผลลัพธ์ของการแบ่งกลุ่มลูกค้าระหว่างวิธีแบบดั้งเดิม (RFM) และ K-Means clustering รวมถึงวิเคราะห์ความแตกต่างเชิงพฤติกรรมและมูลค่าทางธุรกิจของลูกค้าในแต่ละกลุ่ม

3. เพื่อเชื่อมโยงผลการแบ่งกลุ่มลูกค้าไปสู่ข้อเสนอเชิงกลยุทธ์สำหรับการบริหารลูกค้าในธุรกิจค้าปลีกออนไลน์ โดยมุ่งเน้นการเพิ่มมูลค่าลูกค้าและการตัดสินใจเชิงการตลาดที่ขับเคลื่อนด้วยข้อมูล

3. กรอบแนวคิดงานวิจัย

กรอบแนวคิดการวิจัยนี้พัฒนาขึ้นเพื่ออธิบายความสัมพันธ์ระหว่างข้อมูลธุรกรรมลูกค้า วิธีการประมวลผลข้อมูล และเทคนิคการแบ่งกลุ่มลูกค้า โดยเริ่มจากการใช้ข้อมูลธุรกรรมค้าปลีกออนไลน์เป็นข้อมูลนำเข้า ซึ่งผ่านกระบวนการทำความสะอาดข้อมูล การรวมข้อมูลในระดับลูกค้า และการสร้างตัวแปรเชิงพฤติกรรมในรูปแบบ RFM ได้แก่ Recency, Frequency และ Monetary จากนั้นข้อมูลจะถูกนำไปปรับสเกลเพื่อเตรียมความพร้อมสำหรับการวิเคราะห์เชิง Machine Learning

ในขั้นตอนการแบ่งกลุ่มลูกค้า งานวิจัยนี้ใช้ทั้งแนวทางแบบดั้งเดิมและแนวทางเชิงข้อมูล โดยแนวทางแรกคือการใช้ RFM scoring ซึ่งช่วยให้ได้กลุ่มลูกค้าที่สามารถตีความได้ง่ายในเชิงการตลาด ขณะที่แนวทางที่สองคือการใช้ K-Means clustering ซึ่งสามารถค้นหาโครงสร้างเชิงลึกของข้อมูลและสะท้อนความแตกต่างของลูกค้าได้อย่างละเอียดมากขึ้น ผลลัพธ์จากทั้งสองวิธีจะถูกนำมาบูรณาการและเปรียบเทียบ เพื่อวิเคราะห์ความสอดคล้องและความแตกต่างของกลุ่มลูกค้า

สุดท้าย ผลการแบ่งกลุ่มจะนำไปสู่การสร้างองค์ความรู้เกี่ยวกับพฤติกรรมและมูลค่าของลูกค้าในแต่ละกลุ่ม และสามารถนำไปใช้สนับสนุนการกำหนดกลยุทธ์ทางการตลาด เช่น การรักษาลูกค้า การเพิ่ม

มูลค่าลูกค้า และการออกแบบการตลาดแบบเฉพาะบุคคล ซึ่งสะท้อนถึงการเชื่อมโยงระหว่างการวิเคราะห์ข้อมูลและการนำไปใช้เชิงธุรกิจอย่างเป็นรูปธรรม

4. วรรณกรรมที่เกี่ยวข้อง

4.1 แนวคิดการแบ่งกลุ่มลูกค้า (Customer Segmentation)

การแบ่งกลุ่มลูกค้า (Customer segmentation) เป็นแนวคิดพื้นฐานในสาขาการตลาดที่มุ่งจัดกลุ่มผู้บริโภคให้มีความคล้ายคลึงกันภายในกลุ่มและแตกต่างกันข้ามกลุ่ม เพื่อสนับสนุนการกำหนดกลยุทธ์การตลาดที่มีประสิทธิภาพ (Wedel and Kamakura, 2000) แนวคิดนี้ตั้งอยู่บนหลักการว่าความหลากหลายของตลาด (Market heterogeneity) ทำให้องค์กรไม่สามารถใช้กลยุทธ์เดียวตอบโจทย์ลูกค้าทั้งหมดได้อย่างเท่าเทียม แนวทางดั้งเดิมมักแบ่งกลุ่มลูกค้าจากข้อมูลพื้นฐาน เช่น ประชากรศาสตร์ พฤติกรรมการซื้อ หรือการวิเคราะห์ RFM (Recency-Frequency-Monetary) ซึ่งได้รับการใช้แพร่หลายในการจัดการฐานข้อมูลลูกค้า (Blattberg et al., 2008) อย่างไรก็ตาม วิธีแบบดั้งเดิมมีข้อจำกัดในการจัดการกับข้อมูลหลายมิติและไม่สามารถจับรูปแบบพฤติกรรมที่ซับซ้อนในบริบทดิจิทัลได้ สมการที่ใช้ในการคำนวณค่า RFM แสดงดังสมการต่อไปนี้

$$RFM_i = (R_i)U(F_i)U(M_i) \quad \text{สมการที่ 1}$$

แทนค่า R F และ M จะแสดงดังสมการที่ 2

$$RFM_i = (AD_i - LPD_i)U(F_i)U(Q_i \times U_i) \quad \text{สมการที่ 2}$$

หลังจากนั้นจึงทำการให้คะแนนแต่ละกลุ่มโดย คะแนนจะมีสเกลตั้งแต่ 1 ถึง 5 แสดงดังสมการที่ 3

$$R_{score}, F_{score}, M_{score} \in \{1, 2, 3, 4, 5\} \quad \text{สมการที่ 3}$$

โดยที่

i = ลูกค้ารายที่ i

AD = วันที่อ้างอิงในการวิเคราะห์

LPD = วันที่ซื้อครั้งล่าสุดของลูกค้า

F = จำนวนครั้งที่ทำธุรกรรม

Q = จำนวนสินค้าที่สั่งซื้อในแต่ละครั้ง

U = ราคาต่อหน่วยของสินค้า

ในยุคดิจิทัลและพาณิชย์อิเล็กทรอนิกส์ การแบ่งกลุ่มลูกค้าพัฒนาไปสู่การใช้วิธีเชิงข้อมูล (Data-driven segmentation) ที่อาศัยเทคนิคการเรียนรู้ของเครื่อง (Machine learning) เพื่อค้นหา รูปแบบที่ซ่อนอยู่ในข้อมูลหลายมิติ (Wedel and Kannan, 2016) การแบ่งกลุ่มในลักษณะนี้ไม่เพียง ต้องมีความแตกต่างเชิงสถิติเท่านั้น แต่ยังสามารถ นำไปใช้งานเชิงกลยุทธ์ (Managerially actionable) ในการปรับแต่งข้อเสนอและกิจกรรมทางการตลาดให้ เหมาะสมกับลูกค้าแต่ละกลุ่มได้ด้วย เพื่อให้การแบ่ง กลุ่มมีความน่าเชื่อถือ งานวิจัยด้านการวิเคราะห์กลุ่ม จึงให้ความสำคัญกับการประเมินความเสถียรของกลุ่ม (Cluster stability) และตัวชี้วัดคุณภาพของการ จัดกลุ่ม (Cluster validity) อย่างเป็นระบบ (Hennig, 2007) ซึ่งในบริบทของค้าปลีกออนไลน์ที่ข้อมูลมี ความซับซ้อนและมีปริมาณมาก การประเมินเชิงวิธี วิทยาที่เข้มแข็งถือเป็นองค์ประกอบสำคัญของการ สร้างโมเดลแบ่งกลุ่มที่ใช้งานได้จริง

4.2 การแบ่งกลุ่มลูกค้าด้วยการเรียนรู้ของเครื่อง (Machine learning)

การแบ่งกลุ่มลูกค้าด้วยวิธี Machine learning กลายเป็นแนวทางสำคัญในยุคที่ธุรกิจค้าปลีกออนไลน์ มีข้อมูลลูกค้าปริมาณมหาศาล เทคนิคการจัดกลุ่มแบบ ไม่มีผู้สอน (Unsupervised learning) เช่น K-means, DBSCAN, และ Spectral clustering ถูกนำมาใช้

เพื่อระบุพฤติกรรมลูกค้าที่ซับซ้อนจากข้อมูลธุรกรรม และพฤติกรรมบนแพลตฟอร์มดิจิทัล (Kajornkasirat et al., 2024) งานวิจัยล่าสุดยังมีการผสมผสานเทคนิค แบบไฮบริดเพื่อเพิ่มประสิทธิภาพการแบ่งกลุ่ม เช่น การรวมหลายอัลกอริทึมเข้าด้วยกันเพื่อจัดกลุ่มลูกค้า แบบเรียลไทม์บนแพลตฟอร์มอีคอมเมิร์ซ ซึ่งช่วยให้ สามารถจัดการกับความผันผวนของพฤติกรรมลูกค้า ได้ดียิ่งขึ้น (Mahboob et al., 2025)

นอกจากนี้ ยังมีงานศึกษาที่ประยุกต์ ML กับ ข้อมูลรีวิวของลูกค้าเพื่อขยายมิติในการแบ่งกลุ่ม โดยใช้ข้อมูลข้อความ (Text) และการวิเคราะห์ เชิงความรู้สึก (Sentiment analysis) ร่วมกับ clustering เพื่อหารูปแบบพฤติกรรมและความ คิดเห็นของลูกค้า ตัวอย่างเช่น การศึกษาที่ใช้วิธี วิเคราะห์รีวิวสินค้าและจัดกลุ่มลูกค้าจากข้อมูลรีวิว ออนไลน์ ซึ่งช่วยให้ธุรกิจเข้าใจความชอบและแรงจูงใจ ของลูกค้าได้ชัดเจนขึ้น (Das and Biswas, 2025) งานวิจัยเชิงปริมาณเกี่ยวกับค้าปลีกออนไลน์ยังพบว่า การใช้ RFM ร่วมกับ Clustering สามารถแยก Segment ได้อย่างมีนัยสำคัญ โดยมีกลุ่มลูกค้าที่ แสดงพฤติกรรมซื้อและมูลค่าที่ต่างกันอย่างชัดเจน ซึ่งเป็นข้อมูลสำคัญสำหรับการวางแผนการตลาดและ กลยุทธ์การสื่อสารแบบ Personalized (Wong et al., 2024)

โดยรวม งานวิจัยในช่วงที่ผ่านมาแสดงให้เห็นถึงการพัฒนาของ Machine learning สำหรับ Segmentation ในบริบทอีคอมเมิร์ซอย่างต่อเนื่อง ทั้งในด้านการปรับปรุงอัลกอริทึม การเพิ่มความเข้าใจเชิงพฤติกรรมลูกค้าจากข้อมูลต่าง ๆ รวมถึงการประยุกต์ใช้แนวทางหลายมิติ เช่น การวิเคราะห์รีวิวและพฤติกรรมการซื้อ เพื่อให้การแบ่งกลุ่มลูกค้ามีความหมายและสามารถนำไปใช้เชิงกลยุทธ์ได้จริง

4.3 งานวิจัยที่เกี่ยวข้อง

การแบ่งกลุ่มลูกค้า (Customer segmentation) เป็นกลไกสำคัญที่ช่วยให้องค์กรสามารถทำความเข้าใจความหลากหลายของพฤติกรรมผู้บริโภค และออกแบบกลยุทธ์ทางการตลาดที่สอดคล้องกับลักษณะเฉพาะของแต่ละกลุ่มได้อย่างมีประสิทธิภาพ ในบริบทของธุรกิจค้าปลีกออนไลน์ที่มีข้อมูลธุรกรรมและพฤติกรรมลูกค้าจำนวนมาก การประยุกต์ใช้เทคนิคการวิเคราะห์ข้อมูลและ Machine learning ได้เข้ามามีบทบาทสำคัญในการพัฒนากระบวนการแบ่งกลุ่มลูกค้าให้มีความแม่นยำและสะท้อนพฤติกรรมจริงมากยิ่งขึ้น วรรณกรรมในช่วงทศวรรษที่ผ่านมาแสดงให้เห็นถึงการเปลี่ยนผ่านจากแนวทางดั้งเดิมที่อาศัยตัวแปรพื้นฐาน เช่น RFM ไปสู่การประยุกต์ใช้อัลกอริทึมการจัดกลุ่มข้อมูลหลายรูปแบบ รวมถึงการผสานเทคนิค Data mining เพื่อเพิ่มความสามารถในการอธิบายลักษณะของแต่ละกลุ่มลูกค้า อย่างไรก็ตาม แม้ว่างานวิจัยจะมีจำนวนเพิ่มขึ้น แต่ยังคงมีข้อจำกัดในด้านความทนทานของโมเดล การประเมินเสถียรภาพของผลการจัดกลุ่ม และการเชื่อมโยงผลลัพธ์สู่การประยุกต์ใช้เชิงกลยุทธ์อย่างเป็นระบบ งานวิจัยที่เกี่ยวข้องอธิบายดังรายละเอียดต่อไปนี้

Wedel and Kannan (2016) อธิบายบทบาทของ marketing analytics ในยุคที่องค์กรต้องเผชิญกับข้อมูลจำนวนมหาศาลจากแพลตฟอร์มดิจิทัล โดยชี้ให้เห็นว่าการวิเคราะห์ข้อมูลลูกค้า เช่น การแบ่งกลุ่มลูกค้า (Segmentation) และการทำ Personalization

ไม่ควรถูกมองเป็นเพียงเครื่องมือทางเทคนิค แต่เป็นความสามารถเชิงกลยุทธ์ที่ช่วยสร้างความได้เปรียบทางการแข่งขันอย่างยั่งยืน ผู้เขียนเสนอกรอบแนวคิดที่เชื่อมโยงข้อมูลขนาดใหญ่ เทคโนโลยีการวิเคราะห์ และกระบวนการตัดสินใจทางการตลาดเข้าด้วยกัน พร้อมอภิปรายประเด็นด้านความเป็นส่วนตัวและการกำกับดูแลข้อมูล แม้งานวิจัยนี้จะไม่ได้นำเสนออัลกอริทึมการจัดกลุ่มโดยตรง แต่ได้วางรากฐานทางทฤษฎีที่สำคัญสำหรับการพัฒนา Machine Learning-based customer segmentation โดยเน้นว่าการแบ่งกลุ่มลูกค้าที่มีประสิทธิภาพควรถูกออกแบบให้สอดคล้องกับเป้าหมายเชิงกลยุทธ์ขององค์กร และสามารถนำผลลัพธ์ไปใช้สร้างคุณค่าทางธุรกิจได้อย่างเป็นรูปธรรม

Wong et al. (2024) ศึกษาการแบ่งกลุ่มลูกค้าในธุรกิจอีคอมเมิร์ซโดยใช้โมเดล RFM (Recency-Frequency-Monetary) ร่วมกับอัลกอริทึมการจัดกลุ่มหลายรูปแบบ ได้แก่ K-means, K-medoids และ DBSCAN เพื่อเปรียบเทียบประสิทธิภาพของแต่ละวิธีในการจำแนกลูกค้าจากข้อมูลธุรกรรมจริง งานวิจัยมุ่งวิเคราะห์ว่าอัลกอริทึมที่แตกต่างกันส่งผลต่อโครงสร้างและคุณลักษณะของกลุ่มลูกค้าอย่างไร โดยประเมินคุณภาพของคลัสเตอร์ด้วยตัวชี้วัดมาตรฐาน เช่น Silhouette score ผลการศึกษาพบว่าแต่ละวิธีให้ผลลัพธ์ที่ต่างกัน อย่างมีนัยสำคัญ โดย K-means ให้กลุ่มที่มีความกระชับและตีความได้ง่าย ขณะที่ DBSCAN มีความสามารถในการจัดการข้อมูลที่มีความหนาแน่นไม่สม่ำเสมอและตรวจจับลูกค้าที่มีพฤติกรรมผิดปกติ (Outliers) ได้ดีกว่า งานวิจัยชี้ให้เห็นว่าการเลือกอัลกอริทึมการจัดกลุ่มมีผลโดยตรงต่อรูปแบบของ Segment ที่ได้และการนำผลลัพธ์ไปใช้กำหนดกลยุทธ์ทางการตลาดในบริบทอีคอมเมิร์ซอย่างมีประสิทธิภาพ

Kajornkasirat et al. (2024) ศึกษาการแบ่งกลุ่มลูกค้าโดยใช้ข้อมูลธุรกรรมค้าปลีกจริงในประเทศไทยช่วงปี 2021–2023 โดยประยุกต์

ใช้โมเดล RFM ร่วมกับอัลกอริทึมการจัดกลุ่มหลายรูปแบบ ได้แก่ K-means, Fuzzy C-means และ Self-Organizing neural network เพื่อเปรียบเทียบประสิทธิภาพของแต่ละวิธีในการจำแนกลูกค้า นอกจากการจัดกลุ่มเชิงสถิติแล้ว งานวิจัยยังต่อยอดด้วยการใช้เทคนิค Association rule mining (FP-Growth) เพื่อค้นหารูปแบบการซื้อสินค้าที่โดดเด่นภายในแต่ละกลุ่ม ทำให้สามารถอธิบายลักษณะเชิงพฤติกรรมของลูกค้าได้ชัดเจนยิ่งขึ้น ผลการศึกษาชี้ให้เห็นว่าการผสมผสาน Clustering กับการวิเคราะห์กฎความสัมพันธ์ช่วยเพิ่มความหมายเชิงธุรกิจของ Segment ที่ได้ และสนับสนุนการนำผลลัพธ์ไปใช้กำหนดกลยุทธ์ทางการตลาดได้อย่างมีประสิทธิภาพมากกว่าการจัดกลุ่มเพียงอย่างเดียว

Awaliyah et al. (2024) ศึกษาการแบ่งกลุ่มลูกค้าในธุรกิจค้าปลีกออนไลน์โดยใช้ข้อมูลธุรกรรมจริงร่วมกับโมเดล RFM และอัลกอริทึม K-means เพื่อจำแนกลูกค้าตามรูปแบบพฤติกรรมการซื้อ งานวิจัยกำหนดจำนวนกลุ่มที่เหมาะสมด้วยวิธี Elbow method และประเมินคุณภาพของคลัสเตอร์ด้วยค่า Silhouette coefficient เพื่อให้มั่นใจว่ากลุ่มที่ได้มีความกระชับภายในและแตกต่างกันอย่างชัดเจน ผลการวิเคราะห์สามารถแยกลูกค้าออกเป็นหลายกลุ่มที่สะท้อนระดับมูลค่าและความถี่ในการซื้อที่แตกต่างกัน เช่น กลุ่มลูกค้าที่มีมูลค่าสูงและภักดีต่อแบรนด์ กลุ่มลูกค้าที่ซื้อบ่อยแต่มูลค่าต่อครั้งต่ำ และกลุ่มลูกค้าที่มีแนวโน้มเลิกซื้อ งานนี้เน้นการประยุกต์ใช้ผล Segmentation เพื่อสนับสนุนการตัดสินใจเชิงการตลาด โดยแสดงให้เห็นว่าการใช้ RFM ร่วมกับ K-means สามารถช่วยให้องค์กรเข้าใจพฤติกรรมลูกค้าและกำหนดกลยุทธ์ที่เหมาะสมกับแต่ละกลุ่มได้อย่างเป็นระบบ

จากการสังเคราะห์วรรณกรรมพบว่า แม้ทั้งสองแนวทางจะมีจุดแข็งที่ต่างกัน แต่ยังไม่

งานวิจัยจำนวนมากที่สามารถบูรณาการข้อดีของทั้ง RFM และ ML ได้อย่างเป็นระบบ โดย RFM มีความโดดเด่นด้านการตีความและการนำไปใช้เชิงกลยุทธ์ ขณะที่ Clustering มีความสามารถในการค้นหารูปแบบเชิงลึกของข้อมูล แต่ขาดความชัดเจนในการตีความเชิงธุรกิจ นอกจากนี้ งานวิจัยส่วนใหญ่ยังมุ่งเน้นการเปรียบเทียบอัลกอริทึม มากกว่าการพัฒนา “กรอบการแบ่งกลุ่มแบบผสม (Hybrid framework)” ที่สามารถใช้งานได้จริงในบริบทธุรกิจ รวมถึงยังขาดการเชื่อมโยงผลการแบ่งกลุ่มไปสู่การประยุกต์ใช้เชิงกลยุทธ์อย่างเป็นรูปธรรม ดังนั้น งานวิจัยนี้จึงมุ่งเติมเต็มช่องว่างดังกล่าว โดยพัฒนากรอบการแบ่งกลุ่มลูกค้าแบบผสมที่บูรณาการ RFM และ K-Means clustering เข้าด้วยกัน เพื่อให้ได้ทั้งความเข้าใจเชิงธุรกิจและความแม่นยำเชิงข้อมูล พร้อมทั้งใช้ข้อมูลธุรกรรมจริงในบริบทค้าปลีกออนไลน์ของประเทศไทย และเชื่อมโยงผลลัพธ์ไปสู่ข้อเสนอเชิงกลยุทธ์สำหรับการบริหารลูกค้า ซึ่งช่วยยกระดับทั้งในเชิงวิสัยทัศน์และการประยุกต์ใช้ในภาคธุรกิจ

5. วิธีดำเนินการวิจัย

วิธีดำเนินการวิจัยสามารถอธิบายโดยละเอียดแยกเป็นรายข้อดังต่อไปนี้

5.1 รูปแบบการวิจัยและกรณีศึกษา: การวิจัยนี้เป็นการวิจัยเชิงปริมาณ โดยใช้ข้อมูลธุรกรรมของธุรกิจค้าปลีกออนไลน์เป็นกรณีศึกษา ธุรกิจดังกล่าวเป็นร้านค้าออนไลน์ที่ตั้งอยู่ในกรุงเทพมหานครและมีลูกค้ากระจายอยู่ในหลายจังหวัดทั่วประเทศ ข้อมูลที่ใช้ครอบคลุมช่วงเวลาตั้งแต่วันที่ 18 เมษายน 2567 ถึง 26 เมษายน 2568 โดยกำหนดหน่วยวิเคราะห์ในระดับลูกค้า (CustomerID) เพื่อสะท้อนพฤติกรรมการซื้อซ้ำและมูลค่าของลูกค้าแต่ละรายอย่างเหมาะสม

5.2 ลักษณะและขอบเขตของข้อมูล: ข้อมูลที่ใช้เป็นข้อมูลธุรกรรม (transaction data) ซึ่งประกอบด้วย

ตัวแปรสำคัญ ได้แก่ InvoiceNo, StockCode, Description, Quantity, InvoiceDate, UnitPrice, CustomerID และ Province โดยข้อมูลแต่ละแถวแทนรายการสินค้าในใบสั่งซื้อหนึ่งรายการ และสามารถคำนวณมูลค่าการซื้อ (Monetary) ได้จาก $Quantity \times UnitPrice$ ข้อมูลดิบก่อนการคัดกรองมีจำนวน 541,390 แถว

5.3 การเตรียมและประมวลผลข้อมูล แบ่งเป็น 3 ขั้นตอนได้แก่

1. การทำความสะอาดข้อมูล: การทำความสะอาดข้อมูลเริ่มจากดำเนินการตรวจสอบข้อมูลสูญหาย (missing values) และข้อมูลผิดปกติ เช่น จำนวนสินค้าติดลบ (สะท้อนการคืนสินค้า) และค่าที่ไม่สมเหตุสมผล จากการตรวจสอบพบว่าตัวแปร CustomerID มีค่าสูญหายจำนวนมาก จึงทำการลบแถวที่ไม่มี CustomerID ออกจากชุดข้อมูล เนื่องจากไม่สามารถเชื่อมโยงธุรกรรมกับลูกค้าได้ซึ่งเป็นเงื่อนไขสำคัญของการวิเคราะห์ในระดับลูกค้า อย่างไรก็ตามการลบข้อมูลส่วนนี้ส่งผลให้ธุรกรรมของลูกค้าที่ไม่สามารถระบุตัวตนได้ถูกตัดออก ซึ่งเป็นข้อจำกัดของการศึกษา

2. การแปลงข้อมูลระดับลูกค้า: ข้อมูลธุรกรรมถูกแปลงให้อยู่ในระดับลูกค้า โดยคำนวณตัวแปร RFM ได้แก่ Recency: จำนวนวันตั้งแต่การซื้อครั้งล่าสุด Frequency: จำนวนครั้งในการซื้อ Monetary: มูลค่าการใช้จ่ายรวม โดยหลังการเตรียมข้อมูล เหลือข้อมูลที่ใช้วิเคราะห์จำนวน 406,573 แถว

3. การปรับขนาดข้อมูล: ทำการปรับสเกลข้อมูล RFM ด้วยวิธี Standardization เพื่อให้ตัวแปรมีช่วงค่าใกล้เคียงกัน และลดอิทธิพลของตัวแปรที่มีค่ามากในการคำนวณระยะห่างของ K-means

5.4 วิธีการแบ่งกลุ่มลูกค้า แบ่งเป็น 2 แบบได้แก่

1. การแบ่งกลุ่มแบบ RFM (Traditional Approach) ใช้วิธีการให้คะแนน RFM โดยแปลงค่า

Recency, Frequency และ Monetary เป็นคะแนนระดับ 1-5 และจัดกลุ่มลูกค้าออกเป็น 5 กลุ่ม ได้แก่ Lost customers, At risk, Loyal customers, Potential loyalists และ Champions ซึ่งเป็นแนวทางที่สามารถตีความได้ง่ายและเหมาะกับการนำไปใช้เชิงธุรกิจ

2. การแบ่งกลุ่มด้วย ML (K-Means) ใช้ K-Means clustering เพื่อค้นหาโครงสร้างของข้อมูลลูกค้าโดยอาศัยความใกล้เคียงเชิงสถิติ ขั้นตอนเริ่มจากการทดสอบจำนวนคลัสเตอร์ตั้งแต่ 1 ถึง 10 กลุ่ม และใช้ Elbow method โดยพิจารณา ค่า Within-Cluster sum of squares (WCSS) เพื่อเลือกจำนวนคลัสเตอร์ที่เหมาะสม ผลการวิเคราะห์เลือกใช้จำนวน 4 กลุ่ม เนื่องจากให้ความสมดุลระหว่างความแม่นยำและการตีความเชิงธุรกิจ

5.5 การประเมินคุณภาพและความน่าเชื่อถือของโมเดล:

เพื่อให้ผลลัพธ์มีความน่าเชื่อถือ มีการประเมินคุณภาพของคลัสเตอร์โดยตรวจสอบค่า centroid และลักษณะพฤติกรรมของลูกค้าในแต่ละคลัสเตอร์ รวมถึงการเปรียบเทียบผลลัพธ์ระหว่าง RFM และ K-Means เพื่อยืนยันความสอดคล้องของโครงสร้างกลุ่มลูกค้า

5.6 การวิเคราะห์ผลและการนำไปใช้:

ผลการแบ่งกลุ่มลูกค้าจะถูกนำมาวิเคราะห์ลักษณะเฉพาะของแต่ละกลุ่ม โดยพิจารณาความแตกต่างของค่า RFM และพฤติกรรมการซื้อ จากนั้นเชื่อมโยงผลลัพธ์ไปสู่การกำหนดกลยุทธ์ทางการตลาด เช่น การรักษาลูกค้า การเพิ่มมูลค่าลูกค้า และการทำการตลาดแบบเฉพาะบุคคล เพื่อให้การวิเคราะห์มีความสอดคล้องกับการนำไปใช้จริงในธุรกิจค้าปลีกออนไลน์

6. ผลการวิจัย

งานวิจัยนี้ใช้ข้อมูลการสั่งซื้อสินค้าจากร้านค้าออนไลน์ที่ตั้งอยู่ในจังหวัดกรุงเทพมหานครตั้งแต่ 18-04-2024 ถึง 26-04-2025 โดยมีลูกค้ากระจายอยู่ใน 37 จังหวัด

รวมถึงกรุงเทพ ขั้นตอนเริ่มจากการทำความสะอาดข้อมูลเบื้องต้น โดยเริ่มจากการตรวจสอบข้อมูลที่บันทึกผิดหรือมีค่าที่ผิดปกติ ข้อมูลที่เป็นค่าสูญหาย (Missing values) ข้อมูลที่มีค่าติดลบ โดยข้อมูล

เหล่านี้จะถูกลบออกจากชุดข้อมูลเพื่อให้ขั้นตอนการวิเคราะห์ข้อมูลเป็นไปอย่างถูกต้องและมีประสิทธิภาพ โดยการตรวจสอบเบื้องต้นพบว่าข้อมูลที่เป็นค่าสูญหายสรุปดังตารางที่ 1

ตารางที่ 1 สรุปค่าสูญหายในชุดข้อมูล

หัวข้อ	ความหมาย	จำนวนข้อมูลสูญหาย
InvoiceNo	หมายเลขใบแจ้งหนี้	0
StockCode	รหัสสินค้าใช้ระบุประเภทหรือรายการสินค้าแต่ละชนิด	0
Description	รายละเอียดหรือชื่อของสินค้า	1,454
Quantity	จำนวนสินค้าที่ขายในแต่ละรายการ (หน่วย)	0
InvoiceDate	วันที่และเวลาที่ออกใบแจ้งหนี้หรือทำรายการขาย	0
UnitPrice	ราคาต่อหน่วยของสินค้า (บาท)	0
CustomerID	รหัสประจำตัวของลูกค้าใช้ระบุลูกค้าแต่ละราย	134,817
Province	จังหวัดของลูกค้าที่ทำการสั่งซื้อ	0

ข้อมูลก่อนการลบแถวที่มีค่าสูญหายจะมีอยู่ 541,390 แถว หลังการลบเหลืออยู่ 406,573 แถว

6.1 การแบ่งกลุ่มลูกค้าด้วยวิธีการให้คะแนน

คำนวณของลูกค้า 5 รายเรียงตาม CustomerID แสดงดังตารางที่ 2

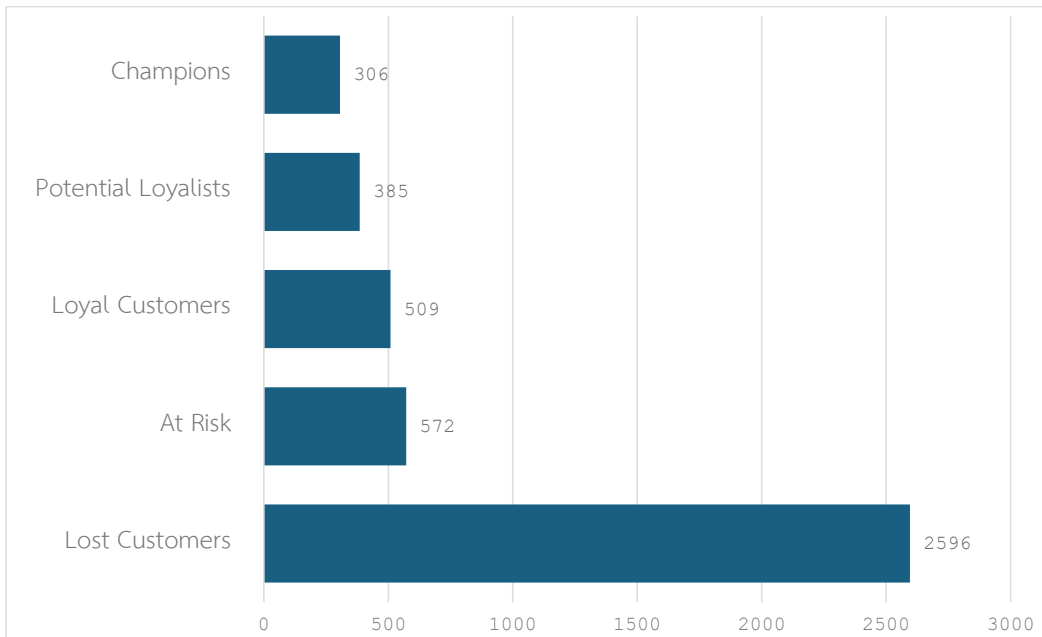
สำหรับขั้นตอนต่อไปจะทำการคำนวณค่า RFM โดยใช้สมการที่ 1-3 ตัวอย่างผลลัพธ์ของการ

ตารางที่ 2 ตัวอย่างคะแนน RFM ของลูกค้า 5 รายเรียงตาม CustomerID

CustomerID	Recency	Frequency	Monetary	R_score	F_score	M_score	RFM_score
12346	325	2	308.88	1	2	1	121
12347	1	7	142,230.00	5	4	5	545
12348	74	4	59,308.92	2	3	4	234
12349	18	1	57,999.15	4	1	4	414
12350	309	1	11,035.20	1	1	2	112

หลังจากนั้นเราสามารถแบ่งกลุ่มของลูกค้าออกเป็น 5 กลุ่ม ได้แก่ 1) ลูกค้าที่เลิกซื้อไปนานแล้ว (Lost customers) 2) ลูกค้าที่กำลังจะหายไป (At risk) 3) ลูกค้าประจำ (Loyal customers)

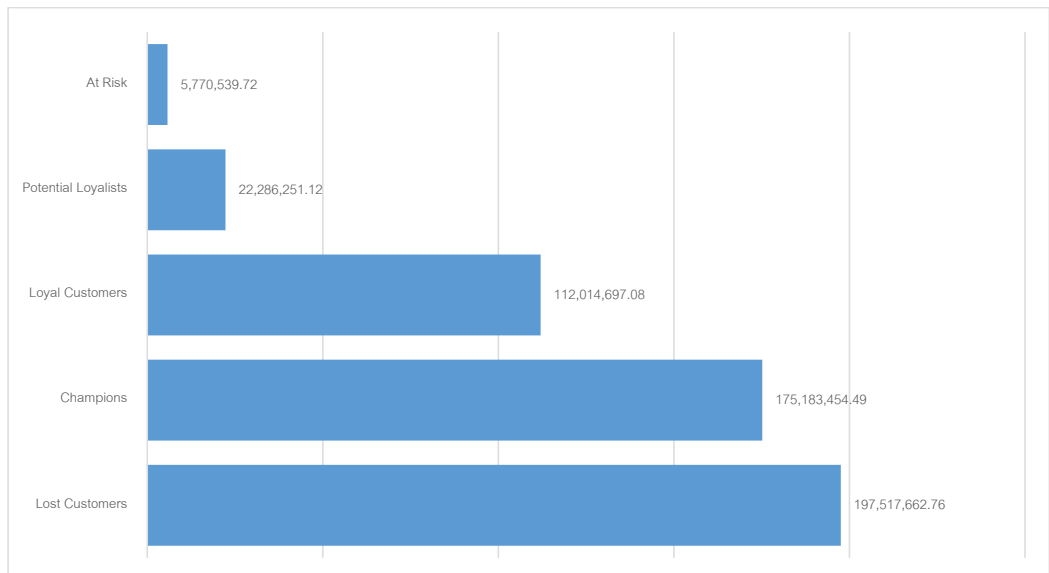
4) ลูกค้าที่มีแนวโน้มจะเป็นลูกค้าประจำ (Potential loyalists) 5) กลุ่มลูกค้าที่ดีที่สุด (Champion) โดยที่จำนวนลูกค้าในแต่ละกลุ่มแสดงดังภาพที่ 1



ภาพที่ 1 ชื่อกลุ่มและจำนวนของลูกค้าในแต่ละกลุ่ม

ยอดขายของลูกค้าแต่ละกลุ่มแสดงดังภาพที่ 2 พบว่าลูกค้ากลุ่ม Lost customers จะมียอดขายมากที่สุด รองลงมาคือกลุ่ม Champion กลุ่ม Loyal customers กลุ่ม Potential loyalists และกลุ่ม At risk ตามลำดับ เมื่อพิจารณาเชิงลึก พบว่ากลุ่ม Champions และ Loyal Customers มีลักษณะเด่นในด้านความถี่และมูลค่าการใช้จ่ายที่สูง ซึ่งสะท้อนถึงพฤติกรรมลูกค้าที่มีความภักดีและสร้างรายได้อย่างต่อเนื่อง ในขณะที่กลุ่ม At Risk และ Lost Customers มีค่า Recency สูง แสดงถึงการห่างหายจากการ

ซื้อสินค้าเป็นระยะเวลานาน ซึ่งบ่งชี้ถึงความเสี่ยงในการสูญเสียลูกค้า อย่างไรก็ตาม แม้ว่าวิธี RFM จะสามารถแยกกลุ่มลูกค้าได้อย่างชัดเจนในเชิงการตลาด แต่การจัดกลุ่มในลักษณะ rule-based ทำให้ลูกค้าที่มีพฤติกรรมแตกต่างกันในเชิงปริมาณอาจถูกจัดอยู่ในกลุ่มเดียวกัน เช่น ลูกค้ากลุ่ม Champions ที่มีระดับการใช้จ่ายแตกต่างกันอย่างมากแต่ยังคงถูกจัดอยู่ในกลุ่มเดียวกัน ซึ่งสะท้อนข้อจำกัดของวิธีนี้ในการอธิบายความแตกต่างเชิงลึกของพฤติกรรมลูกค้า

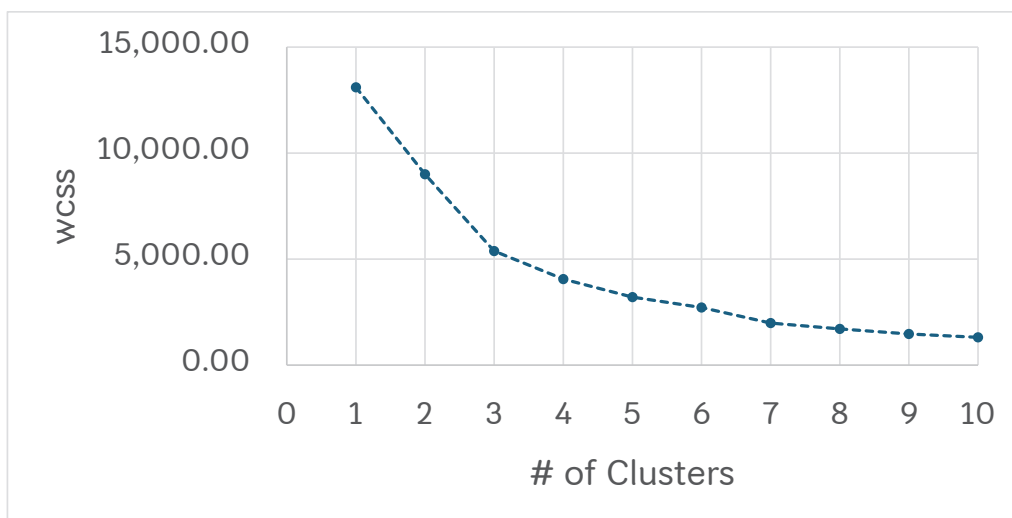


ภาพที่ 2 ยอดขายของลูกค้าในแต่ละกลุ่ม

6.2 การแบ่งกลุ่มลูกค้าด้วยวิธีการเรียนรู้ของเครื่อง (ML)

ในส่วนนี้จะใช้วิธี K-Means clustering ในการแบ่งกลุ่มลูกค้า โดยทำการนำข้อมูล RFM ที่ได้จากขั้นตอนที่แล้วมาทำการปรับสเกลเพื่อให้ค่าของข้อมูลมีรูปแบบสเกลเดียวกัน หลังจากนั้นจึงทำการเลือก

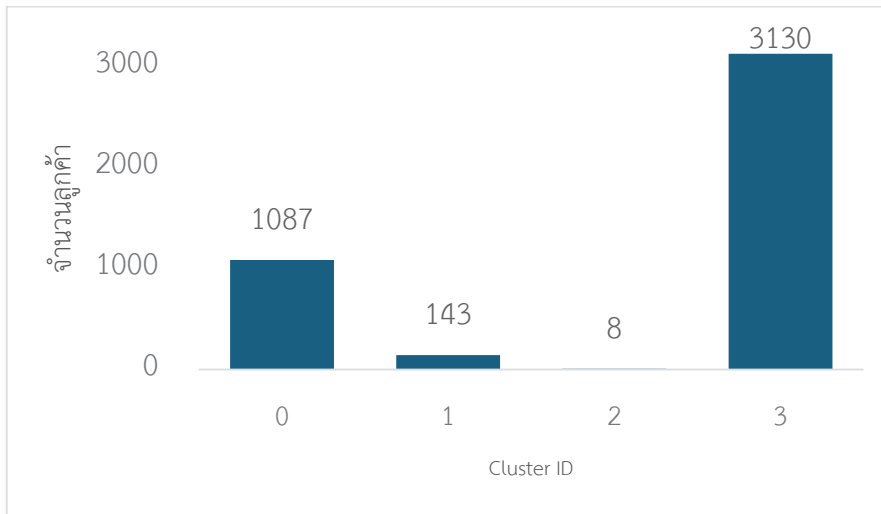
จำนวนกลุ่มของลูกค้าที่ดีที่สุดโดยใช้วิธี Elbow method โดยทำการคำนวณค่า Within-Cluster sum of squares (WCSS) สำหรับค่าจำนวนกลุ่มที่ต่างกัน จาก 1-10 กลุ่ม ผลการคำนวณค่า WCSS แสดงดังภาพที่ 3



ภาพที่ 3 กราฟของวิธี Elbow method

จากกราฟในภาพที่ 3 เราสามารถเลือกใช้ค่าจำนวนกลุ่ม 3 หรือ 4 กลุ่ม เพื่อความสอดคล้องกับวิธีการแบ่งกลุ่มแบบแรก ผู้วิจัยจึงเลือกใช้จำนวนกลุ่ม

เท่ากับ 4 กลุ่ม เมื่อทำการวิเคราะห์ผลลัพธ์ที่ได้พบว่าสามารถแบ่งกลุ่มออกเป็น 4 กลุ่มแสดงดังภาพที่ 4



ภาพที่ 4 ผลการจัดกลุ่มลูกค้า

โดยจุด centroid ของแต่ละกลุ่มลูกค้าแสดงดังตารางที่ 3

ตารางที่ 3 จุด Centroid ของแต่ละกลุ่มลูกค้า

กลุ่มลูกค้า	Recency	Frequency	Monetary
0	247.09	1.82	38,969.05
1	8.66	34	862,840.02
2	15.75	102.88	8,398,546.53
3	40.84	4.64	89,405.35

การวิเคราะห์เชิงลึกพบว่า กลุ่มลูกค้ามูลค่าสูงมาก (Ultra-high-value) แม้จะมีจำนวนเพียงเล็กน้อย แต่มีค่า Monetary และ Frequency สูงกว่ากลุ่มอื่นอย่างมีนัยสำคัญ สะท้อนลักษณะของลูกค้าประเภท “Key revenue drivers” หรือ “Whales” ที่มีผลต่อรายได้ของธุรกิจอย่างไม่สมส่วน (Disproportionate impact) ในขณะที่กลุ่ม Inactive

customers มีค่า Recency สูงและ Frequency ต่ำ แสดงถึงพฤติกรรมที่ห่างหายจากระบบอย่างชัดเจน นอกจากนี้ กลุ่มลูกค้าระดับปานกลางซึ่งมีจำนวนมากที่สุด แสดงถึงฐานลูกค้าหลักของธุรกิจที่ยังมีศักยภาพในการพัฒนา โดยมีลักษณะการซื้อที่สม่ำเสมอแต่ยังไม่ถึงระดับมูลค่าสูง การค้นพบนี้สะท้อนให้เห็นว่า K-means สามารถเปิดเผย

ผ่านการรับรองคุณภาพจากศูนย์ดัชนีการอ้างอิงวารสารไทย (TCI.) อยู่ในกลุ่ม 2 สาขามนุษยศาสตร์และสังคมศาสตร์

“โครงสร้างเชิงลำดับของมูลค่าลูกค้า” ได้ดีกว่าวิธี RFM

6.3 เปรียบเทียบการแบ่งกลุ่มทั้งสองแบบ

เราสามารถเปรียบเทียบการแบ่งกลุ่มทั้งสองแบบจากความเหมือนและความแตกต่างในการระบุกลุ่มลูกค้าดังรายละเอียดต่อไปนี้

1. ความสอดคล้องของกลุ่มมูลค่าสูง/ต่ำ ทั้งสองวิธีสามารถระบุกลุ่มลูกค้ามูลค่าสูง (เช่น กลุ่ม Champions ใน Traditional RFM และ Cluster 1 ของ K-Means) และกลุ่มลูกค้ามูลค่าต่ำหรือกลุ่มเสี่ยงจะหายไป (เช่น At risk และ Lost Customers ซึ่งส่วนใหญ่สอดคล้องกับ Cluster 0 ของ K-Means) ได้อย่างมีประสิทธิภาพ

2. ความสามารถของ K-Means ในการแยกความแตกต่างเชิงลึก

K-Means (เมื่อกำหนด $k = 4$) สามารถแยกกลุ่ม “มูลค่าสูงมากเป็นพิเศษ” (Cluster 2 – “Whales”) ที่มีค่า Monetary และ Frequency สูงโดดเด่นออกมาได้อย่างชัดเจน ซึ่งใน Traditional RFM อาจยังรวมอยู่ในกลุ่ม Champions แสดงให้เห็นว่า K-Means มีจุดแข็งในการค้นหาความแตกต่างที่ละเอียดและซับซ้อนด้วยข้อมูล

3. การจัดการกลุ่มระดับกลาง

กลุ่ม Loyal Customers ใน Traditional RFM ไม่ได้มีคัสเตอร์ที่ตรงตัวเพียงกลุ่มเดียวใน K-Means แสดงให้เห็นว่า K-Means ใช้ระยะทางทางสถิติในการจัดกลุ่ม ทำให้ลูกค้าถูกกระจายไปตามค่าจริงของ RFM มากกว่าการแบ่งตามช่วงคะแนนที่กำหนดไว้ล่วงหน้า เราสามารถวิเคราะห์ความทับซ้อนของลูกค้าจากการแบ่งกลุ่มทั้งสองวิธีดังแสดงในภาพที่ 5



ภาพที่ 5 การทับซ้อนกันของกลุ่มลูกค้า

จากภาพที่ 5 แสดงการทับซ้อนระหว่าง Traditional RFM Segmentation และ K-Means Clustering พบว่า การจัดกลุ่มของทั้งสองวิธีมีความสอดคล้องกันในบางส่วน โดยเฉพาะการระบุลูกค้า

กลุ่มมูลค่าต่ำ (Lost Customers) ซึ่งส่วนใหญ่กระจุกตัวอยู่ใน Cluster 3 และ Cluster 0 ของ K-Means แสดงให้เห็นถึงความสามารถของทั้งสองวิธีในการระบุลูกค้าที่ไม่ active ได้อย่างสอดคล้องกัน อย่างไรก็ตาม

ผ่านการรับรองคุณภาพจากศูนย์ดัชนีการอ้างอิงวารสารไทย (TCI.) อยู่ในกลุ่ม 2 สาขามนุษยศาสตร์และสังคมศาสตร์

กลุ่มลูกค้ามูลค่าสูง เช่น Champions ไม่ได้ถูกรวมอยู่ในคลัสเตอร์เดียวกันทั้งหมด แต่กระจายไปยังหลายคลัสเตอร์ โดยเฉพาะ Cluster 1 และ Cluster 3 สะท้อนให้เห็นว่า K-Means ใช้หลักความใกล้เคียงทางสถิติในการแบ่งกลุ่ม มากกว่าการใช้เกณฑ์คะแนนที่กำหนดไว้ล่วงหน้าแบบ RFM นอกจากนี้ยังมีจำนวนลูกค้าสูงที่สุดและครอบคลุมหลายกลุ่ม RFM แสดงให้เห็นว่า K-Means อาจรวมลูกค้าที่มีพฤติกรรมใกล้เคียงกันทางสถิติ แม้จะอยู่คนละช่วงคะแนนในระบบโดยสรุป Traditional RFM มีความชัดเจนและตีความง่าย ขณะที่ K-Means ให้มุมมองเชิงโครงสร้างข้อมูลทีละเอียดและสะท้อนความแตกต่างภายในกลุ่มได้ดีกว่า

6.4 การนำไปใช้ในการวางกลยุทธ์ทางการตลาด

จากผลการแบ่งกลุ่มลูกค้าด้วยวิธี K-Means พบว่าสามารถจำแนกลูกค้าออกเป็น 4 กลุ่มหลัก ได้แก่ Cluster 0, Cluster 1, Cluster 2 และ Cluster 3 ซึ่งแต่ละกลุ่มมีลักษณะพฤติกรรมที่แตกต่างกันอย่างชัดเจน ดังนั้น การกำหนดกลยุทธ์ทางการตลาดจึงควรออกแบบให้เหมาะสมกับลักษณะเฉพาะของแต่ละคลัสเตอร์ เพื่อให้เกิดประสิทธิภาพสูงสุดในการบริหารความสัมพันธ์ลูกค้า

กลุ่มแรกคือ Cluster 0 ซึ่งมีค่า Recency สูงมาก แสดงว่าลูกค้าไม่ได้ทำรายการมาเป็นเวลานาน อีกทั้งยังมีค่า Frequency และ Monetary ค่อนข้างต่ำ กลุ่มนี้จึงสามารถตีความได้ว่าเป็นลูกค้าที่หายไปหรือใกล้หลุดจากระบบ กลยุทธ์ที่เหมาะสมควรมุ่งเน้นการกระตุ้นให้กลับมาซื้ออีกครั้ง เช่น การเสนอส่วนลดเฉพาะบุคคล การทำแคมเปญสื่อสารเพื่อสร้างความผูกพัน หรือการให้สิทธิพิเศษแบบจำกัดเวลา อย่างไรก็ตาม หากลูกค้าใน Cluster 0 ไม่ตอบสนองต่อการกระตุ้นซ้ำหลายครั้ง ธุรกิจอาจต้องพิจารณาปรับลดความสำคัญทางการตลาดของกลุ่มนี้

Cluster 1 เป็นกลุ่มลูกค้ามูลค่าสูงที่มีค่า Recency ต่ำ หมายถึงเพิ่งซื้อสินค้าเมื่อไม่นานมานี้

อีกทั้งยังมีค่า Frequency และ Monetary สูง แสดงถึงความถี่ในการซื้อและมูลค่าการใช้จ่ายที่โดดเด่น กลุ่มนี้ถือเป็นฐานรายได้หลักขององค์กร ดังนั้น กลยุทธ์ที่เหมาะสมควรเน้นการรักษาความสัมพันธ์ในระยะยาว เช่น โปรแกรมสมาชิกระดับพรีเมียม การมอบสิทธิพิเศษเฉพาะกลุ่ม หรือการสร้างประสบการณ์ที่เหนือกว่าลูกค้าทั่วไป เพื่อเพิ่มความรักดีและลดความเสี่ยงในการย้ายไปยังคู่แข่ง

ในขณะเดียวกัน Cluster 2 เป็นกลุ่มที่มีลักษณะพิเศษที่สุด แม้จะมีจำนวนลูกค้าไม่มาก แต่มีค่า Monetary และ Frequency สูงกว่ากลุ่มอื่นอย่างชัดเจน กลุ่มนี้จึงจัดได้ว่าเป็นลูกค้ามูลค่าสูงมากเป็นพิเศษ หรือกลุ่ม VIP เชิงกลยุทธ์ การดูแลลูกค้ากลุ่มนี้ควรใช้แนวทางที่แตกต่างจากการตลาดทั่วไป เช่น การจัดผู้ดูแลเฉพาะราย การเสนอเงื่อนไขพิเศษ หรือการสร้างความร่วมมือระยะยาวในลักษณะพันธมิตรทางธุรกิจ การรักษาลูกค้าใน Cluster 2 ไว้ได้จะส่งผลต่อเสถียรภาพทางรายได้ขององค์กรอย่างมีนัยสำคัญ

สุดท้ายคือ Cluster 3 ซึ่งเป็นกลุ่มที่มีจำนวนลูกค้ามากที่สุด มีค่า Recency อยู่ในระดับปานกลาง และมีค่า Frequency และ Monetary ในระดับกลางเช่นกัน กลุ่มนี้ถือเป็นฐานลูกค้าหลักที่ยังมีศักยภาพในการเติบโต หากได้รับการส่งเสริมอย่างเหมาะสม ธุรกิจสามารถใช้กลยุทธ์เพิ่มมูลค่าตลอดช่วงชีวิตลูกค้า เช่น การทำ Cross-Selling, Up-Selling หรือโปรแกรมสะสมแต้ม เพื่อกระตุ้นให้ลูกค้าซื้อบ่อยขึ้นและเพิ่มยอดใช้จ่ายในระยะยาว

7. สรุปและอภิปรายผล

ผลการวิจัยครั้งนี้แสดงให้เห็นว่าการแบ่งกลุ่มลูกค้าด้วยวิธี Traditional RFM และ K-Means clustering ให้มุมมองที่แตกต่างกันแต่เสริมกันในเชิงการวิเคราะห์พฤติกรรมลูกค้า โดย RFM สามารถจำแนกลูกค้าได้อย่างชัดเจนในเชิงการตลาดและง่ายต่อการตีความ ขณะที่ K-means สามารถสะท้อน

โครงสร้างเชิงลึกของข้อมูลและความแตกต่างของลูกค้าในเชิงปริมาณได้ละเอียดมากขึ้น โดยเฉพาะการระบุกลุ่มลูกค้ามูลค่าสูงมากที่มีผลต่อรายได้ของธุรกิจอย่างไม่สมส่วน

เมื่อเชื่อมโยงกับแนวคิดทางทฤษฎีด้าน Customer segmentation (Wedel and Kamakura, 2000) ซึ่งเน้นว่าการแบ่งกลุ่มที่ดีควรมีความแตกต่างระหว่างกลุ่ม (Between-group heterogeneity) และความคล้ายคลึงภายในกลุ่ม (Within-group homogeneity) พบว่า K-Means clustering สามารถตอบโจทย์ในเชิงสถิติได้ดีกว่า เนื่องจากใช้หลักการวัดระยะห่างของข้อมูลจริง อย่างไรก็ตาม RFM ยังคงมีความสอดคล้องกับแนวคิดด้าน Database marketing (Blattberg et al., 2008) ที่มุ่งเน้นการใช้งานจริงในเชิงธุรกิจ เนื่องจากสามารถตีความและนำไปใช้กำหนดกลยุทธ์ได้ทันที

ผลการศึกษาที่ยังสอดคล้องกับงานของ Wedel และ Kannan (2016) ที่ชี้ให้เห็นว่าการใช้ข้อมูลขนาดใหญ่ร่วมกับเทคนิคการวิเคราะห์ขั้นสูงสามารถเพิ่มความแม่นยำในการทำ Segmentation และสร้างความได้เปรียบทางการแข่งขัน ในขณะเดียวกันผลลัพธ์ยังสนับสนุนงานของ Wong et al. (2024) และ Kajornkasirat et al. (2024) ที่พบว่าเทคนิค Clustering สามารถเปิดเผยโครงสร้างพฤติกรรมลูกค้าที่ซ่อนอยู่ได้ดีกว่าวิธีแบบดั้งเดิม อย่างไรก็ตามงานวิจัยก่อนหน้านี้ส่วนใหญ่ยังมุ่งเน้นการเปรียบเทียบอัลกอริทึมเป็นหลัก ขณะที่การศึกษานี้ได้แสดงให้เห็นว่าการผสมผสานแนวทางทั้งสองสามารถเพิ่มคุณค่าเชิงวิเคราะห์ได้มากกว่าการใช้วิธีใดวิธีหนึ่งเพียงอย่างเดียว

ในเชิงพฤติกรรมลูกค้า ผลการวิเคราะห์ชี้ให้เห็นว่าลูกค้าไม่ได้กระจายตัวอย่างสม่ำเสมอ แต่มีลักษณะ “Skewed distribution” โดยมีกลุ่ม

ลูกค้ามูลค่าสูงจำนวนน้อยที่สร้างรายได้ส่วนใหญ่ ซึ่งสอดคล้องกับแนวคิด Pareto principle (80/20 rule) ในการตลาด นอกจากนี้ การที่ K-means สามารถแยกกลุ่มลูกค้าระดับกลางออกเป็นหลายกลุ่มย่อยสะท้อนว่าลูกค้ากลุ่มนี้มีความหลากหลายสูงและเป็นกลุ่มที่มีศักยภาพในการพัฒนา ซึ่งเป็นประเด็นที่วิธี RFM ไม่สามารถแสดงให้เห็นได้อย่างชัดเจน

ในด้านข้อจำกัดของวิธีการ พบว่า RFM มีข้อจำกัดในการจับความแตกต่างเชิงลึกของลูกค้าเนื่องจากใช้เกณฑ์คะแนนแบบตายตัว ขณะที่ K-means มีข้อจำกัดด้านความไวต่อพารามิเตอร์และความยากในการตีความเชิงธุรกิจ ซึ่งสอดคล้องกับข้อค้นพบในวรรณกรรมก่อนหน้านี้ การบูรณาการทั้งสองวิธีจึงช่วยลดข้อจำกัดดังกล่าว โดยใช้ RFM เป็นเครื่องมือในการตีความ และใช้ K-means ในการค้นหารูปแบบเชิงลึกของข้อมูล

ในการศึกษาครั้งต่อไป ควรพิจารณาเพิ่มตัวแปรอื่นนอกเหนือจาก Recency Frequency และ Monetary เช่น ประเภทสินค้า ช่องทางการซื้อ หรือพฤติกรรมตอบสนองต่อโปรโมชั่น เพื่อให้การแบ่งกลุ่มลูกค้ามีความละเอียดและครอบคลุมมากยิ่งขึ้น นอกจากนี้ ควรทดลองเปรียบเทียบจำนวนคลัสเตอร์ที่หลากหลาย และใช้ตัวชี้วัดเพิ่มเติม เช่น Silhouette score เพื่อประเมินความเหมาะสมของผลลัพธ์อย่างรอบด้าน อีกทั้งอาจนำวิธีการจัดกลุ่มอื่น ๆ มาเปรียบเทียบ เช่น Hierarchical clustering หรือ DBSCAN เพื่อดูความแตกต่างของโครงสร้างกลุ่มลูกค้า และควรเชื่อมโยงผลการแบ่งกลุ่มกับตัวชี้วัดทางธุรกิจ เช่น Customer lifetime value เพื่อประเมินประโยชน์ในเชิงปฏิบัติจริง การพัฒนาโมเดลให้สามารถอัปเดตข้อมูลแบบต่อเนื่องก็เป็นอีกแนวทางหนึ่งที่ช่วยให้การวิเคราะห์สอดคล้องกับพฤติกรรมลูกค้าที่เปลี่ยนแปลงอยู่เสมอ

เอกสารอ้างอิง

เอกสารอ้างอิง

- Awaliyah, D. A., Prasetyo, B., Muzayanah, R., & Lestari, A. D. (2024). Optimizing customer segmentation in online retail transactions through the implementation of the K-means clustering algorithm. *Scientific Journal of Informatics*, 11(2), 539-548. <https://doi.org/10.15294/sji.v11i2.6137>
- Blattberg, R. C., Kim, B.-D., & Neslin, S. A. (2008). *Database marketing: Analyzing and managing customers*. Springer.
- Cool, B., Aksoy, L., & Keiningham, T. L. (2007). Approaches to Customer Segmentation. *Journal of Relationship Marketing*, 6(3-4), 9-39. https://doi.org/10.1300/J366v06n03_02
- Das, S., & Biswas, B. (2025). Machine learning segment customers based on their reviews in e-commerce portals: A fresh inquiry. *The Lumbini Journal of Business and Economics*, 12(2), 120-130. <https://doi.org/10.3126/ljbe.v12i2.77422>
- Hennig, C. (2007). Cluster-wise assessment of cluster stability. *Computational Statistics and Data Analysis*, 52(1), 258-271. <https://doi.org/10.1016/j.csda.2006.11.025>
- Kajornkasirat, S., Gunglin, P., Puangsuwan, K., & Kaewsuwan, N. (2024). Customer segmentation using association rule mining on retail transaction data. *IAES International Journal of Artificial Intelligence*, 14(3), 1919-1929. <http://doi.org/10.11591/ijai.v14.i3.pp1919-1929>
- Mahboob Basha, S. K., Akhila, N., Deepak, A., Sindhu, J., & Rohith, K. (2025). Real time customer segmentation using hybrid models for smarter e commerce. *International Journal of Engineering Research and Science and Technology*, 21(3(1)), 1459-1469. [https://doi.org/10.62643/ijerst.v21.n3\(1\).pp1459-1469](https://doi.org/10.62643/ijerst.v21.n3(1).pp1459-1469)
- Wedel, M., & Kannan, P. K. (2016). Marketing analytics for data-rich environments. *Journal of Marketing*, 80(6), 97-121. <https://doi.org/10.1509/jm.15.0413>
- Wedel, M., & Kamakura, W. A. (2000). *Market segmentation: Conceptual and methodological foundations* (2nd ed.). Kluwer Academic Publishers.
- Wong, C.-G., Tong, G.-K., & Haw, S.-C. (2024). Exploring Customer Segmentation in E-Commerce using RFM Analysis with Clustering Techniques. *Journal of Telecommunications and the Digital Economy*, 12(3), 97-125. <https://doi.org/10.18080/jtde.v12n3.978>