

การทำนายยอดขายปูนซีเมนต์ ด้วยวิธีสมการเส้นถดถอยเชิงเส้นตรง:
กรณีศึกษากลุ่มร้านจำหน่ายวัสดุก่อสร้างในอำเภอรัตภูมิ จังหวัดสงขลา ช่วงปี 2565-2567
Prediction of Cement Sales Using the Linear Regression Model Equation
Method (2022-2024): A Case Study of Building Material Retail Stores in
Rattaphum District, Songkhla Province

กัณฑ์มน สุขกระจ่าง¹, ธนะรัตน์ รัตนกุล^{1*}, ฉัตรชัย แก้วดี², วีรพล ปานศรีนวล² และ วีระยุทธ สุตสมบุญ²
Kantamon Sukkrajang¹, Tanarat Rattanakool^{1*}, Chatchai Kaewdee², Weeraphol Pansrinual² and
Weerayute Sudsomboon²

¹ คณะเทคโนโลยีอุตสาหกรรม มหาวิทยาลัยราชภัฏสงขลา

² หลักสูตรปรัชญาดุษฎีบัณฑิต สาขาวิชานวัตกรรมเทคโนโลยีอุตสาหกรรม คณะเทคโนโลยีอุตสาหกรรม
มหาวิทยาลัยราชภัฏนครศรีธรรมราช

* ผู้ประสานงานเผยแพร่ (Corresponding Author), E-mail: tanarat.ra@skru.ac.th

วันที่รับบทความ: 2 เมษายน 2568 ; วันที่ทบทวนบทความ: 11 เมษายน 2568 ; วันที่ตอบรับบทความ: 23 เมษายน 2568

วันที่เผยแพร่ออนไลน์: 30 เมษายน 2568

บทคัดย่อ: การวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาและประเมินแบบจำลองการถดถอยเชิงเส้นตรงอย่างง่ายสำหรับพยากรณ์ยอดขายปูนซีเมนต์รายเดือน กลุ่มร้านจำหน่ายวัสดุก่อสร้างในอำเภอรัตภูมิ จังหวัดสงขลา โดยใช้ข้อมูลตั้งแต่เดือนมกราคม พ.ศ. 2565 ถึง กันยายน พ.ศ. 2567 (n=33) ผลการวิเคราะห์พบว่าสมการที่เหมาะสมที่สุด คือ $\hat{Y}_t = 1143.4 - 7.35X$ แสดงแนวโน้มยอดขาย ลดลงเฉลี่ย 7.35 กระสอบต่อเดือน อย่างมีนัยสำคัญทางสถิติ ($p < 0.001$) ความสัมพันธ์ระหว่างเวลาและยอดขายเป็น เชิงลบระดับปานกลาง ($r = -0.6613$) อย่างไรก็ตามแบบจำลองมีความสามารถในการอธิบายความแปรปรวนของยอดขายได้เพียง 43.7% ($R^2 = 0.437$) บ่งชี้ว่าปัจจัยอื่นที่ไม่ได้รวมในแบบจำลอง เช่น สภาพเศรษฐกิจ ฤดูกาล การแข่งขัน มีอิทธิพลอย่างมาก นอกจากนี้ การตรวจสอบข้อตกงเบื้องต้นพบ ข้อจำกัดสำคัญ โดยเฉพาะปัญหาสหสัมพันธ์อัตโนมัติ (Autocorrelation) ซึ่งมักพบในการใช้แบบจำลองพื้นฐานกับข้อมูลอนุกรมเวลา สมการที่ได้ถูกนำไปใช้พยากรณ์ยอดขายในอนาคต ซึ่งผลการพยากรณ์นี้ แม้จะมีข้อจำกัดสามารถใช้เป็นข้อมูลเบื้องต้นประกอบการตัดสินใจในการวางแผนการผลิต การจัดการสินค้าคงคลัง และการกำหนดกลยุทธ์ทางการตลาด โดยควรพิจารณาพร้อมกับปัจจัยอื่นและข้อจำกัดของแบบจำลอง

คำสำคัญ: พยากรณ์ยอดขาย ถดถอยเชิงเส้น ปูนซีเมนต์

Abstract: This research aimed to develop and evaluate a simple linear regression model for forecasting monthly cement sales in Rattaphum District, Songkhla Province, using data from January 2022 to September 2024 (n=33). The analysis revealed the most suitable equation to be $\hat{Y}_t = 1143.4 - 7.35X$ indicating a statistically significant decreasing trend ($p < 0.001$) in sales by an average of 7.35 sacks per month. The relationship between time and sales showed a

moderate negative correlation ($r = -0.6613$). However, the model's explanatory power was limited, with $R^2 = 0.437$ (43.7%), suggesting that other factors not included (e.g., economic conditions, seasonality, competition) significantly influence sales. Furthermore, checks of the underlying assumptions revealed important limitations, particularly significant positive autocorrelation, a common issue when applying basic models to time series data. The derived equation was used to forecast future sales. While subject to these limitations, these forecasts can provide preliminary information for decision-making in production planning, inventory management, and marketing strategy development, provided they are considered alongside other factors and the model's constraints.

Key words: Sales Forecasting, Linear Regression, Cement

1. บทนำ

อุตสาหกรรมก่อสร้างถือเป็นหนึ่งในภาคเศรษฐกิจที่สำคัญของประเทศไทย โดยมีปูนซีเมนต์เป็นวัสดุก่อสร้างหลักที่มีบทบาทสำคัญต่อการเติบโตของภาคส่วนนี้ การพยากรณ์ยอดขายปูนซีเมนต์จึงมีความสำคัญอย่างยิ่งต่อผู้ประกอบการ ผู้ผลิต และผู้วางแผนด้านเศรษฐกิจ เพื่อช่วยในการวางแผนการผลิต การจัดการสินค้าคงคลัง การบริหารต้นทุน และการกำหนดยุทธศาสตร์ทางธุรกิจ ให้สอดคล้องกับความต้องการของตลาด ความแม่นยำในการพยากรณ์ยอดขายจะช่วยลดความเสี่ยงจากสินค้าล้นตลาดหรือขาดตลาด เพิ่มประสิทธิภาพในการดำเนินงาน และสร้างความได้เปรียบในการแข่งขันทางธุรกิจได้ [1] เช่นเดียวกับอุตสาหกรรมอื่น ๆ การพยากรณ์ยอดขายที่มีประสิทธิภาพเป็นสิ่งจำเป็นต่อการดำเนินธุรกิจอย่างยั่งยืน ดังที่ Chen., et al. [2] ได้กล่าวถึงความสำคัญของการพยากรณ์ยอดขายในอุตสาหกรรมแพชั่น ซึ่งช่วยในการบริหารจัดการสินค้าคงคลัง เพิ่มความพึงพอใจของลูกค้า และเพิ่มประสิทธิภาพในการดำเนินงาน

ปัจจัยที่มีอิทธิพลต่อยอดขายปูนซีเมนต์มีความหลากหลายและซับซ้อน ตั้งแต่ปัจจัยภายใน เช่น ราคา วัตถุดิบ ต้นทุนการผลิต และนโยบายการขาย ไปจนถึงปัจจัยภายนอก เช่น สภาพเศรษฐกิจ ฤดูกาล การแข่งขันในตลาด และนโยบายของรัฐบาล [3] ยิ่งไปกว่านั้น

การเปลี่ยนแปลงทางเทคโนโลยีและนวัตกรรมในอุตสาหกรรมก่อสร้าง รวมถึงการเปลี่ยนแปลงพฤติกรรมผู้บริโภค ล้วนส่งผลต่อความต้องการปูนซีเมนต์ ดังที่ Uras., et al. [4] ได้ศึกษาถึงพฤติกรรมราคา Bitcoin ที่มีความผันผวนสูง ซึ่งสะท้อนให้เห็นถึงความท้าทายในการพยากรณ์ในตลาดที่มีความไม่แน่นอนสูง ในบริบทร้านจำหน่ายวัสดุก่อสร้างแห่งหนึ่งในอำเภอรัตนภูมิ จังหวัดสงขลา ซึ่งเป็นกรณีศึกษาของงานวิจัยนี้ กำลังประสบปัญหาความผันผวนของยอดขายปูนซีเมนต์ในแต่ละเดือน โดยไม่มีเครื่องมือช่วยในการวางแผนการผลิตหรือจัดการสินค้าคงคลังที่เป็นระบบ ทำให้เกิดปัญหาสินค้าคงคลังเกินหรือขาด ส่งผลต่อต้นทุนและประสิทธิภาพในการบริหารจัดการธุรกิจ ดังนั้น การพัฒนาแบบจำลองการพยากรณ์ที่มีความแม่นยำในระดับหนึ่งและสามารถนำไปประยุกต์ใช้ได้จริงในบริบทของผู้ประกอบการรายย่อยจึงเป็นสิ่งจำเป็น เพื่อให้สามารถใช้ในการตัดสินใจเชิงกลยุทธ์ได้อย่างมีประสิทธิภาพมากขึ้น และช่วยลดความเสี่ยงจากความไม่แน่นอนของตลาด ซึ่งเป็นความต้องการที่งานวิจัยนี้มุ่งตอบสนอง

ด้วยเหตุนี้ งานวิจัยนี้จึงมุ่งเน้นศึกษาการทำนายยอดขายปูนซีเมนต์ ด้วยวิธีสมการเส้นถดถอยเชิงเส้นตรง (Linear regression model) สำหรับกรณีศึกษาร้านค้า

ในอำเภอรัษฎา จังหวัดสงขลา โดยมีวัตถุประสงค์เพื่อสร้างแบบจำลองการพยากรณ์ เบื้องต้น ที่สามารถนำไปประยุกต์ใช้ในการวางแผนธุรกิจได้ วิธีการถดถอยเชิงเส้นตรงเป็นวิธีการทางสถิติที่ได้รับความนิยมเนื่องจากมีความเรียบง่าย สามารถทำความเข้าใจและนำไปคำนวณได้ไม่ยาก อีกทั้งยังสามารถอธิบายแนวโน้มความสัมพันธ์ระหว่างยอดขายกับเวลาได้อย่างชัดเจน [5] ซึ่งเป็นคุณสมบัติที่เหมาะสมสำหรับผู้ประกอบการที่อาจไม่มีความเชี่ยวชาญด้านสถิติขั้นสูง อย่างไรก็ตาม วิธีการนี้มีข้อจำกัดในการจับรูปแบบที่ซับซ้อนของข้อมูล ซึ่งอาจส่งผลกระทบต่อความแม่นยำในการพยากรณ์ ดังที่ Nelvidawati [6] พบว่าแบบจำลองถดถอยเชิงเส้นมีความคลาดเคลื่อนสูงในการพยากรณ์ปริมาณน้ำฝน เนื่องจากไม่สามารถพิจารณาปัจจัยที่มีความสัมพันธ์ที่ไม่เป็นเชิงเส้นได้ครบถ้วน

อย่างไรก็ตาม การศึกษาเปรียบเทียบแบบจำลองต่าง ๆ เช่น งานวิจัยของ Mohamad Adam Bujang, et al. [7] ที่เปรียบเทียบการถดถอยเชิงเส้นอย่างง่าย (Simple linear regression) และ แนวโน้มเชิงเส้นของโฮลท์ (Holt's linear trend) หรือ Nazir., et al. [8] ที่ใช้ ขั้นตอนวิธีการตามพีชพรรณวิทยา (Phenology based algorithm) และ แบบจำลองการถดถอยเชิงเส้น (Linear regression model) ในการประเมินผลผลิตข้าว หรือ Yu [9] ที่เปรียบเทียบป่าสุ่ม (Random forest) การถดถอยเชิงเส้น (Linear regression) และการถดถอยแบบแลนโซ่ (Lasso regression) ในการพยากรณ์ยอดขาย Walmart หรือ Madhukumar., et al. [10] ที่เปรียบเทียบแบบจำลองหลายรูปแบบในการพยากรณ์การใช้ไฟฟ้า จากงานวิจัยที่ผ่านมาล้วนมีการนำวิธีสมการเส้นถดถอยเชิงเส้นตรง (Linear regression model) ไปเป็นส่วนหนึ่งในการค้นหาแบบจำลองที่เหมาะสมเพื่อนำเสนอข้อเสนอนะในการนำผลการวิจัยไปใช้ประโยชน์และแนวทางในการพัฒนาแบบจำลองให้มีประสิทธิภาพยิ่งขึ้นในอนาคตต่อไปได้

2. วัตถุประสงค์การวิจัย

2.1 เพื่อสร้างสมการความสัมพันธ์เชิงเส้นตรง ซึ่งแสดงความสัมพันธ์ของยอดขายปูนซีเมนต์ในแต่ละเดือน ด้วยวิธีสมการเส้นถดถอยเชิงเส้นตรง กรณีศึกษา กลุ่มร้านจำหน่ายวัสดุก่อสร้างในอำเภอรัษฎา จังหวัดสงขลา

2.2 เพื่อวัดระดับความน่าเชื่อถือของความสัมพันธ์ของยอดขายปูนซีเมนต์

2.3 เพื่อทำนายค่าในอนาคตโดยใช้ค่าที่มีอยู่ด้วยวิธีฟอว์แคสชีส (Forecast sheet)

3. ขอบเขตการวิจัย

3.1 ด้านเนื้อหา คือ การวิจัยครั้งนี้มุ่งศึกษาการทำนายยอดขายปูนซีเมนต์ ด้วยวิธีสมการเส้นถดถอยเชิงเส้นตรง

3.2 ด้านพื้นที่ พื้นที่ที่ทำการศึกษา คือ กลุ่มร้านจำหน่ายวัสดุก่อสร้างในอำเภอรัษฎา จังหวัดสงขลา

3.3 ด้านเวลา คือ มกราคม พ.ศ. 2565 ถึง กันยายน พ.ศ. 2567

4. วิธีดำเนินการวิจัย

ขั้นตอนการวิจัยสำหรับการทำนายยอดขายปูนซีเมนต์กลุ่มร้านจำหน่ายวัสดุก่อสร้างในอำเภอรัษฎา จังหวัดสงขลา โดยใช้แบบจำลองการถดถอยเชิงเส้นตรง (Linear regression model) สามารถปรับปรุงให้มีความชัดเจนและสอดคล้องกับงานวิจัยได้ ดังนี้

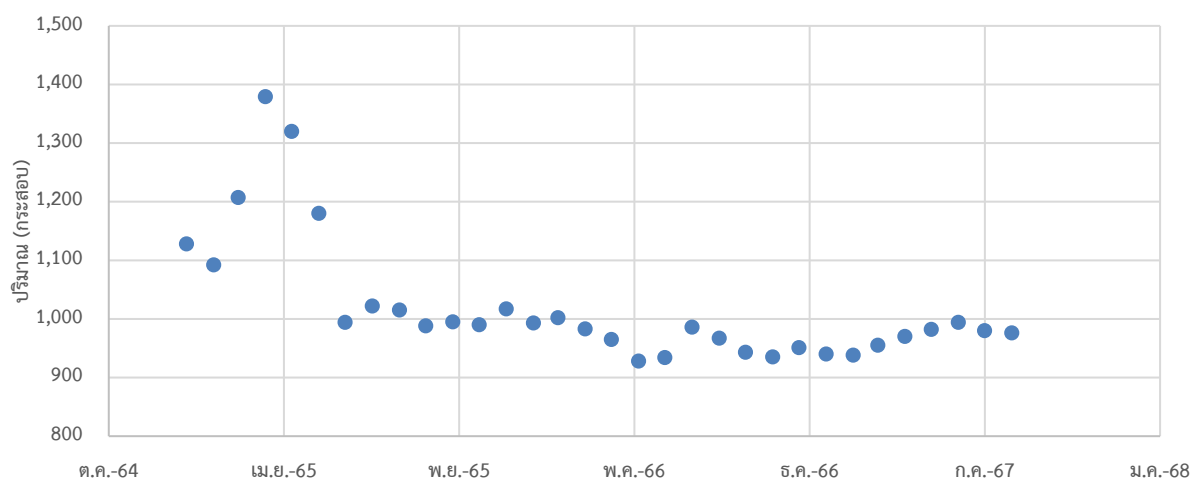
4.1 นิยามปัญหาและกำหนดตัวแปรตาม กำหนดปัญหาการวิจัยให้ชัดเจนว่า มีวัตถุประสงค์เพื่อพยากรณ์ยอดขายปูนซีเมนต์ของร้านจำหน่ายวัสดุก่อสร้างในอำเภอรัษฎา จังหวัดสงขลา โดยใช้วิธีถดถอยเชิงเส้นตรง (Linear regression model) ในการวิเคราะห์ข้อมูล โดยประชากรและกลุ่มตัวอย่าง ที่ใช้ในการวิจัยครั้งนี้ คือ ยอดขายปูนซีเมนต์ของกลุ่มร้านจำหน่ายวัสดุก่อสร้างในอำเภอรัษฎา จังหวัดสงขลา

4.2 ตัวแปรอิสระและรวบรวมข้อมูล ระบุตัวแปรอิสระ (Independent variable) คือ ตั้งแต่ มกราคม พ.ศ. 2565 ถึง กันยายน พ.ศ. 2567 ดังตารางที่ 1

ตารางที่ 1 ยอดขายปูนซีเมนต์ระยะเวลา 3 ปี ตั้งแต่ มกราคม พ.ศ. 2565 ถึง กันยายน พ.ศ. 2567

เดือน	2565	2566	2567
ม.ค.	1,128	1,017	940
ก.พ.	1,092	993	938
มี.ค.	1,207	1,002	955
เม.ย.	1,379	983	970
พ.ค.	1,320	965	982
มิ.ย.	1,180	928	994
ก.ค.	994	934	980
ส.ค.	1,022	986	976
ก.ย.	1,015	967	
ต.ค.	988	943	
พ.ย.	995	935	
ธ.ค.	990	951	

จากตารางที่ 1 ซึ่งแสดงข้อมูลยอดขายปูนซีเมนต์รายเดือนตั้งแต่ มกราคม พ.ศ. 2565 ถึง กันยายน พ.ศ. 2567 ได้นำข้อมูลดังกล่าวมาสร้างกราฟการกระจาย (Scatter plot) เพื่อสำรวจความสัมพันธ์เบื้องต้นระหว่างลำดับเวลา (X) กับยอดขาย (Y) ผลการพล็อตข้อมูล ดังภาพที่ 1



ภาพที่ 1 กราฟการกระจายของยอดขายปูนซีเมนต์รายเดือนเทียบกับลำดับเวลา (มกราคม 2565 ถึง กันยายน 2567)

จากภาพที่ 1 จากการสร้างกราฟการกระจายเบื้องต้น (Scatter plot) โดยใช้ข้อมูลยอดขายรายเดือน (ตารางที่ 1) เทียบกับลำดับเวลา พบว่าข้อมูลแสดงแนวโน้มความสัมพันธ์เชิงลบ ที่มีลักษณะใกล้เคียงเส้นตรง (Approximate linear trend) แม้จะมีการแกว่งตัวของข้อมูลในแต่ละเดือนก็ตาม ลักษณะเบื้องต้นนี้จึงสนับสนุนการเลือกใช้แบบจำลองการถดถอยเชิงเส้นตรงเพื่อวิเคราะห์ในขั้นตอนต่อไป

4.3 สร้างแบบจำลองการถดถอยเชิงเส้น ใช้ข้อมูลที่รวบรวมได้ในการสร้างแบบจำลองการถดถอยเชิงเส้นตรง [11] โดยใช้เทคนิคทางสถิติ คือ Least square method เพื่อหาค่าสัมประสิทธิ์ในสมการถดถอย

ตรวจสอบสมมติฐานทางสถิติของแบบจำลองดังสมการที่ (1)

$$y = a + bx \dots\dots\dots (1)$$

y แทน ตัวแปรตาม (ตัวแปรที่ต้องการพยากรณ์)

x แทน ตัวแปรอิสระ

a แทน จุดตัดแกน y (Y-intercept) หรือค่า y เมื่อ x = 0

b แทน ค่าความชันของเส้นตรง (Slope)

ค่าคงที่ a และ b สามารถคำนวณได้จากสมการที่ (2) และ (3) เมื่อ n = จำนวนข้อมูล

$$b = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2} \dots\dots\dots (2)$$

$$a = \frac{\sum y - b \sum x}{n} \dots\dots\dots (3)$$

4.4 การตรวจสอบข้อตกลงเบื้องต้นของแบบจำลองการถดถอยเชิงเส้นตรง

เพื่อยืนยันความน่าเชื่อถือของผลการวิเคราะห์และการนำแบบจำลอง $Y_t = a + bX$ ไปใช้ได้ทำการตรวจสอบข้อตกลงเบื้องต้นของการถดถอยเชิงเส้นตรงโดยพิจารณาจากค่าความคลาดเคลื่อน (Residuals, e_i) ซึ่งคำนวณจาก [12] ดังสมการ

$$e_i = Y_i - \hat{Y}_i = Y_i - (a + bX_i) \dots\dots\dots (5)$$

โดย

Y_i	คือ	ค่าสังเกตจริง
$\hat{Y}_i$	คือ	ค่าพยากรณ์ ณ เวลา X_i
a, b	คือ	ค่าพารามิเตอร์จากแบบจำลอง

ข้อตกลงที่ตรวจสอบตามแนวทางมาตรฐาน [12], [13] ดังนี้

4.4.1 ความเป็นเส้นตรง (Linearity)

ตรวจสอบว่าความสัมพันธ์ระหว่าง X และ Y เป็นเส้นตรง โดยพิจารณาจากกราฟการกระจาย (Scatter Plot) ภาพที่ 1 และ กราฟค่าความคลาดเคลื่อนเทียบกับค่าพยากรณ์ (Residuals vs. Predicted Values Plot) หากเป็นเส้นตรง จุดใน Residual Plot ควรกระจายแบบสุ่มรอบศูนย์ ไม่มีรูปแบบโค้งงอ [12]

4.4.2 ความเป็นอิสระของค่าความคลาดเคลื่อน (Independence) ตรวจสอบค่าความคลาดเคลื่อน (e_i) ณ เวลาต่างๆ เป็นอิสระต่อกัน (ไม่มี Autocorrelation) โดยใช้ สถิติดอร์บิน-วัตสัน (Durbin-Watson, DW) ซึ่งคำนวณจาก [12] ดังสมการ

$$DW = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2} \dots\dots\dots (6)$$

โดยค่า DW ที่เข้าใกล้ 2 บ่งชี้ว่าค่าความคลาดเคลื่อนค่อนข้างเป็นอิสระต่อกัน

4.4.3 การแจกแจงปกติของค่าความคลาดเคลื่อน (Normality) ตรวจสอบความคลาดเคลื่อน (e_i) มีการแจกแจงใกล้เคียงปกติ โดยพิจารณาจากฮิสโทแกรม (Histogram) และ กราฟ Q-Q Plot (Normal probability plot) ของค่า e_i [12]

4.4.4 ความคงที่ของความแปรปรวน (Homoscedasticity) ตรวจสอบความแปรปรวนของค่าความคลาดเคลื่อน (e_i) คงที่ตลอดช่วงค่า X หรือ $\hat{Y}$ โดยพิจารณาจาก กราฟค่าความคลาดเคลื่อนเทียบกับค่าพยากรณ์ (Residuals vs. predicted values plot) การกระจายตัวของจุดควรมีความกว้างสม่ำเสมอ [12]

4.5 ประเมินความเหมาะสมของแบบจำลอง โดยใช้ค่าสถิติต่าง ๆ ดังนี้

4.5.1 สัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficient, r) ดังสมการที่ (7)

$$r = \frac{n \sum xy - \sum x \sum y}{\sqrt{[n \sum x^2][n \sum y^2 - (\sum y)^2]}} \dots\dots\dots (7)$$

โดยที่ค่าสัมประสิทธิ์ (r) จะอยู่ระหว่าง -1.00 ถึง 1.00

ถ้า r เข้าใกล้ค่า 1.00 จะเป็นค่ายืนยันตัวแปรตาม (ความต้องการแปรผันตรงกับตัวแปรต้น)

ถ้า r เข้าใกล้ค่า -1.00 จะเป็นการยืนยันว่าตัวแปรตามจะแปรผกผันกับตัวแปรต้น

ถ้า r เข้าใกล้ 0 บ่งชี้ว่าตัวแปรต้นกับตัวแปรตามไม่ได้มีความสัมพันธ์เชิงเส้นระหว่างกัน แต่อาจมีความสัมพันธ์ในลักษณะอื่นได้

4.3.2 การทดสอบนัยสำคัญทางสถิติ (Statistical Significance) สามารถทดสอบทางสถิติ ดังนี้

1) การทดสอบนัยสำคัญของสัมประสิทธิ์ความชัน (t-test for slope coefficient, b) ใช้ทดสอบสมมติฐานว่าค่าความชันที่แท้จริงในประชากร (β) แตกต่างจากศูนย์หรือไม่ หากปฏิเสธ H_0 แสดงว่าตัวแปรเวลา (X) มีอิทธิพลต่อยอดขาย (Y) อย่างมีนัยสำคัญทางสถิติในเชิงเส้นตรง การทดสอบนี้อาศัยค่า t-statistic ซึ่งคำนวณจาก [12] ดังสมการที่ 8

$$t = b/SE(b) \dots\dots\dots(8)$$

โดย $SE(b)$ คือ ค่าความคลาดเคลื่อนมาตรฐานของความชัน b

2) การทดสอบนัยสำคัญของแบบจำลองโดยรวม (F-test for overall significance) สำหรับการถดถอยเชิงเส้นตรงอย่างง่าย ผลการทดสอบ F-test จะเทียบเท่ากับการทดสอบ t-test ของค่าความชัน ใช้ทดสอบสมมติฐานว่าแบบจำลองโดยรวมมีนัยสำคัญหรือไม่ (H_0 : แบบจำลองไม่มีนัยสำคัญ vs H_1 : แบบจำลองมีนัยสำคัญ) การทดสอบนี้อาศัยค่า F-statistic ซึ่งคำนวณจากอัตราส่วนความแปรปรวนที่อธิบายได้โดยแบบจำลองต่อความแปรปรวนที่อธิบายไม่ได้

4.6 พยากรณ์ยอดขายปูนซีเมนต์ ใช้แบบจำลองที่ผ่านการประเมินแล้ว ในการพยากรณ์ยอดขายปูนซีเมนต์ในช่วงเวลาที่ต้องการ โดยแทนค่าตัวแปรอิสระที่ต้องการพยากรณ์ลงในสมการถดถอย ด้วยฟอร์มแคสชีส (Forecast sheet)

5. ผลการวิจัย

ผลการการทำนายยอดขายปูนซีเมนต์ ด้วยวิธีสมการเส้นถดถอยเชิงเส้นตรง (Linear regression model) กรณีศึกษา อ.รัตนภูมิ จ.สงขลา มีรายละเอียดการวิจัย ดังนี้

5.1 ผลสร้างสมการความสัมพันธ์เชิงเส้นตรง ซึ่งแสดงความสัมพันธ์ของยอดขายปูนซีเมนต์ในแต่ละเดือน ดังตารางที่ 2

ตารางที่ 2 แสดงข้อมูลการคำนวณเพื่อหาสมการความสัมพันธ์เชิงเส้นตรง

ปี	เดือน	x^2	กระสอบ	y^2	xy
	(x)		(y)		
65	ม.ค.	1	1,128	1,272,384	1128
	ก.พ.	4	1,092	1,192,464	2184
	มี.ค.	9	1,207	1,456,849	3621
	*	*	*	*	*
	*	*	*	*	*
	ธ.ค.	144	990	980,100	11,880
66	ม.ค.	169	1,017	1,034,289	13,221
	ก.พ.	196	993	986,049	13,902
	มี.ค.	225	1,002	1,004,004	15,030
	*	*	*	*	*
	ธ.ค.	576	951	904,401	22,824
67	ม.ค.	625	940	883,600	23,500
	ก.พ.	676	938	879,844	24,388
	มี.ค.	729	955	912,025	25,785
	*	*	*	*	*
	*	*	*	*	*
	ส.ค.	1,089	960	921,600	31,680
รวม	561	12,529	33,609	34,601,013	549,364

จากตารางที่ 2 นำไปสู่สร้างสมการความสัมพันธ์เชิงเส้นตรง เพื่อแสดงความสัมพันธ์ระหว่างตัวแปร และทำนายยอดขายในอนาคต ดังนี้

หาค่า b จากสมการที่ 2

$$b = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2} \dots\dots\dots(9)$$

$$= \frac{33(549,364) - (561)(33,609)}{33(12,529) - (561)^2}$$

$$= -7.35$$

1) หาค่า a จากสมการที่ 1

$$a = \frac{\sum y - b \sum x}{n} \dots\dots\dots(10)$$

$$= \frac{33,609 - (-7.35)(561)}{33}$$

$$= 1,143.4$$

- 2) นำค่า a และ b ที่คำนวณได้ แทนค่าในสมการที่ 1 (สมมติต้องการพยากรณ์เดือนที่ 34)

$$\begin{aligned} y_{34} &= 1,143.4 + (-7.35)(34) \\ &= 893.5 \end{aligned}$$

จากสมการแสดงความสัมพันธ์ อธิบายได้ว่าจากสมการแนวโน้มเชิงเส้น y_{34} หมายถึง ค่าที่ทำนายของตัวแปร y ณ เวลาหรือจุดข้อมูลที่ 34 โดยที่ 1,143.4 คือค่าคงที่หรือจุดตัดแกน y (ค่าของ y เมื่อ $x = 0$) และ -7.35 คือ ค่าความชันของเส้นแนวโน้ม ซึ่งแสดงถึงอัตราการเปลี่ยนแปลงของ y ต่อการเปลี่ยนแปลง 1 หน่วยของ x ส่วน 34 หมายถึง ค่าของตัวแปร x ซึ่งในที่นี้จะเป็นเวลา ลำดับข้อมูล หรือตัวแปรอิสระอื่นๆ สรุปได้ว่าสมการความสัมพันธ์เชิงเส้นตรงของยอดขายปูนซีเมนต์ระยะเวลา 3 ปี ในช่วงเดือนมกราคม พ.ศ. 2565 ถึงกันยายน พ.ศ. 2567 สามารถเขียนเป็นสมการเส้นถดถอยเชิงเส้นตรงได้ ดังนี้

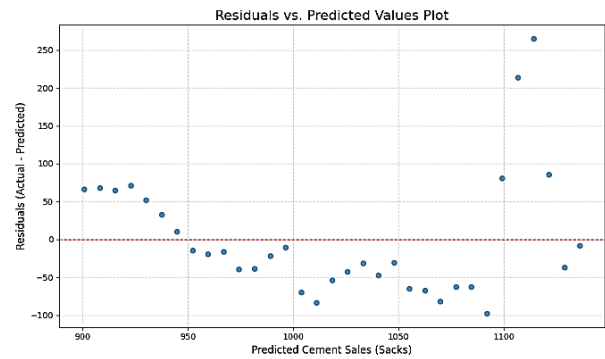
$$\hat{Y}_t = 1143.4 - 7.35X \dots\dots\dots (11)$$

จากสมการนี้แสดงให้เห็นว่าค่าของ y มีแนวโน้มลดลง เนื่องจากความชันเป็นลบ (-7.35) นั่นหมายความว่าทุก ๆ ครั้งที่ x เพิ่มขึ้น 1 หน่วย ค่าของ y จะลดลง 7.35 หน่วย และเมื่อ $x = 34$, ค่าที่ทำนายของ y คือ 893.5

5.2 ผลการตรวจสอบข้อตกลงเบื้องต้น

เพื่อให้มั่นใจในความน่าเชื่อถือของสมการถดถอย การตรวจสอบข้อตกลงเบื้องต้นตามวิธีการกล่าวไว้ข้างต้น มีผลรายการตรวจสอบ ดังนี้

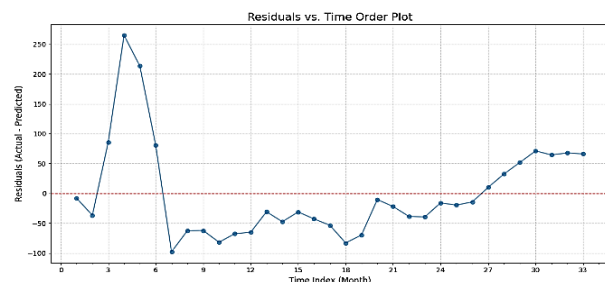
5.2.1 ผลความเป็นเส้นตรง (Linearity) สำหรับข้อตกลงด้านความเป็นเส้นตรง (Linearity) ได้ทำการตรวจสอบจากกราฟค่าความคลาดเคลื่อนเทียบกับค่าพยากรณ์ (Residuals vs. predicted values plot) ดังแสดงใน ภาพที่ 2



ภาพที่ 2 ค่าความคลาดเคลื่อนเทียบกับค่าพยากรณ์ (Residuals vs. Predicted values plot)

จากภาพที่ 2 ใช้ประเมินความเป็นเส้นตรง โดยหลักการแล้ว จุดควรกระจายแบบสุ่มรอบเส้นศูนย์ จากกราฟ พบว่าการกระจายตัว ไม่เป็นแบบสุ่มอย่างสมบูรณ์ โดยมีแนวโน้มที่ค่าความคลาดเคลื่อนจะเป็นบวก (แบบจำลองพยากรณ์ต่ำไป) เมื่อค่าพยากรณ์มีค่าต่ำหรือสูง และกระจายตัวค่อนข้างใกล้ศูนย์ในช่วงกลาง ลักษณะนี้อาจบ่งชี้ถึง ความโค้งเล็กน้อย (Slight curve) ในความสัมพันธ์ที่แท้จริง ดังนั้น ข้อตกลงด้านความเป็นเส้นตรงจึงอาจ ไม่สมบูรณ์นัก แต่ยังพอรับได้สำหรับแบบจำลองพื้นฐานนี้ โดยถือเป็นข้อจำกัดสำคัญของแบบจำลอง

5.2.2 ความเป็นอิสระของค่าความคลาดเคลื่อน (Independence) การตรวจสอบความเป็นอิสระของค่าความคลาดเคลื่อน ซึ่งสำคัญสำหรับข้อมูลอนุกรมเวลา ได้ดำเนินการโดยใช้กราฟค่าความคลาดเคลื่อนเทียบกับลำดับเวลา (Residuals vs. Time Order Plot) ดังภาพที่ 3 และค่าสถิติดอร์บิน-วัตสัน (DW)



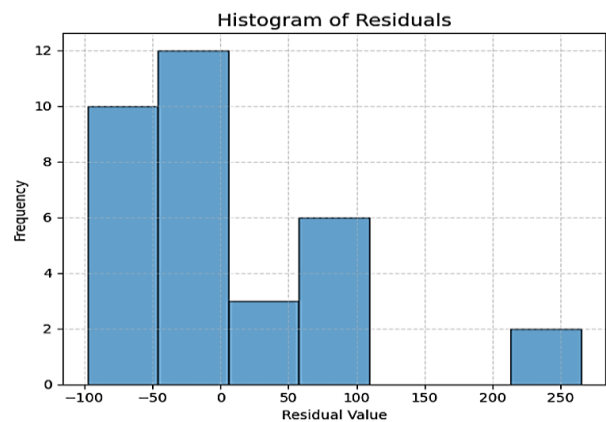
ภาพที่ 3 ค่าความคลาดเคลื่อนเทียบกับลำดับเวลา (Residuals vs. Time Order Plot)

จากภาพที่ 3 พบว่าค่าความคลาดเคลื่อน มีลักษณะการแกว่งตัวที่ไม่สุ่มรอบเส้นศูนย์ โดยมีช่วงที่ ค่าบวกหรือลบเกาะกลุ่มกัน ซึ่งอาจบ่งชี้ถึง สหสัมพันธ์ อัตโนมิติ (Autocorrelation) ในข้อมูล นอกจากนี้ ค่าสถิติ Durbin-Watson (DW) ที่คำนวณได้ คือ 0.5280 ซึ่งมีค่าต่ำกว่า 2 อย่างมาก ยืนยันถึงการมีอยู่ของ สหสัมพันธ์อัตโนมิติเชิงบวกที่ชัดเจน ในค่าความ คลาดเคลื่อน การมีอยู่ของ Autocorrelation นี้แสดงว่า ข้อตกลงด้านความเป็นอิสระของค่าความคลาดเคลื่อน สำหรับแบบจำลองการถดถอยเชิงเส้นตรงอย่างง่าย ไม่ได้ รับการตอบสนองอย่างสมบูรณ์ สำหรับข้อมูลชุดนี้ ซึ่งเป็น ข้อจำกัดทางสถิติที่ทราบกันดีเมื่อใช้แบบจำลองพื้นฐาน กับข้อมูลอนุกรมเวลา และอาจส่งผลต่อความแม่นยำของ การทดสอบนัยสำคัญบางประการ

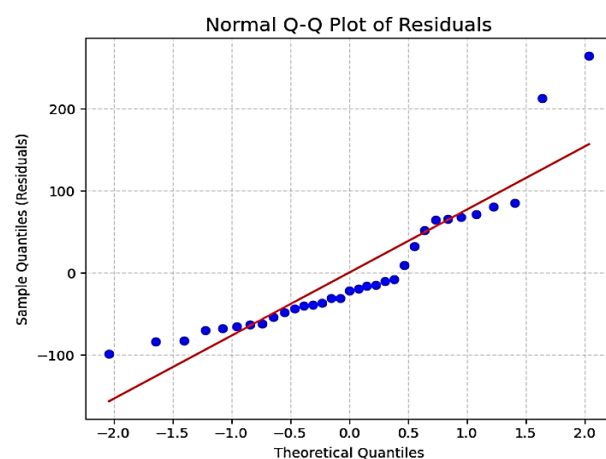
อย่างไรก็ตาม ดังที่กล่าวไว้ในระเบียบวิธีวิจัย วัตถุประสงค์หลักของการศึกษานี้คือการสร้างและ ประเมินแบบจำลองการถดถอยเชิงเส้นตรงอย่างง่าย เพื่อ เป็นเครื่องมือพื้นฐานในการทำความเข้าใจแนวโน้มหลัก และเป็นจุดเริ่มต้นสำหรับการพยากรณ์เบื้องต้น ที่ผู้ประกอบการอาจนำไปประยุกต์ใช้ได้สะดวก แม้จะพบ ปัญหา Autocorrelation จึงยังคงนำเสนอผลการ วิเคราะห์จากแบบจำลองเชิงเส้นตรงอย่างง่ายนี้ ต่อไป เพื่อให้เห็นถึงประสิทธิภาพและข้อจำกัดของมัน ในบริบทจริง แต่ผู้ใช้งานควร ตระหนักถึงข้อจำกัดนี้ อย่างยิ่ง และทราบว่าการพยากรณ์หรือการทดสอบ สมมติฐานจากแบบจำลองนี้อาจมีความคลาดเคลื่อน สูงกว่าที่ควรจะเป็น การใช้เทคนิคการวิเคราะห์อนุกรม เวลาที่ ซับซ้อนขึ้น (เช่น ARIMA หรือการปรับแก้ Autocorrelation) จึงเป็นแนวทางที่แนะนำสำหรับการ วิเคราะห์ในเชิงลึกหรือการพยากรณ์ที่ต้องการความ แม่นยำสูงขึ้นในอนาคต

5.2.3 การแจกแจงปกติ ของค่าความ คลาดเคลื่อน (Normality) ข้อตกลงด้านการแจกแจงปกติ ของค่าความคลาดเคลื่อน (Normality of errors) ได้รับการตรวจสอบโดยใช้ฮิสโตแกรม (Histogram) และกราฟ

Q-Q Plot (Normal probability plot) ของค่าความ คลาดเคลื่อน ดังภาพที่ 4 และ 5 รวมถึงการทดสอบทาง สถิติ Shapiro-Wilk ดังตารางที่ 3



ภาพที่ 4 ฮิสโตแกรมของค่าความคลาดเคลื่อน



ภาพที่ 5 กราฟ Q-Q Plot ของค่าความคลาดเคลื่อน

ตารางที่ 3 ผลการทดสอบ Shapiro-Wilk สำหรับการ แจกแจงปกติของค่าความคลาดเคลื่อน

สถิติทดสอบ (Statistic)	ค่า P-value	ระดับนัยสำคัญ (α)	ผลการ ตัดสินใจ
0.8407	0.0002	0.05	ปฏิเสธ H_0

จากภาพที่ 4 และตารางที่ 3 พบว่าการกระจาย ตัว ไม่สอดคล้องกับการแจกแจงปกติอย่างชัดเจน โดยมี ลักษณะเบ้ขวาและจุดใน Q-Q Plot เบี่ยงเบนออกจาก แนวเส้นตรง ผลนี้ได้รับการยืนยันจากการทดสอบ Shapiro-Wilk test ซึ่งให้ค่า p-value = 0.0002 ต่ำกว่า ระดับนัยสำคัญ 0.05 อย่างมีนัยสำคัญ แสดงว่าข้อตกลง

ด้านการแจกแจงปกติไม่ได้รับการตอบสนอง การที่ค่าความคลาดเคลื่อนไม่เป็นไปตามการแจกแจงปกติอาจส่งผลกระทบต่อความแม่นยำของการทดสอบนัยสำคัญทางสถิติ (t-test, F-test) และช่วงความเชื่อมั่นของพารามิเตอร์ที่ได้จากแบบจำลองนี้

อย่างไรก็ตาม ตามวัตถุประสงค์หลักของงานวิจัยนี้ที่มุ่งเน้นการพัฒนาแบบจำลองเชิงเส้นตรงอย่างง่าย เพื่ออธิบายแนวโน้มหลักของยอดขายปูนซีเมนต์ตามเวลา และเพื่อให้ได้สมการพยากรณ์เบื้องต้นที่สามารถนำไปประยุกต์ใช้ได้โดยง่าย แม้จะพบการละเมิดข้อตกลงด้านการแจกแจงปกตินี้ก็ตาม ค่าประมาณของสัมประสิทธิ์การถดถอย (a และ b) ที่ได้จากวิธี Least Squares ยังคงเป็นค่าประมาณที่ไม่เอนเอียง (Unbiased) และสามารถให้ข้อมูลเกี่ยวกับทิศทางและขนาดของแนวโน้มได้ ดังนั้นจึงยังคงใช้แบบจำลองเชิงเส้นตรงที่ได้ ($Y_t = 1143.4 - 7.35X$) ในการวิเคราะห์แนวโน้มและการพยากรณ์ต่อไปตามวัตถุประสงค์ที่ตั้งไว้

5.2.4 ความคงที่ของความแปรปรวน (Homoscedasticity) การตรวจสอบความคงที่ของความแปรปรวน โดยใช้กราฟค่าความคลาดเคลื่อนเทียบกับค่าพยากรณ์ (ภาพที่ 2) พบว่าการกระจายตัวของค่าความคลาดเคลื่อนในแนวตั้งมีแนวโน้มกว้างขึ้น เมื่อค่าพยากรณ์สูงขึ้น ซึ่งบ่งชี้ถึงปัญหา ความแปรปรวนไม่คงที่ (Heteroscedasticity) อย่างไรก็ตาม สำหรับวัตถุประสงค์ของงานวิจัยนี้ที่มุ่งเน้นการสร้างแบบจำลองพื้นฐานเพื่ออธิบายแนวโน้มหลัก และพิจารณาถึงความเรียบง่ายในการนำไปใช้ แม้จะพบข้อบ่งชี้ถึง Heteroscedasticity การเปลี่ยนแปลงความแปรปรวนอาจยังไม่รุนแรงมากพอที่จะทำให้แบบจำลองเชิงเส้นตรงที่ได้ ($Y_t = 1143.4 - 7.35X$) ไม่สามารถให้ภาพรวมของแนวโน้มได้จึงยังคงใช้แบบจำลองเชิงเส้นตรงนี้ในการวิเคราะห์ต่อไป แต่ตระหนักว่าปัญหา Heteroscedasticity ที่พบนี้เป็นข้อจำกัดหนึ่ง ซึ่งอาจส่งผลกระทบต่อความแม่นยำของช่วงความเชื่อมั่นและการทดสอบนัยสำคัญบางประการ และเป็น

ประเด็นที่ควรพิจารณาปรับปรุงในงานวิจัยหรือการพัฒนาแบบจำลองในอนาคต

5.3 การประเมินระดับความน่าเชื่อถือของความสัมพันธ์ระหว่างยอดขายปูนซีเมนต์กับเวลา

หลังจากได้สมการถดถอยเชิงเส้นตรงและตรวจสอบข้อตกลงเบื้องต้นแล้ว ได้ทำการประเมินความเหมาะสมของแบบจำลองเพื่อวัดระดับความสัมพันธ์และนัยสำคัญทางสถิติ ดังนี้

5.3.1 สัมประสิทธิ์สหสัมพันธ์ (Correlation coefficient, r) จากผลการวิเคราะห์ข้อมูลด้วยวิธีถดถอยเชิงเส้นตรง พบว่าค่าสัมประสิทธิ์สหสัมพันธ์ (r) มีค่าเท่ากับ -0.6613 แสดงให้เห็นถึงความสัมพันธ์เชิงลบระดับปานกลางระหว่างยอดขายปูนซีเมนต์กับระยะเวลา ซึ่งหมายความว่าเมื่อเวลาผ่านไป ยอดขายมีแนวโน้มลดลงอย่างสม่ำเสมอ นอกจากนี้ ค่า R-squared (r^2) เท่ากับ 0.437 บ่งชี้ว่าแบบจำลองสามารถอธิบายความแปรปรวนของยอดขายได้ร้อยละ 43.7 ส่วนที่เหลืออาจเกิดจากปัจจัยอื่นนอกเหนือจากที่นำมาวิเคราะห์ เช่น สภาพเศรษฐกิจ ฤดูกาล หรือการแข่งขันในตลาด

5.3.2 ผลการทดสอบนัยสำคัญทางสถิติ (Statistical significance) ผลการทดสอบนัยสำคัญทางสถิติของแบบจำลองการถดถอยเชิงเส้นตรง การวิเคราะห์นัยสำคัญของสัมประสิทธิ์ความชันและแบบจำลองโดยรวม ดังตารางที่ 4

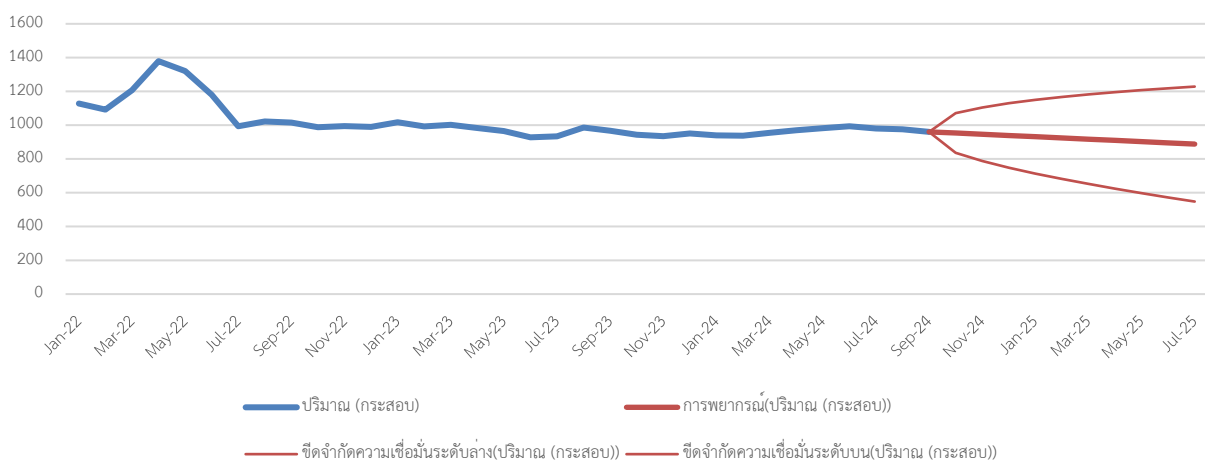
ตารางที่ 4 ผลการทดสอบ t-test สำหรับสัมประสิทธิ์ของตัวแปรเวลา (Time index)

ตัวแปร (Predictor)	สัมประสิทธิ์ (b)	Std. Error (SE)	t-value	P-value
ค่าคงที่ (Intercept)	1142.9678	29.3917	38.888	< 0.001
เวลา (Time index)	-7.3118	1.5084	-4.847	< 0.001
สถิติแบบจำลอง:				
$R^2 = 0.431$				
$F(1, 31) = 23.50$				
$p (F\text{-test}) < 0.001$				

จากตารางที่ 4 แสดงให้เห็นนัยสำคัญของสัมประสิทธิ์ความชัน (Slope Coefficient Significance) ผลการทดสอบ t-test สำหรับสัมประสิทธิ์ของตัวแปรเวลา (Time Index) แสดงค่า $t = -4.847$ และ $p\text{-value} < 0.001$. ค่า $p\text{-value}$ ที่ต่ำกว่าระดับนัยสำคัญ $\alpha = 0.05$ อย่างมีนัยสำคัญนี้ บ่งชี้ถึงการปฏิเสธสมมติฐานหลัก ($H^0: \beta_{TimeIndex} = 0$) ดังนั้น จึงสรุปได้ว่า ตัวแปรเวลา มีความสัมพันธ์เชิงลบกับยอดขายปูนซีเมนต์อย่างมีนัยสำคัญทางสถิติ และนัยสำคัญของแบบจำลองโดยรวม (Overall Model Significance) ด้วยผลการทดสอบ F-test ให้ค่า $F(1, 31) = 23.50$ โดยมีค่า $p\text{-value} < 0.001$. ค่า $p\text{-value}$ ที่ต่ำกว่าระดับนัยสำคัญ $\alpha = 0.05$ อย่างมีนัยสำคัญ ยืนยันการปฏิเสธสมมติฐานหลัก (H_0 : แบบจำลองไม่มีนัยสำคัญ)

ลดลงอย่างชัดเจนในช่วงกลางปี 2565 โดยข้อมูลที่ใช้ในการพยากรณ์ครอบคลุมช่วงเวลาตั้งแต่เดือนมกราคม พ.ศ. 2565 ถึงกันยายน พ.ศ. 2567 และทำการพยากรณ์ต่อเนื่องไปจนถึงเดือนกรกฎาคม พ.ศ. 2568

ผลการพยากรณ์ (เส้นสีแดง) แสดงค่าคาดการณ์ยอดขายในหน่วยกระสอบ โดยมีช่วงความเชื่อมั่นที่ระดับบนและล่าง (เส้นสีแดงบาง) ครอบคลุมช่วงที่อาจเกิดขึ้นได้จริง ซึ่งเป็นประโยชน์ต่อการวางแผนกำลังการผลิต การบริหารสินค้าคงคลัง และการคาดการณ์รายได้ในอนาคต เส้นสีน้ำเงินแสดงปริมาณยอดขายจริงที่ผ่านมา เส้นแดงตรงกลางคือ การพยากรณ์ยอดขายในอนาคต เส้นแดงด้านบนและด้านล่างแสดงช่วงค่าความเชื่อมั่น (Confidence Interval) ซึ่งช่วยให้เห็นความไม่แน่นอนที่อาจเกิดขึ้นดังภาพที่ 6



ภาพที่ 6 กราฟแสดงผลการพยากรณ์ยอดขายปูนซีเมนต์ (หน่วย: กระสอบ) พร้อมช่วงความเชื่อมั่นจากข้อมูลช่วง มกราคม 2565 – กรกฎาคม 2568

สรุปได้ว่าแบบจำลองการถดถอยเชิงเส้นตรงโดยรวม มีความสามารถในการอธิบายความแปรปรวนของยอดขายปูนซีเมนต์ได้อย่างมีนัยสำคัญทางสถิติ

5.4 การพยากรณ์ ยอดขายปูนซีเมนต์ ในอนาคต

การวิเคราะห์ข้อมูลการขายปูนซีเมนต์ในอดีต ด้วยเทคนิค Forecast Sheet จากโปรแกรม Microsoft Excel แสดงให้เห็นแนวโน้มยอดขายในช่วงปี พ.ศ. 2565 ถึงกลางปี 2567 มีลักษณะค่อนข้างคงที่หลังจากมีการ

6. การสรุปผลและอภิปรายผล

การวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาและประเมินแบบจำลองการถดถอยเชิงเส้นตรงอย่างง่ายสำหรับพยากรณ์ยอดขายปูนซีเมนต์รายเดือน ในอำเภอรัษฎา จังหวัดสงขลา โดยใช้ข้อมูลอนุกรมเวลาตั้งแต่เดือนมกราคม พ.ศ. 2565 ถึง กันยายน พ.ศ. 2567 ผลการวิเคราะห์พบว่าสมการถดถอยที่เหมาะสมที่สุด คือ

$$\hat{Y}_t = 1143.4 - 7.35X \dots\dots\dots(12)$$

ซึ่งบ่งชี้ถึง แนวโน้มยอดขายที่ลดลงเฉลี่ย 7.35 กระสอบต่อเดือน อย่างมีนัยสำคัญทางสถิติ ($p < 0.001$ สำหรับทั้ง t-test ของความชัน และ F-test ของแบบจำลองโดยรวม) โดยมีความสัมพันธ์เชิงลบใน ระดับปานกลาง ระหว่างเวลาและยอดขาย ($r = -0.6613$) ผลการประเมินแบบจำลองเผยให้เห็นข้อจำกัดที่สำคัญหลายประการ ประการแรก ค่าสัมประสิทธิ์การตัดสินใจ (R^2) อยู่ในระดับค่อนข้างต่ำเพียง 0.437 (43.7%) ซึ่งหมายความว่าตัวแปรเวลาเพียงอย่างเดียวสามารถอธิบายความผันแปรของยอดขายปูนซีเมนต์ได้ไม่ถึงครึ่ง ยังมีปัจจัยอื่นอีกมาก (56.3%) ที่ส่งผลกระทบต่อยอดขายซึ่งแบบจำลองนี้ไม่ได้รวมไว้ ข้อค้นพบนี้สอดคล้องกับงานวิจัยของ Liu [2] ที่เน้นย้ำความสำคัญของการพิจารณาปัจจัยหลายตัวแปรในการสร้างแบบจำลองพยากรณ์ที่แม่นยำ แสดงให้เห็นว่าปัจจัยภายนอก เช่น สภาพเศรษฐกิจ ฤดูกาล โครงการก่อสร้างในพื้นที่ และการแข่งขันทางการตลาด น่าจะมีบทบาทสำคัญต่อยอดขายปูนซีเมนต์ในพื้นที่ศึกษา และการขาดปัจจัยเหล่านี้เป็นสาเหตุหลักของค่า R^2 ที่ต่ำ ประการที่สอง การตรวจสอบข้อตกลงเบื้องต้นของการถดถอยเชิงเส้นตรง พบว่าแบบจำลอง ไม่ตอบสนองต่อข้อตกลงที่สำคัญหลายข้อ โดยเฉพาะอย่างยิ่งปัญหา สหสัมพันธ์อัตโนมัติ เชิงบวกที่ชัดเจน (Clear positive autocorrelation) ซึ่งยืนยันจากทั้งลักษณะการเกาะกลุ่มของค่าความคลาดเคลื่อนในกราฟ Residuals vs. Time Order (รูปที่ 3) และค่า Durbin-Watson ที่ต่ำมาก ($DW \approx 0.528$) นอกจากนี้ยังพบว่า ค่าความคลาดเคลื่อนไม่มีการแจกแจงปกติ (Non-normality) อย่างชัดเจน (p -value ≈ 0.0002 จาก Shapiro-Wilk test; รูปที่ 4) และมีข้อบ่งชี้ถึงปัญหา ความแปรปรวนไม่คงที่ (Heteroscedasticity) โดยความแปรปรวนมีแนวโน้มเพิ่มขึ้นเมื่อค่าพยากรณ์สูง รวมถึงข้อสังเกตว่า ความเป็นเส้นตรง (Linearity) อาจไม่สมบูรณ์นัก การที่แบบจำลองไม่เป็นไปตามข้อตกลงเบื้องต้นเหล่านี้ โดยเฉพาะ Autocorrelation และ Non-normality เป็นปัญหาที่พบ

บ่อยเมื่อประยุกต์ใช้แบบจำลองถดถอยเชิงเส้นตรงอย่างง่ายกับข้อมูลอนุกรมเวลาโดยตรง ดังที่ตำราเศรษฐมิติมาตรฐานได้กล่าวถึง [12], [13] ปัญหาเหล่านี้ส่งผลกระทบอย่างมีนัยสำคัญต่อความน่าเชื่อถือของผลการทดสอบนัยสำคัญทางสถิติ (p -values), ค่าความคลาดเคลื่อนมาตรฐาน (Standard errors) และช่วงความเชื่อมั่นที่คำนวณได้ ซึ่งหมายความว่าแม้ค่าสัมประสิทธิ์แนวโน้ม (-7.35) จะยังคงเป็นค่าประมาณที่ไม่เอนเอียง แต่ระดับความเชื่อมั่นทางสถิติที่รายงานอาจไม่ถูกต้องนัก

ข้อจำกัดที่พบนี้ยังสะท้อนถึงความซับซ้อนในการพยากรณ์ยอดขาย และชี้ให้เห็นว่าแบบจำลองที่ซับซ้อนกว่าอาจมีความเหมาะสมมากกว่า ดังเช่นงานวิจัยของ Md Rokibul Hasan [14] และ Taufiqh and Ambarwati [15] ที่พบว่าเทคนิค Machine learning เช่น Gradient boosting หรือ Neural network สามารถให้ผลการพยากรณ์ที่แม่นยำกว่า Linear regression สำหรับข้อมูลสินค้าประเภทอื่น ซึ่งอาจเป็นเพราะความสามารถในการจับรูปแบบที่ไม่เป็นเชิงเส้นหรือความสัมพันธ์ที่ซับซ้อนได้ดีกว่า นอกจากนี้ การไม่ได้พิจารณาปัจจัยด้าน แนวโน้มที่ซับซ้อนและฤดูกาล ตามที่ Naresh Kumar Reddy [16] เน้นย้ำ ก็อาจเป็นสาเหตุสำคัญของทั้งค่า R^2 ที่ต่ำและปัญหา Autocorrelation ที่พบในการศึกษานี้ ถึงแม้ว่าการนำเสนอ Least Square Method ในงานของ Juliana., et al. [17] จะแสดงให้เห็นว่าวิธีการนี้ยังคงเป็นเครื่องมือพื้นฐานที่น่าสนใจ แต่ผลการวิจัยนี้ชี้ให้เห็นว่าการประยุกต์ใช้ในรูปแบบเชิงเส้นตรงอย่างง่ายสำหรับข้อมูลอนุกรมเวลาชุดนี้ มีข้อจำกัดทางสถิติที่สำคัญซึ่งต้องตระหนักถึง

อย่างไรก็ตาม สำหรับกรณีศึกษาของกลุ่มร้านค้าวัสดุก่อสร้างในอำเภอรัตนภูมิ ซึ่งอาจต้องการเครื่องมือพยากรณ์เบื้องต้นที่เข้าใจง่ายและนำไปใช้งานได้โดยไม่ซับซ้อน แบบจำลองการถดถอยเชิงเส้นตรงที่พัฒนาขึ้นนี้ ยังคงมีคุณค่าในทางปฏิบัติ มันสามารถให้ภาพรวมของแนวโน้มยอดขายหลัก (การลดลง) ได้อย่างชัดเจน ซึ่งเป็นข้อมูลสำคัญสำหรับการวางแผนเชิงกลยุทธ์เบื้องต้น

เช่น การปรับสต็อกสินค้า หรือการพิจารณากลยุทธ์การตลาดเพื่อกระตุ้นยอดขาย แม้ว่าความแม่นยำในการพยากรณ์จุดต่อจุด (Point forecast accuracy) อาจมีจำกัด และการอนุมานทางสถิติควรทำด้วยความระมัดระวังเนื่องจากการไม่ตอบสนองต่อข้อตกลงเบื้องต้นบางประการ แต่ในฐานะเครื่องมือตั้งต้น แบบจำลองนี้สามารถช่วยให้ผู้ประกอบการเห็นทิศทางและเริ่มนำข้อมูลมาใช้ประกอบการตัดสินใจได้ดีกว่าการคาดการณ์โดยไม่มีหลักเกณฑ์

โดยสรุปแบบจำลองการถดถอยเชิงเส้นตรงอย่างง่ายที่ได้พัฒนาขึ้นนี้ เป็นเครื่องมือที่มีประโยชน์ในระดับหนึ่งสำหรับกรณีศึกษา ในการบ่งชี้แนวโน้มระยะกลางและเป็นจุดเริ่มต้นสำหรับการวางแผนทางธุรกิจ แม้จะมีข้อจำกัดทางสถิติซึ่งควรตระหนักถึง และเปิดโอกาสสำหรับการพัฒนาไปสู่แบบจำลองที่ซับซ้อนและแม่นยำยิ่งขึ้นในอนาคต

7. ข้อเสนอแนะ

7.1 การนำไปใช้งาน

ข้อเสนอแนะการนำผลการวิจัยการพยากรณ์ยอดขายปูนซีเมนต์กลุ่มร้านจำหน่ายวัสดุก่อสร้างในอำเภอรัตนภูมิ จังหวัดสงขลาไปใช้งาน สามารถนำข้อมูลการพยากรณ์มาปรับแผนการผลิตและการจัดการสินค้าคงคลังให้สอดคล้องกับความต้องการ ลดความเสี่ยงสินค้าขาดตลาดหรือสินค้าคงคลังเกิน รวมถึงกำหนดกลยุทธ์ทางการตลาด เช่น การส่งเสริมการขาย การกำหนดราคา และการจัดจำหน่ายได้อย่างมีประสิทธิภาพ นอกจากนี้ ควรศึกษาปัจจัยที่มีผลต่อยอดขายที่ไม่ได้รวมอยู่ในแบบจำลอง เช่น ฤดูกาล สภาพเศรษฐกิจ และการแข่งขัน เพื่อปรับปรุงความแม่นยำ และควรพัฒนาแบบจำลองโดยทดลองใช้วิธีการ Machine Learning อื่นๆ เช่น Gradient Boosting, Random Forest, หรือ Neural Network เพื่อเปรียบเทียบประสิทธิภาพ รวมถึงรวบรวมข้อมูลเพิ่มเติม เช่น ข้อมูลสภาพอากาศ กิจกรรมส่งเสริมการขายของคู่แข่ง และนโยบายรัฐบาล และปรับปรุง

แบบจำลองให้ทันสมัยอยู่เสมอเมื่อมีข้อมูลใหม่ๆ เพื่อให้การพยากรณ์แม่นยำและสะท้อนสถานการณ์ปัจจุบัน สุดท้าย ควรใช้ข้อมูลการพยากรณ์ร่วมกับข้อมูลจากแหล่งอื่นๆ และพิจารณาของของผู้เชี่ยวชาญประกอบการตัดสินใจทางธุรกิจ

7.2 ข้อเสนอแนะเพื่อเป็นแนวทางการพัฒนาเพื่อขยายผลต่อไป

เพื่อต่อยอดการวิจัยการพยากรณ์ยอดขายปูนซีเมนต์นี้ ควรขยายขอบเขตการศึกษาไปยังพื้นที่อื่นๆ เช่น จังหวัดหรือภูมิภาคอื่น ๆ เพื่อเปรียบเทียบปัจจัยที่มีผลต่อยอดขายและพัฒนากลยุทธ์ที่เหมาะสมกับแต่ละพื้นที่ รวมถึงศึกษาปัจจัยเชิงลึกเพิ่มเติมที่อาจมีผลต่อยอดขายปูนซีเมนต์ เช่น นโยบายภาครัฐ โครงการก่อสร้างขนาดใหญ่ ราคาบ้านและที่ดิน เป็นต้น และพฤติกรรมผู้บริโภค ควรพัฒนาแบบจำลองให้มีความซับซ้อนมากขึ้น เช่น การใช้ Deep Learning หรือแบบจำลองที่สามารถรองรับข้อมูลที่ไม่เป็นเชิงเส้น เป็นต้น และพิจารณาการใช้ข้อมูลที่มีความถี่สูงขึ้น เช่น ข้อมูลรายวันหรือรายสัปดาห์ เพื่อให้จับรูปแบบการเปลี่ยนแปลงของยอดขายได้แม่นยำยิ่งขึ้น เป็นต้น นอกจากนี้ ควรมีการพัฒนาเว็บหรือเครื่องมือที่ใช้งานง่าย เช่น Dashboard เพื่อให้ผู้ใช้งานสามารถเข้าถึงและนำข้อมูลการพยากรณ์ไปใช้ประโยชน์ได้สะดวก รวมถึง ควรมีการศึกษาวิจัยเชิงคุณภาพเพิ่มเติม เช่น การสัมภาษณ์ผู้เชี่ยวชาญ หรือกลุ่มตัวอย่าง เพื่อให้เข้าใจถึงปัจจัยที่มีผลต่อยอดขายในเชิงลึก และนำข้อมูลเชิงคุณภาพมาประกอบการพัฒนารูปแบบจำลองให้มีประสิทธิภาพมากขึ้น และสุดท้าย ควรเผยแพร่ผลการวิจัยและแบบจำลองที่พัฒนาขึ้น เพื่อให้ผู้ประกอบการหน่วยงานภาครัฐ และผู้ที่สนใจสามารถนำไปประยุกต์ใช้ได้อย่างกว้างขวาง และเกิดประโยชน์สูงสุด

8. เอกสารอ้างอิง

- [1] Suwanno, T. 2023. *Situation of Cement Industry in Thailand 2023*. [Online]. Available: <https://www.dpim.go.th/service/download?articleid=17266&F=39760> Accessed 17 January 2024.
- [2] Chen, D. and et al. 2022. "Sales Forecasting for Fashion Products Considering Lost Sales," *Applied Sciences*. 12(14): pp. 7081.
- [3] Qu, F. and et al. 2022. "Forecasting of automobile sales based on support vector regression optimized by the grey wolf optimizer algorithm," *Mathematics*. 10(13): pp. 2234.
- [4] Uras, N. and et al. 2020. "Forecasting Bitcoin closing price series using linear regression and neural networks models," *PeerJ Computer Science*. 6: pp. e279.
- [5] Haviluddin, A. S. and Yuniarti, S. 2014. "A Comparison between Simple Linear Regression and Radial Basis Function Neural Network (RBFNN) Models for Predicting Students' Achievement," In *Proceedings of International Conference on Education*, 2 April 2014, pp. 99–308. [Online]. Available: https://repository.unmul.ac.id/bitstream/handle/123456789/1673/file_1031900082.pdf?sequence=1. Accessed 16 November 2024.
- [6] Nelvidawati, N. 2024. "Simple Regression Model Analysis of the Effect of Temperature on Rainfall in Padang City Using Scikit-Learn," *International Journal of Wireless and Multimedia Communications*. 1(2): pp. 1–6.
- [7] Bujang, M. A. and et al. 2009. "Prediction number of deaths by occurrence in Malaysia: a comparison between simple linear regression model and holt's linear trend model," *Statistics Malaysia—Journal of the Department of Statistics, Malaysia*. 2: pp. 25–37.
- [8] Nazir, A. and et al. 2021. "Estimation and forecasting of rice yield using phenology-based algorithm and linear regression model on sentinel-ii satellite data," *Agriculture*. 11(10): pp. 1026.
- [9] Yu, C. 2024. "Walmart Sales Forecasting using Different Models," *Highlights in Science, Engineering and Technology*. 92: pp. 302–307.
- [10] Madhukumar, M. and et al. 2022. "Regression model-based short-term load forecasting for university campus load," *IEEE Access*. 10: pp. 8891–8905.
- [11] Lila, B. 2010. "Analysis and control of inventory," in *Production Planning and Control*. Bangkok, Thailand: Top Publishing Co. Ltd. pp.167-168.
- [12] Montgomery, D. C., Peck, E. A. and Vining, G. G. 2021. *Introduction to linear regression analysis*. [Online]. Available: <https://www.google.com/books?hl=en&lr=&id=tClgEAAAQBAJ&oi=fnd&pg=PR13&dq=Introduction+to+Linear+Regression+Analysis&ots=lgvbWtm3Mo&sig=Ac-2Xtp6Lh3IM7Stq-z0JYXuJXk>. Accessed: 12 April 2025.

- [13] Gujarati, D. N. and Porter, D. C. 2009.
Basic econometrics. [Online]. Available:
<https://thuvienso.hoasen.edu.vn/handle/123456789/8914>. Accessed: 12 April 2025.
- [14] Hasan, R. 2024. "Addressing Seasonality and Trend Detection in Predictive Sales Forecasting: A Machine Learning Perspective," *JBMS*. 6(2): pp. 100–109.
- [15] Taufiqih, R. and Ambarwati, R. 2024.
"Enhancing Sales Prediction for MSMEs: A Comparative Analysis of Neural Network and Linear Regression Algorithms," *JTMI*. 10(1): pp. 81–91.
- [16] Naresh Kumar Reddy, P. 2023.
"Forecasting E-Commerce Trends: Utilizing Linear Regression, Polynomial Regression, Random Forest, and Gradient Boosting for Accurate Sales and Demand Prediction," *International Journal of HRM and Organizational Behavior*. 11(3): pp. 11–26.
- [17] Lim, J., Azriani, N. and Nor, M. 2016.
"Simulation by Queuing System at Immigration Department," *International Academic Research Journal of Social Science*. 2(1): pp. 112–119.