



การพยากรณ์ยอดขายและการจัดการสินค้าอายุสั้นในธุรกิจร้านกาแฟ

SALES FORECASTING AND SHORT-LIVED INVENTORY MANAGEMENT IN COFFEE

RETAIL



นัยนภัค สถาพรพินิจกิจ

บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ

2567

การพยากรณ์ยอดขายและการจัดการสินค้าอายุสั้นในธุรกิจร้านกาแฟ



การค้นคว้าอิสระนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
วิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาการข้อมูล
คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ
ปีการศึกษา 2567
ลิขสิทธิ์ของมหาวิทยาลัยศรีนครินทรวิโรฒ

SALES FORECASTING AND SHORT-LIVED INVENTORY MANAGEMENT IN COFFEE
RETAIL



NAIYAPAK SATAPORNPINIKIT

An Independent Study Submitted in Partial Fulfillment of the Requirements
for the Degree of MASTER OF SCIENCE
(Data Science)

Faculty of Science, Srinakharinwirot University

2024

Copyright of Srinakharinwirot University

สารนิพนธ์
เรื่อง
การพยากรณ์ยอดขายและการจัดการสินค้าอายุสั้นในธุรกิจร้านกาแฟ
ของ
นายนันท์ค สถาวรพินิจกิจ

ได้รับอนุมัติจากบัณฑิตวิทยาลัยให้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาการข้อมูล
ของมหาวิทยาลัยศรีนครินทรวิโรฒ

(รองศาสตราจารย์ นายแพทย์ฉัตรชัย เอกปัญญาสกุล)
คณบดีบัณฑิตวิทยาลัย

คณะกรรมการสอบปากเปล่าสารนิพนธ์

.....ที่ปรึกษาหลัก
(อาจารย์ ดร.ศุภร คนธภาคี)

.....ประธาน
(ผู้ช่วยศาสตราจารย์ ดร.จันตรี ผลประเสริฐ)
.....กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.ศิริสรวพ เหล่าหะเกียรติ)

ชื่อเรื่อง	การพยากรณ์ยอดขายและการจัดการสินค้าอายุสั้นในธุรกิจร้านกาแฟ
ผู้วิจัย	ณัณภัค สถาพรพินิจกิจ
ปริญญา	วิทยาศาสตร์มหาบัณฑิต
ปีการศึกษา	2567
อาจารย์ที่ปรึกษา	อาจารย์ ดร. ศุภร คนธภักดี

ในการดำเนินธุรกิจร้านกาแฟซึ่งต้องเผชิญกับความผันผวนของความต้องการในแต่ละวัน การพยากรณ์ยอดขายรายวันจึงมีความสำคัญอย่างยิ่งต่อการวางแผนการจัดซื้อวัตถุดิบ การบริหารบุคลากร และการควบคุมต้นทุน วัตถุดิบที่มีอายุการเก็บรักษาสั้น เช่น นมสด โยเกิร์ต หรือเบเกอรี่ หากสั่งซื้อไม่สอดคล้องกับความต้องการจริงอาจก่อให้เกิดของเสียหรือการสูญเสียโอกาสทางรายได้ นอกจากนี้ พฤติกรรมผู้บริโภคยังมีแนวโน้มเปลี่ยนแปลงตามวันในสัปดาห์ ช่วงเวลา สภาพอากาศ และเทศกาลต่าง ๆ ส่งผลให้วิธีการพยากรณ์แบบดั้งเดิมไม่สามารถตอบสนองต่อสภาวะจริงได้อย่างแม่นยำ งานวิจัยนี้จึงมีวัตถุประสงค์ในการพัฒนาแบบจำลองการพยากรณ์ยอดขายรายวันของธุรกิจร้านกาแฟโดยประยุกต์ใช้โมเดล Prophet ซึ่งเป็นแบบจำลองการพยากรณ์อนุกรมเวลาที่สามารถจัดการกับแนวโน้มแบบไม่เป็นเชิงเส้น ความแปรผันตามฤดูกาล และวันหยุดได้อย่างมีประสิทธิภาพ โดยผลลัพธ์จากแบบจำลองสามารถนำไปใช้สนับสนุนการวางแผนการดำเนินงานรายวัน ลดต้นทุนของเสีย เพิ่มประสิทธิภาพการจัดการ และยกระดับคุณภาพการบริการให้สอดคล้องกับพฤติกรรมผู้บริโภคที่เปลี่ยนแปลงอยู่ตลอดเวลา

คำสำคัญ : การพยากรณ์ยอดขาย, ร้านกาแฟ, ข้อมูลอนุกรมเวลา, Prophet, การจัดการวัตถุดิบ, ความแม่นยำในการทำนาย

Title	SALES FORECASTING AND SHORT-LIVED INVENTORY MANAGEMENT IN COFFEE RETAIL
Author	NAIYAPAK SATAPORNPINIKIT
Degree	MASTER OF SCIENCE
Academic Year	2024
Thesis Advisor	dr. Subhorn Khonthapagdee

In operating a coffee shop business that faces fluctuating daily demand, daily sales forecasting plays a vital role in planning raw material procurement, workforce management, and cost control. Perishable ingredients such as fresh milk, yogurt, or bakery items, if purchased without accurate demand estimation, may result in spoilage or lost sales opportunities. Furthermore, consumer behavior tends to vary by day of the week, time of day, weather conditions, holidays, making traditional forecasting methods insufficiently responsive to actual conditions. This research aims to develop a daily sales forecasting model for coffee shop businesses by applying the Prophet model, a time series forecasting tool capable of handling nonlinear trends, seasonal variations, and special events or holidays efficiently. The model's outcomes can be applied to support daily operational planning, reduce waste costs, improve resource management, and enhance service quality to align with the constantly changing behavior of consumers.

Keyword : sales forecasting, coffee shop, time series data, Prophet, inventory management, prediction accuracy

กิตติกรรมประกาศ

การจัดทำสารนิพนธ์ฉบับนี้ไม่อาจสำเร็จลุล่วงได้ หากปราศจากคำแนะนำ การชี้แนะแนวทาง และกำลังใจจากหลายภาคส่วน ข้าพเจ้าขอแสดงความขอบคุณอย่างสูงต่ออาจารย์ที่ปรึกษา อาจารย์ ดร. ศุภกร คนธภักดิ์ ผู้ให้คำแนะนำอย่างใกล้ชิดทั้งด้านวิชาการและแนวทางการทำวิจัยอย่างเป็นระบบ จนทำให้ข้าพเจ้าสามารถดำเนินงานได้ตามเป้าหมายที่ตั้งไว้

ข้าพเจ้าขอขอบคุณคุณคณาจารย์ทุกท่านในภาควิชาวิทยาการข้อมูล คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ที่ได้ถ่ายทอดองค์ความรู้ตลอดระยะเวลาการศึกษา ส่งเสริมให้เกิดความเข้าใจเชิงลึกในแต่ละวิชา ซึ่งมีส่วนอย่างยิ่งต่อการนำความรู้ไปประยุกต์ใช้ในงานวิจัยชิ้นนี้ รวมถึงคณะกรรมการสอบสารนิพนธ์ที่ได้ให้คำแนะนำอันมีคุณค่า ช่วยให้ข้าพเจ้าปรับปรุงและพัฒนาเนื้อหาให้สมบูรณ์ยิ่งขึ้น

นอกจากนี้ ข้าพเจ้าขอขอบคุณครอบครัวที่เป็นกำลังใจสำคัญตลอดมา รวมถึงเพื่อน ๆ ทั้งในระดับปริญญาตรีและปริญญาโท สาขาวิทยาการข้อมูล โดยเฉพาะพีวี พีตัน พีภู พีक्रम พีบอล และเฟิร์น ที่ร่วมเดินทางตลอดสองปีแห่งการเรียนรู้ในปริญญาโท คอยสนับสนุน ช่วยเหลือ และให้กำลังใจในทุกช่วงเวลา

ท้ายที่สุด ข้าพเจ้าหวังเป็นอย่างยิ่งว่าสารนิพนธ์ฉบับนี้จะเป็นประโยชน์ต่อผู้ที่ศึกษาในสาขาเดียวกัน หรือต่อผู้ที่สนใจการประยุกต์ใช้เทคนิคการวิเคราะห์ข้อมูลเพื่อการพยากรณ์ในบริบทของธุรกิจ

นัยนภัค สถาพรพินิจกิจ

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	ง
บทคัดย่อภาษาอังกฤษ	จ
กิตติกรรมประกาศ	ฉ
สารบัญ	ช
สารบัญตาราง	ญ
สารบัญรูปภาพ	ฎ
บทที่ 1 บทนำ	1
1.1 ภูมิหลัง	1
1.2 ความมุ่งหมายของการวิจัย	2
1.3 ความสำคัญของงานวิจัย	3
1.4 ขอบเขตการวิจัย	4
1.4.1 ประชากรที่ใช้ในการวิจัย	4
1.4.2 กลุ่มตัวอย่างที่ใช้ในการวิจัย	4
1.4.3 ตัวแปรที่ศึกษา	4
1.5 นิยามศัพท์เฉพาะ	5
1.6 กรอบแนวคิดงานวิจัย	6
1.7 สมมติฐานในการวิจัย	8
บทที่ 2 ทบทวนวรรณกรรม	9
2.1 การพยากรณ์ยอดขาย (Sales Forecasting)	9
2.1.1 ความสำคัญของการพยากรณ์ยอดขาย	9
2.1.2 เทคนิคการพยากรณ์ยอดขาย	10

2.2 แบบจำลองที่ใช้ในงานวิจัย	10
2.2.1 prophet	10
2.2.2 Tableau Forecast	12
2.2.3 Mean Absolute Error (MAE)	12
2.2.4 Mean Squared Error (MSE)	13
2.2.5 R^2 (Coefficient of Determination / R-Squared / R^2 Score)	14
2.3 งานวิจัยที่เกี่ยวข้อง	15
บทที่ 3 วิธีดำเนินการวิจัย	19
3.1 การออกแบบกระบวนการดำเนินการวิจัย	20
3.2 การนำเข้าข้อมูล	21
3.3 การประมวลผลข้อมูล	22
3.4 การสร้างแบบจำลอง	34
บทที่ 4 ผลการดำเนินการวิจัย	47
ผลลัพธ์ของการทดลองที่ 1	49
ผลลัพธ์ของการทดลองที่ 2	50
ผลลัพธ์ของการทดลองที่ 3	51
ผลลัพธ์ของการทดลองที่ 4	52
ผลลัพธ์ของการทดลองที่ 5	53
ผลลัพธ์ของการทดลองที่ 6	55
บทที่ 5 สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ	58
สรุปผลการวิจัย	58
อภิปรายผลการวิจัย	59
ข้อเสนอแนะ	59
บรรณานุกรม	61

ประวัติผู้เขียน	63
-----------------------	----



สารบัญตาราง

หน้า

ตาราง 1 แสดงข้อมูลตัวอย่างที่ใช้ในการพยากรณ์.....	21
---	----



สารบัญรูปภาพ

หน้า

ภาพประกอบ 1 แสดงการออกแบบกระบวนการดำเนินการวิจัย	19
ภาพประกอบ 2 Preview Data Frame เพื่อดูข้อมูล 5 แถวแรก	22
ภาพประกอบ 3 แสดงข้อมูลหลังจากมีการแปลงข้อมูลให้เหมาะสม	23
ภาพประกอบ 4 แสดง Trend ของข้อมูล	24
ภาพประกอบ 5 แสดงแนวโน้มยอดขายรายวันพร้อมจุดที่น่าสนใจ (Interesting Points)	25
ภาพประกอบ 6 แนวโน้มยอดขายรายวันของ JUICE & SMOOTHIES และ COFFEE	26
ภาพประกอบ 7 กราฟแท่ง Average Monthly Sales – Juice & Smoothies vs Coffee	27
ภาพประกอบ 8 Boxplot Monthly Sales Distribution – Juice & Smoothies vs Coffee	27
ภาพประกอบ 9 กราฟเปรียบเทียบ Weekday vs Weekend Sales	28
ภาพประกอบ 10 Heatmap Correlation Matrix – Juice & Smoothies vs Coffee	29
ภาพประกอบ 11 แสดงแนวโน้มยอดขายรายวันพร้อมจุดที่น่าสนใจ (Interesting Points)	29
ภาพประกอบ 12 Boxplot ของยอดขายรายวัน (Boxplot of Daily Sales)	31
ภาพประกอบ 13 การกระจายของยอดขายรายเดือน (Monthly Sales Distribution)	31
ภาพประกอบ 14 แนวโน้มยอดขายรายวันและค่าเฉลี่ยเคลื่อนที่ 30 วัน	32
ภาพประกอบ 15 การแบ่งชุดข้อมูลยอดขายรายวันออกเป็นชุดฝึกและชุดทดสอบ	34
ภาพประกอบ 16 แสดงการแบ่งชุดข้อมูล (Train-Test Split)	35
ภาพประกอบ 17 เปลี่ยนชื่อคอลัมน์ DATE และ ALL_SALE ให้เป็น ds และ y	36
ภาพประกอบ 18 องค์ประกอบของโมเดล Prophet แสดง แนวโน้มและฤดูกาลรายสัปดาห์	37
ภาพประกอบ 19 การปรับแต่งพารามิเตอร์ด้วย Grid Search	39
ภาพประกอบ 20 ตั้งค่า changepoint range = 0.9, n changepoints = 25 และ daily seasonality=True	40

ภาพประกอบ 21 ตั้งค่า changepoint range = 0.8, n changepoints = 10 และ daily seasonality=True	41
ภาพประกอบ 22 การกำหนดช่วงเวลาและโครงสร้างแบบจำลองการพยากรณ์ใน Tableau โดยใช้ แนวโน้มและฤดูกาลแบบส่วนเสริม	43
ภาพประกอบ 23 การกำหนดช่วงเวลาและโครงสร้างแบบจำลองการพยากรณ์ใน Tableau โดยใช้ แนวโน้มและฤดูกาลแบบทวิคูณ	45
ภาพประกอบ 24 ผลลัพธ์จากการทดลองที่ 1	49
ภาพประกอบ 25 ผลลัพธ์จากการทดลองที่ 2	50
ภาพประกอบ 26 ผลลัพธ์จากการทดลองที่ 3	52
ภาพประกอบ 27 ผลลัพธ์จากการทดลองที่ 4	53
ภาพประกอบ 28 ผลการพยากรณ์ยอดขายรวมจาก Tableau ใช้แนวโน้มและฤดูกาลแบบส่วนเสริม	54
ภาพประกอบ 29 กราฟแสดงข้อมูลพยากรณ์ยอดขายรวมจาก Tableau ใช้แนวโน้มและฤดูกาลแบบ ส่วนเสริม	55
ภาพประกอบ 30 ผลการพยากรณ์ยอดขายรวมจาก Tableau ใช้แนวโน้มและฤดูกาลแบบทวิคูณ ...	56
ภาพประกอบ 31 กราฟแสดงข้อมูลพยากรณ์ยอดขายรวมจาก Tableau ใช้แนวโน้มและฤดูกาลแบบ ทวิคูณ	57

บทที่ 1

บทนำ

1.1 ภูมิหลัง

ในยุคที่ภาคธุรกิจต้องเผชิญกับการแข่งขันที่เข้มข้นและสภาวะแวดล้อมที่เปลี่ยนแปลงอย่างรวดเร็ว ความสามารถในการวางแผนและตัดสินใจอย่างแม่นยำกลายเป็นปัจจัยสำคัญที่กำหนดความอยู่รอดและการเติบโตขององค์กร หนึ่งในกระบวนการสำคัญที่ช่วยเสริมสร้างศักยภาพด้านการบริหารจัดการ คือ “การพยากรณ์ยอดขาย” (Sales Forecasting) ซึ่งทำหน้าที่คาดการณ์ความต้องการของตลาดในอนาคต เพื่อสนับสนุนการวางแผนด้านการผลิต การจัดซื้อ การบริหารทรัพยากร และการกำหนดกลยุทธ์ทางธุรกิจในภาพรวมโดยเฉพาะอย่างยิ่งในธุรกิจบริการที่มีรูปแบบการดำเนินงานแบบวันต่อวัน (day-to-day operations) เช่น ธุรกิจร้านอาหาร ซึ่งมีลักษณะเฉพาะที่ต้องพึ่งพาการวางแผนวัตถุดิบสด เช่น นมสด โยเกิร์ต หรือเบเกอรี่ ที่มีอายุการเก็บรักษาจำกัด การจัดซื้อวัตถุดิบที่ไม่สอดคล้องกับความต้องการจริงสามารถนำไปสู่ความสูญเสียที่ชัดเจนในรูปของของเสีย ต้นทุนจม (sunk cost) หรือการสูญเสียโอกาสทางรายได้หากสินค้าขาดแคลนในช่วงเวลาที่ลูกค้าเข้ามาใช้บริการเป็นจำนวนมากความไม่แน่นอนทางยอดขายในแต่ละวันยังส่งผลกระทบต่อการบริหารจัดการทรัพยากรบุคคล เช่น การจัดตารางพนักงาน การเตรียมความพร้อมของหน้าร้าน และการบริหารเวลาให้บริการ หากไม่สามารถประเมินล่วงหน้าได้ว่าในแต่ละช่วงเวลาจะมีลูกค้าเข้าใช้บริการมากน้อยเพียงใด ก็จะเกิดทั้งความเสี่ยงด้านคุณภาพการบริการในช่วงที่ลูกค้าหนาแน่น และต้นทุนแรงงานส่วนเกินในช่วงเวลาที่ลูกค้าบางตา อีกทั้งทั้งพฤติกรรมของผู้บริโภคในธุรกิจร้านอาหารยังมีแนวโน้มที่เปลี่ยนแปลงตามวันในสัปดาห์ ช่วงเวลาในแต่ละวัน รวมถึงปัจจัยด้านสภาพอากาศ วันหยุดนักขัตฤกษ์ หรือเทศกาลต่าง ๆ ทำให้การวางแผนโดยอิงจากค่าเฉลี่ยในอดีตเพียงอย่างเดียวไม่สามารถตอบสนองต่อสถานการณ์จริงได้อย่างแม่นยำ ความสามารถในการคาดการณ์ยอดขายในระดับ “รายวัน” หรือ “รายชั่วโมง” จึงมีความจำเป็นอย่างยิ่ง เพื่อใช้ในการเตรียมการล่วงหน้า เช่น การสั่งซื้อวัตถุดิบในช่วงบ่ายสำหรับวันถัดไป การจัดรอบอบเบเกอรี่ หรือแม้แต่การตั้งเป้าหมายยอดขายของพนักงานรายวัน

ในอดีต การพยากรณ์ยอดขายมักอาศัยเทคนิคทางสถิติเชิงเส้น เช่น การถ่วงเฉลี่ยเคลื่อนที่ (Moving Average) หรือการวิเคราะห์แนวโน้ม (Trend Analysis) ซึ่งเหมาะสำหรับข้อมูลที่มีความต่อเนื่องเรียบง่าย อย่างไรก็ตาม เทคนิคเหล่านี้มีข้อจำกัดในการจับความเปลี่ยนแปลงที่เกิดขึ้นอย่างฉับพลัน หรือกรณีที่มีปัจจัยภายนอกเข้ามาแทรกซ้อน เช่น ฤดูกาล วันหยุด หรือกิจกรรมส่งเสริมการขายที่ส่งผลกระทบต่อพฤติกรรมการซื้อของลูกค้าในช่วงทศวรรษที่ผ่านมา เทคโนโลยีด้านการ

วิเคราะห์ข้อมูลขั้นสูง โดยเฉพาะ "การเรียนรู้ของเครื่อง" (Machine Learning) ได้กลายเป็นแนวทางใหม่ที่สามารถพยากรณ์ยอดขายได้อย่างแม่นยำมากขึ้น เนื่องจากสามารถเรียนรู้รูปแบบเชิงซ้อนจากข้อมูลในอดีตได้โดยไม่ต้องกำหนดสมการตายตัว โมเดลหนึ่งที่ได้รับการยอมรับอย่างกว้างขวาง คือ Prophet ซึ่งพัฒนาโดยบริษัท Meta (Facebook) โดยมีคุณสมบัติเด่นด้านการพยากรณ์ข้อมูลอนุกรมเวลา (time series) ที่สามารถจัดการกับความไม่เป็นเชิงเส้นของแนวโน้ม (non-linear trends) การแปรผันตามฤดูกาล (seasonality) และวันหยุดหรือเหตุการณ์พิเศษต่าง ๆ ได้อย่างมีประสิทธิภาพ

แบบจำลอง Prophet ถูกออกแบบให้ใช้งานง่ายแม้ไม่มีความรู้เชิงลึกด้านสถิติ และสามารถให้ผลลัพธ์ที่เข้าใจได้ง่ายผ่านภาพกราฟและการแสดงแนวโน้มแยกตามองค์ประกอบต่าง ๆ จึงเหมาะสมอย่างยิ่งกับธุรกิจร้านค้าแพที่ต้องเผชิญกับความผันผวนในระดับรายวัน และต้องการเครื่องมือพยากรณ์ที่สามารถนำไปใช้งานได้จริงในระดับปฏิบัติการ เช่น การจัดการวัตถุดิบรายวัน หรือการวางแผนทรัพยากรในช่วงวันหยุดยาว

การศึกษาวิจัยครั้งนี้จึงมีเป้าหมายเพื่อพัฒนาแบบจำลองการพยากรณ์ยอดขายรายวันของธุรกิจร้านค้าแพ โดยประยุกต์ใช้โมเดล Prophet เพื่อให้สามารถวางแผนการดำเนินงานล่วงหน้าได้อย่างแม่นยำมากยิ่งขึ้น เพิ่มประสิทธิภาพในการจัดซื้อ ลดต้นทุนจากของเสีย และยกระดับการให้บริการให้ตอบโจทย์กับพฤติกรรมของผู้บริโภคที่เปลี่ยนแปลงอยู่ตลอดเวลา ซึ่งเป็นองค์ประกอบสำคัญในการเสริมสร้างความสามารถในการแข่งขันของธุรกิจในระยะยาว

1.2 ความมุ่งหมายของการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์หลักเพื่อพัฒนาแนวทางการพยากรณ์ยอดขายรายวันในธุรกิจร้านค้าแพ โดยมุ่งหวังให้เกิดประโยชน์เชิงปฏิบัติในการวางแผนการดำเนินงานและการบริหารทรัพยากรอย่างมีประสิทธิภาพ โดยผู้วิจัยได้กำหนดความมุ่งหมายของงานวิจัยไว้ ดังนี้

1.2.1 เพื่อพัฒนาเทคนิคหรือกระบวนการที่สนับสนุนการตัดสินใจในธุรกิจร้านค้าแพ โดยมุ่งเน้นการวางแผนการสั่งซื้อสินค้า การจัดสรรพนักงานให้เหมาะสมกับปริมาณลูกค้าในแต่ละวัน ลดความเสี่ยงจากการมีสินค้ามากเกินไปหรือน้อยเกินไป และเพิ่มระดับความพึงพอใจของลูกค้าในการได้รับบริการที่รวดเร็วและสอดคล้องกับความต้องการ

1.2.2 เพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองการพยากรณ์ยอดขายระหว่างโมเดล Machine Learning (Prophet) กับการพยากรณ์จากฟังก์ชันในซอฟต์แวร์ Tableau โดยใช้เกณฑ์ด้านความแม่นยำ ความเหมาะสมในการใช้งานจริง และความสะดวกในการประยุกต์ใช้ในบริบทของร้านค้าแพ

1.3 ความสำคัญของงานวิจัย

ในบริบทของธุรกิจร้านกาแฟ การบริหารจัดการวัตถุดิบสด เช่น นมสด โยเกิร์ต และผลไม้ นับเป็นองค์ประกอบสำคัญของกระบวนการควบคุมต้นทุน เนื่องจากวัตถุดิบประเภทดังกล่าวมีอายุการเก็บรักษาสั้น การสั่งซื้อวัตถุดิบในปริมาณที่ไม่เหมาะสมอาจส่งผลให้เกิดต้นทุนจม (sunk cost) จากของเสียที่ไม่สามารถนำมาใช้ได้ทันเวลา หรือในทางตรงกันข้าม หากมีการจัดซื้อวัตถุดิบในปริมาณที่ต่ำกว่าความต้องการ อาจส่งผลให้ไม่สามารถตอบสนองความต้องการของลูกค้าได้อย่างเพียงพอ ส่งผลกระทบต่อรายได้และความพึงพอใจของผู้บริโภคโดยตรง

ด้วยเหตุนี้ การพยากรณ์ยอดขายรายวันอย่างแม่นยำจึงมีบทบาทสำคัญในกระบวนการวางแผนการจัดซื้อและการผลิต ตลอดจนการบริหารทรัพยากรมนุษย์ เพื่อให้สอดคล้องกับความต้องการที่เกิดขึ้นจริงในแต่ละวัน อย่างไรก็ตาม การพยากรณ์ยอดขายในลักษณะดังกล่าวมีความซับซ้อนเนื่องจากพฤติกรรมผู้บริโภคมีความแปรผันตามปัจจัยหลายประการ ได้แก่ ความแตกต่างของวันในสัปดาห์ วันหยุดนักขัตฤกษ์ ช่วงเวลาของเดือน ตลอดจนปัจจัยด้านสภาพอากาศและกิจกรรมส่งเสริมการขาย ซึ่งล้วนส่งผลต่อปริมาณการซื้อสินค้าและรูปแบบการบริโภคเทคนิคการวิเคราะห์เชิงสถิติแบบดั้งเดิม เช่น การถ่วงเฉลี่ยเคลื่อนที่ หรือการวิเคราะห์แนวโน้มเชิงเส้น แม้จะสามารถนำมาใช้ในบางกรณี แต่ยังไม่สามารถรองรับความซับซ้อนและความแปรผันของข้อมูลในลักษณะดังกล่าวได้อย่างมีประสิทธิภาพมากนัก ปัจจุบัน เทคโนโลยีการเรียนรู้ของเครื่อง (Machine Learning) จึงได้รับการยอมรับอย่างแพร่หลายในฐานะเครื่องมือที่สามารถวิเคราะห์ข้อมูลจำนวนมากและเรียนรู้รูปแบบพฤติกรรมที่ซับซ้อนในข้อมูลอนุกรมเวลาได้อย่างมีประสิทธิภาพ

แบบจำลอง Prophet ซึ่งพัฒนาโดยบริษัท Meta (Facebook) เป็นหนึ่งในแบบจำลองที่ได้รับความนิยมในการพยากรณ์ข้อมูลอนุกรมเวลา ด้วยคุณสมบัติที่สามารถรองรับแนวโน้มไม่เป็นเชิงเส้น (non-linear trend), ความผันผวนตามฤดูกาล (seasonality), และการปรับค่าจากวันหยุดหรือเหตุการณ์พิเศษต่าง ๆ ได้อย่างยืดหยุ่น ขณะเดียวกัน ซอฟต์แวร์ Tableau ซึ่งเป็นเครื่องมือสำหรับการวิเคราะห์ข้อมูลในเชิงภาพ (visual analytics) ยังมีฟังก์ชันการพยากรณ์ในตัวที่สามารถประมวลผลและแสดงผลได้อย่างรวดเร็ว โดยเหมาะสมกับการใช้งานเบื้องต้นในระดับธุรกิจ

การเปรียบเทียบประสิทธิภาพของแบบจำลอง Prophet กับการพยากรณ์ใน Tableau จึงเป็นแนวทางที่สำคัญในการประเมินความเหมาะสมของเครื่องมือแต่ละประเภทในเชิงประจักษ์ โดยพิจารณาทั้งในด้านความแม่นยำ ความยืดหยุ่นในการปรับใช้ และความสะดวกในการนำไปใช้งานจริงในสภาพแวดล้อมของธุรกิจร้านกาแฟ

งานวิจัยฉบับนี้จึงมีความสำคัญในฐานะที่เป็นการประยุกต์ใช้เทคโนโลยีการพยากรณ์สมัยใหม่เพื่อลดความสูญเสียจากอุบัติเหตุที่ไม่เหมาะสม เพิ่มประสิทธิภาพในการวางแผนการดำเนินงาน ตลอดจนส่งเสริมการตัดสินใจเชิงกลยุทธ์บนพื้นฐานของข้อมูลจริง โดยมีเป้าหมายเพื่อยกระดับขีดความสามารถในการแข่งขันของธุรกิจร้านค้ากาแฟ และสามารถนำผลลัพธ์ที่ได้ไปประยุกต์ใช้กับธุรกิจอื่นที่มีลักษณะคล้ายคลึงกันต่อไป

1.4 ขอบเขตการวิจัย

1.4.1 ประชากรที่ใช้ในการวิจัย

ประชากรที่ใช้ในการวิจัยนี้คือข้อมูลยอดขายของธุรกิจร้านค้ากาแฟรายวันที่ได้จากฐานข้อมูลทางธุรกิจ โดยคัดเลือกข้อมูลที่ครอบคลุมระยะเวลาการขายตั้งแต่ช่วงเดือนพฤษภาคม 2567 - มิถุนายน 2568

1.4.2 กลุ่มตัวอย่างที่ใช้ในการวิจัย

กลุ่มตัวอย่างที่นำมาใช้ในการวิจัยคือข้อมูลยอดขายรายวัน โดยข้อมูลที่ใช้จะครอบคลุมข้อมูลยอดขายในช่วงวันธรรมดาและวันหยุด เพื่อให้ได้กลุ่มข้อมูลที่หลากหลายและสามารถนำไปใช้ในการพยากรณ์ยอดขายได้อย่างแม่นยำ

1.4.3 ตัวแปรที่ศึกษา

ตัวแปรที่นำมาใช้ในการวิจัยคือข้อมูลยอดขายรายวันของธุรกิจร้านค้ากาแฟ โดยแบ่งรายละเอียดของตัวแปรดังนี้

1.4.3.1 ตัวแปรต้น

วันที่ของข้อมูลรายวันโดยเริ่มจากวันที่ 1 พฤษภาคม 2566 – 30 มิถุนายน 2567 โดยข้อมูลดังกล่าวจะถูกนำมาเป็นปัจจัยนำเข้าในการทำแบบจำลองการพยากรณ์

1.4.3.2 ตัวแปรตาม

ยอดขายในแต่ละวัน โดยจะนำมาเป็นตัวแปรที่ต้องการสำหรับการพยากรณ์ในงานวิจัยครั้งนี้

1.4.3.3 ตัวแปรร่วม

ปัจจัยที่อาจมีผลต่อยอดขาย เช่น วันหยุดนักขัตฤกษ์, ช่วงวันทำงานในสัปดาห์สุดท้ายของเดือนซึ่งเป็นช่วงเร่งปิดยอดขายประจำเดือน, วันก่อนวันหยุดนักขัตฤกษ์ที่ร้านค้าอาจมีการสต็อกสินค้าไว้รองรับนักท่องเที่ยว หรือวันหลังวันหยุดนักขัตฤกษ์ที่สินค้าอาจจะหมดเร็วจากการบริการนักท่องเที่ยวในช่วงวันหยุด เป็นต้น

1.5 นิยามศัพท์เฉพาะ

1.5.1 การพยากรณ์ยอดขาย (Sales Forecasting)

หมายถึง กระบวนการคาดการณ์ปริมาณยอดขายในอนาคตโดยอาศัยข้อมูลยอดขายในอดีต ข้อมูลปัจจัยแวดล้อม และพฤติกรรมผู้บริโภค เพื่อช่วยในการวางแผนการจัดซื้อ การผลิต การจัดสรรทรัพยากร และการบริหารงานขายของธุรกิจให้มีประสิทธิภาพ

1.5.2 ข้อมูลอนุกรมเวลา (Time Series Data)

หมายถึง ชุดข้อมูลที่มีการบันทึกตามช่วงเวลาอย่างต่อเนื่อง เช่น รายวัน รายเดือน หรือรายปี ซึ่งข้อมูลจะถูกนำมาวิเคราะห์เพื่อหาแนวโน้ม ความผันผวน และรูปแบบตามฤดูกาล เพื่อนำไปใช้ในการพยากรณ์เหตุการณ์ในอนาคต

1.5.3 แบบจำลอง Prophet

หมายถึง อัลกอริธึมการพยากรณ์ที่พัฒนาโดย Meta (Facebook) ซึ่งถูกออกแบบมาให้สามารถจัดการกับข้อมูลอนุกรมเวลาที่มีแนวโน้มตามฤดูกาลและความเปลี่ยนแปลงซับซ้อนได้ดี มีจุดเด่นที่ใช้งานง่าย ไม่ต้องปรับพารามิเตอร์มาก และเหมาะสำหรับข้อมูลที่มีความซับซ้อนในระดับปานกลางถึงสูง

1.5.4 Tableau

หมายถึง เครื่องมือสำหรับการแสดงผลข้อมูล (Data Visualization) ที่สามารถเชื่อมโยงกับแหล่งข้อมูลและสร้างแดชบอร์ดแบบอินเทอร์แอคทีฟ รวมถึงมีฟังก์ชันในการพยากรณ์แบบง่าย (Forecast) โดยใช้เทคนิคเชิงสถิติเบื้องต้น เช่น Exponential Smoothing

1.5.5 สินค้าที่เน่าเสียได้ง่าย (Perishable Goods)

หมายถึง วัตถุดิบหรือสินค้าที่มีอายุการใช้งานสั้น และต้องได้รับการจัดเก็บและใช้งานภายในระยะเวลาที่กำหนด เช่น นมสด โยเกิร์ต ผลไม้ หรือขนมอบสด หากเก็บไว้นานเกินไปอาจทำให้เกิดความเสียหาย ส่งผลให้ต้นทุนจมและขาดทุนทางธุรกิจ

1.5.6 ความแม่นยำในการพยากรณ์ (Forecast Accuracy)

หมายถึง ความสามารถของแบบจำลองในการคาดการณ์ผลลัพธ์ที่ใกล้เคียงกับค่าจริง โดยวัดผ่านค่าตัวชี้วัด เช่น Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), และ Mean Absolute Percentage Error (MAPE)

1.5.7 วันหยุดนักขัตฤกษ์ (Public Holidays)

หมายถึง วันที่กำหนดโดยภาครัฐให้เป็นวันหยุดราชการ ซึ่งมีแนวโน้มส่งผลต่อพฤติกรรม การบริโภคและอุปสงค์ของสินค้าในธุรกิจร้านค้าปลีก เช่น ความหนาแน่นของลูกค้า หรือการสต็อก วัตถุดิบมากกว่าปกติ

1.5.8 ความผันผวนของยอดขาย (Sales Volatility)

หมายถึง ระดับความไม่แน่นอนหรือการเปลี่ยนแปลงของยอดขายในแต่ละวัน ซึ่งอาจเกิด จากปัจจัยภายนอก เช่น สภาพอากาศ กิจกรรมส่งเสริมการขาย หรือพฤติกรรมของลูกค้า

1.6 กรอบแนวคิดงานวิจัย

งานวิจัยนี้เน้นศึกษาการพยากรณ์ยอดขายรายวัน โดยใช้ข้อมูลยอดขายที่เก็บรวบรวมใน แต่ละวัน ด้วยการใช้เทคนิคการวิเคราะห์ข้อมูลที่ทันสมัยอย่าง Machine Learning และ เทคนิค การพยากรณ์จาก Tableau เพื่อให้ผลลัพธ์ที่ได้มีความแม่นยำมากที่สุด โดยมีองค์ประกอบหลัก ของกรอบแนวคิด ดังนี้

1.6.1 ข้อมูลนำเข้า (Input Data)

ข้อมูลยอดขายที่เก็บรวบรวมในแต่ละวันตั้งแต่วันที่ 1 พฤษภาคม 2566 – 30 มิถุนายน 2567 ข้อมูลเหล่านี้จะถูกนำมาวิเคราะห์เพื่อพยากรณ์ยอดขายรายวัน 30 วันล่วงหน้าโดยใช้ข้อมูล ยอดขายในอดีตเป็นข้อมูลนำเข้า

1.6.2 การประมวลผลข้อมูล (Data Processing)

1.6.2.1 การรวมไฟล์และจัดรูปแบบข้อมูลนำข้อมูลยอดขายรายวันจากหลาย ไฟล์ Excel มารวมกันเป็น Data Frame เดียวและแปลงค่าวันที่จากพุทธศักราช (พ.ศ.) เป็น คริสต์ศักราช (ค.ศ.) เพื่อให้สามารถประมวลผลได้อย่างถูกต้อง

1.6.2.2 การทำความสะอาดข้อมูล (Data Cleaning) ลบหรือแทนค่าที่หายไป (Missing Values) เช่น ค่า null หรือ "-" ด้วยค่า 0 ตรวจสอบค่าผิดปกติ (Outliers) ด้วยวิธี Boxplot และ Histogram และพิจารณาการจัดการค่าผิดปกติให้เหมาะสม

1.6.2.3 การแปลงข้อมูล (Data Transformation) แปลงค่าที่เป็นข้อความให้อยู่ ในรูปแบบตัวเลข สร้างตัวแปรใหม่ เช่น ค่าเฉลี่ย 7 วัน(7-day moving average), ดัชนีวันใน สัปดาห์ (day-of-week)

1.6.2.4 การเลือกคุณลักษณะ (Feature Engineering/Selection) เลือกฟีเจอร์ที่ เกี่ยวข้องกับพฤติกรรมการขาย เช่น ยอดขายวันก่อนหน้า วันในสัปดาห์ วันสิ้นเดือน วันหยุดนักขัต ฤกษ์ เป็นต้น

1.6.3 การสร้างแบบจำลองการพยากรณ์ (Model Building)

1.6.3.1 การเลือกแบบจำลอง Prophet เป็นแบบจำลองเชิงอนุกรมเวลา (Time Series Forecasting Model) ที่สามารถระบุแนวโน้มตามฤดูกาลและการเปลี่ยนแปลงอย่างเฉียบพลันได้ดี เหมาะกับการนำมาใช้พยากรณ์ยอดขายรายวันในธุรกิจร้านค้าแพที่มีรูปแบบยอดขายไม่คงที่

1.6.3.2 การใช้ Tableau ใช้ความสามารถของซอฟต์แวร์ Tableau ในการประเมินและวิเคราะห์แนวโน้มของยอดขายจากข้อมูลในอดีต โดยใช้ฟังก์ชันพยากรณ์ในตัวที่สามารถสร้างเส้นคาดการณ์และช่วงความเชื่อมั่น (Confidence Interval) ได้ในรูปแบบที่เข้าใจง่าย เหมาะสำหรับการใช้งานเชิงธุรกิจ

1.6.3.3 เปรียบเทียบโมเดลทำการเปรียบเทียบประสิทธิภาพของ Prophet และ Tableau เพื่อดูว่าแบบจำลองใดสามารถพยากรณ์ยอดขายได้แม่นยำและเหมาะสมกับการใช้งานจริงในบริบทของร้านค้าแพมากที่สุด

1.6.4 การฝึกฝนและทดสอบแบบจำลอง (Training & Testing)

1.6.4.1 การแบ่งชุดข้อมูล ข้อมูลที่ใช้ในการวิจัยจะถูกแบ่งออกเป็น 2 ส่วน ได้แก่ ชุดข้อมูลสำหรับฝึกฝนแบบจำลอง (Training set) และชุดข้อมูลสำหรับทดสอบแบบจำลอง (Test set) โดยทั่วไปจะใช้สัดส่วนประมาณ 80:20 เพื่อให้สามารถประเมินประสิทธิภาพของแบบจำลองได้อย่างเหมาะสม

1.6.4.2 การฝึกโมเดล Prophet ทำการฝึกแบบจำลอง Prophet โดยใช้ข้อมูล Training set เพื่อเรียนรู้แนวโน้ม (trend), ความเป็นฤดูกาล (seasonality) และองค์ประกอบอื่น ๆ ของยอดขาย จากนั้นจึงนำไปทำนายชุด Test set เพื่อประเมินความแม่นยำ

1.6.4.3 การสร้าง Forecast ด้วย Tableau ใช้ข้อมูลชุดเดียวกันกับ Prophet ใน Tableau เพื่อสร้างแบบจำลองพยากรณ์ และทำการแยกชุดข้อมูลออกเป็นช่วง Training และ Testing เช่นเดียวกัน เพื่อให้สามารถเปรียบเทียบผลลัพธ์ได้อย่างเป็นธรรม

1.6.4.4 การจัดการกับค่า Missing และ Outliers ก่อนทำการฝึกโมเดล ข้อมูลจะถูกตรวจสอบค่าที่หายไปหรือค่าผิดปกติ และทำการจัดการให้เหมาะสม เช่น การเติมค่าเฉลี่ยหรือการลบข้อมูลที่ผิดปกติออก

1.6.4.5 การตรวจสอบความสม่ำเสมอของช่วงเวลาตรวจสอบให้แน่ใจว่าข้อมูลที่ใช้มีความต่อเนื่องในเชิงเวลา เช่น ไม่มีวันที่ขาดหาย เพื่อให้การเรียนรู้ของแบบจำลองมีประสิทธิภาพสูงสุด

1.6.5 การประเมินประสิทธิภาพของแบบจำลอง (Model Evaluation)

1.6.5.1 ตัวชี้วัดประสิทธิภาพ ใช้ดัชนีวัดความแม่นยำหลัก 3 ตัว ได้แก่ Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) และ Mean Absolute Percentage Error (MAPE) เพื่อเปรียบเทียบผลการพยากรณ์ของแต่ละโมเดลในชุดข้อมูลทดสอบ

1.6.5.2 การเปรียบเทียบผลลัพธ์ เปรียบเทียบค่าความคลาดเคลื่อนที่เกิดขึ้นระหว่างค่าจริง (Actual) และค่าที่พยากรณ์ได้ (Forecasted) จากโมเดล Prophet และ Tableau เพื่อประเมินว่าแบบจำลองใดให้ผลการพยากรณ์ที่แม่นยำและเสถียรกว่า

1.6.5.3 การแสดงผลเชิงภาพสร้างกราฟเปรียบเทียบเส้นยอดขายจริงกับเส้นที่พยากรณ์ได้จากแต่ละโมเดล เพื่อวิเคราะห์เชิงคุณภาพถึงแนวโน้ม ความสม่ำเสมอ และการตอบสนองต่อการเปลี่ยนแปลงของข้อมูลจริง

1.6.5.4 การสรุปผลใช้ผลลัพธ์เชิงปริมาณจากการวัดค่าความผิดพลาด และผลลัพธ์เชิงคุณภาพจากกราฟวิเคราะห์แนวโน้ม เพื่อสรุปว่าโมเดลใดเหมาะสมที่สุดต่อการนำไปใช้จริงในบริบทของธุรกิจร้านกาแฟ

1.7 สมมติฐานในการวิจัย

1.7.1 ข้อมูลยอดขายรายวันย้อนหลังสามารถนำมาสร้างแบบจำลองเพื่อพยากรณ์ยอดขายล่วงหน้าได้อย่างแม่นยำ ด้วยการใช้เทคนิคการเรียนรู้ของเครื่อง (Machine Learning)

1.7.2 แบบจำลอง Prophet ซึ่งออกแบบมาเพื่อจัดการข้อมูลอนุกรมเวลา (Time Series) สามารถพยากรณ์ยอดขายของร้านกาแฟได้อย่างแม่นยำมากกว่าการใช้ฟังก์ชันการพยากรณ์ใน Tableau โดยเฉพาะในบริบทที่ต้องวิเคราะห์ความเป็นฤดูกาล (Seasonality) และแนวโน้มในระยะยาว

1.7.3 การพยากรณ์ยอดขายรายวันโดยไม่ใช้ปัจจัยภายนอก (univariate forecasting) สามารถให้ผลลัพธ์ที่มีระดับความแม่นยำเพียงพอต่อการใช้งานในเชิงปฏิบัติ และหากมีการนำตัวแปรร่วมมาใช้อาจช่วยเพิ่มประสิทธิภาพของแบบจำลองได้มากยิ่งขึ้น

บทที่ 2

ทบทวนวรรณกรรม

ในการดำเนินการวิจัยครั้งนี้ ผู้วิจัยศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องสำหรับการทำวิจัยนี้ โดยได้นำเสนอตามหัวข้อดังต่อไปนี้

1. การพยากรณ์ยอดขาย (Sale Forecast)
2. แบบจำลองที่ใช้ในงานวิจัย
3. โปรแกรมสำเร็จรูปที่ใช้ในการพยากรณ์
4. งานวิจัยที่เกี่ยวข้อง

2.1 การพยากรณ์ยอดขาย (Sales Forecasting)

การพยากรณ์ยอดขาย (Sales Forecasting) เป็นกระบวนการสำคัญที่ช่วยให้ธุรกิจดำเนินการกิจการอย่างมีประสิทธิภาพ เพื่อที่สามารถคาดการณ์ยอดขายสำหรับรองรับความต้องการของลูกค้า โดยเฉพาะในสภาวะการแข่งขันทางการตลาดที่มีความผันผวนสูง เช่น ธุรกิจร้านอาหาร ซึ่งต้องบริหารจัดการวัตถุดิบที่มีอายุสั้น เช่น นมสด โยเกิร์ต หรือผลไม้ การพยากรณ์ยอดขายที่แม่นยำในระดับรายวันจึงมีบทบาทในการลดของเสีย ควบคุมต้นทุน และเพิ่มความพึงพอใจให้กับลูกค้าได้โดยตรง

การพยากรณ์ยอดขายยังช่วยส่งเสริมกระบวนการบริหารจัดการในด้านอื่น เช่น การวางแผนการผลิต การจัดการสินค้าคงคลัง และการบริหารทรัพยากรบุคคล

การมีข้อมูลคาดการณ์ล่วงหน้า ทำให้ธุรกิจสามารถตอบสนองต่อความต้องการของตลาดได้อย่างเหมาะสมและรวดเร็ว ลดความเสี่ยงจากการขาดแคลนสินค้า หรือสินค้าค้างสต็อกที่เกินความจำเป็น

2.1.1 ความสำคัญของการพยากรณ์ยอดขาย

2.1.1.1 สนับสนุนการจัดการสินค้าคงคลังให้เหมาะสม ลดโอกาสการเกิดของเสียหรือสินค้าหมดสต็อก

2.1.1.2 ช่วยในการบริหารจัดการพนักงานได้อย่างเหมาะสม เช่น การจัดทำตารางทำงานให้เพียงพอกับความต้องการในช่วงยอดขายสูง

2.1.1.3 เป็นข้อมูลสนับสนุนสำหรับการปรับกลยุทธ์การตลาด เช่น การทำโปรโมชั่น หรือการเปิดสาขาชั่วคราวตามฤดูกาล

2.1.1.4 เป็นพื้นฐานสำหรับการวางแผนการผลิตในลักษณะที่สอดคล้องกับ
แนวโน้มความต้องการของลูกค้า

2.1.2 เทคนิคการพยากรณ์ยอดขาย

เทคนิคที่นิยมใช้ในการพยากรณ์ยอดขายในปัจจุบันครอบคลุมทั้งแบบจำลองเชิงสถิติ
ดั้งเดิม และแบบจำลองที่อิงกับ Machine Learning ซึ่งสามารถรองรับข้อมูลขนาดใหญ่และ
ซับซ้อนได้ดี ได้แก่

2.1.2.1 Exponential Smoothing และ Moving Average ใช้สำหรับการ
คาดการณ์ที่ไม่มีความผันผวนสูง

2.1.2.2 ARIMA และ SARIMA เหมาะสำหรับข้อมูลที่มีแนวโน้มและฤดูกาล
ชัดเจน

2.1.2.3 แบบจำลอง Machine Learning เช่น Random Forest, XG Boost,
Gradient Boosting, และ Linear Regression ซึ่งสามารถจับความสัมพันธ์ที่ไม่เป็นเชิงเส้นได้

2.1.2.4 แบบจำลองอนุกรมเวลาเชิงลึก (Time Series Models) เช่น Prophet
ของ Facebook ซึ่งสามารถจัดการกับข้อมูลที่มีลักษณะไม่ต่อเนื่อง มีวันหยุด หรือความเป็น
ฤดูกาลได้ดี

2.2 แบบจำลองที่ใช้ในงานวิจัย

งานวิจัยนี้เลือกใช้แบบจำลองการพยากรณ์ยอดขาย 2 รูปแบบหลักได้แก่

2.2.1 prophet

เป็นเครื่องมือการพยากรณ์อนุกรมเวลาที่ได้รับการพัฒนาโดยทีมวิจัยของ Facebook ซึ่ง
ออกแบบมาเพื่อให้สามารถใช้งานได้ง่าย มีความยืดหยุ่น และเหมาะกับข้อมูลที่มีฤดูกาลซ้ำๆ
(seasonality) มีแนวโน้มการเปลี่ยนแปลง (trend) และสามารถรับมือกับวันหยุดหรือเหตุการณ์
พิเศษได้ดี โดยเฉพาะชุดข้อมูลยอดขายรายวันของธุรกิจ เช่น ร้านกาแฟ ที่มักมีพฤติกรรมของ
ยอดขายที่ซับซ้อนและมีการเปลี่ยนแปลงตามช่วงเวลา Prophet ทำงานโดยแยกองค์ประกอบของ
อนุกรมเวลาออกเป็น 3 ส่วนหลัก

สมการ

$$y(t) = g(t) + s(t) + h(t) + \epsilon t$$

โดยที่

$g(t)$	คือ แนวโน้ม (trend) ของข้อมูลในระยะยาว เช่น การเพิ่มขึ้นของยอดขายตามฤดูกาล
$s(t)$	คือ ฤดูกาล (seasonality) ซึ่งเป็นรูปแบบที่เกิดซ้ำ เช่น ยอดขายสูงทุกวันศุกร์ หรือทุกสิ้นเดือน
$h(t)$	คือ ผลกระทบจากวันหยุดหรือเหตุการณ์พิเศษ (holiday effects) เช่น ยอดขายเพิ่มในวันสงกรานต์
ϵ_t	คือ ส่วนของความผิดพลาดหรือ noise

2.2.1.1 แนวโน้ม (Trend) Prophet สามารถเลือกใช้ 2 รูปแบบของแนวโน้มได้แก่

2.2.1.1.1 แบบเส้นตรง (linear trend): เหมาะกับข้อมูลที่ยอดขายเติบโตหรือลดลงอย่างสม่ำเสมอ

2.2.1.1.2 แบบลอจิสติก (logistic growth): ใช้ในกรณีที่ยอดขายมีการอิ่มตัวในช่วงปลาย เช่น สินค้ามีข้อจำกัดด้านปริมาณสามารถกำหนด “จุดเปลี่ยนแนวโน้ม” (changepoints) ได้โดยอัตโนมัติหรือระบุด้วยตนเอง เพื่อให้แบบจำลองสามารถจับการเปลี่ยนแปลงสำคัญ เช่น หลังเปิดตัวเมนูใหม่

2.2.1.2 ความเป็นฤดูกาล (Seasonality) Prophet รองรับทั้งฤดูกาลรายวัน รายสัปดาห์ รายเดือน โดยใช้ฟังก์ชันไซน์และโคไซน์ (Fourier series) ซึ่งเหมาะกับข้อมูลที่มีรูปแบบซ้ำๆ เช่น ยอดขายร้านกาแฟสูงช่วงเช้าและลดลงตอนบ่าย

2.2.1.3 ผลกระทบจากวันหยุด (Holiday Effects) สามารถกำหนดวันหยุดพิเศษ เช่น วันสงกรานต์ วันหยุดนักขัตฤกษ์ หรืองานเทศกาล เพื่อให้แบบจำลองเข้าใจว่าช่วงเวลานั้นจะส่งผลต่อยอดขายสูงขึ้นหรือลดลง

2.2.1.4 ความสามารถในการปรับพารามิเตอร์ (Tuning Parameters) Prophet เปิดให้ผู้ใช้งานปรับค่าต่างๆ เพื่อควบคุมความยืดหยุ่นของแบบจำลอง เช่น

1. Changepoint prior scale ควบคุมการตอบสนองต่อการเปลี่ยนแนวโน้ม
 2. Seasonality prior scale ควบคุมความแรงของฤดูกาล
 3. Holidays prior scale ควบคุมผลกระทบจากวันหยุด
 4. Seasonality mode เลือกการรวมฤดูกาลแบบ additive หรือ multiplicative
- ข้อดีของ Prophet

1. ใช้งานง่ายและไม่จำเป็นต้องมีความรู้ทางสถิติขั้นสูง
2. จัดการกับข้อมูลที่มี missing value หรือ gap ได้
3. เหมาะกับธุรกิจที่มีข้อมูลในลักษณะเวลา เช่น ร้านกาแฟที่บันทึกยอดขายเป็นรายวัน
4. สามารถประเมินความไม่แน่นอนของผลพยากรณ์ได้ (uncertainty intervals)

ข้อจำกัดของ Prophet

1. ไม่สามารถเรียนรู้ความสัมพันธ์จากตัวแปรอิสระอื่นได้ (univariate forecasting model)
2. ไม่สามารถพิจารณาความสัมพันธ์หลายตัวแปรพร้อมกันได้เช่นเดียวกับแบบจำลอง Machine Learning อื่นๆ
3. อาจไม่เหมาะกับข้อมูลที่มีความซับซ้อนมากเกินไป เช่น รายชั่วโมงแบบละเอียดหากไม่มีความเป็นฤดูกาลชัดเจน

การประยุกต์ในบริบทของงานวิจัยนี้

ในงานวิจัยนี้ Prophet ถูกนำมาใช้ในการพยากรณ์ยอดขายรายวันของร้านกาแฟ โดยอาศัยข้อมูลย้อนหลังตั้งแต่เดือนพฤษภาคม 2566 – มิถุนายน 2567 เพื่อสร้างแบบจำลอง โดยมีการปรับค่าพารามิเตอร์ผ่าน Grid Search เพื่อให้ได้ผลลัพธ์ที่แม่นยำที่สุด จากนั้นจึงเปรียบเทียบกับวิธี Tableau เพื่อหาวิธีที่เหมาะสมที่สุดสำหรับการใช้งานจริงในธุรกิจร้านกาแฟ

2.2.2 Tableau Forecast

Tableau เป็นซอฟต์แวร์ด้านการวิเคราะห์และแสดงผลข้อมูลเชิงภาพที่มีความสามารถในการสร้างฟังก์ชันการพยากรณ์เบื้องต้น โดยอิงจากวิธีทางสถิติเช่น Exponential Smoothing หรือ Autoregressive Models ผู้ใช้สามารถปรับช่วงการพยากรณ์และดูผลลัพธ์ได้อย่างรวดเร็วผ่านแดชบอร์ด ซึ่งเหมาะกับผู้ใช้งานทั่วไปและธุรกิจที่ต้องการการตัดสินใจแบบเร่งด่วน การนำ Tableau มาใช้พยากรณ์ยอดขายรายวันในงานวิจัยนี้ช่วยให้สามารถประเมินแนวโน้มยอดขายได้ง่าย และสามารถเปรียบเทียบประสิทธิภาพกับแบบจำลอง Prophet ได้อย่างชัดเจน

2.2.3 Mean Absolute Error (MAE)

Mean Absolute Error (MAE) คือ ค่าความผิดพลาดเฉลี่ยแบบสัมบูรณ์ระหว่างค่าที่แบบจำลองพยากรณ์ (ค่าที่คาดการณ์) กับค่าจริงในชุดข้อมูลทดสอบ โดยจะคำนวณความแตกต่างระหว่างค่าจริงกับค่าที่ทำนาย แล้วนำค่าผลต่างที่เป็น "ค่าสัมบูรณ์" (Absolute Value) มาหาค่าเฉลี่ย

MAE จะให้ค่าที่ไม่ขึ้นอยู่กับการทิศทางของความคลาดเคลื่อน กล่าวคือ ไม่ว่าจะเป็นค่าทำนายที่ต่ำกว่าหรือสูงกว่าค่าจริง ค่าความผิดพลาดจะถูกนับเป็นค่าบวกเสมอ (ไม่ลบด่างกัน) จึงสะท้อนขนาดความคลาดเคลื่อนของการพยากรณ์ได้อย่างตรงไปตรงมา

ค่า MAE ที่น้อยแสดงว่าแบบจำลองสามารถทำนายค่าได้ใกล้เคียงกับค่าจริงมากขึ้น ซึ่งมักใช้เปรียบเทียบประสิทธิภาพของโมเดลพยากรณ์ในงานวิเคราะห์อนุกรมเวลา (Time Series Forecasting) อย่างแพร่หลาย

สมการ

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_{\text{actual}} - y_{\text{predicted}}|$$

โดยที่

n

คือ จำนวนข้อมูลทั้งหมดที่ใช้ในการประเมิน

y_{actual}

คือ ค่าจริงที่เกิดขึ้นจริงในชุดข้อมูล

$y_{\text{predicted}}$

คือ ค่าที่ได้จากแบบจำลองพยากรณ์

2.2.4 Mean Squared Error (MSE)

ค่าคลาดเคลื่อนกำลังสองเฉลี่ย เป็นดัชนีชี้วัดที่ใช้ประเมินประสิทธิภาพของแบบจำลองการพยากรณ์ โดยเฉพาะในกรณีที่ต้องการให้แบบจำลองมีความแม่นยำในการทำนายค่าของตัวแปรตามในลักษณะเชิงปริมาณ MSE ทำหน้าที่วัดความแตกต่างระหว่างค่าที่พยากรณ์โดยแบบจำลอง (Predicted Values) กับค่าจริงที่เกิดขึ้น (Actual Values) โดยการนำความแตกต่างนั้นมายกกำลังสอง เพื่อขจัดปัญหาเครื่องหมายลบ (ซึ่งเกิดจากค่าพยากรณ์ต่ำกว่าค่าจริง) แล้วนำค่ากำลังสองของผลต่างทั้งหมดมาเฉลี่ย

ค่าที่ได้จาก MSE เป็นค่าบวกเสมอ (≥ 0) และยังมีค่าเข้าใกล้ 0 มากเท่าใด แสดงว่าแบบจำลองมีความแม่นยำมากขึ้น เนื่องจากค่าพยากรณ์เบี่ยงเบนจากค่าจริงน้อยลง

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_{\text{actual}} - y_{\text{predicted}})^2$$

โดยที่

n

คือ จำนวนข้อมูลทั้งหมดที่ใช้ประเมิน

y_{actual}	คือ ค่าจริง
$y_{\text{predicted}}$	คือ ค่าที่แบบจำลองทำนาย

2.2.5 R^2 (Coefficient of Determination / R-Squared / R^2 Score)

สัมประสิทธิ์การตัดสินใจ เป็นหนึ่งในตัวชี้วัดที่นิยมใช้ในการประเมินความสามารถของแบบจำลองการพยากรณ์ โดยแสดงให้เห็นว่า แบบจำลองสามารถอธิบายความแปรผันของข้อมูลได้มากน้อยเพียงใด โดยมีค่าระหว่าง 0 ถึง 1 (หรืออาจติดลบได้ในกรณีแบบจำลองแย่กว่าใช้ค่าเฉลี่ย)

สมการ

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_{\text{actual}} - y_{\text{predicted}})^2}{\sum_{i=1}^n (y_{\text{actual}} - \bar{y})^2}$$

$\bar{y}$	โดยที่ คือ ค่าเฉลี่ยของค่าจริงทั้งหมด
y_{actual}	คือ ค่าจริง
$y_{\text{predicted}}$	คือ ค่าที่แบบจำลองพยากรณ์ได้

การแปลผลค่า R^2

$R^2 = 1$ หมายถึง แบบจำลองสามารถอธิบายความแปรผันของข้อมูลได้ทั้งหมด กล่าวคือ พยากรณ์ได้ตรงกับค่าจริง 100%

$R^2 = 0$ หมายถึง แบบจำลองไม่สามารถอธิบายความแปรผันของข้อมูลได้เลย ดีพอ ๆ กับการใช้เพียงค่าเฉลี่ยของข้อมูลเป็นตัวพยากรณ์

$R^2 < 0$ หมายถึง แบบจำลองแย่กว่าการใช้ค่าเฉลี่ยของข้อมูลในการพยากรณ์ (มักพบในแบบจำลองที่ overfit หรือตั้งค่าไม่เหมาะสม)

2.3 งานวิจัยที่เกี่ยวข้อง

มีงานวิจัยจำนวนมากที่นำแบบจำลองทางสถิติและ Machine Learning มาประยุกต์ใช้ในการพยากรณ์ยอดขายในหลากหลายบริบท เช่น ธุรกิจค้าปลีก ธุรกิจอาหารและเครื่องดื่ม รวมถึงธุรกิจแฟชั่น โดยเน้นทั้งในด้านความแม่นยำ ความสามารถในการจัดการข้อมูลเชิงเวลา และการประยุกต์ใช้เครื่องมือวิเคราะห์สมัยใหม่

2.3.1 Prediction of Footwear Demand Using Prophet and SARIMA โดย Pablo Negre และคณะ (2024) (Negre et al., 2024)

งานวิจัยนี้มีเป้าหมายเพื่อพัฒนาและเปรียบเทียบความสามารถของแบบจำลองการพยากรณ์ยอดขายรองเท้า โดยใช้สองโมเดลหลัก ได้แก่ Prophet ซึ่งเป็นโมเดลแบบ non-linear statistical approach ที่สามารถจัดการกับแนวโน้มและฤดูกาลได้อย่างยืดหยุ่น กับโมเดล SARIMA ซึ่งเป็นโมเดลเชิงสถิติแบบ linear ที่เน้นการวิเคราะห์ข้อมูลอนุกรมเวลาแบบดั้งเดิม ผู้วิจัยใช้ข้อมูลยอดขายรองเท้าที่มีลักษณะตามฤดูกาลจากฐานข้อมูลจริง พร้อมสร้างโครงสร้างข้อมูลแบบ relational database เพื่อวิเคราะห์ผลการพยากรณ์ระยะยาว โดยประเมินผลลัพธ์ด้วยตัวชี้วัด MAE, MAPE และ RMSE ผลการศึกษาแสดงให้เห็นว่า Prophet มีความแม่นยำสูงกว่า SARIMA โดยสามารถทำความแม่นยำเฉลี่ยได้ถึง 98.8% และค่า MAE เท่ากับ 158.8 MAE ขณะที่ SARIMA มีความแม่นยำ 93% และ MAE เท่ากับ 83.9 ในการวัดค่า KPI ที่เกี่ยวข้องกับยอดขาย พบว่า Prophet มีความสามารถในการประเมินจำนวนสินค้าที่คาดว่าจะขายได้ในอนาคตได้ดีเป็นพิเศษ ขณะที่ SARIMA เหมาะกับการคาดการณ์ KPI รายเดือนที่คงที่ และให้ค่าคลาดเคลื่อนต่ำในบริบทนี้ งานวิจัยนี้เสนอว่าควรเลือกใช้โมเดลให้สอดคล้องกับลักษณะของข้อมูลและวัตถุประสงค์ในการวิเคราะห์ เช่น หากต้องการความยืดหยุ่นในข้อมูลที่มีแนวโน้มไม่คงที่ Prophet จะเหมาะสมกว่า ในขณะที่ SARIMA เหมาะสำหรับข้อมูลที่มีรูปแบบคงที่หรือแนวโน้มที่ชัดเจนในระยะยาว

2.3.2 บทความวิจัยเรื่อง Forecasting Daily Freight Flow in Cold Regions of China Using the Hybrid Prophet Model Considering the Importance of Festivals and Epidemic Prevention Policy โดย Haonan Chang และคณะ (2023)

งานวิจัยนี้มุ่งเน้นการพัฒนาแบบจำลองการพยากรณ์การขนส่งสินค้ารายวันในภูมิภาคหนาวของประเทศจีน โดยพิจารณาถึงปัจจัยสำคัญอย่างเทศกาลประจำชาติและนโยบายการป้องกันการระบาดของโรค ซึ่งส่งผลต่อความแปรปรวนของน้ำหนักการขนส่งในแต่ละวัน จากข้อจำกัดของโมเดล Prophet แบบดั้งเดิมที่ไม่สามารถคัดเลือกตัวแปรสำคัญได้อย่างเหมาะสม ผู้วิจัยจึงเสนอการผสมผสานระหว่างโมเดล Prophet กับ XGBoost-SHAP เพื่อให้สามารถคัดเลือก

ตัวแปรที่มีนัยสำคัญและพยากรณ์น้ำหนักรถขนส่งล่วงหน้าได้อย่างแม่นยำ โดยใช้ชุดข้อมูลจริงของน้ำหนักสินค้าที่ขนส่งในแต่ละวันจากบริษัทขนส่งในภูมิภาค Heijiang ทางตอนเหนือของจีน ผลการทดลองแสดงให้เห็นว่าโมเดลไฮบริดที่เสนอมีประสิทธิภาพเหนือกว่าโมเดล LSTM, ARIMA, SARIMA และ Prophet แบบดั้งเดิม โดยสามารถระบุความสัมพันธ์ระหว่างน้ำหนักรถขนส่งกับเหตุการณ์พิเศษ เช่น เทศกาลและมาตรการล็อกดาวน์ได้ชัดเจนมากขึ้น ข้อค้นพบสำคัญชี้ว่าเทศกาลและนโยบายป้องกันโรคระบาดมีผลกระทบสูงต่อรูปแบบการขนส่ง และการรวมปัจจัยเหล่านี้เข้าไปในโมเดลสามารถเพิ่มประสิทธิภาพในการวางแผนเชิงกลยุทธ์ของระบบขนส่งได้อย่างมีนัยสำคัญ

งานวิจัยนี้เสนอว่าโมเดล Prophet ที่ผนวกกับตัวแปรเหตุการณ์พิเศษสามารถตอบสนองวัตถุประสงค์เชิงกลยุทธ์ในบริบทของการจัดการทรัพยากรการขนส่งได้ดีกว่าโมเดลทั่วไป (Chang et al., 2025)

2.3.3 บทความวิจัยเรื่อง A Switching Based Forecasting Approach for Forecasting Sales Data in Supply Chains โดย Supriyo Ahmed และคณะ (2023)

งานวิจัยนี้มีวัตถุประสงค์เพื่อเสนอแนวทางใหม่ในการพยากรณ์ยอดขายในระบบห่วงโซ่อุปทาน (Supply Chains) โดยเน้นการแก้ไขข้อจำกัดของโมเดลแบบดั้งเดิมและโมเดลแบบ AI/ML ที่มักให้ผลลัพธ์แตกต่างกันตามบริบทของข้อมูลในแต่ละช่วงเวลา ผู้วิจัยจึงพัฒนาเทคนิคการพยากรณ์แบบใหม่ที่เรียกว่า Switching Based Forecasting Approach (SBFA) ซึ่งออกแบบมาเพื่อสลับเลือกโมเดลการพยากรณ์ที่เหมาะสมที่สุดในแต่ละช่วงเวลาของการพยากรณ์ โดยอิงจากประสิทธิภาพของแต่ละโมเดล

วิธีการดำเนินการเริ่มจากการใช้โมเดลการพยากรณ์แบบหลากหลาย ทั้งแบบดั้งเดิม เช่น Probabilistic และแบบ AI/ML เช่น Deep Learning (DL) และ Machine Learning (ML) เพื่อทำการพยากรณ์ยอดขายจากข้อมูลในอดีต จากนั้นจึงเปรียบเทียบประสิทธิภาพของโมเดลแต่ละตัว และระบบ SBFA จะทำการสลับไปใช้โมเดลที่ให้ผลลัพธ์แม่นยำที่สุดในช่วงเวลานั้นโดยอัตโนมัติ ทั้งนี้ ผู้วิจัยออกแบบ supply chain model แบบสองชั้น (two-echelon supply chain model) ซึ่งประกอบด้วยสินค้าที่มีคุณลักษณะแตกต่างกัน เพื่อประเมินผลกระทบของการใช้ SBFA ต่อประสิทธิภาพโดยรวมของระบบ

ผลการทดลองแสดงให้เห็นว่า การสลับโมเดลแบบไดนามิกตามลักษณะของข้อมูลสามารถปรับปรุงความแม่นยำในการพยากรณ์ได้อย่างมีนัยสำคัญ โดยเฉพาะในกรณีที่ไม่มีโมเดลใดที่เหมาะสมกับทุกช่วงของข้อมูลอย่างสม่ำเสมอ นอกจากนี้ ยังพบว่า จุดสั่งซื้อใหม่ (Reorder

Points) ที่คำนวณจาก SBFA มีแนวโน้มที่จะลดต้นทุนชีพพลายเซนโดยรวมได้มากกว่าการใช้โมเดลแบบตายตัว (fixed model) ไม่ว่าจะเป็นแบบ conventional, ML หรือ DL

ข้อเสนอแนะจากงานวิจัยชิ้นนี้คือ ในบริบทที่มีความซับซ้อนและมีความผันผวนสูง เช่นชีพพลายเซนขององค์กร การใช้แนวทาง SBFA ซึ่งมีความสามารถในการปรับเปลี่ยนตามลักษณะของข้อมูลจริงในแต่ละช่วงเวลา จะช่วยให้สามารถวางแผนการผลิตและการบริหารสินค้าคงคลังได้อย่างมีประสิทธิภาพมากขึ้น ลดต้นทุน และตอบสนองต่อความต้องการของตลาดได้อย่างทันทั่วทั้งนี้ แนวทาง SBFA ยังสามารถนำไปประยุกต์ใช้ในบริบทอื่นๆ ที่มีความซับซ้อนของข้อมูลได้อย่างยืดหยุ่น (Ahmed et al., 2024)

2.3.4 บทความวิจัยเรื่อง Comparative Analysis of ARIMA and Prophet Algorithms in Bitcoin Price Forecasting โดย Michael David Angelo และคณะ (2023) (Angelo et al., 2023)

งานวิจัยชิ้นนี้มีจุดมุ่งหมายในการเปรียบเทียบประสิทธิภาพของแบบจำลอง ARIMA และ Prophet ในการพยากรณ์ราคาบิตคอยน์ โดยใช้ชุดข้อมูลราคาบิตคอยน์ย้อนหลังเป็นระยะเวลา 2 ปี ตั้งแต่เดือนกุมภาพันธ์ 2019 ถึงปี 2021 ซึ่งจำแนกออกเป็นสามช่วงเวลา ได้แก่ รายวัน รายสัปดาห์ และรายเดือน ทั้ง ARIMA และ Prophet ถูกนำไปใช้เป็นแบบจำลองแบบ univariate ซึ่งใช้เพียง 2 ตัวแปรในกระบวนการฝึกโมเดล โดยในการดำเนินการนั้น ARIMA ถูกนำไปปรับแต่งค่าพารามิเตอร์หลายชุดในแต่ละช่วงเวลา เพื่อหาค่าที่ให้ผลลัพธ์ที่ดีที่สุด ส่วน Prophet ใช้ทั้งวิธีการปรับค่าพารามิเตอร์แบบอัตโนมัติและแบบแมนนวล พร้อมนำค่าที่ได้จากแต่ละโมเดลมาเปรียบเทียบกันเพื่อหาค่าที่เหมาะสมที่สุด จากนั้นจึงประเมินประสิทธิภาพของโมเดลโดยใช้เกณฑ์ที่วัด Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Mean Squared Error (MSE) และ Rooted Mean Squared Error (RMSE) ผลการทดลองพบว่าแบบจำลอง Prophet ให้ค่าความแม่นยำที่ดีกว่า ARIMA อย่างชัดเจนในช่วงข้อมูลรายวันและรายสัปดาห์ ในขณะที่แบบจำลอง ARIMA กลับให้ผลที่ดีกว่าในช่วงข้อมูลรายเดือน ซึ่งแสดงให้เห็นว่าลักษณะของช่วงเวลาในข้อมูลมีอิทธิพลต่อความสามารถในการทำนายของแต่ละโมเดลอย่างมีนัยสำคัญ

งานวิจัยนี้จึงเสนอแนะว่าการเลือกใช้แบบจำลองพยากรณ์ควรพิจารณาตามลักษณะของข้อมูลในแต่ละช่วงเวลา เพื่อให้สามารถเลือกใช้โมเดลที่เหมาะสมกับลักษณะข้อมูลมากที่สุด ซึ่งจะช่วยเพิ่มความแม่นยำในการพยากรณ์และการนำไปใช้จริงในเชิงธุรกิจได้อย่างมีประสิทธิภาพมากยิ่งขึ้น

จากการทบทวนวรรณกรรมข้างต้นสะท้อนให้เห็นว่า Prophet ได้รับความนิยมอย่างแพร่หลายในงานวิจัยด้านการพยากรณ์ยอดขาย โดยเฉพาะเมื่อข้อมูลมีความเป็นคาบและแนวโน้ม ส่วน Tableau มักใช้เพื่อการนำเสนอผลลัพธ์และสนับสนุนการตัดสินใจในภาคธุรกิจ แม้จะมีข้อจำกัดด้านความยืดหยุ่นในการปรับแต่งแบบจำลอง แต่ก็ยังเป็นเครื่องมือที่ทรงพลังสำหรับการสื่อสารข้อมูลเชิงภาพได้อย่างมีประสิทธิภาพ

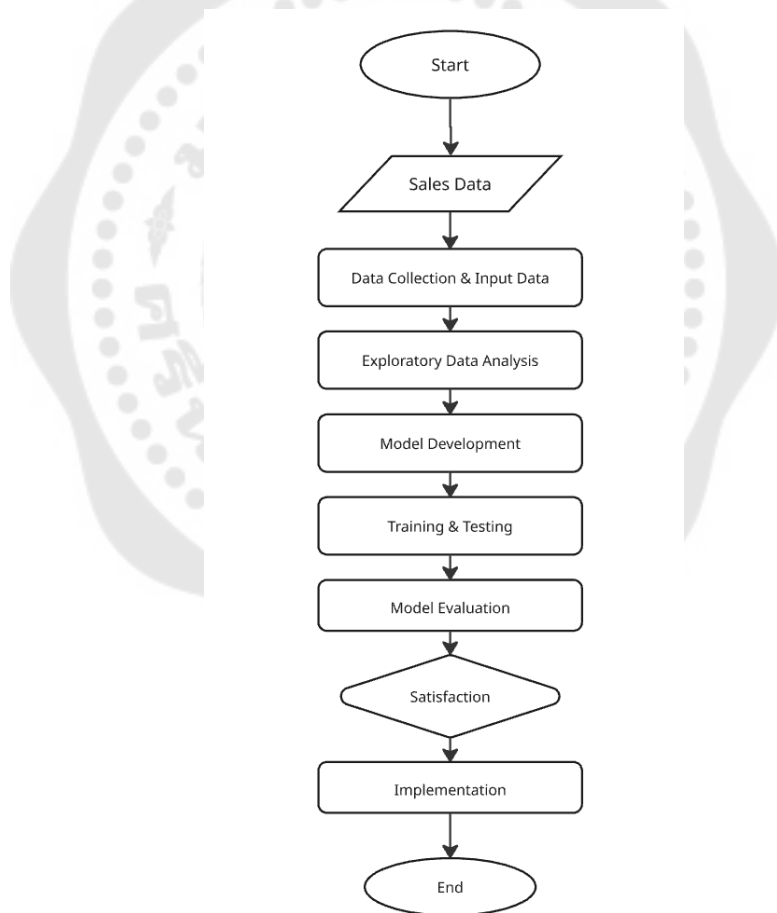


บทที่ 3

วิธีดำเนินการวิจัย

ในการวิจัยครั้งนี้ผู้วิจัยได้ดำเนินการตามขั้นตอนดังนี้

1. การออกแบบกระบวนการดำเนินการวิจัย
2. การนำเข้าข้อมูล (Input Data)
3. การประมวลผลข้อมูล (Data Processing)
4. การสร้างแบบจำลอง (Model Building)
5. การประเมินประสิทธิภาพแบบจำลอง (Model Evaluation)



ภาพประกอบ 1 แสดงการออกแบบกระบวนการดำเนินการวิจัย

3.1 การออกแบบกระบวนการดำเนินการวิจัย

การดำเนินการวิจัยครั้งนี้ได้รับการออกแบบให้เป็นกระบวนการที่มีลำดับขั้นตอนที่ชัดเจน ตั้งแต่การเตรียมข้อมูลไปจนถึงการนำแบบจำลองไปประยุกต์ใช้งานจริง โดยเน้นการพยากรณ์ ยอดขายรายวันของร้านกาแฟ ซึ่งมีความเฉพาะตัวในเรื่องความผันผวนของยอดขายตามฤดูกาล วันหยุด และพฤติกรรมผู้บริโภค การออกแบบกระบวนการนี้จึงต้องสอดคล้องกับลักษณะ ข้อมูลและเป้าหมายเชิงธุรกิจของร้านโดยตรงกระบวนการเริ่มต้นจากการกำหนดเป้าหมายของการพยากรณ์ยอดขาย และดำเนินการรวบรวมข้อมูลยอดขายรายวันจากระบบ ณ จุดขาย (POS) ของร้านกาแฟ ซึ่งถือเป็นข้อมูลหลักที่ใช้ในการสร้างแบบจำลอง หลังจากได้ข้อมูลมาแล้ว ขั้นตอนต่อไปคือการนำเข้าสู่ข้อมูลเข้าสู่ระบบเพื่อประมวลผล พร้อมทั้งตรวจสอบความครบถ้วนและความถูกต้องของข้อมูลในเบื้องต้น ต่อจากนั้นจะเข้าสู่กระบวนการวิเคราะห์ข้อมูลเบื้องต้น (Exploratory Data Analysis) เพื่อศึกษาลักษณะของข้อมูล เช่น แนวโน้มในแต่ละช่วงเวลา ความผันผวนตามฤดูกาล หรือช่วงเวลาที่มียอดขายสูงผิดปกติ ซึ่งการทำ EDA นี้มีความสำคัญเพราะช่วยให้เข้าใจ บริบทของข้อมูลมากยิ่งขึ้น และสามารถเตรียมความพร้อมให้กับการสร้างแบบจำลองในขั้นถัดไปได้ อย่างมีประสิทธิภาพ เมื่อข้อมูลได้รับการทำความสะอาดและวิเคราะห์เบื้องต้นแล้ว ขั้นตอนต่อไปคือการพัฒนาแบบจำลอง (Model Development) โดยใช้โมเดล Prophet ซึ่งมีความสามารถในการจัดการกับแนวโน้ม ความเป็นฤดูกาล และปัจจัยวันหยุด ซึ่งมักเกิดขึ้นในธุรกิจร้านกาแฟ หลังจากเลือกโมเดลและตั้งค่าพารามิเตอร์เบื้องต้นแล้ว จะทำการแบ่งข้อมูลออกเป็น ชุดฝึกและชุดทดสอบ และดำเนินการฝึกแบบจำลองเพื่อให้สามารถเรียนรู้จากข้อมูลในอดีต

หลังจากแบบจำลองได้รับการฝึกฝนแล้ว จะเข้าสู่ขั้นตอนของการประเมินประสิทธิภาพแบบจำลอง โดยใช้ตัวชี้วัด เช่น ค่า MAE, RMSE หรือ MAPE เพื่อพิจารณาความแม่นยำของผลการพยากรณ์ ในกรณีที่ผลการประเมินยังไม่เป็นที่น่าพอใจ อาจต้องมีการปรับแต่งแบบจำลองใหม่ เช่น การเปลี่ยนค่าพารามิเตอร์ หรือเพิ่มองค์ประกอบฤดูกาลที่เหมาะสม แล้วจึงทำการฝึกและประเมินซ้ำอีกครั้ง เมื่อแบบจำลองให้ผลลัพธ์ที่มีความแม่นยำและเสถียรภาพเพียงพอ จึงเข้าสู่ขั้นตอนของการนำไปใช้งานจริง โดยนำแบบจำลองที่ผ่านการทดสอบไปประยุกต์ใช้ในการพยากรณ์ยอดขายรายวันสำหรับสาขาร้านกาแฟที่ต้องการ ทั้งในเชิงการวางแผนสต็อกสินค้า การสั่งซื้อวัตถุดิบที่มีอายุสั้น เช่น โยเกิร์ต และการจัดตารางพนักงานให้เหมาะสมกับปริมาณลูกค้าที่คาดการณ์ได้

กระบวนการทั้งหมดสิ้นสุดลงเมื่อแบบจำลองได้รับการนำไปใช้จริงในบริบทของธุรกิจ และสามารถประเมินผลได้ว่าสามารถช่วยสนับสนุนการตัดสินใจในระดับปฏิบัติการของร้านกาแฟได้

อย่างมีประสิทธิภาพ ซึ่งจะนำไปสู่การลดของเสีย เพิ่มกำไร และยกระดับประสบการณ์ของลูกค้าได้ในที่สุด

3.2 การนำเข้าข้อมูล

งานวิจัยนี้ใช้ข้อมูลยอดขายรายวันของร้านกาแฟในเครือธุรกิจแฟรนไชส์แห่งหนึ่งในประเทศไทย โดยข้อมูลที่ใช้ครอบคลุมระยะเวลา 12 เดือน ตั้งแต่วันที่ พฤษภาคม 2567 - มิถุนายน 2568 ข้อมูลประกอบไปด้วย 2 คอลัมน์ ซึ่งครอบคลุมทั้งวันที่ทำการขาย และยอดขายในแต่ละวัน

ข้อมูลที่ได้ถูกรวบรวมจากระบบ POS (Point of Sale) ของร้าน ซึ่งบันทึกข้อมูลการขายแบบรายวัน เพื่อให้สามารถนำมาวิเคราะห์แนวโน้มและสร้างแบบจำลองพยากรณ์ยอดขายรายวันได้อย่างเหมาะสม รายละเอียดตัวอย่างข้อมูลแสดงไว้ในตารางที่ 1

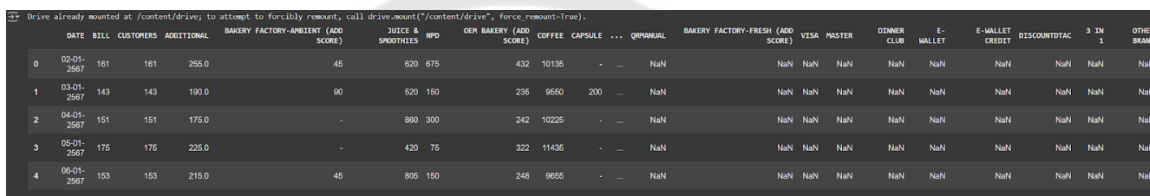
ตาราง 1 แสดงข้อมูลตัวอย่างที่ใช้ในการพยากรณ์

ชื่อคอลัมน์	คำอธิบายข้อมูล	ตัวอย่างข้อมูล
Date	วันที่	03/07/2023
JUICE & SMOOTHIES	ยอดขายของน้ำผลไม้	620
COFFEE	ยอดขายของกาแฟ	10,135
All Sale	ยอดขายในแต่ละวัน	15,000

ในการดำเนินการวิจัยครั้งนี้ ผู้วิจัยได้ใช้ภาษาไพทอน (Python) เป็นเครื่องมือหลักสำหรับการสร้างแบบจำลองการเรียนรู้ของเครื่อง (Machine Learning) และการประมวลผลข้อมูล โดยดำเนินการผ่านแพลตฟอร์ม Google Colab (Collaboratory) ซึ่งเป็นเครื่องมือสำหรับเขียนโค้ดภาษา Python บนเว็บเบราว์เซอร์ที่สามารถใช้งานได้โดยไม่เสียค่าใช้จ่าย และมีความสะดวกต่อการใช้งานอย่างมาก ขั้นตอนเริ่มต้นประกอบด้วยการ นำเข้าไลบรารี (Import Library) ซึ่งเป็นชุดคำสั่งหรือฟังก์ชันสำเร็จรูปที่ช่วยลดความซับซ้อนในการเขียนโค้ดใหม่ทั้งหมด จากนั้นจึงทำการนำเข้าข้อมูล (Import Data) ที่ใช้ในการวิเคราะห์ โดยผู้วิจัยได้จัดเก็บข้อมูลไว้ใน Google Drive และตั้งค่าเส้นทางการเข้าถึงข้อมูล (Set Path) ไว้ล่วงหน้า เพื่อความสะดวกและลดข้อผิดพลาดที่อาจเกิดจากการระบุเส้นทางใหม่ซ้ำ ๆ ในแต่ละขั้นตอนของกระบวนการวิเคราะห์

3.3 การประมวลผลข้อมูล

ภายหลังจากการนำเข้าข้อมูลเรียบร้อยแล้ว ขั้นตอนถัดไปคือการประมวลผลข้อมูลเบื้องต้น โดยเริ่มจากการตรวจสอบความถูกต้องและคุณลักษณะของข้อมูลดิบ เพื่อประเมินคุณภาพของข้อมูลก่อนเข้าสู่กระบวนการวิเคราะห์ในเชิงลึก การดำเนินการในขั้นตอนนี้ประกอบด้วย การแสดงโครงสร้างของชุดข้อมูล เช่น จำนวนแถว จำนวนคอลัมน์ ประเภทของตัวแปร และค่าที่ขาดหาย (missing values) เพื่อให้ผู้วิจัยเข้าใจลักษณะของข้อมูลโดยรวม ทั้งนี้สามารถดูรายละเอียดประกอบได้ตามภาพประกอบที่ 2 และภาพประกอบที่ 3 ซึ่งแสดงผลลัพธ์ของการตรวจสอบข้อมูลในเบื้องต้น



	DATE	BILL	CUSTOMERS	ADDITIONAL	BAKERY FACTORY-AMBIENT (ADD SCORE)	THREE B SMOOTHIES	MPD	DEB BAKERY (ADD SCORE)	COFFEE CAPSULE	...	QIPANNUAL	BAKERY FACTORY-FRESH (ADD SCORE)	VISA	PASTER	OSTMER CLUB	E-WALLET	E-WALLET CREDIT	DISCOUNTS	3 IN 1	OTHER BRAND
0	02-01-2567	161	161	255.0	45	820	675	432	10135	-	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
1	03-01-2567	143	143	190.0	90	520	150	235	9550	200	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	04-01-2567	151	151	175.0	-	980	300	242	10225	-	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	05-01-2567	175	175	225.0	-	420	75	322	11435	-	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	06-01-2567	153	153	215.0	45	805	150	248	9555	-	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

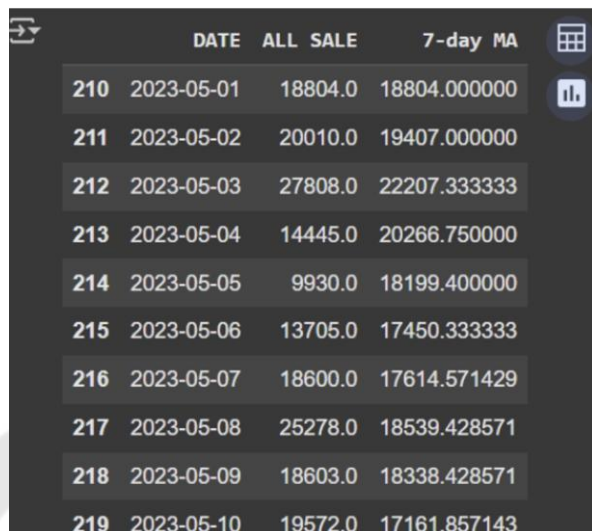
ภาพประกอบ 2 Preview Data Frame เพื่อดูข้อมูล 5 แถวแรก

ในการเตรียมข้อมูลเบื้องต้นก่อนการวิเคราะห์เชิงอนุกรมเวลา ผู้วิจัยได้ดำเนินการจัดการกับคอลัมน์ ALL SALE ซึ่งเดิมมีรูปแบบข้อมูลเป็นข้อความ (string) ที่ประกอบด้วยเครื่องหมายจุลภาค (,) เพื่อแบ่งหลักพัน และเครื่องหมายขีดกลาง (-) ซึ่งใช้แทนค่าที่ขาดหายไป ส่งผลให้ไม่สามารถนำข้อมูลดังกล่าวไปใช้ในการคำนวณได้โดยตรง ดังนั้นจึงมีการดำเนินการแปลงข้อมูลโดยการลบเครื่องหมายจุลภาคออก และแทนที่ค่าขาดหายไปด้วยค่าศูนย์ จากนั้นแปลงข้อมูลให้เป็นชนิดตัวเลขแบบทศนิยม (float) เพื่อให้สามารถนำไปประมวลผลและวิเคราะห์ได้อย่างถูกต้อง

ต่อมา ผู้วิจัยได้จัดการกับคอลัมน์ DATE ซึ่งอยู่ในรูปแบบวัน-เดือน-ปี ตามระบบพุทธศักราช (พ.ศ.) โดยทำการแปลงปีพุทธศักราชให้เป็นคริสต์ศักราช (ค.ศ.) ด้วยการลบออก 543 ปี และดำเนินการแปลงชนิดข้อมูลเป็น datetime เพื่อให้สามารถนำไปใช้ในการจัดเรียงตามลำดับเวลา และคำนวณทางสถิติที่เกี่ยวข้องได้อย่างมีประสิทธิภาพ

เมื่อได้ข้อมูลวันที่ในรูปแบบที่เหมาะสมแล้ว ผู้วิจัยได้ดำเนินการจัดเรียงข้อมูลตามลำดับเวลา และคำนวณค่าเฉลี่ยเคลื่อนที่แบบ 7 วัน (7-day Moving Average) เพื่อใช้ในการวิเคราะห์แนวโน้มยอดขายรายวัน โดยวิธีการดังกล่าวมีวัตถุประสงค์เพื่อทำให้เห็นทิศทางและรูปแบบของข้อมูลได้อย่างชัดเจนมากขึ้น แม้ยอดขายในแต่ละวันจะมีความผันผวนสูงก็ตาม นอกจากนี้ ผู้วิจัยยังได้ดำเนินการเติมค่าที่ขาดหายไป (missing values) ด้วยวิธีการแทนค่าข้อมูลล่าสุดที่มีอยู่

(forward fill) เพื่อให้ชุดข้อมูลมีความต่อเนื่องและพร้อมสำหรับการวิเคราะห์เพิ่มเติมทั้งในเชิงสถิติ และการทำนายยอดขายในอนาคต



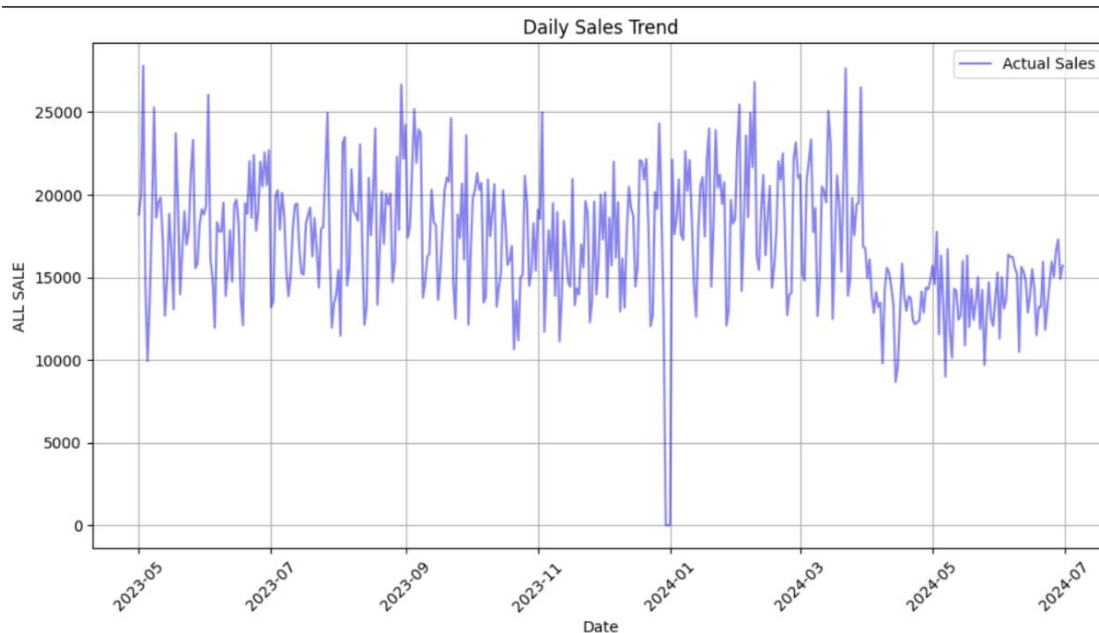
	DATE	ALL SALE	7-day MA
210	2023-05-01	18804.0	18804.000000
211	2023-05-02	20010.0	19407.000000
212	2023-05-03	27808.0	22207.333333
213	2023-05-04	14445.0	20266.750000
214	2023-05-05	9930.0	18199.400000
215	2023-05-06	13705.0	17450.333333
216	2023-05-07	18600.0	17614.571429
217	2023-05-08	25278.0	18539.428571
218	2023-05-09	18603.0	18338.428571
219	2023-05-10	19572.0	17161.857143

ภาพประกอบ 3 แสดงข้อมูลหลังจากมีการแปลงข้อมูลให้เหมาะสม

เพื่อให้ข้อมูลมีความพร้อมสำหรับการวิเคราะห์เชิงอนุกรมเวลา ผู้วิจัยได้ดำเนินการสร้างช่วงวันที่ที่ต่อเนื่องรายวัน (daily continuous date range) โดยกำหนดช่วงเวลาให้ครอบคลุมตั้งแต่วันแรกจนถึงวันสุดท้ายของข้อมูลจริงที่ปรากฏในคอลัมน์ DATE ผลลัพธ์ที่ได้คือชุดข้อมูลวันที่แบบต่อเนื่อง (full_dates) ซึ่งใช้เป็นฐานสำหรับสร้างข้อมูลที่ไม่ขาดช่วงในลำดับเวลา จากนั้นได้ทำการสร้างตารางข้อมูลใหม่ในรูปแบบ DataFrame ชื่อว่า df_full ซึ่งประกอบด้วยเพียงคอลัมน์วันที่ เพื่อใช้เป็นแม่แบบสำหรับจัดโครงสร้างข้อมูลยอดขายที่สมบูรณ์ในแต่ละวัน ขั้นตอนต่อมาคือการรวมตาราง df_full เข้ากับตารางข้อมูลยอดขายที่ผ่านการคำนวณค่าเฉลี่ยเคลื่อนที่ 7 วันแล้ว (final_df) โดยใช้วิธีการเชื่อมตารางแบบ left join ซึ่งช่วยให้สามารถคงลำดับของช่วงวันที่ต่อเนื่องไว้ได้ แม้ในบางวันจะไม่มีข้อมูลยอดขายจริงก็ตามผลลัพธ์ที่ได้คือ Data Frame ใหม่ชื่อว่า df_merged ซึ่งประกอบด้วยข้อมูลที่เรียงตามลำดับวันอย่างครบถ้วน พร้อมทั้งค่าของยอดขาย (ALL SALE) และค่าเฉลี่ยเคลื่อนที่ 7 วัน (7-day Moving Average) ในแต่ละวัน ทั้งนี้ ในกรณีที่พบค่าที่ขาดหายไป (NaN) เช่น ในวันที่ไม่มีการบันทึกยอดขายจริง ระบบจะดำเนินการแทนค่าดังกล่าวด้วยเลข 0 ทั้งในคอลัมน์ยอดขายและค่าเฉลี่ย เพื่อป้องกันปัญหาที่อาจเกิดขึ้นในการวิเคราะห์หรือการคำนวณทางสถิติในขั้นตอนถัดไป

กระบวนการดังกล่าวช่วยให้มั่นใจได้ว่าข้อมูลมีความครบถ้วนและต่อเนื่องในเชิงเวลา ซึ่งเป็นปัจจัยสำคัญในการวิเคราะห์แนวโน้มยอดขาย การพยากรณ์ข้อมูลในอนาคต หรือการตรวจสอบความผิดปกติของรูปแบบข้อมูลในลำดับเวลาได้อย่างถูกต้องและมีประสิทธิภาพ

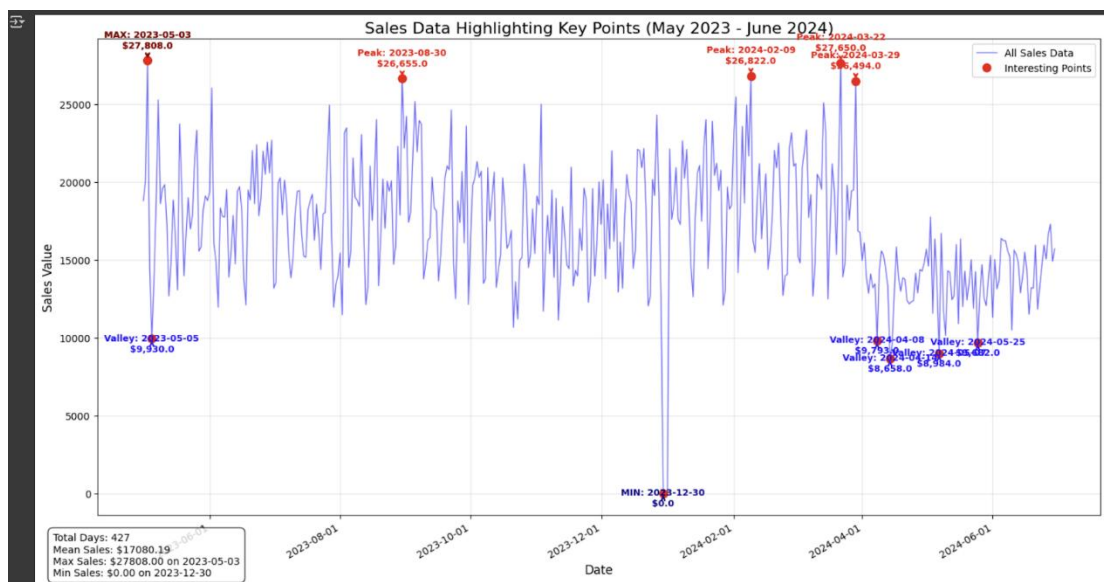
หลังจากที่นำเข้าสู่ข้อมูลแล้ว ในลำดับต่อมาจะเป็นการประมวลผล โดยลำดับแรกจะเป็นการตรวจสอบข้อมูลเบื้องต้น หลังจากนั้นจึงนำข้อมูลไปทำการ EDA เพื่อหาข้อมูลเชิงลึกตามภาพประกอบ ที่ 4



ภาพประกอบ 4 แสดง Trend ของข้อมูล

จากการวิเคราะห์แนวโน้มยอดขายรายวันตามภาพที่แสดง พบว่าข้อมูลยอดขายมีลักษณะไม่เป็นเชิงเส้นตรง (non-linear pattern) และแสดงให้เห็นถึงการเกิดความเป็นฤดูกาล (seasonality) ในบางช่วงของเวลา โดยเฉพาะในช่วงต้นปีและกลางปีที่มีความผันผวนของยอดขายเพิ่มขึ้นหรือลดลงอย่างมีนัยสำคัญ ลักษณะเช่นนี้สะท้อนถึงความซับซ้อนของพฤติกรรม การซื้อขาย ซึ่งไม่สามารถอธิบายได้ด้วยโมเดลเชิงเส้นทั่วไป ด้วยเหตุนี้ การเลือกใช้แบบจำลองที่สามารถรองรับการเปลี่ยนแปลงของแนวโน้ม (trend) และฤดูกาล (seasonality) จึงมีความเหมาะสม โดยเฉพาะโมเดลที่ออกแบบมาเพื่อการวิเคราะห์ข้อมูลอนุกรมเวลาโดยเฉพาะ เช่น Prophet ซึ่งสามารถจัดการกับข้อมูลที่มีความผันผวนสูงและมีแนวโน้มไม่สม่ำเสมอได้อย่างมีประสิทธิภาพ นอกจากนี้ การเติมข้อมูลให้ครบถ้วนทุกวัน พร้อมแทนค่าที่ขาดหายไปด้วยศูนย์ (0)

ยังช่วยให้ชุดข้อมูลมีความต่อเนื่องในเชิงเวลา (temporal continuity) ซึ่งเป็นเงื่อนไขสำคัญสำหรับการสร้างแบบจำลองข้อมูลอนุกรมเวลา (Time Series Modeling) ที่ต้องการลำดับเวลาแบบไม่ขาดตอน เพื่อให้การวิเคราะห์และการพยากรณ์เป็นไปอย่างแม่นยำและเชื่อถือได้

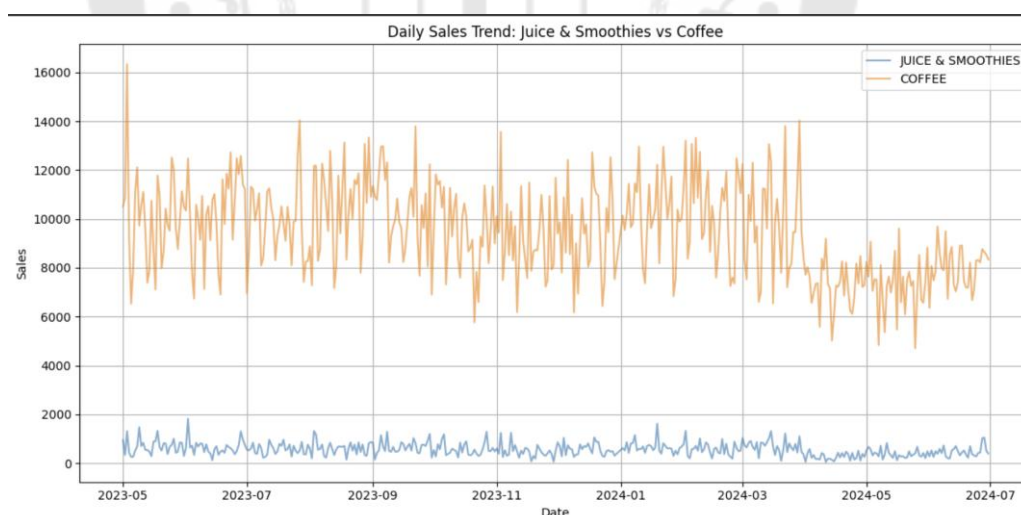


ภาพประกอบ 5 แสดงแนวโน้มยอดขายรายวันพร้อมจุดที่น่าสนใจ (Interesting Points)

จากการวิเคราะห์ข้อมูลยอดขายรายวันในช่วงระหว่างเดือนพฤษภาคม พ.ศ. 2566 ถึงเดือนมิถุนายน พ.ศ. 2567 (รวมทั้งสิ้น 427 วัน) โดยใช้เทคนิคการวิเคราะห์เชิงสถิติเพื่อระบุจุดที่น่าสนใจ (Interesting Points) ในชุดข้อมูล พบว่าข้อมูลมีความผันผวนค่อนข้างสูง และมีลักษณะของจุดพีค (peak) และจุดต่ำสุดชั่วคราว (valley) เกิดขึ้นอย่างต่อเนื่องตลอดช่วงเวลา จุดที่มียอดขายสูงที่สุด (Maximum) เกิดขึ้นในวันที่ 3 พฤษภาคม พ.ศ. 2566 โดยมียอดขายรวมสูงถึง 27,808.0 บาท ซึ่งถือเป็นจุดสูงสุดในช่วงระยะเวลาที่ทำการศึกษา ขณะที่จุดที่มียอดขายต่ำที่สุด (Minimum) ปรากฏในวันที่ 30 ธันวาคม พ.ศ. 2566 โดยมียอดขายเท่ากับ 0.0 บาท ซึ่งมีความเป็นไปได้ว่าเกิดจากเหตุการณ์พิเศษ เช่น การปิดร้านชั่วคราว หรือการขาดข้อมูลบันทึกยอดขายในวันดังกล่าว นอกจากจุดสูงสุดและต่ำสุดแล้ว ยังพบจุดพีค (peaks) อื่น ๆ ที่เกิดขึ้นในช่วงที่ยอดขายเพิ่มขึ้นอย่างมีนัยสำคัญเมื่อเปรียบเทียบกับวันก่อนและหลัง เช่น ในวันที่ 30 สิงหาคม พ.ศ. 2566 และวันที่ 9 กุมภาพันธ์ พ.ศ. 2567 จุดเหล่านี้แสดงให้เห็นถึงช่วงเวลาที่อาจมีการส่งเสริมการขาย กิจกรรมพิเศษ หรือความต้องการของผู้บริโภคที่เพิ่มขึ้นแบบเฉียบพลัน ในทางกลับกัน ยังพบจุดต่ำสุดชั่วคราว (valleys) ที่สะท้อนถึงการลดลงของยอดขายอย่างชัดเจน เช่น ใน

วันที่ 5 พฤษภาคม พ.ศ. 2566 และวันที่ 8 เมษายน พ.ศ. 2567 ซึ่งอาจเกิดจากปัจจัยด้านฤดูกาล หรือพฤติกรรมผู้บริโภคในช่วงเวลาดังกล่าว จากการสังเกตภาพรวม พบว่าแนวโน้มยอดขายมีลักษณะผันผวนตลอดช่วงเวลา โดยมียอดขายเพิ่มขึ้นและลดลงเป็นระยะอย่างสม่ำเสมอ จุดที่มียอดขายสูงสุดและช่วงที่เกิดจุดพีคส่วนใหญ่มักปรากฏในช่วงกลางถึงปลายปี ขณะที่ในช่วงต้นปี พ.ศ. 2567 เริ่มมีแนวโน้มยอดขายลดลงอย่างต่อเนื่อง ทั้งนี้ ค่าเฉลี่ยของยอดขายรายวันในช่วงเวลาทั้งหมดอยู่ที่ประมาณ 17,080 บาท ซึ่งสามารถใช้เป็นค่ากลางในการอ้างอิงเพื่อตรวจสอบความผิดปกติ หรือใช้ประกอบการวิเคราะห์แนวโน้มและการวางแผนการดำเนินงานของธุรกิจได้ต่อไป

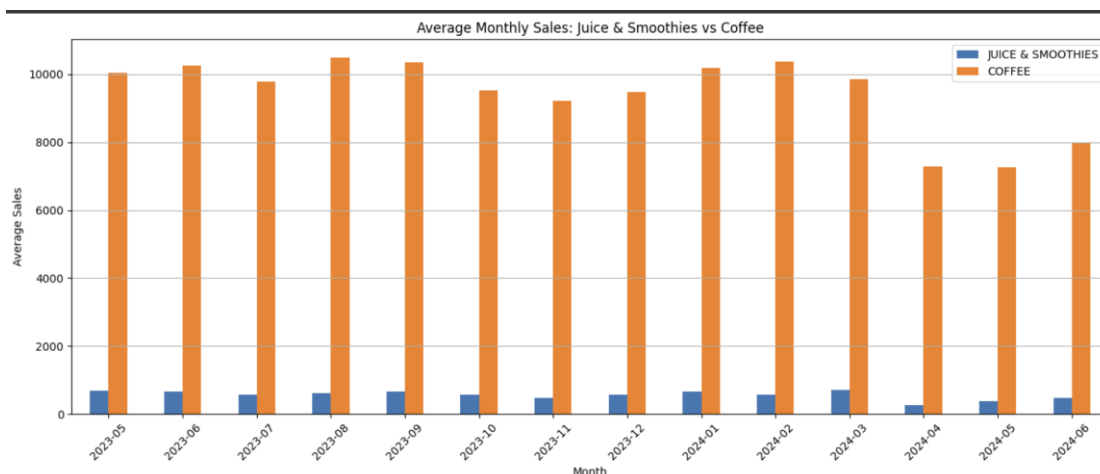
อย่างไรก็ตาม เพื่อให้การวิเคราะห์มีความลึกยิ่งขึ้น จึงได้แยกวิเคราะห์แนวโน้มยอดขายของสินค้ารายหมวดที่สำคัญ ได้แก่ หมวด JUICE & SMOOTHIES และหมวด COFFEE ซึ่งมีลักษณะการบริโภคและต้นทุนที่แตกต่างกัน โดยเฉพาะหมวด JUICE & SMOOTHIES ที่ใช้วัตถุดิบที่เน่าเสียง่ายอย่างโยเกิร์ตซึ่งต้องสั่งซื้อรายสัปดาห์ การแยกวิเคราะห์จึงช่วยให้สามารถประเมินพฤติกรรมการขายเฉพาะกลุ่มได้แม่นยำยิ่งขึ้น ช่วยสนับสนุนการวางแผนจัดการวัตถุดิบลดของเสีย และพัฒนาแบบจำลองการพยากรณ์ในระดับหมวดหมู่ได้อย่างมีประสิทธิภาพมากขึ้น



ภาพประกอบ 6 แนวโน้มยอดขายรายวันของ JUICE & SMOOTHIES และ COFFEE

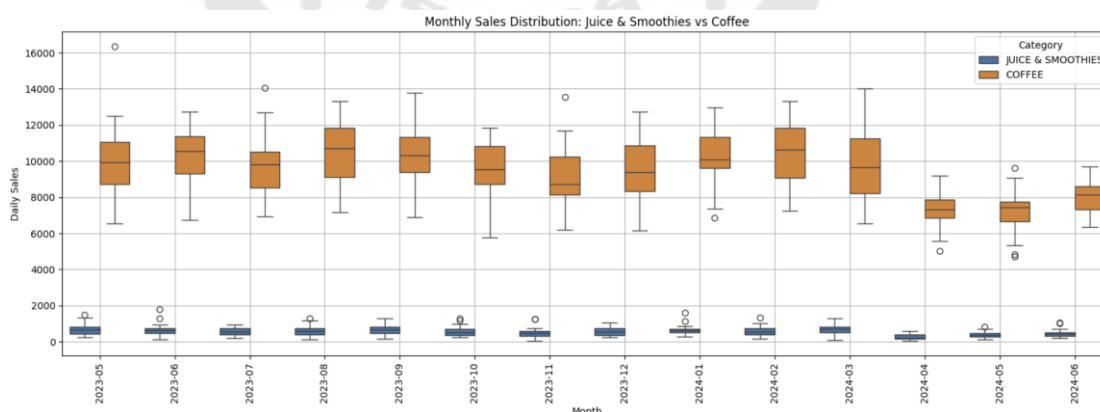
จากกราฟแนวโน้มยอดขายรายวันของสินค้า 2 หมวด พบว่า ยอดขายกาแฟ (COFFEE) มีค่าที่สูงกว่าและแปรผันมากกว่าอย่างชัดเจนเมื่อเทียบกับ ยอดขายของหมวด JUICE & SMOOTHIES ตลอดช่วงเวลาที่วิเคราะห์ยอดขายกาแฟเริ่มมีความผันผวนสูงในช่วงต้น โดยเฉพาะช่วงกลางปี 2023 จากนั้นเริ่มลดลงอย่างชัดเจนในไตรมาสที่ 2 ของปี 2024 ซึ่งอาจสะท้อนถึง

ฤดูกาลหรือการเปลี่ยนแปลงเชิงกลยุทธ์ในการขายหรือโปรโมชั่นในขณะที่ยอดขายของหมวดหมู่นี้มีแนวโน้มคงที่ แต่ค่อนข้างต่ำกว่ากาแฟอย่างมากตลอดทั้งปี โดยยอดขายในหมวดนี้มีความเสถียรมากกว่าและไม่มีช่วงพีคหรือจุดตกต่ำอย่างชัดเจน



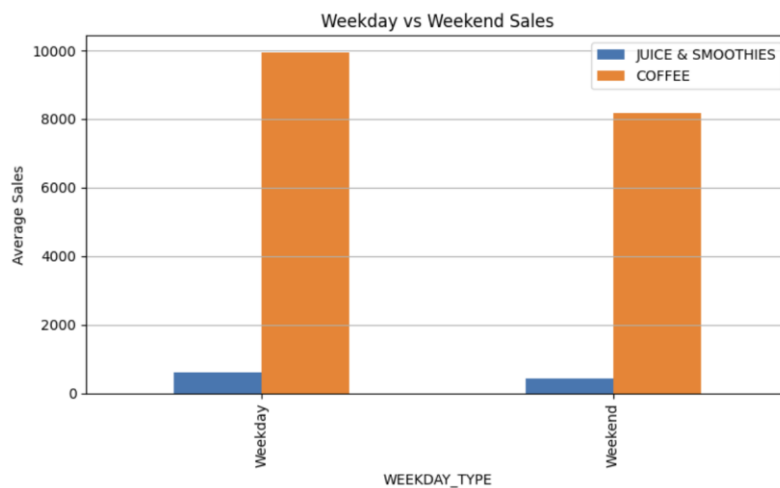
ภาพประกอบ 7 กราฟแท่ง Average Monthly Sales – Juice & Smoothies vs Coffee

จากการวิเคราะห์ยอดขายเฉลี่ยรายเดือน พบว่าในทุกช่วงเวลา หมวดกาแฟมียอดขายเฉลี่ยสูงกว่าหมวดสมูทตี้อย่างมีนัยสำคัญ ส่วนหมวดสมูทตี้มีแนวโน้มลดลงในช่วงต้นปี 2024 ซึ่งอาจเกี่ยวข้องกับพฤติกรรมผู้บริโภคหรือฤดูกาล



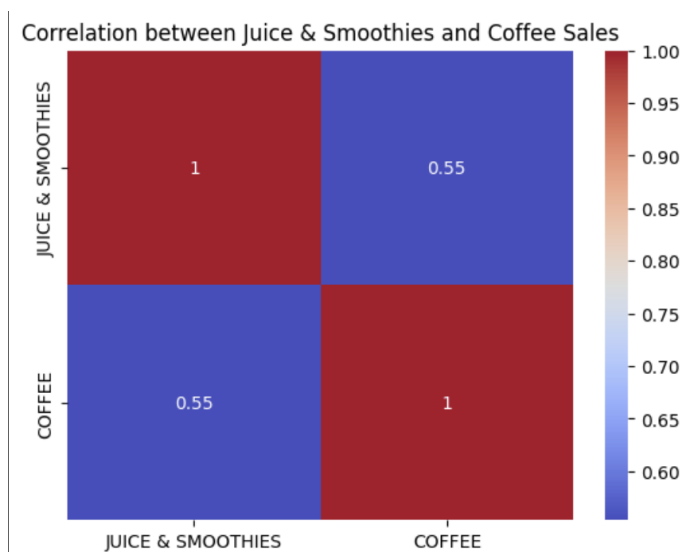
ภาพประกอบ 8 Boxplot Monthly Sales Distribution – Juice & Smoothies vs Coffee

จาก boxplot รายเดือน พบว่า หมวดกาแฟมีการกระจายของยอดขายรายวันสูงกว่า และพบ outliers จำนวนมาก โดยเฉพาะในช่วงครึ่งหลังของปี 2023 ขณะที่หมวดสมูทตี้มียอดขายแปรผันน้อยกว่าและค่ากลางต่ำกว่าอย่างต่อเนื่อง



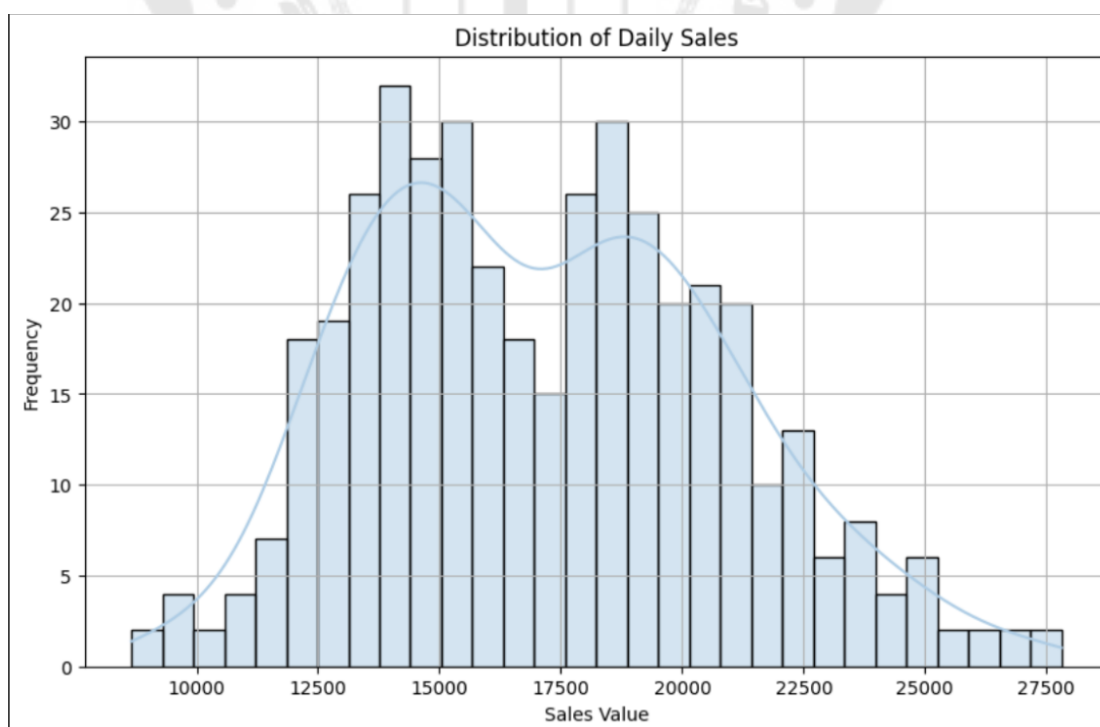
ภาพประกอบ 9 กราฟเปรียบเทียบ Weekday vs Weekend Sales

การวิเคราะห์ยอดขายระหว่างวันธรรมดาและวันหยุดพบว่า หมวดกาแฟมียอดขายเฉลี่ยลดลงในช่วงวันหยุดอย่างชัดเจน สะท้อนถึงความเชื่อมโยงกับพฤติกรรมของกลุ่มผู้บริโภคที่ซื้อกาแฟระหว่างวันทำงาน ในขณะที่หมวดสมูทตี้ก็ได้รับผลกระทบ แต่ในระดับที่น้อยกว่า



ภาพประกอบ 10 Heatmap Correlation Matrix – Juice & Smoothies vs Coffee

ค่า correlation ระหว่างยอดขายของกาแฟและสมูทตี้เท่ากับ 0.55 ซึ่งถือว่าอยู่ในระดับปานกลาง สะท้อนว่ามีบางช่วงที่ยอดขายทั้งสองหมวดเคลื่อนไหวไปในทิศทางเดียวกัน เช่น ช่วงโปรโมชั่นหรือฤดูกาลพิเศษ แต่อาจไม่ได้เกิดขึ้นอย่างต่อเนื่องทุกช่วงเวลา



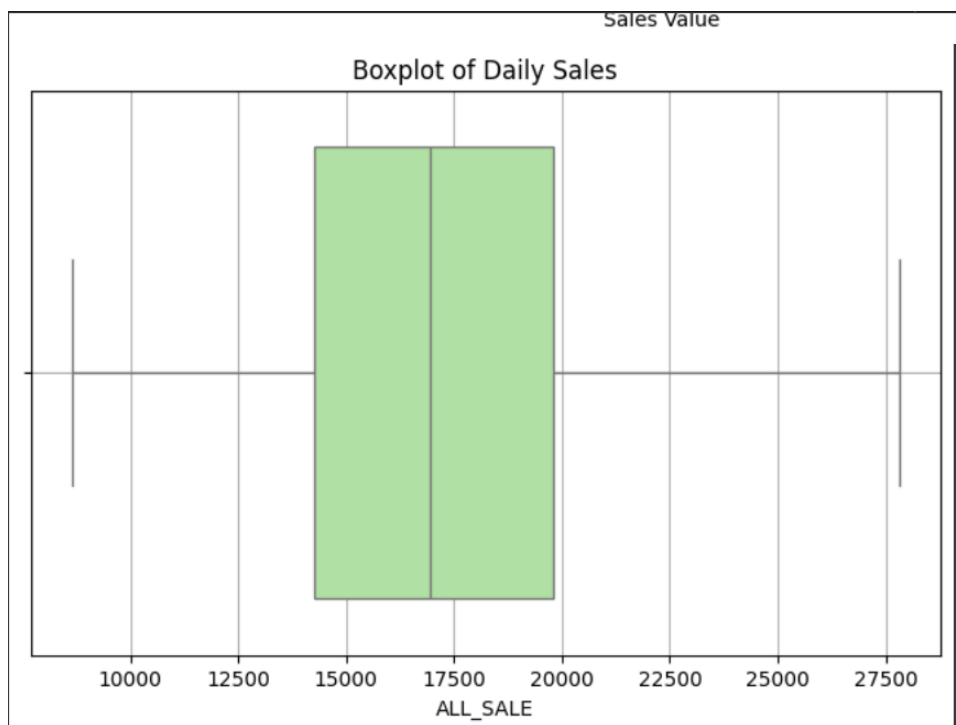
ภาพประกอบ 11 แสดงแนวโน้มยอดขายรายวันพร้อมจุดที่น่าสนใจ (Interesting Points)

แสดงการแจกแจงความถี่ของยอดขายรายวันในรูปแบบ Histogram โดยแกนแนวนอน (X) แสดงค่าของยอดขาย (Sales Value) และแกนแนวตั้ง (Y) แสดงจำนวนวันที่มียอดขายอยู่ในช่วงค่าดังกล่าว จากการสังเกตพบว่าการกระจายของข้อมูลมีลักษณะคล้ายการแจกแจงแบบเบ้ขวา (right-skewed) เล็กน้อย ซึ่งหมายความว่าแม้ยอดขายส่วนใหญ่จะอยู่ในช่วงกลาง แต่ยังมีบางวันที่มียอดขายสูงกว่าปกติอย่างชัดเจน อันอาจเกิดจากช่วงเวลาเฉพาะ เช่น การจัดโปรโมชั่นหรือเทศกาลสำคัญ

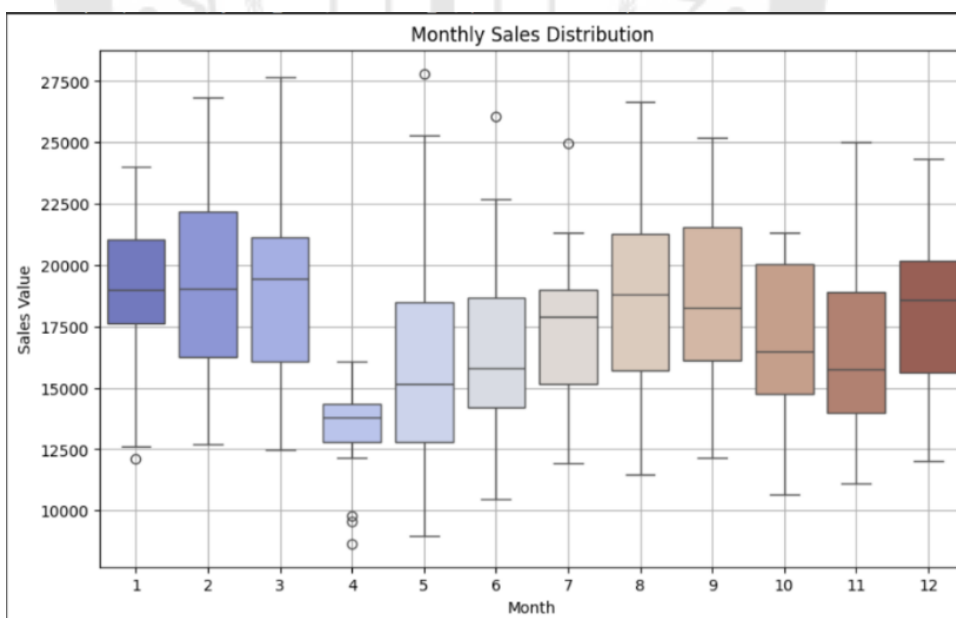
ยอดขายรายวันส่วนใหญ่อยู่ในช่วงประมาณ 14,000 ถึง 20,000 บาท โดยมีจุดศูนย์กลางของการกระจาย (mode) อยู่ในช่วง 13,000–20,000 บาท เส้นโค้งสีน้ำเงินที่ปรากฏบนกราฟคือเส้นประมาณค่าความหนาแน่น (Kernel Density Estimation: KDE) ซึ่งช่วยแสดงแนวโน้มของการกระจายข้อมูลอย่างต่อเนื่อง ทำให้เห็นภาพรวมของการกระจายตัวได้ชัดเจนยิ่งขึ้น

ลักษณะของการกระจายข้อมูลบ่งชี้ถึงความหลากหลายของพฤติกรรมการขายในแต่ละวัน โดยค่าเฉลี่ยของยอดขายรายวันอยู่ที่ประมาณ 17,201 บาท และค่ามัธยฐานอยู่ที่ 16,932 บาท ซึ่งใกล้เคียงกันอย่างมาก แสดงให้เห็นว่าไม่มีค่าผิดปกติ (outliers) รุนแรงที่ส่งผลกระทบต่อความสมมาตรของข้อมูล นอกจากนี้ ค่ายอดขายต่ำสุดที่สังเกตได้ในชุดข้อมูลอยู่ที่ประมาณ 8,600 บาท ซึ่งถือว่าค่อนข้างสูงเมื่อเทียบกับค่าต่ำสุดในหลายธุรกิจ สะท้อนให้เห็นว่าร้านแทบไม่มีวันหยุดการขาย ยกเว้นในบางกรณีที่อาจเติมค่าศูนย์สำหรับวันปิดร้าน

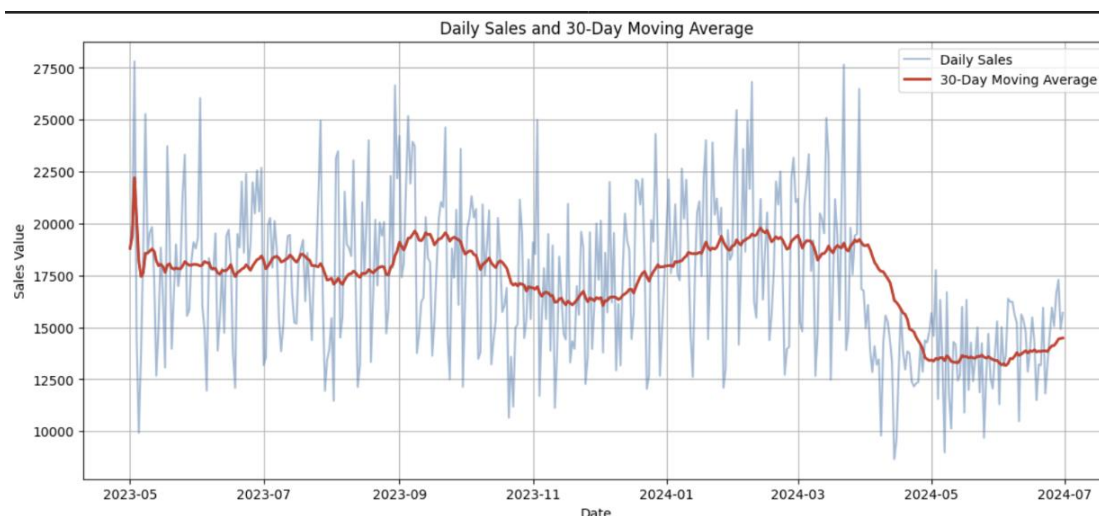
จากข้อมูลทั้งหมดนี้สามารถสรุปได้ว่ารูปแบบการกระจายของยอดขายมีลักษณะเสถียรในระดับหนึ่ง และสามารถนำไปใช้ประกอบในการสร้างแบบจำลองทางสถิติหรือการพยากรณ์ยอดขายต่อไปได้อย่างมีประสิทธิภาพ โดยควรให้ความสนใจกับวันที่มียอดขายต่ำหรือเป็นศูนย์เนื่องจากอาจเป็นปัจจัยที่มีผลต่อความแม่นยำของแบบจำลองในขั้นตอนการวิเคราะห์เชิงเวลา



ภาพประกอบ 12 Boxplot ของยอดขายรายวัน (Boxplot of Daily Sales)



ภาพประกอบ 13 การกระจายของยอดขายรายเดือน (Monthly Sales Distribution)



ภาพประกอบ 14 แนวโน้มยอดขายรายวันและค่าเฉลี่ยเคลื่อนที่ 30 วัน

ผลการวิเคราะห์ข้อมูลเชิงสำรวจ (Exploratory Data Analysis: EDA) ในการวิเคราะห์ข้อมูลยอดขายรายวันของร้านกาแฟในช่วงเดือนพฤษภาคม พ.ศ. 2566 ถึงเดือนมิถุนายน พ.ศ. 2567 ผู้วิจัยได้ดำเนินการตรวจสอบและจัดการข้อมูลเบื้องต้นเพื่อให้เหมาะสมต่อการวิเคราะห์เชิงเวลา โดยเริ่มจากการแปลงค่าคอลัมน์ ALL SALE ให้อยู่ในรูปแบบตัวเลขชนิดทศนิยม (float) และปรับรูปแบบวันที่จากพุทธศักราช (พ.ศ.) เป็นคริสต์ศักราช (ค.ศ.) เพื่อให้สามารถประมวลผลทางสถิติได้อย่างถูกต้อง จากนั้นได้คำนวณค่าเฉลี่ยเคลื่อนที่รายสัปดาห์ (7-day Moving Average) เพื่อใช้เปรียบเทียบแนวโน้มในเชิงเวลาอย่างมีประสิทธิภาพ ต่อมา ได้มีการเติมวันให้ครบถ้วนตลอดช่วงเวลาที่ทำการศึกษา แม้ในบางวันจะไม่มีการขายจริง เช่น ในวันหยุดร้าน โดยใช้เทคนิคการเติมค่าว่างด้วยศูนย์ (0) เพื่อรักษาความต่อเนื่องของลำดับเวลา พร้อมทั้งจัดทำดัชนีเวลาแบบสม่ำเสมอ ซึ่งมีความสำคัญต่อการวิเคราะห์ข้อมูลอนุกรมเวลาในขั้นตอนถัดไป จากการวาดกราฟแนวโน้มยอดขายรายวัน (Daily Sales Trend) พบว่า ยอดขายมีความผันผวนค่อนข้างสูงตลอดปี พ.ศ. 2566 โดยเฉพาะในช่วงกลางปี และเริ่มมีแนวโน้มลดลงอย่างชัดเจนตั้งแต่เดือนมีนาคม พ.ศ. 2567 ซึ่งอาจจะท่อนถึงฤดูกาลหรือเหตุการณ์เฉพาะ เช่น วันหยุดยาว หรือช่วงเวลาที่ผู้บริโภคลดพฤติกรรมการซื้อ ทั้งนี้ ยังพบวันที่ยอดขายเป็นศูนย์อย่างชัดเจนในวันที่ 30 ธันวาคม พ.ศ. 2566 ซึ่งอาจสันนิษฐานได้ว่าเป็นวันที่ร้านปิดทำการ

การวิเคราะห์จุดสำคัญ (Key Points) จากกราฟแนวโน้มยอดขายรายวันชี้ให้เห็นว่า จุดที่มียอดขายสูงสุด (Maximum) เกิดขึ้นในวันที่ 3 พฤษภาคม พ.ศ. 2566 โดยมียอดขาย 27,808 บาท ขณะที่จุดต่ำสุด (Minimum) เกิดขึ้นในวันที่ 30 ธันวาคม พ.ศ. 2566 ซึ่งมียอดขายเป็นศูนย์บาท

นอกจากนี้ ยังตรวจพบจุดพีค (Local Peaks) และจุดต่ำเฉพาะจุด (Local Valleys) ซึ่งมีค่าที่แตกต่างจากค่าเฉลี่ยอย่างมีนัยสำคัญ เช่น วันที่ 30 สิงหาคม พ.ศ. 2566 และวันที่ 9 กุมภาพันธ์ พ.ศ. 2567 เป็นช่วงที่ยอดขายพุ่งสูงผิดปกติ ส่วนวันที่ 8 เมษายน พ.ศ. 2567 เป็นหนึ่งในช่วงที่ยอดขายลดลงอย่างชัดเจน

จากการวิเคราะห์สถิติเชิงสรุป (Summary Statistics) พบว่ายอดขายเฉลี่ยรายวันอยู่ที่ประมาณ 17,201 บาท และค่ามัธยฐานอยู่ที่ 16,932 บาท ซึ่งทั้งสองค่านี้ใกล้เคียงกันมาก แสดงให้เห็นว่าไม่มีค่าผิดปกติ (outliers) รุนแรงในชุดข้อมูล และการแจกแจงของข้อมูลมีลักษณะเบ้ขวา (right-skewed) เล็กน้อยตามที่ปรากฏในกราฟฮิสโตแกรม ค่าความเบี่ยงเบนมาตรฐาน (Standard Deviation) อยู่ที่ประมาณ 3,720 บาท ซึ่งสะท้อนถึงระดับความผันผวนของยอดขายในแต่ละวัน

ผลจากการวิเคราะห์กล่อง (boxplot) แสดงให้เห็นว่าช่วงยอดขายส่วนใหญ่อยู่ระหว่าง 14,000–20,000 บาทต่อวัน โดยมีค่าต่ำสุดและสูงสุดไม่เกินขอบเขต whisker มากนัก ซึ่งสนับสนุนข้อสังเกตที่ว่าไม่มี outlier ที่ผิดปกติในระดับสูงเมื่อพิจารณาค่าเฉลี่ยเคลื่อนที่ราย 30 วัน (30-Day Moving Average) จะพบแนวโน้มภาพรวมของยอดขายที่ชัดเจนมากขึ้น โดยเส้นแนวโน้มมีลักษณะลดลงอย่างต่อเนื่องหลังจากไตรมาสแรกของปี พ.ศ. 2567 ซึ่งสอดคล้องกับผลที่ได้จากกราฟแนวโน้มรายวันสำหรับการวิเคราะห์ในเชิงเดือน (Monthly Sales Distribution) โดยใช้ boxplot พบว่า เดือนที่มียอดขายเฉลี่ยสูงได้แก่ เดือนมกราคม กุมภาพันธ์ และมีนาคม ขณะที่เดือนเมษายนและพฤษภาคมมียอดขายเฉลี่ยต่ำที่สุดในรอบปี ซึ่งอาจเกี่ยวข้องกับวันหยุดยาวหรือฤดูกาลที่มีผลกระทบต่อพฤติกรรมของผู้บริโภค

โดยสรุป การวิเคราะห์ข้อมูลเชิงสำรวจ (Exploratory Data Analysis: EDA) ในครั้งนี้มีจุดมุ่งหมายเพื่อทำความเข้าใจโครงสร้างและพฤติกรรมของข้อมูลยอดขายรายวันในภาพรวมก่อนการพัฒนาแบบจำลองพยากรณ์ โดยได้วิเคราะห์ทั้งในด้านสถิติพื้นฐาน การแจกแจงข้อมูล แนวโน้มระยะยาว ความผันผวนในแต่ละช่วงเวลา รวมถึงรูปแบบฤดูกาลที่อาจส่งผลกระทบต่อยอดขาย เช่น ช่วงปลายปีที่มีพฤติกรรมการซื้อที่เพิ่มขึ้น และช่วงต้นปีที่ยอดขายมีแนวโน้มลดลงอย่างต่อเนื่อง ผลการวิเคราะห์พบว่า ยอดขายรวมรายวัน (ALL_SALE) มีความผันผวนค่อนข้างมาก มีบางช่วงที่มียอดขายพุ่งสูงเป็นพิเศษ อันอาจเกิดจากกิจกรรมส่งเสริมการขายหรือวันพิเศษ ขณะเดียวกันก็พบการลดลงของยอดขายอย่างชัดเจนในบางช่วงเวลา นอกจากนี้ยังพบลักษณะของฤดูกาลที่ส่งผลต่อพฤติกรรมการซื้อของลูกค้าในบางเดือนอย่างมีนัยสำคัญ

ด้วยเหตุนี้ ผู้วิจัยจึงตัดสินใจเลือกใช้ยอดขายรวม (ALL_SALE) เป็นตัวแปรหลักในการสร้างแบบจำลอง Prophet เพื่อการพยากรณ์ เนื่องจากสามารถสะท้อนภาพรวมของการดำเนินธุรกิจได้ชัดเจนและครอบคลุมกว่าการใช้ยอดขายเฉพาะหมวด (JUICE & SMOOTHIES หรือ COFFEE) ที่มีความผันผวนเฉพาะกลุ่มและอาจไม่สะท้อนแนวโน้มของธุรกิจโดยรวม อีกทั้งยังช่วยลดความซับซ้อนของแบบจำลองและเหมาะสมกับวัตถุประสงค์ของการวางแผนเชิงกลยุทธ์ เช่น การจัดโปรโมชั่นในช่วงยอดขายต่ำ หรือการเตรียมวัตถุดิบในช่วงยอดขายสูง เพื่อเพิ่มประสิทธิภาพในการบริหารจัดการธุรกิจในอนาคตได้อย่างเป็นระบบและยั่งยืน

3.4 การสร้างแบบจำลอง

```
import matplotlib.pyplot as plt

# วาดกราฟ Train-Test Split เห็น 0 ชัด
plt.figure(figsize=(16, 8))

# Plot training data พร้อมจุด marker
plt.plot(train_data['DATE'], train_data['ALL_SALE'],
         color='#005b96', linewidth=2.5, marker='o', markersize=3, label='Train Series')

# Plot testing data พร้อมจุด marker
plt.plot(test_data['DATE'], test_data['ALL_SALE'],
         color='#ff7b25', linewidth=2.5, marker='o', markersize=3, label='Test Series')

# เส้นแบ่งช่วง Train/Test
plt.axvline(x=train_data['DATE'].iloc[-1], color='gray', linestyle='--', alpha=0.7)

# ไม่ขีด ylim = ปลดปล่อยแกน y เริ่ม 0 อัตโนมัติ
# (ถ้าอยาก fix เพิ่มเองก็ใส่ plt.ylim(0, 30000) ประมาณนี้ได้)

plt.xlabel('Date', fontsize=14)
plt.ylabel('Sales Value', fontsize=14)
plt.title('Time Series Train-Test Split (Markers and Zero Visible)', fontsize=16)
plt.legend(fontsize=14)
plt.grid(True, alpha=0.3)
plt.xticks(rotation=45)
plt.tight_layout()

plt.show()
```

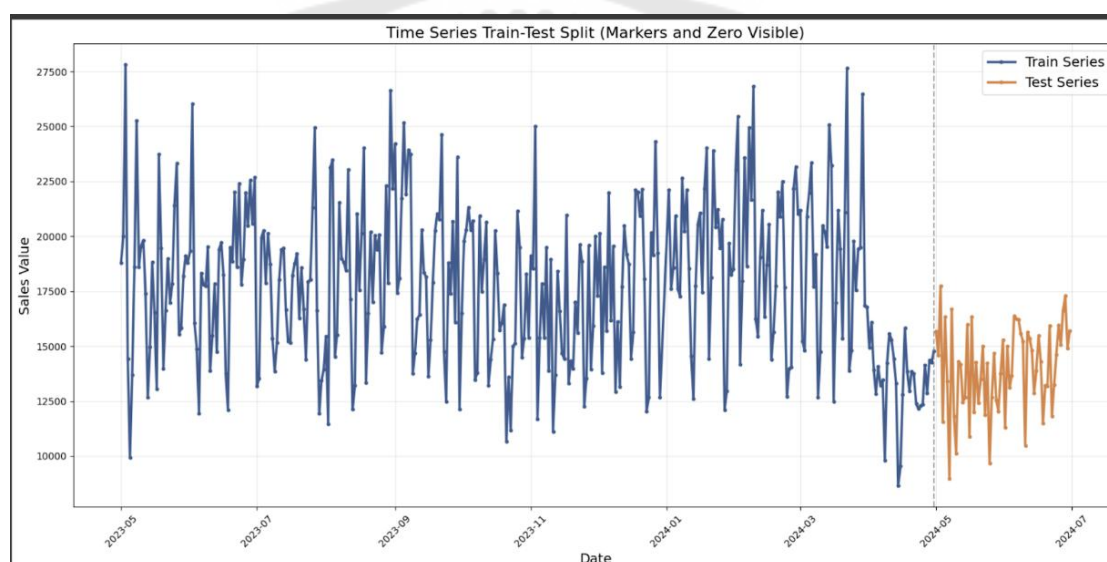
ภาพประกอบ 15 การแบ่งชุดข้อมูลยอดขายรายวันออกเป็นชุดฝึกและชุดทดสอบ

(Train-Test Split)

ในการดำเนินการวิเคราะห์ข้อมูล ผู้วิจัยได้ทำการแบ่งชุดข้อมูลยอดขายรวมรายวัน (ALL_SALE) ออกเป็นสองช่วง ได้แก่ ช่วงข้อมูลสำหรับฝึกโมเดล (training set) และช่วงข้อมูลสำหรับทดสอบความแม่นยำของโมเดล (test set) โดยใช้หลักการแบ่งแบบอนุกรมเวลา (time series split) เพื่อรักษาลำดับของข้อมูลตามเวลา ไม่ให้ข้อมูลอนาคตไปรบกวนการฝึกโมเดลในอดีต ซึ่งเหมาะสมกับลักษณะของข้อมูลเชิงเวลา จากนั้นจึงทำการแสดงผลลัพธ์ของการแบ่งข้อมูลด้วยกราฟเส้น โดยใช้ข้อมูลจากแต่ละช่วงวาดแยกกันเพื่อเปรียบเทียบได้ชัดเจน โดยชุดข้อมูลฝึก

จะแสดงด้วยเส้นสีน้ำเงินเข้ม ส่วนชุดข้อมูลทดสอบจะแสดงด้วยเส้นสีส้ม พร้อมทั้งมีการใส่สัญลักษณ์จุด (marker) ในแต่ละวันเพื่อให้เห็นพฤติกรรมรายวันของยอดขายอย่างชัดเจน และเพิ่มเส้นแนวตั้งสีเทาในตำแหน่งวันสุดท้ายของชุดฝึกเพื่อเป็นตัวแบ่งระหว่างสองช่วง

การแสดงผลในลักษณะนี้ช่วยให้เห็นภาพรวมของข้อมูลได้ชัดเจนยิ่งขึ้น โดยเฉพาะการตรวจสอบพฤติกรรมเปลี่ยนแปลงของยอดขายระหว่างช่วงฝึกและช่วงทดสอบ เช่น การลดลงหรือผันผวนของยอดขายช่วงปลายชุดข้อมูล ตลอดจนจุดที่มียอดขายเป็นศูนย์ ซึ่งอาจเกิดจากวันหยุดหรือความผิดปกติของการเก็บข้อมูล โดยข้อมูลเหล่านี้มีความสำคัญต่อการพิจารณาความเหมาะสมของข้อมูลก่อนนำไปฝึกแบบจำลองพยากรณ์ และช่วยให้การวิเคราะห์มีความแม่นยำและเป็นระบบมากยิ่งขึ้น



ภาพประกอบ 16 แสดงการแบ่งชุดข้อมูล (Train-Test Split)

ก่อนดำเนินการสร้างแบบจำลองเพื่อพยากรณ์ยอดขายรายวัน ผู้วิจัยได้ดำเนินการเตรียมข้อมูลในเชิงโครงสร้างเวลาให้พร้อมสำหรับการเรียนรู้ของแบบจำลอง โดยมีขั้นตอนสำคัญประการหนึ่งคือ การแบ่งชุดข้อมูล (Data Splitting) ซึ่งเป็นกระบวนการแยกข้อมูลออกเป็น ชุดฝึกสอน (Training Set) และ ชุดทดสอบ (Testing Set) เพื่อประเมินประสิทธิภาพของแบบจำลองได้อย่างเป็นกลางและแม่นยำ

การแบ่งข้อมูลในงานวิจัยนี้มีจุดมุ่งหมายเพื่อจำลองสถานการณ์จริงที่โมเดลได้รับข้อมูลในอดีตเพื่อเรียนรู้ และทดสอบความสามารถในการพยากรณ์ข้อมูลในอนาคตที่ยังไม่เคยเห็นมาก่อน ซึ่งเหมาะสมกับลักษณะข้อมูลประเภทอนุกรมเวลา (Time Series) โดยมีการกำหนด วันที่ 30 เมษายน พ.ศ. 2567 เป็นจุดสิ้นสุดของชุดข้อมูลฝึกสอน จากนั้นจึงแบ่งข้อมูลออกเป็น 2 ส่วน คือ

train data ประกอบด้วยข้อมูลยอดขายรายวันตั้งแต่เดือนพฤษภาคม พ.ศ. 2566 ถึง 30 เมษายน พ.ศ. 2567 ซึ่งใช้สำหรับฝึกโมเดลและ Test data ประกอบด้วยข้อมูลตั้งแต่ 1 พฤษภาคม พ.ศ. 2567 เป็นต้นไป สำหรับใช้ประเมินผลการพยากรณ์ของโมเดล

หลังจากการแบ่งข้อมูลเสร็จสิ้น ได้มีการตรวจสอบขนาดของแต่ละชุดพบว่า ชุดฝึกสอนครอบคลุมข้อมูลประมาณ 360 วัน ขณะที่ชุดทดสอบครอบคลุมช่วงประมาณ 60 วัน ซึ่งเป็นสัดส่วนที่เหมาะสมสำหรับการประเมินแบบจำลองในบริบทของ Time Series ทั้งนี้ แนวทางการแบ่งข้อมูลตามลำดับเวลา (chronological order) ช่วยหลีกเลี่ยงการนำข้อมูลในอนาคตมาใช้ในการฝึกโมเดลโดยไม่ตั้งใจ ซึ่งจะส่งผลให้ผลการพยากรณ์มีความแม่นยำสมจริง และสามารถนำไปใช้งานในเชิงปฏิบัติได้อย่างมีประสิทธิภาพ

ขั้นตอนการแบ่งข้อมูลนี้จึงถือเป็นจุดเริ่มต้นสำคัญของการสร้างแบบจำลอง ที่ไม่เพียงแต่เตรียมข้อมูลสำหรับการเรียนรู้ แต่ยังเป็นกลไกที่ช่วยประเมินความสามารถเชิงคาดการณ์ของแบบจำลองได้อย่างมีนัยสำคัญอีกด้วย

ในการศึกษาครั้งนี้ ผู้วิจัยได้วางแผนทำการทดลองการใช้แบบจำลอง ออกเป็น 6 การทดลอง โดยมีรายละเอียดดังนี้

การทดลองที่ 1 มีวัตถุประสงค์เพื่อประเมินความแม่นยำเบื้องต้นของแบบจำลอง Prophet ในการพยากรณ์ยอดขายรายวัน โดยใช้โครงสร้างพื้นฐานของโมเดลแบบไม่ปรับแต่งค่าพารามิเตอร์เพิ่มเติม เพื่อใช้เป็นจุดอ้างอิงสำหรับเปรียบเทียบกับผลการทดลองในขั้นต่อไปที่มีการปรับแต่งโมเดลเพิ่มเติม

ก่อนเริ่มการสร้างแบบจำลอง ผู้วิจัยได้ดำเนินการเตรียมข้อมูลให้เหมาะสมกับข้อกำหนดของ Prophet โดยเปลี่ยนชื่อคอลัมน์ DATE และ ALL_SALE ให้เป็น ds และ y ตามลำดับ ซึ่งเป็นโครงสร้างมาตรฐานที่ Prophet ใช้ โดย ds แทนวันที่ในรูปแบบ datetime และ y แทนค่าตัวแปรเป้าหมายที่ต้องการพยากรณ์ (ยอดขายรายวัน)

```
# Import ที่ต้องไป
from prophet import Prophet
from sklearn.metrics import mean_absolute_error, mean_squared_error
import matplotlib.pyplot as plt
import numpy as np

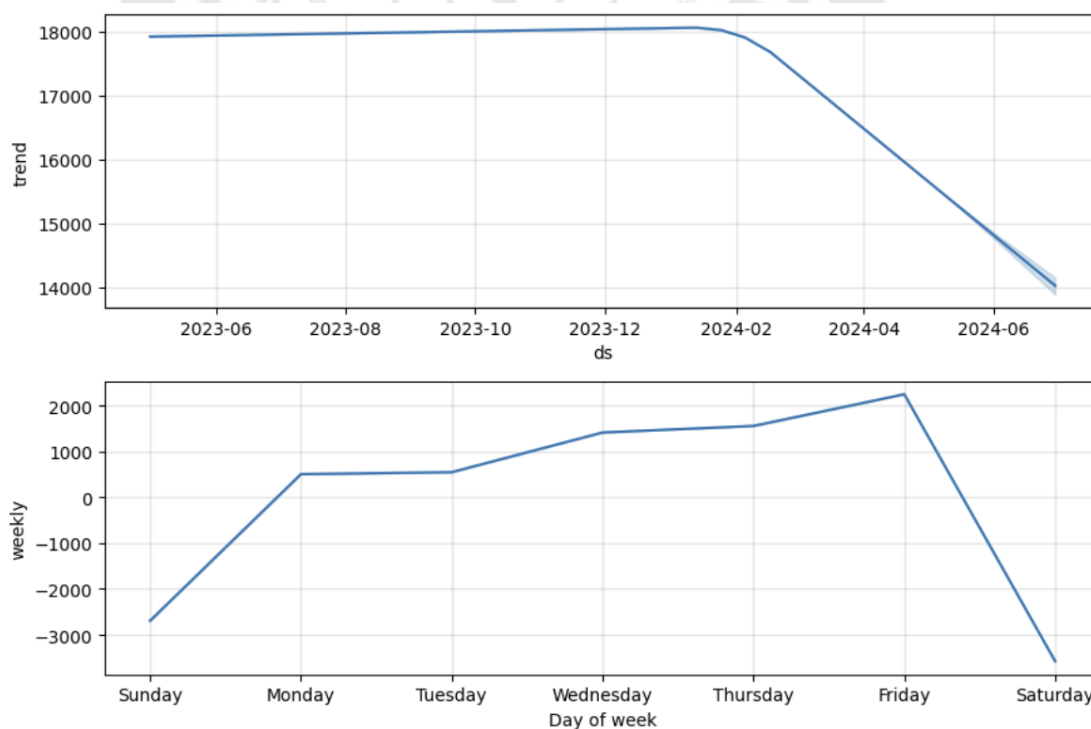
# ===== PROPHET MODEL =====

# เตรียมข้อมูลให้ Prophet (ต้องมี 'ds' และ 'y')
prophet_train = train_data.rename(columns={'DATE': 'ds', 'ALL_SALE': 'y'})
prophet_test = test_data.rename(columns={'DATE': 'ds', 'ALL_SALE': 'y'})
```

ภาพประกอบ 17 เปลี่ยนชื่อคอลัมน์ DATE และ ALL_SALE ให้เป็น ds และ y

หลังจากนั้นได้มีการแบ่งชุดข้อมูลออกเป็น 2 ส่วน ได้แก่ ชุดข้อมูลฝึกสอน (Training Set) ซึ่งครอบคลุมข้อมูลตั้งแต่เดือนพฤษภาคม พ.ศ. 2566 จนถึงวันที่ 30 เมษายน พ.ศ. 2567 และชุดข้อมูลทดสอบ (Testing Set) ซึ่งประกอบด้วยข้อมูลตั้งแต่วันที่ 1 พฤษภาคม พ.ศ. 2567 เป็นต้นไป การแบ่งข้อมูลนี้สอดคล้องกับหลักการของการวิเคราะห์ข้อมูลอนุกรมเวลา ซึ่งต้องรักษาลำดับเวลาตามจริงเพื่อให้การประเมินโมเดลสะท้อนผลลัพธ์ที่เกิดขึ้นในอนาคตได้อย่างแท้จริง นอกจากนี้ ยังมีการใช้ `.copy()` เพื่อแยกข้อมูลทั้งสองชุดออกจาก DataFrame ต้นฉบับอย่างชัดเจน ป้องกันไม่ให้เกิดปัญหาการเปลี่ยนแปลงข้อมูลร่วมกันโดยไม่ตั้งใจ

ในการสร้างแบบจำลอง ผู้วิจัยได้ใช้โมเดล Prophet โดยไม่กำหนดพารามิเตอร์เพิ่มเติม กล่าวคือใช้โครงสร้างแบบตั้งต้นที่ Prophet กำหนดไว้ เพื่อให้โมเดลเรียนรู้แนวโน้มและฤดูกาลตามค่าเริ่มต้นที่ระบบสร้างขึ้น จากนั้นได้ทำการฝึกโมเดลด้วยชุดข้อมูลฝึกสอน และใช้ฟังก์ชัน `make_future_dataframe()` เพื่อสร้างช่วงวันในอนาคตสำหรับการพยากรณ์ ซึ่งมีจำนวนวันเท่ากับช่วงของข้อมูลทดสอบ จากนั้นจึงใช้ฟังก์ชัน `predict()` เพื่อสร้างค่าพยากรณ์ยอดขายสำหรับแต่ละวันในช่วงที่ต้องการ



ภาพประกอบ 18 องค์ประกอบของโมเดล Prophet แสดง แนวโน้มและฤดูกาลรายสัปดาห์

หลังการฝึกด้วยข้อมูลยอดขายรายวันของร้านกาแฟ โดยมีวัตถุประสงค์เพื่อแสดงโครงสร้างภายในของโมเดลที่เกี่ยวข้องกับแนวโน้มหลักและรูปแบบฤดูกาลรายสัปดาห์ ซึ่งเป็นองค์ประกอบสำคัญในการพยากรณ์ยอดขายในอนาคต

กราฟด้านบนแสดงแนวโน้มหลักของยอดขายรายวัน (trend component) ซึ่งแสดงให้เห็นว่ายอดขายมีลักษณะคงที่ตั้งแต่ช่วงกลางปี พ.ศ. 2566 ไปจนถึงต้นปี พ.ศ. 2567 ก่อนจะเริ่มลดลงอย่างชัดเจนตั้งแต่ประมาณเดือนกุมภาพันธ์ พ.ศ. 2567 เป็นต้นไป แนวโน้มที่ลดลงนี้สอดคล้องกับผลการวิเคราะห์เชิงสำรวจที่พบว่ายอดขายในช่วงต้นปีมีการชะลอตัวลง ซึ่งอาจมีความเชื่อมโยงกับปัจจัยด้านฤดูกาล สภาพเศรษฐกิจ หรือพฤติกรรมของผู้บริโภคในช่วงเวลาดังกล่าว

ขณะที่กราฟด้านล่างแสดงฤดูกาลรายสัปดาห์ (weekly seasonality) ซึ่งแสดงความแตกต่างของยอดขายในแต่ละวันของสัปดาห์ โดยพบว่ายอดขายต่ำที่สุดในวันอาทิตย์ และมีแนวโน้มเพิ่มขึ้นเรื่อย ๆ ตั้งแต่วันจันทร์จนถึงจุดสูงสุดในวันศุกร์ จากนั้นลดลงอีกครั้งในวันเสาร์ ลักษณะการเปลี่ยนแปลงนี้สะท้อนถึงพฤติกรรมของการซื้อของลูกค้าที่อาจมีแนวโน้มจับจ่ายมากขึ้นในช่วงปลายสัปดาห์ และพักการบริโภคหรือลดกิจกรรมในช่วงวันหยุดอย่างวันอาทิตย์

ภาพรวมขององค์ประกอบทั้งสองนี้ช่วยให้เข้าใจโครงสร้างภายในของแบบจำลอง Prophet ได้ชัดเจนยิ่งขึ้น อีกทั้งยังสามารถนำไปใช้สนับสนุนการตัดสินใจเชิงกลยุทธ์ในทางธุรกิจ เช่น การกำหนดวันจัดกิจกรรมส่งเสริมการขาย หรือการจัดสรรทรัพยากรบุคคลให้เหมาะสมกับปริมาณยอดขายที่คาดการณ์ไว้ในแต่ละวันของสัปดาห์

การทดลองที่ 2 การปรับแต่งโมเดล Prophet ด้วย Grid Search เพื่อเพิ่มความแม่นยำของการพยากรณ์ยอดขายรายวัน การทดลองที่ 2 ได้ดำเนินการปรับแต่งโมเดล Prophet โดยใช้เทคนิค Grid Search ซึ่งเป็นกระบวนการค้นหาค่าพารามิเตอร์ที่เหมาะสมที่สุดอย่างเป็นระบบ โดยมีเป้าหมายเพื่อให้โมเดลสามารถเรียนรู้แนวโน้มและฤดูกาลของข้อมูลได้อย่างลึกซึ้งมากขึ้นกว่าการตั้งค่าเริ่มต้น การทดลองนี้ครอบคลุมชุดค่าพารามิเตอร์ที่แตกต่างกันทั้งหมด 128 ชุด ซึ่ง

แต่ละชุดประกอบด้วยค่าที่หลากหลายของพารามิเตอร์สำคัญของ Prophet ได้แก่ changepoint_prior_scale, seasonality_prior_scale, holidays_prior_scale และ seasonality_mode ค่าพารามิเตอร์ changepoint_prior_scale มีบทบาทในการควบคุมความไวของโมเดลต่อการเปลี่ยนแปลงแนวโน้มของข้อมูล หากค่าดังกล่าวต่ำ โมเดลจะตอบสนองต่อจุดเปลี่ยนแปลงของแนวโน้มได้ช้ากว่า แต่หากตั้งค่าสูงเกินไป อาจเกิดความเสี่ยงต่อการ overfitting ขณะที่ค่าของ seasonality_prior_scale และ holidays_prior_scale ใช้สำหรับกำหนดระดับความยืดหยุ่นที่โมเดลจะใช้ในการเรียนรู้รูปแบบฤดูกาลและผลกระทบจากวันหยุดตามลำดับ ซึ่งใน

ธุรกิจร้านค้าแฟ้มักได้รับผลกระทบจากวันหยุดราชการหรือเทศกาลอย่างมีนัยสำคัญ ส่วน seasonality_mode เป็นตัวกำหนดรูปแบบของฤดูกาลว่าจะมีผลในลักษณะการบวกแบบคงที่ (additive) หรือในลักษณะสัดส่วน (multiplicative) ในการทดลองนี้ โมเดลจะถูกฝึกซ้ำด้วยค่าพารามิเตอร์แต่ละชุด โดยใช้ข้อมูลฝึกสอนเดิมจากการทดลองที่ 1 และทำการพยากรณ์ข้อมูลในชุดทดสอบ จากนั้นเปรียบเทียบผลลัพธ์ของแต่ละชุดพารามิเตอร์ด้วยตัวชี้วัดความแม่นยำ เช่น Mean Absolute Percentage Error (MAPE) เพื่อระบุชุดค่าที่ให้ผลลัพธ์ที่ดีที่สุด กระบวนการนี้ช่วยให้สามารถประเมินผลกระทบของพารามิเตอร์แต่ละตัวที่มีต่อประสิทธิภาพของการพยากรณ์ และคัดเลือกโมเดลที่เหมาะสมที่สุดสำหรับชุดข้อมูลจริง

ผลการปรับแต่งพารามิเตอร์ด้วย Grid Search ถือเป็นส่วนสำคัญของกระบวนการ fine-tuning ที่ช่วยให้โมเดล Prophet สามารถตอบสนองต่อลักษณะเฉพาะของข้อมูลอนุกรมเวลาจากธุรกิจร้านค้าแฟ้มได้ดียิ่งขึ้น ไม่ว่าจะเป็นการเปลี่ยนแปลงของแนวโน้มในช่วงเวลาต่าง ๆ ความสม่ำเสมอของฤดูกาล หรือความผันผวนจากปัจจัยภายนอก

```
# Import ที่ต้องใช้
import itertools
import numpy as np
from prophet import Prophet
from sklearn.metrics import mean_absolute_error, mean_squared_error

param_grid = {
    'changepoint_prior_scale': [0.001, 0.01, 0.1, 0.5],
    'seasonality_prior_scale': [0.01, 0.1, 1.0, 10.0],
    'holidays_prior_scale': [0.01, 0.1, 1.0, 10.0],
    'seasonality_mode': ['additive', 'multiplicative']
}
```

ภาพประกอบ 19 การปรับแต่งพารามิเตอร์ด้วย Grid Search

การทดลองที่ 3 การประเมินประสิทธิภาพของโมเดล Prophet หลังการปรับจูนภายหลังการดำเนินการปรับแต่งโมเดล Prophet ด้วยเทคนิค Grid Search เพื่อค้นหาชุดค่าพารามิเตอร์ที่เหมาะสมที่สุดแล้ว การทดลองที่ 3 มุ่งเน้นไปที่การนำพารามิเตอร์ที่ได้จากกระบวนการปรับจูนดังกล่าวมาประยุกต์ใช้ในการฝึกและประเมินโมเดลอีกครั้ง โดยมีการกำหนดค่าเพิ่มเติมเพื่อเพิ่มความสามารถในการพยากรณ์ของแบบจำลองให้เหมาะสมกับลักษณะเฉพาะของข้อมูลยอดขายรายวันมากยิ่งขึ้นในการทดลองครั้งนี้ ผู้วิจัยได้ตั้งค่า changepoint_range เท่ากับ 0.9 เพื่อให้โมเดลสามารถพิจารณาจุดเปลี่ยนแนวโน้ม

(changepoints) ได้ครอบคลุมมากขึ้นในช่วงต้นของข้อมูล ซึ่งช่วยให้โมเดลตรวจจับการเปลี่ยนแปลงในเชิงพฤติกรรมของยอดขายได้แม่นยำยิ่งขึ้น นอกจากนี้ยังเพิ่มจำนวน n_changepoints เป็น 25 จุด เพื่อให้โมเดลมีความยืดหยุ่นในการตอบสนองต่อการเปลี่ยนแปลงแนวโน้มในช่วงเวลาต่าง ๆ ได้ละเอียดมากขึ้น อีกทั้งได้มีการเปิดใช้งานฤดูกาลรายวัน (daily_seasonality=True) เพื่อให้โมเดลสามารถเรียนรู้รูปแบบที่เกิดขึ้นซ้ำเป็นประจำทุกวัน ซึ่งมีความสำคัญในบริบทของธุรกิจร้านค้าปลีกที่มีพฤติกรรมการซื้อของลูกค้าซ้ำ ๆ ในแต่ละวันของสัปดาห์

```
# ✅ Train final model with optimized parameters and additional settings
best_prophet_model = Prophet(
    changepoint_prior_scale=best_params['changepoint_prior_scale'],
    seasonality_prior_scale=best_params['seasonality_prior_scale'],
    holidays_prior_scale=best_params['holidays_prior_scale'],
    seasonality_mode=best_params['seasonality_mode'],
    changepoint_range=0.9, # Allow changepoints up to 90% of the training data
    n_changepoints=25,     # Set number of changepoints
    daily_seasonality=True # Enable daily seasonality explicitly
)
```

ภาพประกอบ 20 ตั้งค่า changepoint range = 0.9, n changepoints = 25 และ daily seasonality=True

โมเดล Prophet ที่ปรับค่าดังกล่าวจึงได้รับการฝึกซ้ำใหม่โดยใช้ชุดข้อมูลฝึกสอนเดิม จากนั้นจึงทำการพยากรณ์ข้อมูลในชุดทดสอบ และประเมินผลลัพธ์ที่ได้ โดยเปรียบเทียบกับข้อมูลจริงเพื่อคำนวณค่าตัวชี้วัด เช่น Mean Absolute Percentage Error (MAPE) และค่าเฉลี่ยความคลาดเคลื่อนในภาพรวม การทดลองนี้มีเป้าหมายเพื่อยืนยันว่าการปรับพารามิเตอร์ส่งผลให้โมเดลมีประสิทธิภาพสูงขึ้นอย่างมีนัยสำคัญ และสามารถตอบสนองต่อรูปแบบที่ซับซ้อนของข้อมูลยอดขายได้ดีกว่าค่าตั้งต้นที่ใช้ในการทดลองก่อนหน้านี้

ในการทดลองนี้ได้มีการปรับค่า changepoint_range จากค่าเริ่มต้นที่ 0.8 เป็น 0.9 ซึ่งมีวัตถุประสงค์เพื่อเพิ่มขอบเขตที่โมเดล Prophet สามารถพิจารณาการเปลี่ยนแปลงของแนวโน้มได้ครอบคลุมมากยิ่งขึ้น โดยเฉพาะในช่วงท้ายของชุดข้อมูลฝึกสอน การตั้งค่า changepoint_range ที่สูงขึ้นทำให้โมเดลสามารถตรวจจับจุดเปลี่ยนแปลงของแนวโน้มที่เกิดขึ้นใกล้กับปลายชุดข้อมูลได้อย่างแม่นยำ ซึ่งในกรณีของข้อมูลยอดขายร้านค้าปลีกพบว่ามีลักษณะความผันผวนและแนวโน้มที่เปลี่ยนแปลงอย่างชัดเจนในช่วงเดือนเมษายน พ.ศ. 2567 เป็นต้นไป

หากคงค่า changepoint_range ไว้ที่ระดับเริ่มต้น 0.8 โมเดลจะจำกัดการค้นหาคจุดเปลี่ยนอยู่เพียง 80% แรกของข้อมูล ซึ่งอาจทำให้ละเลยแนวโน้มที่เปลี่ยนแปลงอย่างสำคัญในช่วงปลายก่อนเข้าสู่ชุดทดสอบ โดยเฉพาะช่วงที่ยอดขายเริ่มลดลงหรือฟื้นตัว ซึ่งเป็นข้อมูลสำคัญต่อการพยากรณ์อนาคต การเพิ่ม changepoint_range จึงถือเป็นการเพิ่มความยืดหยุ่นให้กับโมเดล ช่วยให้ Prophet สามารถจับการเปลี่ยนแปลงเหล่านี้ได้แม่นยำยิ่งขึ้น

แนวทางนี้เหมาะสมอย่างยิ่งสำหรับข้อมูลที่มีลักษณะไม่คงที่ หรือมีความเป็นไปได้ที่จะเกิดเหตุการณ์สำคัญในช่วงปลายของชุดฝึกสอน เช่น การเปลี่ยนแปลงเชิงนโยบายของร้าน การจัดโปรโมชั่น การปรับราคาสินค้า หรือพฤติกรรมของลูกค้าที่เปลี่ยนแปลงตามฤดูกาล การปรับค่าดังกล่าวจึงเป็นองค์ประกอบสำคัญที่ส่งเสริมให้โมเดล Prophet มีความสามารถในการพยากรณ์ที่ตอบสนองต่อพลวัตของข้อมูลจริงได้ดียิ่งขึ้น และลดความคลาดเคลื่อนของผลการพยากรณ์โดยรวม

การทดลองที่ 4 การควบคุม Overfitting ด้วยการลดจำนวนจุดเปลี่ยนแนวโน้มในขณะทำการทดลองก่อนหน้าเน้นการเพิ่มความยืดหยุ่นให้กับโมเดล Prophet เพื่อให้สามารถตอบสนองต่อการเปลี่ยนแปลงของแนวโน้มได้อย่างละเอียดและแม่นยำ การทดลองที่ 4 นี้มีวัตถุประสงค์ในทางกลับกัน กล่าวคือ มุ่งเน้นไปที่การ ควบคุมปัญหา Overfitting โดยการลดความซับซ้อนของโมเดล เพื่อให้ผลการพยากรณ์มีความเรียบและเสถียรมากขึ้น การทดลองนี้เริ่มต้นจากการลดจำนวนจุดเปลี่ยนแนวโน้ม (n_changepoints) จากค่าตั้งต้นที่ 25 จุด ให้เหลือเพียง 10 จุด และปรับค่า changepoint_range กลับไปใช้ค่ามาตรฐานที่ 0.8 ซึ่งหมายความว่าโมเดลจะพิจารณาการเปลี่ยนแปลงของแนวโน้มเฉพาะในช่วง 80% แรกของข้อมูลเท่านั้น การลดจำนวน changepoints มีจุดมุ่งหมายเพื่อลดโอกาสที่โมเดลจะตอบสนองต่อความผันผวนของข้อมูลที่เป็น noise หรือความเปลี่ยนแปลงเล็กน้อยที่ไม่สำคัญ ซึ่งมักพบในข้อมูลอนุกรมเวลาที่บันทึกในระดับวัน

```
# ✅ สร้างโมเดล Prophet ที่ลดการ overfit โดยลด n_changepoints
low_cp_model = Prophet(
    changepoint_prior_scale=best_params['changepoint_prior_scale'],
    seasonality_prior_scale=best_params['seasonality_prior_scale'],
    holidays_prior_scale=best_params['holidays_prior_scale'],
    seasonality_mode=best_params['seasonality_mode'],
    changepoint_range=0.8, # กลับมาใช้ค่า default
    n_changepoints=10,     # ลดจำนวน changepoints เพื่อให้โมเดลเรียบ
    daily_seasonality=True
)
```

ภาพประกอบ 21 ตั้งค่า changepoint range = 0.8, n changepoints = 10 และ daily seasonality=True

แนวคิดของการทดลองนี้ตั้งอยู่บนสมมติฐานว่า ยอดขายรายวันอาจมีความผันผวนสูงจากปัจจัยชั่วคราว เช่น สภาพอากาศ พฤติกรรมเฉพาะวัน หรือเหตุการณ์ที่ไม่เกิดซ้ำ การให้โมเดลตอบสนองต่อความเปลี่ยนแปลงเหล่านี้มากเกินไปจะนำไปสู่การ Overfitting ซึ่งหมายถึง โมเดลเรียนรู้ข้อมูลเฉพาะจุดมากเกินไปจนไม่สามารถพยากรณ์ข้อมูลใหม่ได้ดี การลดจำนวน changepoints จึงเป็นวิธีที่ช่วยให้โมเดลสามารถเรียนรู้โครงสร้างแนวโน้มที่มีเสถียรมากขึ้น โดยไม่หลงตามความเปลี่ยนแปลงเล็กน้อยในชุดฝึกสอน

ผลการทดลองที่ 4 จะถูกนำไปเปรียบเทียบกับผลการทดลองก่อนหน้านี้ โดยเฉพาะการทดลองที่ 3 ซึ่งใช้จำนวน changepoints สูงและครอบคลุมจุดเปลี่ยนแนวโน้มได้มากกว่า เพื่อตรวจสอบว่าแบบจำลองที่มีแนวโน้มเรียบกว่า จะส่งผลอย่างไรต่อความแม่นยำในการพยากรณ์ยอดขาย และสามารถป้องกันการ overfit ได้จริงหรือไม่ การเปรียบเทียบนี้จะสะท้อนถึงการ trade-off ระหว่างความยืดหยุ่นของโมเดลและความสามารถในการ generalize ไปยังข้อมูลในอนาคต ซึ่งเป็นประเด็นสำคัญของการออกแบบแบบจำลองอนุกรมเวลาในงานวิจัยนี้

การทดลองที่ 5 การพยากรณ์ยอดขายด้วยเครื่องมือ Forecast ใน Tableau จากการทดลองพยากรณ์ยอดขายรายวันด้วยเครื่องมือ Forecast ใน Tableau ผู้วิจัยได้กำหนดช่วงการพยากรณ์ล่วงหน้าเป็นระยะเวลา 60 วัน โดยครอบคลุมตั้งแต่วันที่ 30 เมษายน พ.ศ. 2567 ถึงวันที่ 28 มิถุนายน พ.ศ. 2567 เพื่อประเมินศักยภาพของ Tableau ในการสร้างแบบจำลองเชิงอนุกรมเวลาอย่างอัตโนมัติในบริบทของข้อมูลจริงจากร้านกาแฟ

ในการสร้างแบบจำลองพยากรณ์ดังกล่าว ได้ใช้ข้อมูลย้อนหลังในช่วงระหว่างวันที่ 31 ธันวาคม พ.ศ. 2566 ถึงวันที่ 29 เมษายน พ.ศ. 2567 ซึ่งเป็นข้อมูลที่อยู่ในช่วงปลายของชุดข้อมูลทั้งหมด และสะท้อนลักษณะของพฤติกรรมผู้บริโภคที่เปลี่ยนแปลงตามช่วงเวลาที่มีความอ่อนไหว เช่น ช่วงต้นปีหรือก่อนฤดูฝน ซึ่งอาจส่งผลต่อปริมาณยอดขายรายวันได้อย่างมีนัยสำคัญ การพยากรณ์ดำเนินการในระดับข้อมูลรายวัน (daily frequency) ซึ่งเหมาะสมกับลักษณะการดำเนินงานของร้านค้าประเภทเครื่องดื่มและอาหารที่มียอดขายเปลี่ยนแปลงทุกวันตามพฤติกรรมของลูกค้า โดย Tableau ได้ใช้โครงสร้างของฤดูกาลแบบรอบสัปดาห์ (weekly seasonality) ซึ่งถือว่าเหมาะสมกับบริบทของธุรกิจร้านกาแฟที่มักมีรูปแบบการซื้อซ้ำในแต่ละวันของสัปดาห์ เช่น วันศุกร์มักมียอดขายสูง ขณะที่วันอาทิตย์อาจมียอดขายลดลงทั้งแนวโน้ม (trend) และฤดูกาล (seasonality) ถูกตั้งค่าให้อยู่ในรูปแบบ "ส่วนเสริม" (additive) ซึ่งหมายความว่าผลของฤดูกาลและแนวโน้มจะถูกเพิ่มเข้ากับค่าคาดการณ์โดยตรงในลักษณะเชิงเส้น การกำหนดโครงสร้างแบบ

additive นี้เหมาะกับกรณีที่ฤดูกาลมีผลกระทบคงที่ต่อยอดขาย ไม่แปรผันตามระดับของยอดขาย ในช่วงต่าง ๆ ซึ่งสอดคล้องกับลักษณะข้อมูลในชุดนี้ที่ไม่ได้มีแนวโน้มเติบโตแบบทวีคูณ

นอกจากนี้ ยังมีการเติมค่าที่หายไปในช่วงข้อมูลต้นทางด้วยค่าเฉลี่ย (mean imputation) ซึ่งเป็นวิธีที่ช่วยให้โมเดลไม่ถูกปิดเบือนจากข้อมูลที่ขาดหาย และช่วยรักษาความต่อเนื่องของลำดับเวลาให้มีเสถียรภาพในเชิงโครงสร้าง ในการแสดงผล ผู้วิจัยได้เลือกให้แสดงช่วงความเชื่อมั่นของการพยากรณ์ที่ระดับ 95% เพื่อประเมินความไม่แน่นอนของแบบจำลองในแต่ละวัน ซึ่งช่วยให้สามารถตัดสินใจเชิงกลยุทธ์บนพื้นฐานของช่วงค่าที่คาดว่าจะเกิดขึ้นจริงได้อย่างรอบคอบมากยิ่งขึ้น

การตั้งค่าดังกล่าวทั้งหมดสะท้อนถึงการใช้ Tableau เพื่อสร้างการพยากรณ์ยอดขายที่มีความโปร่งใส ชัดเจน และเหมาะสมกับผู้ใช้งานเชิงธุรกิจที่ต้องการความสะดวกรวดเร็วในการสร้างแบบจำลองโดยไม่ต้องเขียนโค้ด ในขณะเดียวกันก็สามารถควบคุมองค์ประกอบหลักของแบบจำลองได้อย่างยืดหยุ่นภายใต้กรอบแนวคิดเชิงอนุกรมเวลา

ตัวเลือกการพยากรณ์

ความยาวของการพยากรณ์

☐ หักโนมีนัล 60 วันถัดไป

☐ ตรงทุกประการ 1 ปี

☒ จนถึง 60 วัน

ข้อมูลต้นทาง

รวมโดย: วัน

ละเว้นล่าสุด: 1 วัน

☒ เติมค่าที่หายไปด้วยศูนย์

แบบจำลองการพยากรณ์

กำหนดเอง

แนวโน้ม: ส่วนเสริม

ซีซั่น: ส่วนเสริม

☒ แสดงช่วงการคาดการณ์ 95%

ขณะนี้กำลังใช้แหล่งข้อมูลจาก 31 December 2566 ถึง 29 April 2567 เพื่อสร้างการพยากรณ์ผ่าน 28 June 2567

ค้นหารูปแบบที่อาจเกิดขึ้นในแต่ละช่วงเวลาทุก 7 วัน

เรียนรู้เพิ่มเติมเกี่ยวกับตัวเลือกการพยากรณ์

ตกลง

ภาพประกอบ 22 การกำหนดช่วงเวลาและโครงสร้างแบบจำลองการพยากรณ์ใน Tableau โดยใช้แนวโน้มและฤดูกาลแบบส่วนเสริม

การทดลองที่ 6 การพยากรณ์ยอดขายใน Tableau ด้วยแนวโน้มและฤดูกาลแบบทวีคูณ หลังจากที่ได้ผู้วิจัยได้ดำเนินการพยากรณ์ยอดขายรายวันใน Tableau โดยใช้แนวโน้มและฤดูกาล

แบบ *ส่วนเสริม* (additive) ไปแล้วในการทดลองก่อนหน้านี้ การทดลองที่ 6 นี้จึงมุ่งเน้นไปที่การเปรียบเทียบโครงสร้างของแบบจำลอง โดยเปลี่ยนมาใช้แนวโน้มและฤดูกาลในรูปแบบ *ทวีคูณ* (multiplicative) แทน เพื่อวิเคราะห์ว่าโครงสร้างดังกล่าวจะช่วยให้แบบจำลองสามารถจับความผันแปรของยอดขายที่สัมพันธ์กับระดับยอดขายได้ดีขึ้นหรือไม่

แบบจำลองใน Tableau ได้รับการตั้งค่าให้พยากรณ์ยอดขายล่วงหน้าเป็นระยะเวลา 60 วัน เริ่มตั้งแต่วันที่ 30 เมษายน พ.ศ. 2567 ถึงวันที่ 28 มิถุนายน พ.ศ. 2567 โดยใช้ข้อมูลย้อนหลังตั้งแต่วันที่ 31 ธันวาคม พ.ศ. 2566 ถึงวันที่ 29 เมษายน พ.ศ. 2567 สำหรับการฝึกแบบจำลองระดับของข้อมูลถูกกำหนดไว้เป็นรายวัน (daily) และมีการเติมค่าที่ขาดหายไปด้วยค่าเฉลี่ย (mean imputation) เพื่อรักษาความต่อเนื่องของชุดข้อมูลในส่วนของการสร้างโมเดล ได้เลือกใช้แนวโน้มและฤดูกาลเป็นแบบ *ทวีคูณ* ซึ่งหมายความว่า ผลของฤดูกาลในแต่ละวันจะถูกคูณกับค่าพยากรณ์ที่ได้จากแนวโน้มหลักของยอดขาย โดยสมมติว่าอิทธิพลของฤดูกาลจะมีความรุนแรงมากขึ้นหากยอดขายโดยรวมมีค่ามากขึ้น และลดลงหากยอดขายลดลง ซึ่งต่างจากแบบ *ส่วนเสริม* ที่ฤดูกาลจะมีค่าคงที่ไม่ขึ้นกับระดับของยอดขาย โมเดลแบบทวีคูณจึงมีความยืดหยุ่นมากกว่าในกรณีที่ฤดูกาลมีผลกระทบในเชิงสัดส่วน เช่น ฤดูกาลที่ชัดเจนในช่วงยอดขายสูง หรือช่วงโปรโมชั่นที่เสริมฤดูกาลให้ชัดเจน

นอกจากนี้ ยังมีการกำหนดให้แสดงช่วงความเชื่อมั่นของค่าพยากรณ์ที่ระดับ 95% เพื่อให้สามารถประเมินความไม่แน่นอนของการพยากรณ์ได้อย่างเหมาะสม ทั้งนี้ Tableau ใช้รูปแบบฤดูกาลแบบรายสัปดาห์ (weekly seasonality) ซึ่งจะค้นหารูปแบบที่เกิดขึ้นเป็นรอบทุก 7 วัน และคาดการณ์ค่าตามโครงสร้างดังกล่าวว่าการทดลองที่ 6 นี้จึงมีจุดประสงค์เพื่อเปรียบเทียบกับผลการทดลองที่ 5 โดยตรงว่าการเปลี่ยนรูปแบบแนวโน้มและฤดูกาลจากส่วนเสริมมาเป็นทวีคูณจะส่งผลต่อความแม่นยำหรือความสมจริงของผลการพยากรณ์อย่างไร โดยเฉพาะในกรณีที่ยอดขายมีการเปลี่ยนแปลงตามฤดูกาลอย่างไม่สม่ำเสมอ ซึ่งเป็นลักษณะที่พบได้บ่อยในธุรกิจร้านค้าแฟ การเปรียบเทียบนี้จะนำไปสู่ข้อเสนอแนะเชิงปฏิบัติในการเลือกใช้แบบจำลองในระบบ Tableau ให้เหมาะสมกับลักษณะของข้อมูลในบริบทธุรกิจจริง

ภาพประกอบ 23 การกำหนดช่วงเวลาและโครงสร้างแบบจำลองการพยากรณ์ใน Tableau โดยใช้แนวโน้มและฤดูกาลแบบทวีคูณ

Tableau ใช้กลไกของแบบจำลอง Exponential Smoothing ในการพยากรณ์ข้อมูลอนุกรมเวลา ซึ่งมีลักษณะของแนวโน้มและฤดูกาลที่สามารถกำหนดได้ทั้งในรูปแบบส่วนเสริม (additive) และแบบทวีคูณ (multiplicative) สำหรับข้อมูลในบริบทของธุรกิจร้านค้ากาแฟซึ่งมียอดขายที่แปรผันตามฤดูกาลอย่างต่อเนื่อง เช่น การเพิ่มขึ้นของยอดขายในช่วงวันหยุดสุดสัปดาห์หรือเทศกาลสำคัญ โมเดลแบบทวีคูณของ Tableau จึงสามารถจับลักษณะการเปลี่ยนแปลงในลักษณะสัดส่วนได้อย่างเหมาะสม กล่าวคือ ยิ่งยอดขายโดยรวมเพิ่มขึ้น ฤดูกาลจะยังมีอิทธิพลต่อผลลัพธ์มากขึ้นตามไปด้วย

อย่างไรก็ตาม แม้ว่า Tableau จะมีข้อดีด้านความสะดวกในการใช้งานและเหมาะกับผู้ใช้ทั่วไปในสายธุรกิจที่ต้องการวิเคราะห์ผลเบื้องต้นโดยไม่ต้องเขียนโค้ด แต่แบบจำลองที่ใช้ใน Tableau ยังคงมีข้อจำกัดบางประการ โดยเฉพาะเมื่อเปรียบเทียบกับแบบจำลอง Prophet ที่ผู้วิจัยสามารถปรับพารามิเตอร์ได้อย่างละเอียด Tableau ไม่มีความสามารถในการปรับค่าพารามิเตอร์ขั้นสูง เช่น changepoint prior scale หรือ seasonality priorscale ซึ่งเป็นค่าที่ส่งผลโดยตรงต่อความยืดหยุ่นของโมเดลในการตรวจจับแนวโน้มที่เปลี่ยนแปลงอย่างเฉียบพลัน

นอกจากนี้ Tableau ยังไม่สามารถกำหนดจุดเปลี่ยนแนวโน้ม (changepoints) ได้เองหรือรองรับวันหยุดพิเศษ (holiday effects) ที่อาจมีอิทธิพลต่อยอดขายเฉพาะจุด เช่น วันหยุดนักขัตฤกษ์ วันโปรโมชัน หรือเหตุการณ์ที่ส่งผลกระทบเฉพาะกิจต่อพฤติกรรมของผู้บริโภค ทำให้ในกรณีที่ข้อมูลมีความผันผวนสูงในช่วง หรือมีลักษณะเหตุการณ์เฉพาะเจาะจง โมเดลใน Tableau อาจไม่สามารถตอบสนองได้อย่างแม่นยำเท่ากับ Prophet หรือโมเดล Machine Learning อื่นที่เปิดให้ผู้ใช้งานกำหนดโครงสร้างและกระบวนการเรียนรู้ได้อย่างยืดหยุ่น

การเปรียบเทียบนี้สะท้อนให้เห็นว่า แม้ Tableau จะเหมาะสำหรับการใช้งานเบื้องต้นและการวิเคราะห์ข้อมูลเชิงธุรกิจอย่างรวดเร็ว แต่ในการใช้งานที่ต้องการความแม่นยำสูงหรือการพยากรณ์ในสถานการณ์ที่ซับซ้อน เช่น การวางแผนการผลิต การบริหารสต็อก หรือการตอบสนองต่อเหตุการณ์ไม่ปกติ การเลือกใช้แบบจำลองที่สามารถควบคุมพารามิเตอร์อย่างละเอียดได้จะเป็นทางเลือกที่เหมาะสมกว่า



บทที่ 4

ผลการดำเนินการวิจัย

จากการดำเนินการวิจัยเกี่ยวกับการพยากรณ์ยอดขายรายวันของร้านกาแฟ โดยมีวัตถุประสงค์เพื่อประเมินประสิทธิภาพของแบบจำลองการพยากรณ์ในบริบทของข้อมูลจริงจากร้านค้าประเภทเครื่องดื่มและอาหาร ซึ่งมีลักษณะยอดขายที่ผันผวนในระดับวัน และได้รับอิทธิพลจากแนวโน้มเชิงเวลา รวมถึงฤดูกาลที่เกิดขึ้นซ้ำตามช่วงของสัปดาห์หรือเทศกาล

ในการดำเนินการวิจัย ผู้วิจัยได้ใช้ข้อมูลยอดขายรายวันตั้งแต่เดือนพฤษภาคม พ.ศ. 2566 จนถึงเดือนเมษายน พ.ศ. 2567 ซึ่งผ่านกระบวนการเตรียมข้อมูลเบื้องต้นในบทที่ 3 แล้ว เพื่อให้สอดคล้องกับรูปแบบการนำเข้าและข้อกำหนดของแบบจำลองต่าง ๆ ที่ใช้ในงานวิจัยครั้งนี้ โดยแบ่งออกเป็น 2 ส่วน ได้แก่ ข้อมูลสำหรับการฝึกแบบจำลอง (Training Set) และข้อมูลสำหรับทดสอบผลการพยากรณ์ (Testing Set)

ในการศึกษานี้ได้ดำเนินการทดลองจำนวน 6 การทดลอง เพื่อประเมินแนวทางในการพยากรณ์ยอดขายรายวันของร้านกาแฟ โดยแต่ละการทดลองถูกออกแบบให้แตกต่างกันในด้านวิธีการปรับแต่งแบบจำลองและเครื่องมือที่ใช้ เพื่อเปรียบเทียบแนวทางการวิเคราะห์ที่หลากหลาย

การทดลองที่ 1 เป็นการใช้แบบจำลอง Prophet โดยไม่ได้มีการปรับแต่งพารามิเตอร์เพิ่มเติม ใช้ค่าตั้งต้นที่ระบบกำหนดไว้เพื่อสร้างโมเดล และดำเนินการแบ่งชุดข้อมูลออกเป็นชุดฝึกสอน (training set) และชุดทดสอบ (test set) โดยการพยากรณ์จะครอบคลุมช่วงเวลาที่อยู่ถัดจาก training set โดยตรง โมเดล Prophet ถูกฝึกด้วยข้อมูลที่จัดโครงสร้างใหม่ให้เหมาะสมกับรูปแบบของโมเดล โดยเปลี่ยนชื่อคอลัมน์เป็น ds และ y ตามข้อกำหนด

การทดลองที่ 2 ได้ต่อยอดจากการทดลองแรกโดยมีการปรับพารามิเตอร์ของ Prophet ด้วยกระบวนการ Grid Search ซึ่งเป็นเทคนิคการทดลองแบบครอบคลุม (Exhaustive Search) ในการกำหนดชุดค่าพารามิเตอร์ที่หลากหลาย เพื่อเลือกชุดค่าที่ให้ผลการพยากรณ์ที่ดีที่สุด โดยพารามิเตอร์ที่ใช้ทดสอบได้แก่ ค่า changepoint, seasonality และโหมดของฤดูกาล

การทดลองที่ 3 เป็นการนำพารามิเตอร์ที่ได้จาก Grid Search มาปรับเปลี่ยนเพิ่มเติมอีกครั้งเพื่อให้โมเดลมีความยืดหยุ่นมากขึ้นในการตรวจจับการเปลี่ยนแปลงแนวโน้ม โดยเฉพาะในช่วงท้ายของข้อมูล ซึ่งอาจเกิดเหตุการณ์พิเศษ เช่น วันหยุด หรือยอดขายลดลงกะทันหัน จึงมีการปรับค่า changepoint_range ให้ครอบคลุมช่วงเวลาดังกล่าวมากขึ้น พร้อมเพิ่มจำนวนจุดเปลี่ยนแนวโน้ม (changepoints) และเปิดใช้งานฤดูกาลรายวัน (daily seasonality)

การทดลองที่ 4 มุ่งเน้นการควบคุมความซับซ้อนของโมเดลเพื่อหลีกเลี่ยงปัญหา Overfitting โดยลดจำนวนจุด changepoints ลง และกลับไปใช้ช่วงเวลาของ changepoints ตามค่าตั้งต้น การปรับนี้มีเป้าหมายเพื่อให้โมเดลเรียนรู้เฉพาะแนวโน้มที่มีความสำคัญเท่านั้น ไม่ตอบสนองต่อความผันผวนที่เกิดจาก noise หรือความไม่สม่ำเสมอของข้อมูลในระดับวัน

การทดลองที่ 5 ได้ใช้เครื่องมือ Tableau เพื่อดำเนินการพยากรณ์ยอดขายรายวันในช่วงเวลาเดียวกับที่ใช้ในแบบจำลอง Prophet โดยใช้แบบจำลอง Exponential Smoothing ที่มีการตั้งค่ารูปแบบแนวโน้มและฤดูกาลเป็นแบบ “ส่วนเสริม” (Additive Trend & Additive Seasonality) ลักษณะการทำงานของ Tableau ในการพยากรณ์นี้จะอาศัยข้อมูลย้อนหลังเป็นช่วงเวลาที่ โดยในกรณีนี้เลือกใช้ข้อมูลย้อนหลังประมาณ 120 วันก่อนเริ่มต้นช่วงทำนาย ซึ่งระบบของ Tableau จะทำการสร้างแบบจำลองโดยอัตโนมัติจากข้อมูลในอดีต พร้อมกำหนดฤดูกาลแบบรายสัปดาห์ และแสดงผลการพยากรณ์รวมถึงช่วงความเชื่อมั่น (Confidence Interval) ที่ระดับ 95% วิธีนี้เหมาะสมกับข้อมูลที่มีแนวโน้มและฤดูกาลไม่รุนแรงมากนัก และเน้นความสะดวกในการใช้งานสำหรับการวิเคราะห์เบื้องต้นเชิงธุรกิจ

การทดลองที่ 6 การทดลองนี้เป็นการประเมินความสามารถในการพยากรณ์ยอดขายของ Tableau ภายใต้เงื่อนไขการตั้งค่าแบบจำลองที่แตกต่างจากการทดลองก่อนหน้านี้ โดยเปลี่ยนรูปแบบแนวโน้มและฤดูกาลจากแบบ “ส่วนเสริม” (Additive) เป็นแบบ “ทวีคูณ” (Multiplicative) ซึ่งเหมาะสำหรับข้อมูลที่ยอดขายมีลักษณะเปลี่ยนแปลงตามฤดูกาลในลักษณะของสัดส่วนมากกว่าการเพิ่มหรือลดในระดับคงที่ ตัวอย่างเช่น ยอดขายที่เพิ่มขึ้นอย่างชัดเจนในช่วงวันหยุดหรือช่วงเทศกาลสำคัญ

แบบจำลองที่ใช้ในการทดลองนี้ยังคงอิงกับ Exponential Smoothing เช่นเดียวกับการทดลองที่ 5 โดยมีการตั้งค่าให้พยากรณ์ยอดขายรายวันในช่วงเวลา 60 วันข้างหน้า และใช้ข้อมูลย้อนหลังประมาณ 120 วันเพื่อฝึกโมเดล ระบบของ Tableau จะคำนวณค่าพารามิเตอร์ภายในโดยอัตโนมัติ และเติมค่าที่ขาดหายไปด้วยค่าเฉลี่ย (Mean Imputation) พร้อมทั้งแสดงช่วงคาดการณ์ที่ระดับความเชื่อมั่น 95% เช่นเดิม

การเปลี่ยนแปลงไปใช้แนวโน้มและฤดูกาลแบบทวีคูณในครั้งนี้ มีวัตถุประสงค์เพื่อทดสอบว่ารูปแบบโมเดลที่สามารถสะท้อนการเปลี่ยนแปลงของยอดขายในเชิงอัตราส่วน จะสามารถปรับตัวให้เข้ากับพฤติกรรมการซื้อขายของลูกค้าที่เปลี่ยนไปในเวลาพิเศษได้แม่นยำมากยิ่งขึ้นหรือไม่ โดยเฉพาะในธุรกิจร้านค้ากาแฟที่มีแนวโน้มยอดขายพุ่งขึ้นในวันหยุดสุดสัปดาห์หรือช่วงเทศกาล ซึ่งอาจไม่สามารถอธิบายได้ด้วยแนวโน้มเชิงเส้นเพียงอย่างเดียว

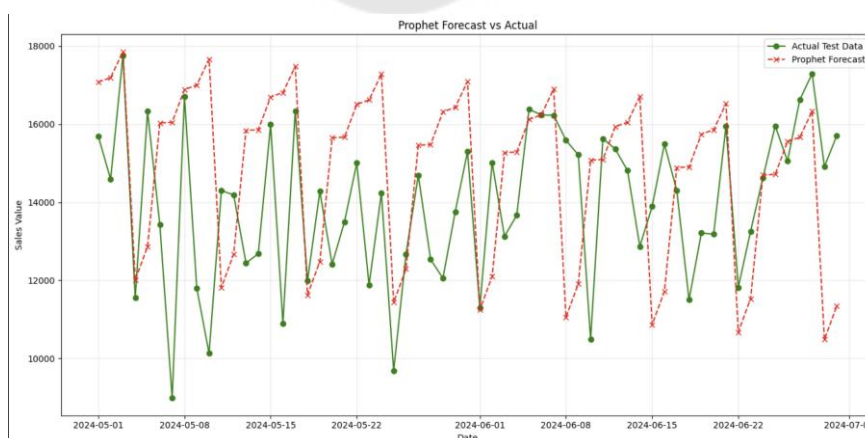
ผลลัพธ์ของการทดลองที่ 1

ในการทดลองที่ 1 ได้ทำการนำแบบจำลอง Prophet มาใช้พยากรณ์ยอดขายรายวันของร้านกาแฟ โดยใช้ข้อมูลในช่วงพฤษภาคม 2023 ถึงเมษายน 2024 เป็นชุดฝึกสอน (training set) และใช้ข้อมูลตั้งแต่พฤษภาคม 2024 เป็นต้นไปเป็นชุดทดสอบ (test set) จากนั้นจึงประเมินผลลัพธ์ของการพยากรณ์ด้วยตัวชี้วัดมาตรฐาน ได้แก่ MAE, RMSE และ MAPE ซึ่งช่วยสะท้อนความแม่นยำของโมเดลได้ในหลายมิติ

ผลการประเมินชี้ว่า โมเดล Prophet สามารถทำนายยอดขายได้ในระดับที่น่าพอใจ โดยมี MAE ประมาณ 2,272 บาทต่อวัน และ MAPE อยู่ที่ 17.73% ซึ่งถือว่าอยู่ในระดับปานกลางถึงดีในบริบทของข้อมูลธุรกิจจริง แสดงให้เห็นว่าโมเดลมีความสามารถในการพยากรณ์ยอดขายได้อย่างเหมาะสม แม้จะยังมีบางช่วงที่ค่าที่พยากรณ์ได้ต่ำกว่าค่าจริง โดยเฉพาะในวันที่มียอดขายพุ่งสูงซึ่งอาจเกิดจากเหตุการณ์พิเศษ เช่น โปรโมชัน หรือวันหยุดพิเศษ

นอกจากนี้ยังพบว่าแนวโน้มยอดขายมีลักษณะลดลงอย่างชัดเจนตั้งแต่เดือนเมษายน 2024 เป็นต้นไป และมีลักษณะฤดูกาลรายสัปดาห์ โดยยอดขายจะสูงในวันศุกร์ และลดลงในวันเสาร์-อาทิตย์ ซึ่งสอดคล้องกับพฤติกรรมผู้บริโภคในร้านกาแฟที่อาจลดลงในช่วงวันหยุดสุดสัปดาห์ ผลจากการแสดงองค์ประกอบของโมเดล (Trend และ Weekly Seasonality) ช่วยอธิบายรูปแบบข้อมูลได้อย่างชัดเจน และสามารถนำไปใช้ปรับกลยุทธ์ทางธุรกิจ เช่น การเสริมกิจกรรมส่งเสริมการขายในช่วงที่ยอดขายตกต่ำได้ในอนาคต

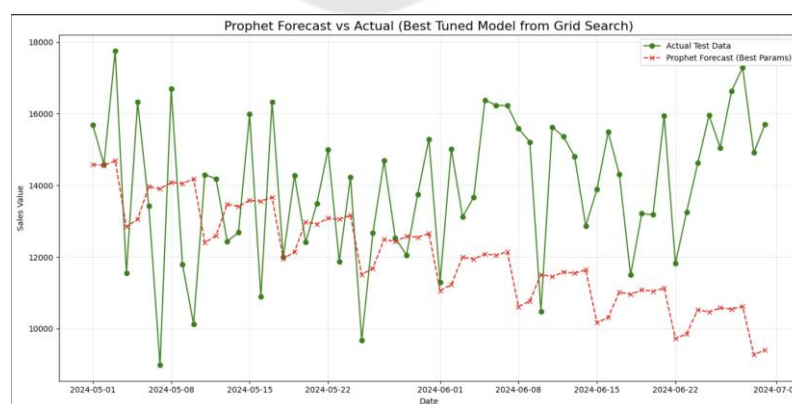
โดยสรุป การทดลองที่ 1 แสดงให้เห็นถึงความเหมาะสมของ Prophet ในการใช้พยากรณ์ยอดขายรายวัน โดยไม่ต้องปรับแต่งพารามิเตอร์เพิ่มเติมมากนัก แต่ยังมีข้อจำกัดบางประการในการจับความเปลี่ยนแปลงอย่างเฉียบพลัน



ภาพประกอบ 24 ผลลัพธ์จากการทดลองที่ 1

ผลลัพธ์ของการทดลองที่ 2

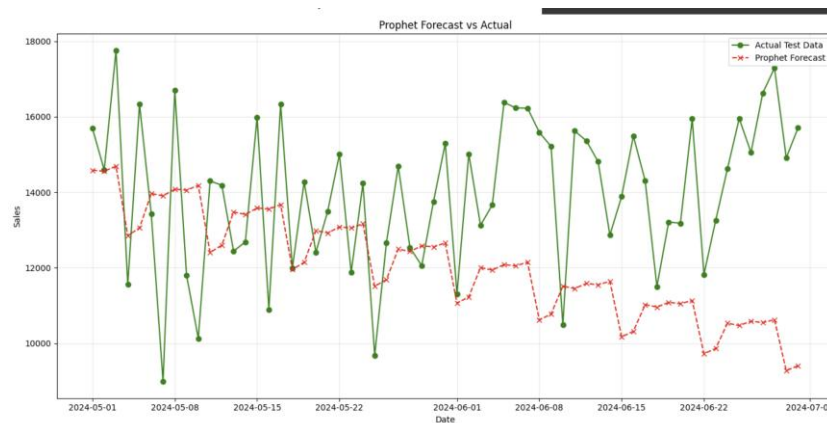
ได้มีการปรับแต่งโมเดล Prophet โดยใช้เทคนิค Grid Search เพื่อค้นหาชุดค่าพารามิเตอร์ที่ให้ผลการพยากรณ์ยอดขายรายวันได้แม่นยำที่สุด โดยครอบคลุมการทดสอบทั้งหมด 128 ชุดค่าพารามิเตอร์ ซึ่งประกอบด้วยการกำหนดค่า changepoint prior scale, seasonality prior scale, holidays prior scale และ seasonality mode โดยค่าพารามิเตอร์เหล่านี้มีบทบาทสำคัญในการควบคุมความยืดหยุ่นของโมเดลต่อการเปลี่ยนแปลงของแนวโน้ม ความแรงของฤดูกาล และความสำคัญของวันหยุดที่อาจส่งผลกระทบต่อยอดขายจากการประเมินผลการพยากรณ์ พบว่าโมเดลที่ใช้พารามิเตอร์ที่ได้จาก Grid Search มีประสิทธิภาพดีขึ้นอย่างชัดเจนเมื่อเปรียบเทียบกับค่าเริ่มต้น (Default) ในการทดลองก่อนหน้านี้ โดยได้ค่า MAE ลดลงมาอยู่ที่ 2,025.21 และ RMSE ลดลงเป็น 2,434.98 ในขณะที่ค่า MAPE ซึ่งสะท้อนความคลาดเคลื่อนในเชิงเปอร์เซ็นต์เมื่อเทียบกับยอดขายจริง ลดลงเหลือ 14.81% แสดงให้เห็นว่าโมเดลมีความแม่นยำในการพยากรณ์สูงขึ้นอย่างมีนัยสำคัญผลจากการปรับแต่งพารามิเตอร์ยังช่วยให้โมเดลสามารถตอบสนองต่อความเปลี่ยนแปลงของแนวโน้มและฤดูกาลได้ดีขึ้น ทำให้เส้นการพยากรณ์มีความสอดคล้องกับพฤติกรรมยอดขายจริงในระยะเวลาทดสอบมากขึ้น โดยเฉพาะในช่วงที่ยอดขายมีความแปรปรวนจากปัจจัยภายนอก เช่น โปรโมชันหรือวันหยุดเทศกาล ทั้งนี้ยังคงรักษาความสามารถในการอธิบายแนวโน้มระยะยาวและรูปแบบฤดูกาลรายสัปดาห์ไว้ได้อย่างครบถ้วน โดยสรุป การทดลองที่ 2 แสดงให้เห็นว่า การปรับแต่งค่าพารามิเตอร์ของ Prophet ด้วยวิธีการ Grid Search สามารถเพิ่มประสิทธิภาพของแบบจำลองได้อย่างมีนัยสำคัญ ทั้งในด้านความแม่นยำและความสามารถในการติดตามพฤติกรรมของยอดขายจริงในเชิงเวลา จึงเป็นแนวทางที่เหมาะสมสำหรับการพัฒนาแบบจำลองในบริบทของข้อมูลยอดขายที่มีความซับซ้อนและผันผวน



ภาพประกอบ 25 ผลลัพธ์จากการทดลองที่ 2

ผลลัพธ์ของการทดลองที่ 3

ได้ดำเนินการฝึกและประเมินโมเดล Prophet อีกครั้งโดยอ้างอิงค่าพารามิเตอร์ที่ได้จากกระบวนการ Grid Search ซึ่งเป็นค่าที่ให้ผลลัพธ์แม่นยำที่สุดก่อนหน้านี้ พร้อมทั้งมีการปรับแต่งเพิ่มเติมในเชิงโครงสร้างของโมเดลเพื่อเพิ่มความยืดหยุ่นในการจับแนวโน้มที่อาจมีการเปลี่ยนแปลงในช่วงท้ายของข้อมูล ตัวอย่างเช่น การกำหนดค่า changepoint range ให้กว้างขึ้นเป็น 0.9 เพื่อให้ Prophet ตรวจจับจุดเปลี่ยนแปลงแนวโน้มในช่วงท้ายได้ดีขึ้น การเพิ่มจำนวน n changepoints เป็น 25 เพื่อให้โมเดลมีความสามารถในการดักจับการเปลี่ยนแปลงแนวโน้มได้มากขึ้น และการเปิดใช้งาน daily seasonality เพื่อให้โมเดลเรียนรู้พฤติกรรมรายวันของยอดขายได้ดียิ่งขึ้นจากผลการประเมินประสิทธิภาพโมเดลที่ได้ฝึกซ้ำ พบว่าค่า MAE อยู่ที่ประมาณ 2,646 บาทต่อวัน ซึ่งแสดงถึงระดับของความคลาดเคลื่อนเฉลี่ยระหว่างค่าทำนายกับค่าจริงที่อยู่ในเกณฑ์ยอมรับได้ในบริบทของข้อมูลยอดขายที่มีความผันผวนสูง ส่วนค่า RMSE อยู่ที่ประมาณ 3,169 ซึ่งสะท้อนว่าแม้จะมีบางวันยอดขายจริงเบี่ยงเบนจากค่าทำนายมากเป็นพิเศษ แต่โมเดลยังสามารถรองรับความแปรปรวนได้ในระดับที่เหมาะสม สำหรับค่า MAPE อยู่ที่ 18.51% แม้ว่าจะเพิ่มขึ้นเล็กน้อยเมื่อเทียบกับโมเดลในผลการทดลองที่ 2 (14.81%) แต่ค่าดังกล่าวเกิดจากการใช้พารามิเตอร์ที่ถูกปรับแต่งให้สอดคล้องกับลักษณะเฉพาะของข้อมูลในเชิงโครงสร้าง ทำให้โมเดลมีแนวโน้มที่จะทำงานได้ดีในสถานการณ์จริงที่อาจมีความผันผวนหรือไม่เป็นไปตามรูปแบบในอดีตโดยสรุป การทดลองที่ 3 ชี้ให้เห็นว่าการปรับแต่งโครงสร้างของโมเดล Prophet ควบคู่กับการใช้พารามิเตอร์ที่ผ่านการเลือกอย่างเหมาะสมสามารถยกระดับความสามารถในการพยากรณ์ยอดขายให้ใกล้เคียงความจริงได้ในระดับที่น่าพอใจ โมเดลที่ได้จึงมีความเหมาะสมต่อการนำไปใช้งานจริงในภาคธุรกิจ เช่น การวางแผนการจัดสรรวัตถุดิบล่วงหน้า หรือการเตรียมการผลิตเครื่องดื่มในร้านกาแฟเพื่อรองรับความต้องการในแต่ละวันอย่างมีประสิทธิภาพ



ภาพประกอบ 26 ผลลัพธ์จากการทดลองที่ 3

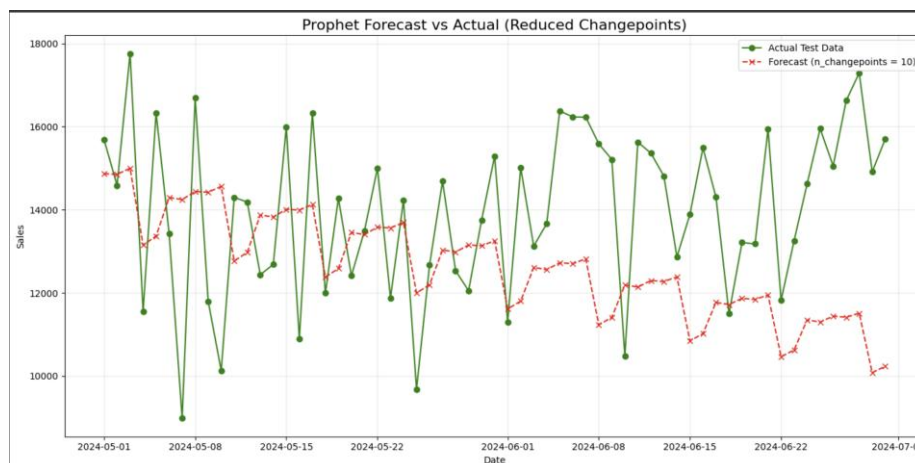
ผลลัพธ์ของการทดลองที่ 4

การควบคุม Overfitting ด้วยการลดจำนวน Changepoints ในโมเดล Prophet การทดลองนี้มีวัตถุประสงค์เพื่อควบคุมปัญหา Overfitting ที่อาจเกิดขึ้นจากการใช้จำนวน changepoints ที่มากเกินไปในโมเดล Prophet ซึ่งอาจทำให้โมเดลตอบสนองต่อความผันผวนเล็กน้อยในข้อมูล (noise) มากเกินความจำเป็น โดยเฉพาะในชุดข้อมูลยอดขายรายวันที่มีความผันผวนสูง การลดจำนวน changepoints ลงเหลือเพียง 10 จุด และใช้ค่า changepoint range = 0.8 ตามค่าตั้งต้นของ Prophet ทำให้โมเดลมีแนวโน้มที่ “เรียบ” ขึ้นและเน้นเรียนรู้เฉพาะแนวโน้มหลักที่มีนัยสำคัญเท่านั้น

เมื่อเปรียบเทียบประสิทธิภาพของโมเดลที่ลดจำนวน changepoints กับโมเดลแบบ Default และโมเดลที่ได้รับการปรับแต่งจาก Grid Search พบว่า โมเดล Reduced CP มีค่า MAE อยู่ที่ 2,309.68 บาท RMSE ที่ 2,763.88 และ MAPE ที่ 16.43% โดยแม้ว่า MAE จะสูงกว่าแบบ Default เล็กน้อย แต่ค่า MAPE กลับลดลง ซึ่งแสดงให้เห็นถึงความแม่นยำในเชิงสัมพัทธ์ที่เพิ่มขึ้น และ RMSE ที่ลดลงก็ชี้ว่าโมเดลสามารถลดผลกระทบจาก Outliers ได้ดีขึ้น โมเดลที่แนวโน้มเรียบขึ้นนี้จึงเหมาะกับกรณีที่ข้อมูลมี noise รายวันสูงแต่ยังคงมีแนวโน้มหลักที่ชัดเจน เช่น ยอดขายของร้านกาแฟที่แปรผันตามพฤติกรรมผู้บริโภคในแต่ละวัน

จากการเปรียบเทียบค่า MAPE ระหว่างโมเดลต่างๆ พบว่าโมเดลที่ให้ค่าต่ำที่สุดคือ Tuned Prophet ที่ได้จากการปรับพารามิเตอร์ด้วย Grid Search โดยมีค่า MAPE เท่ากับ 14.81% จึงสามารถสรุปได้ว่าแม้การลดจำนวน changepoints จะช่วยลดความซับซ้อนและควบคุม overfitting ได้ในระดับหนึ่ง แต่โมเดลที่ใช้จำนวน changepoints 25 และผ่านการเลือก

พารามิเตอร์อย่างเหมาะสมจะยังคงเป็นโมเดลที่ให้ผลลัพธ์ที่ดีที่สุดในเชิงความแม่นยำของการพยากรณ์รายวัน และเหมาะสมที่สุดในการนำไปใช้งานจริงในบริบทของธุรกิจร้านกาแฟ.



ภาพประกอบ 27 ผลลัพธ์จากการทดลองที่ 4

ผลลัพธ์ของการทดลองที่ 5

การทดลองครั้งนี้เป็นการประเมินประสิทธิภาพของการพยากรณ์ยอดขายรายวันด้วยเครื่องมือ Tableau โดยใช้แบบจำลองเชิงอนุกรมเวลาแบบ Exponential Smoothing ซึ่งเป็นเทคนิคดั้งเดิมที่มีความสามารถในการจับแนวโน้มและฤดูกาลของข้อมูลอย่างต่อเนื่อง โดยในการทดลองได้ตั้งค่าการพยากรณ์ให้ครอบคลุมช่วงเวลา 60 วันข้างหน้า ตั้งแต่วันที่ 30 เมษายน ถึง 28 มิถุนายน พ.ศ. 2567 บนระดับข้อมูลรายวัน ซึ่งครอบคลุมพฤติกรรมการขายตามฤดูกาลของธุรกิจร้านกาแฟ

แบบจำลองที่ใช้ใน Tableau กำหนดแนวโน้มและฤดูกาลให้อยู่ในรูปแบบ “ส่วนเสริม” (Additive) พร้อมเติมค่าที่ขาดหายด้วยค่าเฉลี่ย และระดับความมั่นใจของช่วงคาดการณ์ไว้ที่ 95% โดยข้อมูลย้อนหลังที่ใช้ในการฝึกแบบจำลองคือยอดขายตั้งแต่วันที่ 31 ธันวาคม 2566 ถึง 29 เมษายน 2567 ซึ่งครอบคลุมพฤติกรรมรายวันของลูกค้าในช่วงไตรมาสแรกของปี

จากผลลัพธ์การพยากรณ์ พบว่ายอดขายเฉลี่ยที่ได้ ณ วันที่เริ่มต้น (30 เมษายน 2567) อยู่ที่ 12,838 บาท โดยมีช่วงความไม่แน่นอน (confidence interval) $\pm 6,092$ บาท ซึ่งแปลว่ามียอดขายที่อาจเกิดขึ้นจริงระหว่าง 6,746 ถึง 18,930 บาท ในช่วง 60 วันถัดมาแนวโน้มยอดขายมี

การลดลงอย่างต่อเนื่อง โดยภาพรวมลดลงประมาณ 1,331 บาท จากต้นช่วงถึงปลายช่วงการพยากรณ์

เมื่อพิจารณาคุณภาพของแบบจำลองผ่านตัวชี้วัด พบว่า Tableau ให้ค่า MAPE เท่ากับ 14.9% ซึ่งถือว่าอยู่ในระดับแม่นยำปานกลางและใกล้เคียงกับแบบจำลอง Prophet ที่มีค่า MAPE เท่ากับ 14.81% นอกจากนี้ ค่า MAE (2,502) และ RMSE (3,108) บ่งชี้ถึงความคลาดเคลื่อนในเชิงปริมาณของแบบจำลองอยู่พอสมควร โดยเฉพาะในบางวันซึ่งอาจมีความผิดพลาดจากค่าพยากรณ์ที่เบี่ยงเบนจากค่าจริง ส่วนค่า Gamma ที่สูงกว่า Alpha แสดงว่าแบบจำลองให้ความสำคัญกับการอธิบายฤดูกาล (seasonality) มากกว่าแนวโน้ม (trend)

กล่าวโดยสรุป แบบจำลองที่ Tableau ใช้ ซึ่งอยู่บนพื้นฐานของ Exponential Smoothing แบบ Additive มีความเหมาะสมสำหรับข้อมูลยอดขายรายวันที่มีรูปแบบความผันผวนซ้ำ ๆ ตามวันในสัปดาห์ เช่น ยอดขายสูงในวันศุกร์และเสาร์ อย่างไรก็ตาม ความสามารถในการปรับแต่งแบบจำลองใน Tableau ยังจำกัดเมื่อเทียบกับ Prophet เช่น ไม่สามารถตั้งค่า changepoints หรือปรับพารามิเตอร์ได้ จึงอาจไม่เหมาะกับการพยากรณ์ในกรณีที่มีการเปลี่ยนแปลงแนวโน้มอย่างเฉียบพลันหรือเหตุการณ์พิเศษ เช่น โปรโมชันหรือวันหยุดนักขัตฤกษ์ที่มีผลกระทบต่อกิจกรรมผู้บริโภคอย่างมีนัยสำคัญ

อธิบายการพยากรณ์

สรุป แบบจำลอง

ตัวเลือกที่ใช้สร้างการพยากรณ์

อนุกรมเวลา: วันของ Date

การวัดผล ผลรวมของ All Sale

พยากรณ์ล่วงหน้า: 60 วัน (30 April 2567 – 28 June 2567)

พยากรณ์ตาม: 31 December 2566 – 29 April 2567

ละเว้นล่าสุด: 1 วัน (30 April 2567)

รูปแบบตามฤดูกาล: รอบ 7 วัน

ผลรวมของ All Sale

อักษรย่อ	เปลี่ยนจากค่าเริ่มต้น	เอฟเฟกต์ตามฤดูกาล		ผลงาน		คุณภาพ	
30 April 2567	30 April 2567 – 28 June 2567	สูง	ต่ำ	แนวโน้ม	ฤดูกาล		
12,838 ± 6,092	-1,331	28 June 2567 1,901	24 June 2567 -495	62.7%	37.3%	ตกลง	

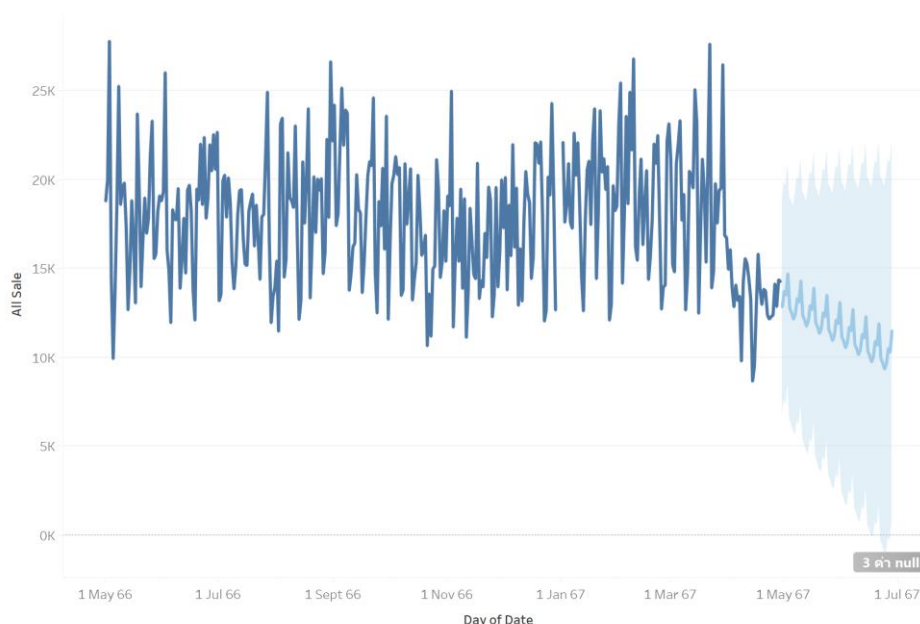
การพยากรณ์ทั้งหมดคำนวณโดยใช้การปรับเรียบแบบเอ็กซ์โปเนนเชียล

ผลรวมของ All Sale

แบบจำลอง		เมตริกคุณภาพ					สัมประสิทธิ์การปรับให้เรียบ			
ระดับ	แนวโน้ม	ฤดูกาล	RMSE	MAE	MASE	MAPE	AIC	Alpha	เบต้า	แกมมา
ส่วนเสริม	ส่วนเสริม	ส่วนเสริม	3,108	2,502	0.79	14.9%	1,922	0.112	0.000	0.350

ภาพประกอบ 28 ผลการพยากรณ์ยอดขายรวมจาก Tableau ใช้แนวโน้มและฤดูกาลแบบส่วน

เสริม



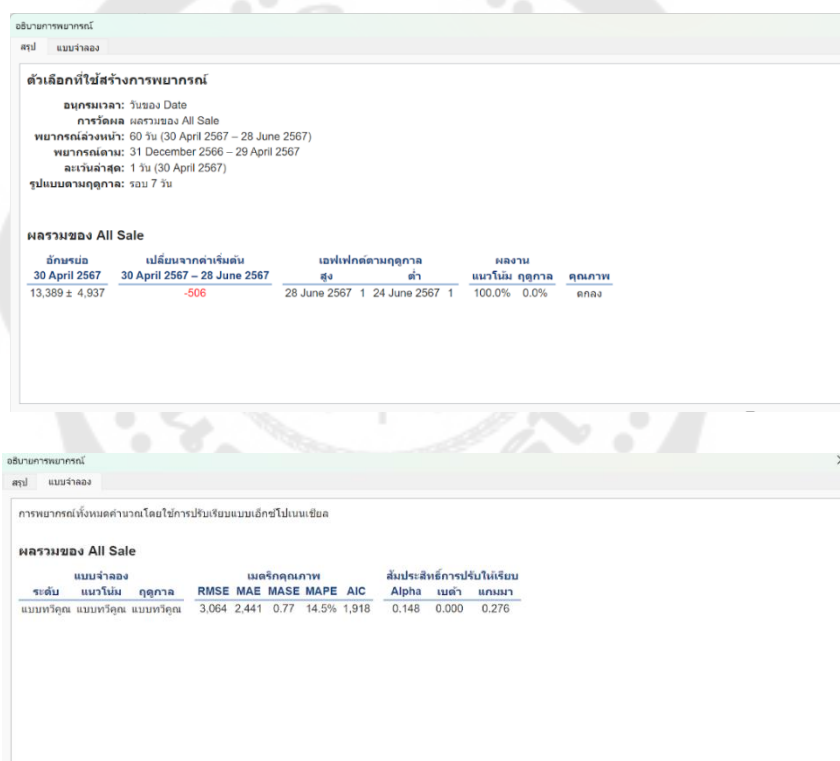
ภาพประกอบ 29 กราฟแสดงข้อมูลพยากรณ์ยอดขายรวมจาก Tableau ใช้แนวโน้มและฤดูกาล
แบบส่วนเสริม

ผลลัพธ์ของการทดลองที่ 6

การทดลองครั้งนี้ใช้ข้อมูลยอดขายรายวันย้อนหลังตั้งแต่วันที่ 31 ธันวาคม 2566 ถึง 29 เมษายน 2567 เพื่อพยากรณ์แนวโน้มยอดขายในช่วง 60 วันถัดไป ครอบคลุมระยะเวลาตั้งแต่วันที่ 30 เมษายน ถึง 28 มิถุนายน 2567 โดยแบบจำลองที่ใช้ใน Tableau เป็นแบบ Exponential Smoothing ซึ่งมีการกำหนดโครงสร้างแนวโน้มและฤดูกาลเป็นรูปแบบ “ทวีคูณ (Multiplicative)” ทั้งสององค์ประกอบ เพื่อสะท้อนถึงพฤติกรรมยอดขายที่มีการเปลี่ยนแปลงตามฤดูกาลอย่างชัดเจน โดยเฉพาะในบริบทของธุรกิจร้านค้าแฟรนไชส์ที่มีการเปลี่ยนแปลงตามฤดูกาลอย่างชัดเจน โดยเฉพาะในบริบทของธุรกิจร้านค้าแฟรนไชส์ที่มีการเปลี่ยนแปลงตามฤดูกาลอย่างชัดเจน

ผลการวิเคราะห์แสดงให้เห็นว่า แบบจำลองใน Tableau ตรวจพบฤดูกาลแบบรายสัปดาห์ (ทุก 7 วัน) ซึ่งสอดคล้องกับพฤติกรรมของลูกค้าประจำที่มีแนวโน้มการเข้าร้านแตกต่างกันระหว่างวันธรรมดาและวันหยุด สำหรับค่าการพยากรณ์ในวันเริ่มต้นของช่วงทำนาย (30 เมษายน 2567) ระบบคาดการณ์ยอดขายที่ระดับ 13,389 หน่วย โดยมีช่วงความไม่แน่นอน $\pm 4,937$ หน่วย ซึ่งแสดงถึงความกว้างของช่วงคาดการณ์ที่มากพอสมควร และอาจเกิดจากความผันผวนในข้อมูลช่วงที่ผ่านมา

เมื่อพิจารณาการเปลี่ยนแปลงของยอดขายโดยรวม พบว่ายอดขายในช่วง 60 วันข้างหน้า มีแนวโน้มลดลงประมาณ 506 หน่วยจากจุดเริ่มต้น โดยวันพยากรณ์ที่คาดว่าจะมียอดขายสูงที่สุด คือวันที่ 28 มิถุนายน 2567 ขณะที่วันที่ 24 มิถุนายน 2567 เป็นวันที่คาดว่าจะมียอดขายต่ำที่สุด สะท้อนให้เห็นถึงฤดูกาลของพฤติกรรมผู้บริโภคที่ยังมีอิทธิพลอย่างต่อเนื่องในช่วงปลายไตรมาสที่ 2 สำหรับตัวชี้วัดประสิทธิภาพของแบบจำลองจาก Tableau ได้แก่ MAPE อยู่ที่ 14.5% ซึ่งถือว่าอยู่ในระดับแม่นยำ “พอใช้ได้” ตามมาตรฐานของงานพยากรณ์ข้อมูลจริง ส่วนค่า RMSE และ MAE ที่ได้คือ 3,064 และ 2,441 ตามลำดับ แสดงถึงค่าความคลาดเคลื่อนเฉลี่ยที่ยังอยู่ในระดับที่เหมาะสมสำหรับข้อมูลประเภทยอดขายรายวันซึ่งมีความแปรปรวนสูง นอกจากนี้ ค่าดัชนี AIC ที่ระดับ 1,918 สะท้อนว่าแบบจำลองมีความสอดคล้องกับข้อมูลในระดับที่ยอมรับได้



ภาพประกอบ 30 ผลการพยากรณ์ยอดขายรวมจาก Tableau ใช้แนวโน้มและฤดูกาลแบบทวีคูณ



ภาพประกอบ 31 กราฟแสดงข้อมูลพยากรณ์ยอดขายรวมจาก Tableau ใช้แนวโน้มและฤดูกาล
แบบทวีคูณ

สรุปได้ว่า แบบจำลอง Exponential Smoothing ที่ Tableau นำมาใช้ในครั้งนี้ มีจุดเด่นใน
ด้านความสะดวกและการตีความเชิงธุรกิจ โดยสามารถจับรูปแบบพฤติกรรมรายสัปดาห์ได้ชัดเจน
อย่างไรก็ตาม การที่ Tableau ไม่สามารถปรับแต่งพารามิเตอร์เชิงลึก เช่น changepoints หรือ
seasonality tuning ได้เหมือนกับโมเดล Prophet อาจส่งผลให้ประสิทธิภาพของแบบจำลองใน
สถานการณ์ที่แนวโน้มเปลี่ยนแปลงกะทันหัน เช่น การส่งเสริมการขาย หรือพฤติกรรมหลังวันหยุด
นักชัตฤดู มีข้อจำกัดในการตอบสนองได้อย่างแม่นยำเทียบเท่าระบบที่ออกแบบโมเดล Machine
Learning ด้วยตนเอง

บทที่ 5

สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

ในการศึกษาครั้งนี้ ผู้วิจัยมีวัตถุประสงค์เพื่อค้นหาแบบจำลองที่สามารถพยากรณ์ยอดขายรายวันของธุรกิจร้านกาแฟได้อย่างแม่นยำและมีความเสถียร โดยใช้ข้อมูลยอดขายรายวันในช่วงระยะเวลา 14 เดือน ตั้งแต่เดือนพฤษภาคม พ.ศ. 2566 ถึงเดือนมิถุนายน พ.ศ. 2567 ซึ่งเป็นข้อมูลจริงจากระบบฐานข้อมูลของธุรกิจ โดยข้อมูลที่ใช้ครอบคลุมทั้งช่วงเวลาปกติ ช่วงเทศกาล และช่วงยอดขายตกต่ำ เพื่อให้ได้ภาพรวมของพฤติกรรมการขายที่สมบูรณ์ที่สุด และสามารถนำไปวิเคราะห์เพื่อใช้วางแผนทางธุรกิจได้จริง การวิเคราะห์ประกอบด้วยการขั้นตอนการทำความสะอาดข้อมูล การสร้างแบบจำลอง การทดสอบ และการเปรียบเทียบประสิทธิภาพของโมเดล โดยใช้ทั้งเทคนิคการเรียนรู้ของเครื่อง (Machine Learning) และแบบจำลองสำเร็จรูปที่มีอยู่ในซอฟต์แวร์ทางธุรกิจ เช่น Prophet และ Tableau Forecasting

สรุปผลการวิจัย

ในการศึกษาครั้งนี้ ผู้วิจัยมีวัตถุประสงค์เพื่อหาแบบจำลองที่สามารถพยากรณ์ยอดขายรายวันของธุรกิจร้านกาแฟได้อย่างแม่นยำ โดยใช้ข้อมูลยอดขายตั้งแต่เดือนพฤษภาคม พ.ศ. 2566 ถึงเดือนมิถุนายน พ.ศ. 2567 ซึ่งเป็นข้อมูลที่ครอบคลุมฤดูกาลต่าง ๆ และรูปแบบพฤติกรรมผู้บริโภคที่หลากหลายในแต่ละช่วงเวลา เพื่อนำมาใช้ในการวางแผนการจัดการภายในร้าน เช่น การจัดการวัตถุดิบ การจัดการตารางพนักงาน และการเตรียมสต็อกสินค้า โดยเฉพาะกลุ่มผลิตภัณฑ์ที่มีอายุสั้น เช่น สมูทตี้ที่ใช้โยเกิร์ตเป็นส่วนผสมหลัก

กระบวนการดำเนินงานเริ่มจากการเตรียมข้อมูลและทำความสะอาดข้อมูล จากนั้นได้ทดลองพยากรณ์ยอดขายโดยใช้แบบจำลอง Prophet ซึ่งเป็นแบบจำลองที่พัฒนาโดย Facebook ที่เหมาะสมสำหรับชุดข้อมูลอนุกรมเวลา (time series) และสามารถจัดการกับข้อมูลที่มีแนวโน้ม (trend) และฤดูกาล (seasonality) ได้เป็นอย่างดี ในขั้นต้นได้ทำการทดลองด้วยโมเดล Prophet แบบค่าเริ่มต้น (default parameters) เพื่อดูประสิทธิภาพพื้นฐาน จากนั้นได้ทำการปรับแต่งพารามิเตอร์โดยใช้เทคนิค Grid Search ซึ่งเป็นการทดสอบค่าพารามิเตอร์ต่าง ๆ ที่เป็นไปได้ทั้งหมดจำนวน 128 ชุด เพื่อหาโมเดลที่มีค่าความคลาดเคลื่อนต่ำที่สุด จากนั้นได้เลือกค่าพารามิเตอร์ที่ดีที่สุดมาสร้างแบบจำลอง Prophet ฉบับปรับแต่ง และทดสอบผลการพยากรณ์ นอกจากนี้ ยังได้เปรียบเทียบประสิทธิภาพของแบบจำลอง Prophet กับแบบจำลองที่สร้างใน Tableau โดยใช้เทคนิค Exponential Smoothing ซึ่งเป็นอีกหนึ่งวิธีที่นิยมใช้ในการพยากรณ์

อนุกรมเวลา ทั้งนี้เพื่อประเมินความเหมาะสมของแต่ละเครื่องมือในบริบทของข้อมูลจริงที่ใช้ในการบริหารธุรกิจร้านค้าแฟ

อภิปรายผลการวิจัย

จากการทดลองที่ดำเนินการทั้งหมด พบว่าแบบจำลอง Prophet ที่ใช้พารามิเตอร์ที่ดีที่สุดจากการทำ Grid Search และมีการตั้งค่าเพิ่มเติม เช่น การกำหนด changepoint_range ที่กว้างขึ้น (0.9), การเพิ่มจำนวนจุดเปลี่ยนแนวโน้ม (n_changepoints = 25), และเปิดใช้งานฤดูกาลรายวัน (daily_seasonality = True) ให้ผลลัพธ์ในการพยากรณ์ที่มีความแม่นยำสูงที่สุดเมื่อพิจารณาจากค่าความคลาดเคลื่อนเฉลี่ยสัมบูรณ์ (MAE = 2,025.21), รากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (RMSE = 2,434.98) และค่าความคลาดเคลื่อนสัมพัทธ์เฉลี่ย (MAPE = 14.81%) ซึ่งต่ำที่สุดเมื่อเทียบกับโมเดลอื่นที่ทดลอง

ในขณะที่แบบจำลอง Tableau ที่ใช้เทคนิค Exponential Smoothing แม้จะมีข้อได้เปรียบด้านความสะดวกในการใช้งานและการแสดงผลที่เข้าใจง่าย แต่ค่าความคลาดเคลื่อนที่ได้ยังสูงกว่า Prophet เล็กน้อย โดย MAPE อยู่ที่ประมาณ 14.5% ถึง 14.9% ขึ้นอยู่กับรูปแบบการตั้งค่าฤดูกาล (แบบ additive หรือ multiplicative) และแนวโน้ม (trend)

ข้อสังเกตที่สำคัญคือ แบบจำลอง Tableau ไม่สามารถปรับค่าพารามิเตอร์ได้ละเอียดเท่า Prophet จึงอาจทำให้การตอบสนองต่อจุดเปลี่ยนแนวโน้มในข้อมูลไม่แม่นยำเท่าที่ควร อย่างไรก็ตาม แบบจำลอง Tableau ก็แสดงผลได้อย่างเหมาะสมเมื่อใช้กับข้อมูลที่มีฤดูกาลคงที่ เช่น ความนิยมที่สูงในช่วงวันหยุดสุดสัปดาห์

การใช้ Grid Search ในการปรับค่าพารามิเตอร์ของ Prophet แสดงให้เห็นถึงความสำคัญของการเลือกค่าพารามิเตอร์ที่เหมาะสม โดยเฉพาะในชุดข้อมูลที่มีความผันผวนรายวันสูง เนื่องจากค่าพารามิเตอร์เริ่มต้นอาจไม่เหมาะสมกับบริบทของข้อมูลในธุรกิจเฉพาะกลุ่ม เช่น ธุรกิจร้านค้าแฟที่มีความแปรปรวนของยอดขายตามวันและช่วงเวลาต่าง ๆ

ข้อเสนอแนะ

จากผลการวิจัย ผู้วิจัยมีข้อเสนอแนะดังนี้

1. การนำแบบจำลอง Prophet ที่ผ่านการปรับแต่งแล้วไปใช้จริง ควรมีการอัปเดตโมเดลอย่างสม่ำเสมอ เมื่อมีข้อมูลใหม่เข้ามา เพื่อให้โมเดลสามารถเรียนรู้แนวโน้มใหม่ ๆ ได้ทัน

2. สำหรับองค์กรที่ไม่มีทีมพัฒนาโมเดล สามารถใช้ Tableau เป็นเครื่องมือพื้นฐานในการพยากรณ์ เพื่อการตัดสินใจเบื้องต้น และหากต้องการความแม่นยำสูงในเชิงกลยุทธ์ ควรใช้แบบจำลอง Prophet ร่วมกับการปรับแต่งพารามิเตอร์

3. ควรพิจารณาเพิ่มตัวแปรแวดล้อม (external regressors) เช่น วันหยุดนักขัตฤกษ์, โปรโมชัน หรือข้อมูลสภาพอากาศ เข้าไปในโมเดล Prophet เพื่อเพิ่มความแม่นยำยิ่งขึ้น โดยเฉพาะในธุรกิจที่มีปัจจัยภายนอกส่งผลกระทบต่อยอดขายอย่างชัดเจน

4. แนะนำให้มีการทดสอบโมเดลในช่วงเวลาต่าง ๆ อย่างน้อยสองฤดูกาล เพื่อประเมินความเสถียรของโมเดลในระยะยาว และลดความลำเอียงจากฤดูกาลใดฤดูกาลหนึ่ง

5. ในงานวิจัยถัดไป อาจพิจารณาทดลองใช้แบบจำลองอื่น ๆ เช่น XGBoost, LightGBM หรือ LSTM เพื่อเปรียบเทียบกับ Prophet โดยเฉพาะเมื่อต้องการพยากรณ์ในระดับชั่วโมงหรือระดับสาขา (branch-level)



บรรณานุกรม

- Ahmed, S., Chakraborty, R. K., Essam, D. L., & Ding, W. (2024). A switching based forecasting approach for forecasting sales data in supply chains. *Applied Soft Computing*, 167, 112419. Retrieved 2024/12/01/, from Retrieved from <https://www.sciencedirect.com/science/article/pii/S1568494624011931>
- Angelo, M. D., Fadhiilrahman, I., & Purnama, Y. (2023). Comparative Analysis of ARIMA and Prophet Algorithms in Bitcoin Price Forecasting. *Procedia Computer Science*, 227, 490-499. Retrieved 2023/01/01/, from Retrieved from <https://www.sciencedirect.com/science/article/pii/S1877050923017179>
- Chang, H., Yang, Z., & Zhang, Y. (2025). Forecasting daily freight flow in cold regions of China using the hybrid Prophet model considering the importance of festivals and epidemic prevention policy. *Research in Transportation Business & Management*, 59, 101294. Retrieved 2025/03/01/, from Retrieved from <https://www.sciencedirect.com/science/article/pii/S2210539525000094>
- Negre, P., Alonso, R. S., Prieto, J., García, Ó., & de-la-Fuente-Valentín, L. (2024). Prediction of footwear demand using Prophet and SARIMA. *Expert Systems with Applications*, 255, 124512. Retrieved 2024/12/01/, from Retrieved from <https://www.sciencedirect.com/science/article/pii/S0957417424013794>

ประวัติผู้เขียน

