

บทที่ 2

แนวคิด ทฤษฎี เครื่องมือและวรรณกรรมที่เกี่ยวข้อง

โครงการเรื่องการวิเคราะห์ข้อมูลเพื่อพยากรณ์มูลค่าการส่งออกสินค้าของวิสาหกิจขนาดกลางและขนาดย่อม โดยใช้เทคนิคการทำเหมืองข้อมูล ในบทนี้เป็นการนำเสนอเกี่ยวกับแนวคิด ทฤษฎี เครื่องมือ และวรรณกรรมที่เกี่ยวข้อง ของการวิเคราะห์ข้อมูลเพื่อพยากรณ์มูลค่าการส่งออกสินค้าของวิสาหกิจขนาดกลางและขนาดย่อม โดยใช้เทคนิคการทำเหมืองข้อมูล ซึ่งได้รวบรวมการศึกษาเอกสารงานวิจัยที่เกี่ยวข้อง เพื่อใช้เป็นแนวทางการศึกษา ประกอบด้วยรายละเอียดตามลำดับ ดังนี้

2.1 แนวคิด

2.1.1 แนวคิดเกี่ยวกับการวิเคราะห์ข้อมูล (Data Analytics)

2.1.2 แนวคิดเกี่ยวกับการประยุกต์ใช้ข้อมูลทางสถิติ

2.1.3 แนวคิดเกี่ยวกับการแสดงผลข้อมูล (Data Visualization)

2.2 ทฤษฎี

2.2.1 การทำเหมืองข้อมูล (Data Mining)

2.2.2 หลักในการทำเหมืองข้อมูล

2.2.3 การวิเคราะห์การถดถอย (Regression Analysis)

2.2.4 ทฤษฎีเกี่ยวข้องกับการสร้างเว็บไซต์

2.3 วรรณกรรมที่เกี่ยวข้อง

2.4 เครื่องมือในการออกแบบและวิเคราะห์ข้อมูล

2.4.1 กระบวนการวิเคราะห์ข้อมูลด้วย (CRISP-DM)

2.4.2 การวิเคราะห์การถดถอยเชิงเส้น (Linear Regression)

2.5 บทสรุป

2.1 แนวคิด

2.1.1 แนวคิดเกี่ยวกับการวิเคราะห์ข้อมูล (Data Analytics) ในการดำเนินงานเรื่องการวิเคราะห์ข้อมูลเพื่อพยากรณ์มูลค่าการส่งออกสินค้าของวิสาหกิจขนาดกลางและขนาดย่อม โดยใช้เทคนิคการทำเหมืองข้อมูล ผู้ศึกษาได้ศึกษาหลักการ และทฤษฎีต่าง ๆ ซึ่งองค์ประกอบที่สำคัญ คือ การวิเคราะห์ข้อมูล ซึ่งมีรายละเอียด ดังนี้

2.1.1.1 การวิเคราะห์ข้อมูล (Data Analytics) คือ การนำข้อมูลตั้งแต่อดีตจนถึงปัจจุบัน มาวิเคราะห์เพื่อคาดการณ์อนาคต และนำผลลัพธ์ไปใช้ในหลายด้าน ซึ่งครอบคลุมทุกอุตสาหกรรม ทั้งภาครัฐ และเอกชน โดยอาศัยเทคโนโลยี หรือซอฟต์แวร์เฉพาะด้าน เพื่อให้การประมวลผล และวิเคราะห์มีความแม่นยำรวมถึงสามารถนำไปใช้งานได้มีประสิทธิภาพ แบ่งออกเป็น 4 ประเภท ดังนี้

1) Diagnostic Analytics คือ การวิเคราะห์เชิงวินิจฉัย ซึ่งเป็นการวิเคราะห์เชิงลึกที่ต่อยอดจาก Descriptive Analytics เพื่อตอบคำถามว่าเหตุการณ์นั้น ๆ จึงเกิดขึ้นโดยใช้เทคนิคต่าง ๆ เช่น Data Discovery หรือ Data Mining เพื่อหาปัจจัย และตัวแปรที่ส่งผล ตัวอย่าง เช่น การวิเคราะห์สาเหตุที่ยอดขายเพิ่มขึ้น การเพิ่มขึ้นของผู้เข้าชมเว็บไซต์ การลดลงของจำนวนลูกค้าที่เข้าร้าน โปรโมชันที่ไม่ค่อยได้รับความนิยม หรือเนื้อหาโฆษณาที่มีอัตราการคลิกสูงกว่าปกติ นอกจากนี้ยังสามารถนำมาใช้วิเคราะห์การทำงานของคอมพิวเตอร์เพื่อตรวจจับความผิดปกติ ข้อมูลที่ได้จะช่วยให้องค์กรเข้าใจความต้องการของตลาด พฤติกรรมของลูกค้า และสาเหตุของปัญหาทางเทคโนโลยี รวมถึงสามารถปรับปรุงวัฒนธรรมองค์กร เพื่อเพิ่มประสิทธิภาพในการทำงาน

2) การวิเคราะห์ข้อมูลในปัจจุบันเป็นสิ่งสำคัญอย่างยิ่งสำหรับองค์กร หรือธุรกิจในการดำเนินงานเชิงกลยุทธ์และตัดสินใจอย่างมีข้อมูล การวิเคราะห์เชิงพยากรณ์ (Predictive Analytics) และการวิเคราะห์เชิงคาดการณ์ (Forecast Analytics) เป็นวิธีการวิเคราะห์ข้อมูลที่ถูกใช้งานเพื่อให้ข้อมูลเชิงลึกที่มีคุณค่าแก่ผู้บริหาร ซึ่งทั้งสองวิธีการ มีความแตกต่างกันในด้านของวัตถุประสงค์ และวิธีการใช้งาน ซึ่งสามารถแยกได้อย่างชัดเจนตามบริบทของการใช้งาน (P3 Adaptive, n.d.) Predictive Analytics คือ การวิเคราะห์ข้อมูลในอดีต และปัจจุบัน เพื่อตรวจสอบแนวโน้ม ทำนาย หรือพยากรณ์สิ่งที่จะเกิดขึ้นตามวัตถุประสงค์ที่กำหนด โดยใช้แบบจำลองทางสถิติร่วมกับเทคโนโลยีปัญญาประดิษฐ์ (AI) ซึ่งสามารถนำไปใช้ประโยชน์

ได้หลายด้าน เช่น การคาดการณ์ความเสี่ยง และโอกาส ยอดขาย ภัยไซเบอร์ สภาพอากาศ การลงทุน หุ้น หรือผลการเลือกตั้ง ทั้งนี้ความแม่นยำของ Predictive Analytics ขึ้นอยู่กับคุณภาพของข้อมูลที่องค์กรควรให้ความสำคัญเป็นอันดับแรก เพื่อให้ผลการวิเคราะห์มีประสิทธิภาพ ลดความผิดพลาด และสร้างประโยชน์ได้อย่างแท้จริง (Blendata, 2024) Forecast Analytics คือ การวิเคราะห์เชิงคาดการณ์ เป็นการวิเคราะห์เชิงพยากรณ์มุ่งเน้นไปที่การทำนายเหตุการณ์ที่เฉพาะเจาะจง การวิเคราะห์เชิงคาดการณ์จะมุ่งเน้นไปที่การคาดการณ์แนวโน้มของเหตุการณ์ที่เกิดขึ้นในระยะยาว โดยใช้ข้อมูลในอดีตเป็นฐานในการคาดการณ์อนาคต ซึ่งมักใช้กับตัวแปรที่เป็นค่าเชิงเวลา (time-series data) เช่น การคาดการณ์ยอดขายรายปี หรือการประเมินแนวโน้มการใช้สินค้าคงคลัง (P3 Adaptive, n.d.)

3) Predictive Analytics และ Forecast Analytics จะเป็นการใช้ข้อมูลในอดีตในการคาดการณ์อนาคต แต่ทั้งสองวิธีการมีวัตถุประสงค์และการใช้งานที่แตกต่างกันอย่างชัดเจน Predictive Analytics มักใช้ในการคาดการณ์พฤติกรรมของบุคคล หรือเหตุการณ์เฉพาะเจาะจง โดยเน้นการวิเคราะห์ในระดับจุลภาค (micro-level) ขณะที่ Forecast Analytics มุ่งเน้นไปที่แนวโน้มในอนาคตในระดับมหภาค (macro-level) และให้ภาพรวมที่กว้างขึ้นสำหรับการตัดสินใจในระยะยาว (One Model, n.d.)

4) Prescriptive Analytics คือ การวิเคราะห์เพื่อให้คำแนะนำ ซึ่งเป็นการวิเคราะห์ที่มีความซับซ้อนที่สุด โดยต่อเนื่องจากการทำ Predictive Analytics เมื่อได้ข้อมูลแนวโน้มที่จะเกิดขึ้น Prescriptive Analytics จะช่วยแนะนำแนวทางปฏิบัติที่เหมาะสมที่สุด รวมถึงวิเคราะห์ผลลัพธ์ที่อาจเกิดขึ้นจากการเลือกแนวทาง หรือแนะนำวิธีการแก้ไขปัญหา เครื่องมือนี้จึงมีความสำคัญอย่างยิ่งในการตัดสินใจที่ซับซ้อนด้วยข้อมูล Prescriptive Analytics ใช้ Big Data อัลกอริทึม Machine Learning และเทคโนโลยี AI ในการวิเคราะห์ข้อมูลจำนวนมาก และซับซ้อน ซึ่งเกินขีดความสามารถของมนุษย์ (Blendata, 2024)

2.1.2 แนวคิดเกี่ยวกับการประยุกต์ใช้ข้อมูลทางสถิติ

สถิติ (Statistic) คือ ศาสตร์ที่ศึกษาเกี่ยวกับข้อมูล การเก็บรวบรวม การวิเคราะห์ข้อมูล การแปลความหมาย และการนำเสนอข้อมูลข้อสรุปของการวิเคราะห์ โดยสถิติจะแสดงข้อมูลข้อเท็จจริงเกี่ยวกับด้านในด้านหนึ่ง โดยประเภทของแบ่งตามวิธีการดำเนินการกับข้อมูลเป็น 2 ประเภท ดังนี้

1) สถิติเชิงพรรณนา (Descriptive Statistic) คือ การอธิบาย บรรยาย หรือนำเสนอข้อมูลแผนภูมิของกลุ่มข้อมูลที่เก็บมาศึกษา เพื่อให้เห็นคุณลักษณะ หรือเข้าใจในกลุ่มนั้น ๆ เช่น ค่าเฉลี่ย (Mean) คือ การหาค่าเฉลี่ย จะใช้กับข้อมูลที่มีมาตรวัดระดับอัตราภาคขึ้นไปเท่านั้น หรือข้อมูลที่เป็นตัวเลขที่ได้จากการวัด การนับ ค่ามัธยฐาน (Median) คือ ค่าที่อยู่กึ่งกลางของชุดข้อมูลเมื่อนำข้อมูลทั้งหมดมาเรียงลำดับแล้ว ดังนั้นจะมีจำนวนข้อมูลครึ่งหนึ่งที่น้อยกว่าค่ามัธยฐาน และอีกครึ่งหนึ่งที่มีค่ามากกว่ามัธยฐาน และค่าฐานนิยม (Mode) คือ ค่าของข้อมูลที่มีจำนวนครั้งที่มากที่สุด หรือมีความถี่สูงสุด ฯลฯ

2) สถิติเชิงอนุมาน (Inferential Statistic) คือ การศึกษาลักษณะข้อมูลในกลุ่มตัวอย่าง (Sample) แล้วนำคุณลักษณะที่พบนั้นไปอธิบาย หรือสรุปอ้างอิงลักษณะของประชากร (Population) (กอบเกียรติ สระอุบล, 2563)

สถิติเป็นศาสตร์ที่ใช้เป็นเครื่องมือช่วยในการตัดสินใจอย่างมีเหตุผล การแก้ปัญหาส่วนใหญ่มีความจำเป็นต้องใช้ข้อมูลสารสนเทศ และกระบวนการทางสถิติมาช่วยในการสรุปผลปัญหาหรืองานต่าง ๆ ในชีวิตประจำวันทั้งในธุรกิจ และส่วนราชการต้องใช้สถิติมาช่วยในการตัดสินใจ รวมถึงการกำหนดยุทธศาสตร์ ซึ่งมีตัวอย่างการใช้ข้อมูลสถิติสำหรับการพัฒนาในด้านต่าง ๆ ที่สำคัญในภาครัฐ ดังนี้

1) ด้านการศึกษา ในการกำหนดนโยบาย และการวางแผนพัฒนาการศึกษา และการกระจายโอกาสทางการศึกษาของประชาชนในระดับการศึกษาต่าง ๆ ข้อมูลสำคัญที่ต้องการใช้ ได้แก่ ประชากรก่อนวัยเรียน และวัยเรียน บุคลากรทางการศึกษา ปริมาณการผลิตและพัฒนาครูในแต่ละสาขา จำนวนสถานศึกษา ค่าใช้จ่ายในแต่ละระดับการศึกษา เป็นต้น

2) ด้านการเกษตร ในการกำหนดนโยบาย และวางแผนพัฒนาทางการเกษตรของประเทศ ข้อมูลที่ต้องการใช้ ได้แก่ ครัวเรือนที่ทำกรเกษตร เนื้อที่การเพาะปลูก ผลผลิตทางการเกษตร จำนวนปศุสัตว์ ราคา สินค้าเกษตรกรรม และเครื่องมือเครื่องใช้ทางการเกษตร เป็นต้น

3) ด้านอุตสาหกรรม ใช้จัดทำแผนงาน หรือกำหนดนโยบาย และส่งเสริมอุตสาหกรรม ส่งเสริมการลงทุน และพัฒนาเทคโนโลยีทางด้านอุตสาหกรรม ได้แก่ ข้อมูลเกี่ยวกับปริมาณการผลิตทางอุตสาหกรรม ต้นทุนการผลิต จำนวนแรงงาน และมูลค่าเพิ่ม เป็นต้น

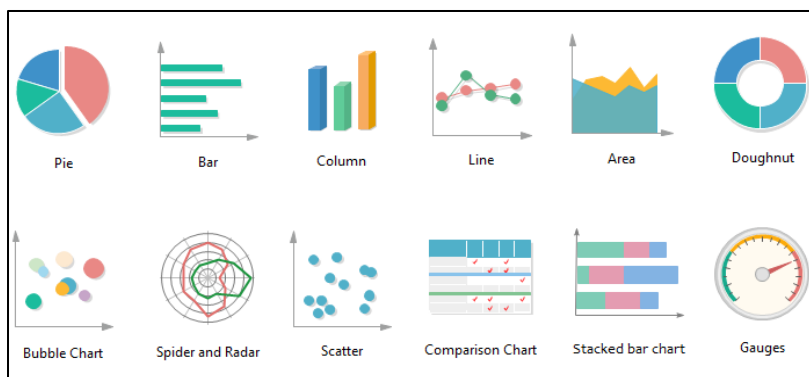
4) ด้านคมนาคม และขนส่ง การปรับปรุงบริการ และพัฒนาทางการคมนาคมขนส่ง และการสื่อสารของประเทศ เพื่อสนับสนุนการพัฒนาในด้านต่าง ๆ และกระจายความเจริญ

ไปสู่ภูมิภาค ข้อมูลที่ใช้ ได้แก่ ปริมาณผู้ใช้บริการในแต่ละเส้นทาง ปริมาณการขนส่งทางถนน ทางน้ำ และทางอากาศ รายละเอียดเส้นทางคมนาคม ข้อมูลเกี่ยวกับการจัดสรรความถี่วิทยุ จำนวนครัวเรือนที่มีเครื่องรับวิทยุ โทรทัศน์ เป็นต้น (กรมวิชาการเกษตร, 2563)

2.1.3 แนวคิดเกี่ยวกับการแสดงผลข้อมูล (Data Visualization)

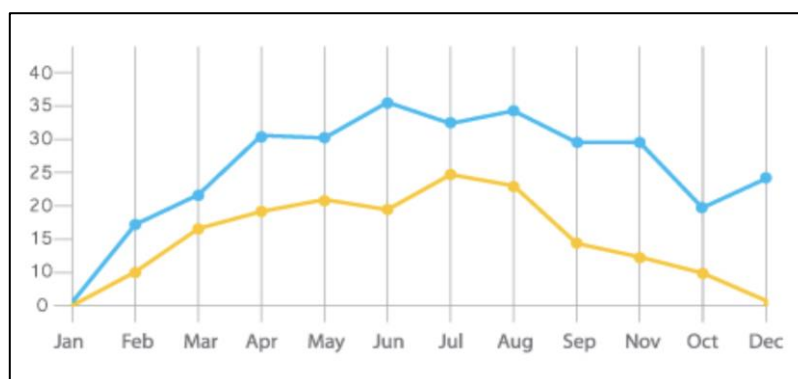
Data Visualization เป็นการนำข้อมูลมาแสดงในรูปแบบของ ภาพกราฟิก แผนภูมิ กราฟ ไดอะแกรม หรืออื่น ๆ (Charts Graphs Diagrams หรือ Maps) เพื่อให้เราเห็น และเข้าใจลักษณะ แนวโน้ม และความสัมพันธ์ของข้อมูล (Patterns trends และ correlations) ได้ง่ายขึ้น ชัดเจนขึ้น เนื่องจากโดยทั่วไปแล้ว ข้อมูลที่เราสนใจมักจะอยู่ในรูปแบบข้อความ อักขระ หรือตัวเลข ซึ่งก็มีอยู่เป็นจำนวนมาก การศึกษาหรือการวิเคราะห์โดยพิจารณา หรือดูจากข้อมูลดังกล่าวตรง ๆ หรือดูจากตารางที่มีข้อมูลจำนวนมาก ๆ จะทำได้ค่อนข้างยาก แต่การใช้ Data Visualization จะช่วยให้การศึกษา และวิเคราะห์ทำได้ง่ายขึ้น เนื่องจากนอกจากจะเห็นภาพรวมของข้อมูลทั้งหมด สามารถสั่งให้แสดงเน้นข้อมูลบางส่วนที่สนใจ แยกสีแยกกลุ่ม ทำให้เข้าใจถึงลักษณะ แนวโน้ม การเปรียบเทียบข้อแตกต่าง และความสัมพันธ์ของข้อมูลได้ง่ายขึ้น ดังคำกล่าวที่ว่า "ภาพหนึ่งภาพ แทนคำได้นับพันคำ" การทำ Data Visualization จะต้องพิจารณาถึงวัตถุประสงค์ ว่าต้องการค้นหา หรือต้องการดูข้อมูลสรุปอะไร เพื่อที่จะเลือกชนิดของกราฟที่เหมาะสม เช่น ต้องการดูจำนวน ความถี่ ใช้กราฟแท่ง Histogram ต้องการดูการกระจาย ต้องการดูความสัมพันธ์ หรือแนวโน้มของข้อมูล ใช้ Scatter และยังคงคำนึงถึงด้วยว่า ข้อมูลที่นำมาพล็อตเป็นข้อมูลชนิดใด ชนิดหนึ่ง ต่อเนื่อง หรือแบบ Category ฯลฯ (กอบเกียรติ สระอุบล, 2563)

1) แผนภูมิ (Charts) เป็นรูปแบบที่นิยมมากที่สุด และเป็นรูปแบบที่มีหลากหลายชนิดที่เหมาะสมกับการนำเสนอข้อมูลที่แตกต่างกันไปตามวัตถุประสงค์ เช่น Pie Chart จะช่วยให้เราเห็นปริมาณความแตกต่างได้ชัดเจน Comparison Chart เหมาะสำหรับการเปรียบเทียบ คุณสมบัติหลาย ๆ ข้อมูลมาตรวัด (Gauges) จะช่วยให้เห็นความเข้มข้น ความรุนแรง หรือน้ำหนัก



ภาพที่ 2.1 ตัวอย่างแผนภูมิรูปแบบต่าง ๆ
(ที่มา: medium.com/@Lynia_Li/)

2) กราฟ (Graphs) คือ Subset หรือประเภทหนึ่งของแผนภูมิ โดยกราฟจะทำหน้าที่แสดงความสัมพันธ์ระหว่างข้อมูล 2 ตัวแปร ผ่านแกนแนวนอน (แกน X) และแกนแนวตั้ง (แกน Y) ช่วยให้แนวโน้มสถานการณ์ประกอบกับบริบทได้เป็นอย่างดี



ภาพที่ 2.2 ตัวอย่างกราฟเส้นตัวอย่าง
(ที่มา: <https://www.def2design.com/Mkt-Corner/81>)

3) ตาราง (Tables) ก็เป็นอีกรูปแบบที่ใช้กันมากเพื่อนำเสนอข้อมูลให้ออกมาได้ง่าย ตารางประกอบไปด้วย 2 ส่วน ได้แก่ คอลัมน์ และแถว ซึ่งช่วยจัดการข้อมูลให้เรียบร้อย ช่วยให้มองเห็นบริบท และความสัมพันธ์ของข้อมูลหลาย ๆ ชุดได้อย่างง่าย

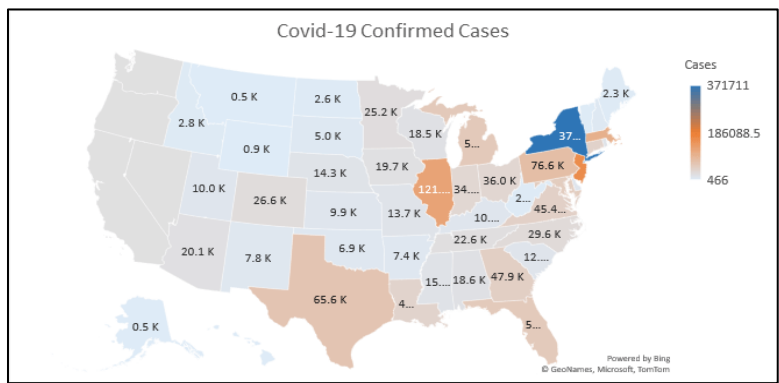
ม.ก. - เม.ย. 55				ม.ก. - เม.ย. 56		การเติบโต (%)
รายการ						
บัญชี	อุตสาหกรรม	กลุ่ม	รวม	รวม		
ตัวแทน	4,717,862	408,911	147,896	5,274,669	4,779,573	10.36
ผู้บริหาร	3,551,398	41,807	24,578	8,862,582	2,842,459	27.28
รวม(ทุกรายการ)	8,248,163	450,718	163,700	8,862,581	7,582,236	16.89

หมายเหตุ : กำหนด 1 หน่วย : 1,000 บาท

ภาพที่ 2.3 ตัวอย่างตารางแสดงข้อมูล

(ที่มา: <https://ms1--q2--57.blogspot.com/p/week7.html>)

4) แผนที่ (Maps) เป็นการนำเสนอข้อมูลบนแผนที่เพื่อแสดงข้อมูลเกี่ยวกับพื้นที่ต่าง ๆ ยกตัวอย่างเช่น การนำเสนอข้อมูลยอดผู้ติดเชื้อ Covid-19 ในแต่ละรัฐของประเทศสหรัฐอเมริกา ซึ่งนอกจากการใส่ข้อมูลลงไปยังพื้นที่ต่าง ๆ แล้ว ยังสามารถใช้สีสีนเพื่อบอกช่วงปริมาณ หรือความหนาแน่นของผู้ติดเชื้ออีกด้วย

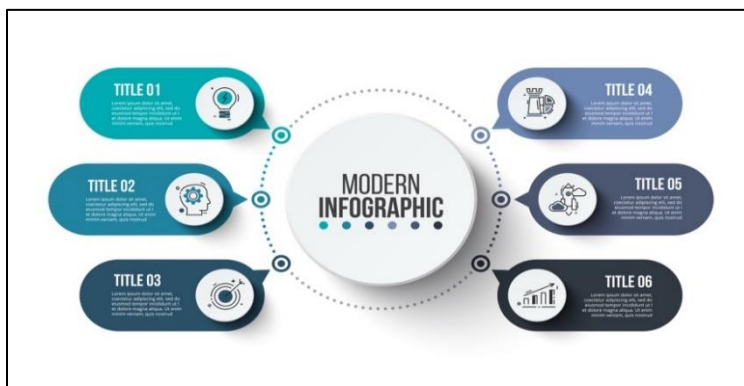


ภาพที่ 2.4 ตัวอย่างแผนที่

(ที่มา: spreadsheetweb.com)

5) อินโฟกราฟิก (Infographic) คือ การนำเสนอสารสนเทศ (Info: information) ด้วยภาพกราฟิก (Graphic) เป็นรูปแบบการนำเสนอข้อมูลที่ใช้ภาพสื่อแทน ทำให้ผู้อ่านข้อมูลเข้าใจข้อมูลได้ง่ายหรือสามารถทำความเข้าใจผ่านภาพแทนที่คุ้นเคย นอกจากนี้ อินโฟกราฟิกยังเป็น

รูปแบบการนำเสนอข้อมูลที่น่าสนใจ มีการนำเทคนิคการเล่าเรื่อง (Storytelling) มาใช้ทำให้ข้อมูลน่าสนใจ น่าดึงดูด จึงมักจะใช้เพื่อนำเสนอเนื้อหา ความรู้ หรือเป็นสื่อการเรียนการสอน



ภาพที่ 2.5 ตัวอย่างอินโฟกราฟิก

(ที่มา: <https://www.datawow.co.th/blogs/what-is-data-visualization>)

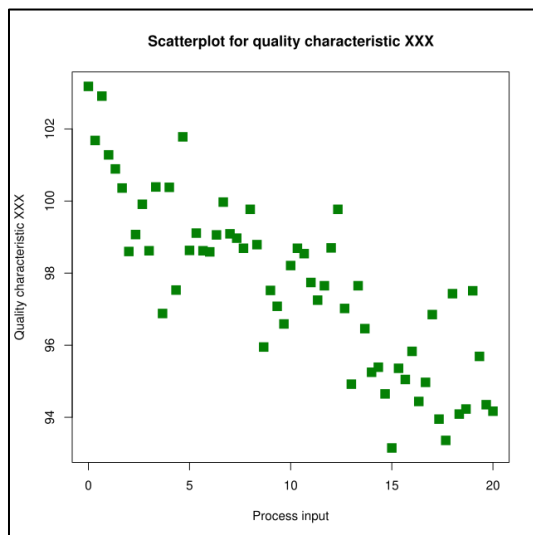
6) แดชบอร์ด (Dashboards) คือ การนำข้อมูลต่าง ๆ มาเรียบเรียง และสรุปเป็นภาพ โดยใช้แผนภูมิและกราฟต่าง ๆ มาใช้นำเสนอ ปัจจุบันแดชบอร์ดเป็น Data Visualization ที่นิยมใช้กับการนำเสนอข้อมูลแบบ Real-time ผ่านซอฟต์แวร์ หรือเครื่องมือจัดการ และวิเคราะห์ข้อมูลต่าง ๆ เช่น เครื่องมือการตลาด เครื่องมือบริหารจัดการข้อมูล เครื่องมือติดตาม และดูแลเว็บไซต์ ฯลฯ (เพ็รสรภาพท์ ดิจิทัล โซลูชั่น, 2563)



ภาพที่ 2.6 ตัวอย่างแดชบอร์ด

(ที่มา: <https://www.datawow.co.th/blogs/what-is-data-visualization>)

7) แผนภูมิจุดแบบกระจัดกระจาย หรือ Scatter plot คือ การสร้างกราฟเพื่อแสดงความสัมพันธ์ระหว่าง 2 ค่าในแนวแกน X และแนวแกน Y (กอบเกียรติ สระอุบล, 2563)



ภาพที่ 2.7 ตัวอย่างแผนภูมิจุดแบบกระจัดกระจาย
(ที่มา: https://en.wikipedia.org/wiki/Scatter_plot)

2.2 ทฤษฎี

2.2.1 การทำเหมืองข้อมูล (Data Mining)

การทำเหมืองข้อมูล (Data Mining) คือ การค้นหาข้อมูลที่มีประโยชน์จากแหล่งข้อมูลที่มีเป็นจำนวนมาก เพื่อดึงข้อมูลที่มีประโยชน์มาทำการวิเคราะห์ ค้นหารูปแบบ หรือความสัมพันธ์ที่เกิดในฐานข้อมูล และจัดทำเป็นสารสนเทศเพื่อใช้ในการวางแผนบริหารจัดการธุรกิจ โดยการแยกข้อมูลที่มีประโยชน์ออกมาใช้งานเปรียบเทียบกับการทำเหมืองแร่ที่จะต้องทำการแยกเศษหินดินทรายที่ไม่มีค่า และมีปริมาณมากออกจากแร่ที่มีมูลค่ามาก และมักจะมีปริมาณน้อย (กรมส่งเสริมอุตสาหกรรม, 2562)

2.2.2 หลักในการทำเหมืองข้อมูล

เหมืองข้อมูลเป็นการวิเคราะห์ความสัมพันธ์ของข้อมูลจากฐานข้อมูลขนาดใหญ่ในอดีต เพื่อการค้นหากฎความสัมพันธ์ของข้อมูล การจัดกลุ่มข้อมูล และจำแนกประเภทข้อมูล เพื่อพยากรณ์สิ่งที่จะเกิดขึ้นในอนาคตโดยมีรายละเอียด ดังต่อไปนี้

1) การค้นหากฎความสัมพันธ์ของข้อมูล (Association Rule) เป็นเทคนิคที่ใช้ในการค้นหารูปแบบกฎความสัมพันธ์ที่สนใจ หรือความสัมพันธ์ที่มีความถี่ร่วมกันบ่อยที่เกิดขึ้นในฐานข้อมูลขนาดใหญ่ที่มีค่าสนับสนุน (Support) และค่าความเชื่อมั่น (Confidence) ที่ผ่านเกณฑ์ขั้นต่ำ (Minimum Threshold) มีวัตถุประสงค์เพื่อวิเคราะห์ตะกร้าทางการตลาด (Marketing basket analysis) จากพฤติกรรมบันทึกการซื้อสินค้า (Transaction Data) ของลูกค้าเพื่อนำผลลัพธ์ที่ได้มาใช้ในการจัดวางชั้นสินค้า หรือจัดโปรโมชั่นส่งเสริมการขาย

2) การจัดกลุ่มข้อมูล (Clustering) เป็นเทคนิคที่ใช้ในการลดขนาดของข้อมูลด้วยวิธีการจัดกลุ่มข้อมูลที่มีลักษณะคล้ายคลึงกันให้อยู่ในกลุ่มเดียวกัน จำนวน k กลุ่ม มีวัตถุประสงค์สำคัญ 2 ประการ คือ

- การจัดกลุ่มหน่วยวิเคราะห์ เช่น การวิเคราะห์รูปแบบและพฤติกรรมกรรมการซื้อของลูกค้าในแต่ละกลุ่มได้

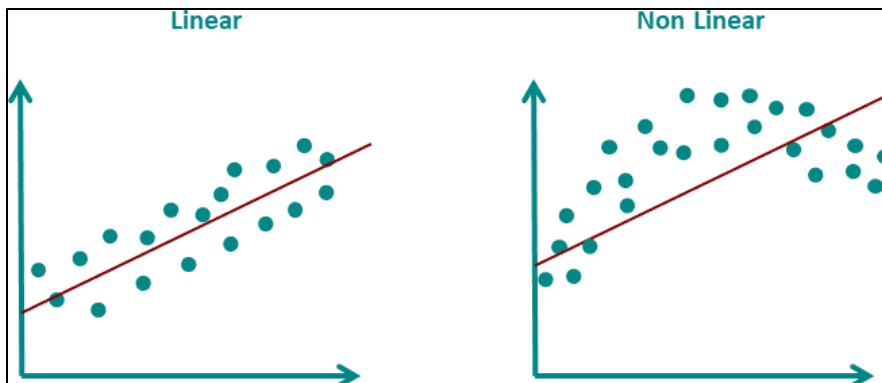
- การจัดกลุ่มตัวแปรลักษณะเฉพาะของลูกค้า เช่น เพศ อาชีพ รายได้ เป็นต้น

3) การจำแนกประเภทข้อมูล (Classification) เป็นเทคนิคที่ใช้ในการจำแนกประเภทข้อมูลด้วยคุณลักษณะต่าง ๆ ที่มีการกำหนดค่าตอบไว้แล้ว (อรรถพล อุดยุตศาสน์, 2565)

2.2.3 การวิเคราะห์การถดถอย (Regression Analysis)

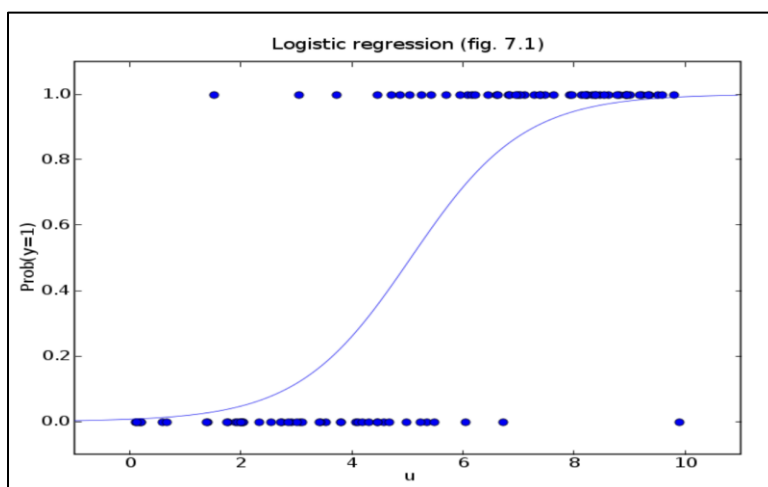
คือ เทคนิคทางสถิติที่ใช้ในการประมาณค่าความสัมพันธ์ระหว่างตัวแปร โดยกำหนดให้ตัวแปรหนึ่งเป็นตัวแปรที่ทราบค่า เรียกว่าตัวแปรอิสระ (Independent variable : X) ในขณะที่อีกตัวแปรหนึ่งเป็นตัวแปรที่ต้องการประมาณค่าเรียกว่าตัวแปรตาม (Dependent variable: Y) ครั้งนี้ประเภทของการวิเคราะห์การถดถอยได้ถูกสรุปไว้โดย (เฉลิมพล จตุพร, 2562) และ (National Institute of Standards and Technology : NIST, n.d.) ไว้ดังนี้

1) การวิเคราะห์การถดถอยเชิงเส้น (Linear Regression) เป็นการหาความสัมพันธ์ระหว่างตัวแปรตาม (Dependent Variable) กับตัวแปรอิสระ (Independent Variable) โดยสมมติว่าความสัมพันธ์ระหว่างตัวแปรทั้งสองเป็นเส้นตรง



ภาพที่ 2.8 ตัวอย่างกราฟแสดงความสัมพันธ์เชิงเส้น
(ที่มา: <https://datatab.net/tutorial/linear-regression>)

2) การวิเคราะห์การถดถอยโลจิสติก (Logistic Regression) ใช้ในการวิเคราะห์ข้อมูลที่มีผลลัพธ์เป็นค่าจำกัด (เช่น ใช่/ไม่ใช่, ผ่าน/ไม่ผ่าน) โดยการหาความสัมพันธ์ระหว่างตัวแปรตามที่เป็นค่าจำกัดกับตัวแปรอิสระ



ภาพที่ 2.9 ความสัมพันธ์ของตัวแปรตาม และตัวแปรทำนายในการวิเคราะห์โลจิสติก
(ที่มา: <https://medium.com/@johnmccool>)

3) การวิเคราะห์การถดถอยพหุคูณ (Multiple Regression) เป็นการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรตามหนึ่งตัวกับตัวแปรอิสระหลายตัวเพื่อให้ได้สมการที่สามารถทำนายค่าของตัวแปรตามได้

4) การวิเคราะห์การถดถอยพหุระดับ (Multilevel Regression) ใช้เมื่อข้อมูลมีโครงสร้างเป็นลำดับชั้น เช่น ข้อมูลนักเรียนในโรงเรียนหลายแห่ง โดยพิจารณาความสัมพันธ์ที่ซับซ้อนมากขึ้นระหว่างตัวแปรตามกับตัวแปรอิสระในแต่ละระดับ

5) การวิเคราะห์การถดถอยไร้เชิงเส้น (Nonlinear Regression) ใช้ในการวิเคราะห์ข้อมูลที่มีความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระไม่เป็นเส้นตรง โดยใช้สมการไร้เชิงเส้น

6) การวิเคราะห์การถดถอยสเปียร์แมน (Spearman Regression) ใช้ในการหาความสัมพันธ์ระหว่างตัวแปรตามที่เป็นอันดับกับตัวแปรอิสระ โดยใช้ค่าสถิติสเปียร์แมนแทนการวิเคราะห์เชิงเส้นตรง

2.2.4 ทฤษฎีเกี่ยวข้องกับการสร้างเว็บไซต์

2.2.4.1 การออกแบบเว็บไซต์ หรือ Web Design คือ กระบวนการในการสร้างหน้าตาและองค์ประกอบภายในเว็บไซต์ ตั้งแต่การจัดวางเนื้อหา การเลือกใช้สี ฟอนต์ การเลือกใช้ภาพหรือกราฟิก และการออกแบบระบบนำทางผู้ใช้งานหน้าเว็บ ให้สวยงามสอดคล้อง และสามารถใช้งานได้อย่างราบรื่น โดยใช้หลักการของ Responsive Web Design เข้ามาช่วยในการออกแบบเพื่อให้สามารถแสดงผลได้ทุก ๆ อุปกรณ์ ซึ่งจะใช้ URL ร่วมกัน แต่การแสดงผลในแต่ละอุปกรณ์จะแตกต่างกันไป การตอบสนองเป็นการปฏิบัติของการสร้างเว็บไซต์ที่เหมาะสมเพื่อทำงานบนทุกอุปกรณ์ และทุกขนาดหน้าจอทั้งขนาดใหญ่ หรือขนาดเล็ก เช่น มือถือ หรือเดสก์ท็อป ฯลฯ การตอบสนองจะเน้นการใช้งานง่าย จึงต้องมีการออกแบบอย่างเป็นระบบ

เว็บไซต์ที่ประสบความสำเร็จ จะต้องสอดคล้องกับแบรนด์ และเป้าหมายทางธุรกิจ สามารถสื่อสารคุณค่าของสินค้ากับบริการได้อย่างโดดเด่น ขณะเดียวกันก็ต้องเข้าใจความต้องการของกลุ่มเป้าหมาย เพื่อมอบประสบการณ์ใช้งานที่เหมาะสมกับผู้ใช้จริง (ดวงพร เกียรติคำ, 2560)

2.2.4.2 องค์ประกอบในการออกแบบเว็บไซต์ การออกแบบเว็บไซต์ที่มีประสิทธิภาพนั้นต้องคำนึงถึงองค์ประกอบสำคัญ ดังต่อไปนี้

1) การออกแบบที่เรียบง่าย (Simplicity) ในยุคที่ผู้คนมีความสนใจในช่วงเวลาสั้นลง (Short attention span) และต้องการข้อมูลที่กระชับ และตรงประเด็น การออกแบบเว็บไซต์แบบมินิมอล (Minimalist) ซึ่งเน้นความเรียบง่ายจึงเป็นแนวคิดที่ควรนำมาใช้ เพื่อให้ผู้ใช้งานสามารถมุ่งเน้นประเด็นสำคัญได้อย่างชัดเจน

ทั้งนี้ ความเรียบง่ายไม่ควรถึงขั้นที่ทำให้ขาดจุดเด่นที่น่าสนใจ เพราะการออกแบบหน้าเว็บไซต์ที่ซับซ้อนเกินไปอาจทำให้ผู้ใช้งานเกิดความสับสน และไม่สามารถค้นหา

ข้อมูลที่ต้องการได้อย่างสะดวก ดังนั้น ควรใช้กราฟิกและข้อความในปริมาณที่พอเหมาะ เน้นที่ความสะอาดตา และการใช้งานที่ราบรื่น โดยยังสามารถสื่อสารข้อมูลสำคัญของหน้าเว็บไซต์ ได้อย่างมีประสิทธิภาพ

2) การเลือกใช้สีที่เข้ากัน (Color Scheme) การเลือกใช้สีเป็นองค์ประกอบ สำคัญในการออกแบบเว็บไซต์ (Web Design) เนื่องจากสีสามารถสื่อถึงอารมณ์ และความรู้สึก รวมทั้งส่งผลต่อการรับรู้ของผู้ใช้งาน หากเลือกใช้สีที่เหมาะสม และสอดคล้องกับลักษณะของ เว็บไซต์ ย่อมสามารถสร้างความโดดเด่น และเอกลักษณ์ได้ดี การใช้สีไม่ควรเกิน 1 – 2 สีหลัก และสีรอง 2 – 3 สี เพื่อให้ดูเป็นระเบียบน่าสนใจ

นอกจากนี้ การเลือกสีควรคำนึงถึงความคมชัดในการแสดงผล ไม่ควรใช้ สีที่ฉูดฉาดเกินไปเพราะอาจทำให้การอ่านเนื้อหาลำบาก ควรเลือกใช้คู่สีที่เป็น Complementary colors หรือใช้ Color palette ที่มีโทนสีใกล้เคียงกัน เพื่อให้เว็บไซต์ดูสวยงามและน่าสบายตา

3) การเลือกฟอนต์ที่อ่านง่าย (Readable Fonts) ฟอนต์ และการจัดวาง ข้อความเป็นองค์ประกอบสำคัญในการออกแบบหน้าเว็บไซต์ (Web Page Design) เนื่องจากมีผล ต่อความสะดวกในการอ่าน ควรเลือกใช้ฟอนต์ที่อ่านง่าย และไม่ซับซ้อน เช่น ฟอนต์ Sans-serif อย่าง Arial Verdana Helvetica สำหรับเนื้อหาทั่วไป หรือฟอนต์ Serif เช่น Times New Roman Georgia ที่เหมาะกับเนื้อหาบทความยาว

ขนาดของฟอนต์ (Font size) และการเว้นระยะห่างระหว่างบรรทัด (Line-Height) ที่เหมาะสมจะช่วยเพิ่มความสะดวกในการอ่าน และนำเสนอข้อมูลได้อย่างมีประสิทธิภาพ

4) การใช้พื้นที่ว่างอย่างเหมาะสม (White Space) หรือพื้นที่ว่าง เป็นส่วนที่ปราศจากข้อความหรือกราฟิก ซึ่งมักใช้เป็นพื้นหลังสีขาว หรือสีอ่อน เพื่อแบ่งแยก เนื้อหาให้เป็นสัดส่วน สร้างความเป็นระเบียบเรียบร้อย และช่วยให้ผู้ใช้เข้าใจเนื้อหาได้ง่ายขึ้น สีพื้น หลังไม่จำเป็นต้องเป็นสีขาว อาจใช้สีอ่อนหรือเส้นที่ไม่ทำให้ดูรกตา สิ่งสำคัญ คือ ต้องใช้ White Space อย่างพอเหมาะ ไม่มาก หรือน้อยเกินไป

5) การออกแบบเว็บไซต์ให้รองรับการแสดงผลบนมือถือ ด้วยการเพิ่มขึ้นของ ผู้ใช้งานอินเทอร์เน็ตผ่านสมาร์ทโฟน และแท็บเล็ต การออกแบบเว็บไซต์ที่รองรับการแสดงผล บนอุปกรณ์พกพา หรือ Responsive Design จึงกลายเป็นปัจจัยสำคัญ Responsive Design คือ

การออกแบบเว็บไซต์ให้ปรับขนาด และการจัดเรียงเนื้อหาได้โดยอัตโนมัติ ไม่ว่าจะเป็นคอมพิวเตอร์ สมาร์ทโฟน หรือแท็บเล็ต ทำให้ผู้ใช้สามารถใช้งานได้อย่างราบรื่น

นักออกแบบต้องคำนึงถึงการแสดงผลทั้งแนวตั้ง (Portrait) และแนวนอน (Landscape) รวมถึงขนาดของปุ่ม และลิงก์ที่ต้องใหญ่พอเพื่อให้กดได้สะดวกบนหน้าจอสัมผัส

6) การใช้ Visual Content ในการตกแต่งเว็บไซต์ Visual Content เช่น รูปภาพ ไอคอน อินโฟกราฟิก และวิดีโอ เป็นองค์ประกอบที่ช่วยเสริมสร้างความน่าสนใจให้กับเว็บไซต์ นอกจากการตกแต่งให้สวยงาม ยังช่วยดึงดูดความสนใจ และอธิบายเนื้อหาที่ยากให้เข้าใจง่ายขึ้น การใช้ Visual Content ควรอยู่ในปริมาณที่เหมาะสม และต้องคำนึงถึงขนาดไฟล์เพื่อไม่ให้กระทบต่อความเร็วในการโหลดเว็บไซต์

7) การวางตำแหน่งเนื้อหาหลักให้เด่นชัด (Highlight Key Content) การวางตำแหน่งเนื้อหาที่ต้องการเน้นให้เห็นได้ชัดเจนเป็นสิ่งสำคัญ เพราะผู้ใช้งานให้ความสนใจกับสิ่งที่เห็นชัดเจนที่สุดก่อน เช่น ข้อความขนาดใหญ่ หรือรูปภาพเด่นชัด ควรใช้หลักการจัดลำดับความสำคัญขององค์ประกอบ (Visual Hierarchy) เพื่อทำให้เนื้อหาสำคัญโดดเด่น และสะดุดตามากขึ้น

8) การออกแบบระบบนำทางเว็บไซต์ (Navigation) การออกแบบระบบนำทางหรือเมนูที่ใช้งานง่ายเป็นหัวใจสำคัญ เพราะช่วยให้ผู้ใช้งานสามารถค้นหา และเข้าถึงข้อมูลได้อย่างรวดเร็ว ควรวางโครงสร้างเมนูให้ชัดเจน และเป็นหมวดหมู่ การใช้ข้อความในเมนูควรกระชับ และเข้าใจง่าย

9) การใส่ Call-to-Action (CTA) เป็นองค์ประกอบที่กระตุ้นให้ผู้ใช้ดำเนินการตามที่เว็บไซต์ต้องการ เช่น การกดสั่งซื้อ หรือการสมัครสมาชิก ปุ่ม CTA ควรมีขนาดใหญ่พอที่จะสังเกตเห็นได้ชัดเจน และข้อความควรกระชับ ชัดเจน เช่น "สั่งซื้อทันที" หรือ "ดาวน์โหลดฟรี"

10) การทดสอบ และปรับปรุงการออกแบบเว็บไซต์อยู่เสมอ การออกแบบเว็บไซต์เป็นกระบวนการที่ต้องมีการทดสอบ และปรับปรุงอย่างต่อเนื่อง เพื่อให้สอดคล้องกับเทคโนโลยีที่เปลี่ยนแปลง และพฤติกรรมของผู้ใช้ การใช้เครื่องมืออย่าง Google Analytics ช่วยให้เห็นพฤติกรรมการใช้งาน และสามารถปรับปรุงเว็บไซต์ให้ตอบโจทย์ความต้องการได้อย่างมีประสิทธิภาพ (ฟอกซ์บิช, 2567)

2.2.4.3 รูปแบบโครงสร้างของเว็บไซต์ การออกแบบโครงสร้างของเว็บไซต์ สามารถทำได้หลากหลายแบบ ซึ่งก็ขึ้นอยู่กับความชอบ และความถนัดของแต่ละบุคคล นอกจากนี้ยังขึ้น

อยู่กับกลุ่มเป้าหมายที่ต้องการนำเสนอ เพราะจะต้องออกแบบให้เหมาะสมกับการใช้งานของกลุ่มเป้าหมายมากที่สุด (วันปีลีฟ, 2560) โดยโครงสร้างของเว็บไซต์ส่วนใหญ่ก็จะประกอบไปด้วย 4 รูปแบบดังนี้

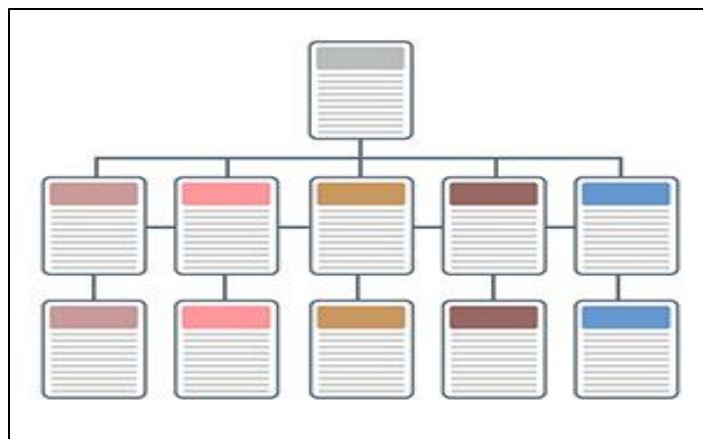
1) โครงสร้างเว็บไซต์แบบเรียงลำดับ จะเป็นโครงสร้างแบบธรรมดาที่นิยมใช้งานกันมากที่สุด เนื่องจากมีความง่ายต่อการจัดระบบข้อมูล และสามารถนำเสนอเรื่องราวตามลำดับได้เป็นอย่างดี เหมาะกับเว็บไซต์ที่มีขนาดเล็ก มีเนื้อหาที่ไม่ซับซ้อน ส่วนใหญ่ก็จะเป็นพวกเว็บไซต์ที่ให้ความรู้ หรือเว็บไซต์องค์กรขนาดย่อม โดยลักษณะการลิ้งค์เนื้อหา ก็จะลิ้งค์ไปที่ละหน้า มีทิศทาง การเข้าสู่เนื้อหาต่าง ๆ ในแบบเส้นตรง ใช้ปุ่มเดินหน้า - ถอยหลังในการกำหนดทิศทาง จึงทำให้การใช้งานเป็นไปอย่างง่าย แต่โครงสร้างเว็บไซต์แบบเรียงลำดับมีข้อเสียคือจะทำให้ผู้ใช้งานต้องเสียเวลาในการเข้าสู่เนื้อหาเพราะไม่สามารถกำหนดทิศทาง การเข้าสู่เนื้อหาด้วยตัวเองได้



ภาพที่ 2.10 แสดงโครงสร้างแบบเรียงลำดับ

(ที่มา: <https://asria.org/website-structure/>)

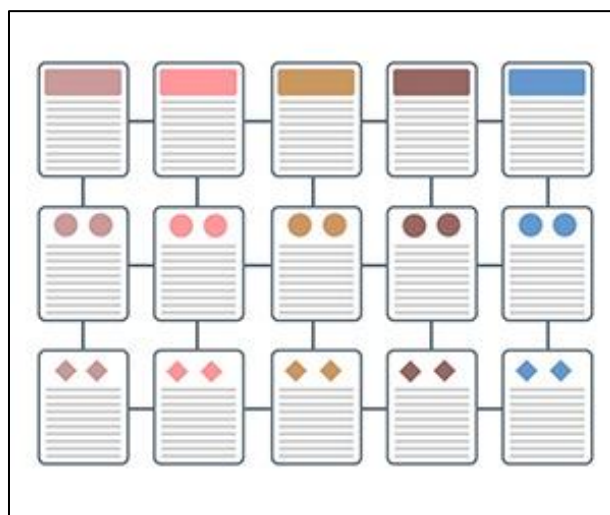
2) โครงสร้างแบบลำดับชั้น นิยมใช้กับเว็บที่มีความซับซ้อนของข้อมูล เพื่อให้สามารถเข้าถึงข้อมูลต่าง ๆ ได้ง่ายขึ้น โดยจะมีการแบ่งเนื้อหาออกเป็น ส่วน ๆ และมีการนำเสนอรายละเอียดค่อย ๆ ที่ลดหลั่นกันมา ทำให้สามารถทำความเข้าใจกับโครงสร้างเนื้อหาได้ง่ายขึ้น โดยจะมีโฮมเพจเป็นจุดเริ่มต้น และจุดรวมจุดเดียวที่จะนำไปสู่การเชื่อมโยงเนื้อหาเป็นลำดับจากบนลงล่าง



ภาพที่ 2.11 แสดงโครงสร้างแบบลำดับขั้น

(ที่มา: <https://asria.org/website-structure/>)

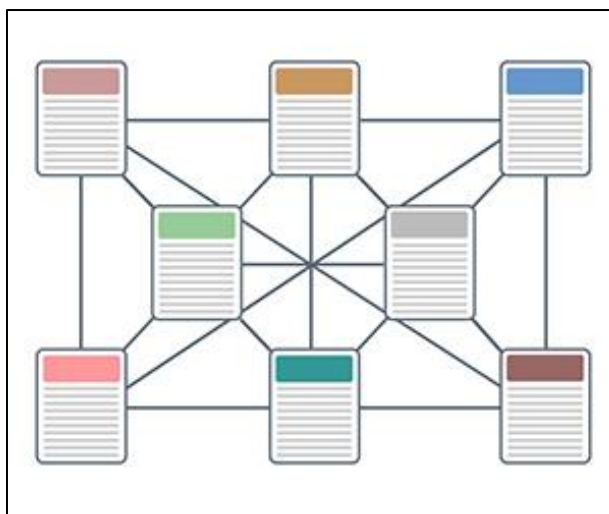
3) โครงสร้างแบบตาราง เป็นโครงสร้างการออกแบบเว็บไซต์ที่มีความซับซ้อน แต่ก็มีคามยืดหยุ่นในระดับหนึ่ง เพื่อให้ผู้ใช้งานสามารถเข้าสู่เนื้อหาต่าง ๆ ได้ง่ายขึ้น การออกแบบในลักษณะนี้จะมีการเชื่อมโยงเนื้อหาในแต่ละส่วนซึ่งกัน และกัน ทำให้ผู้ใช้งานสามารถเปลี่ยนทิศทาง หรือกำหนดทิศทางในการเข้าสู่เนื้อหาด้วยตัวเองได้ จึงไม่ทำให้เสียเวลา และทำให้เว็บไซต์มีความทันสมัยขึ้น



ภาพที่ 2.12 แสดงโครงสร้างแบบตาราง

(ที่มา: <https://asria.org/website-structure/>)

4) โครงสร้างแบบใยแมงมุม เป็นโครงสร้างที่ได้รับความนิยมเป็นอย่างมาก เพราะมีความยืดหยุ่นมากที่สุด โดยทุกหน้าเว็บจะมีการเชื่อมโยงถึงกันหมด ทำให้สามารถเข้าถึงหน้าเว็บเพจต่าง ๆ ที่ต้องการได้อย่างง่าย และมีความอิสระมากขึ้น นอกจากนี้ก็สามารถเชื่อมโยงไปสู่เว็บไซต์ภายนอกได้ดี



ภาพที่ 2.13 แสดงโครงสร้างแบบใยแมงมุม
(ที่มา: <https://asria.org/website-structure/>)

2.2.4.4 ส่วนประกอบหลักในหน้าเว็บไซต์ มีองค์ประกอบหลัก ๆ เช่น ชื่อเว็บไซต์ แถบเมนูนำทาง แถบเมนูย่อย ข้อความหัวเรื่อง ภาพประกอบตกแต่ง หรือภาพเนื้อหา ป้ายโฆษณา และอื่น ๆ ตามที่ต้องการนำเสนอ บางเว็บไซต์อาจจะแสดงภาพกราฟิกของสินค้า หรือบริการ แทนการแสดงข้อความ ทั้งนี้ขึ้นอยู่กับการออกแบบหน้าเว็บไซต์ให้เหมาะสมกับเนื้อหาที่ต้องการนำเสนอ (ดวงพร เกียงคำ, 2560)

ส่วนประกอบบนหน้าเว็บเพจ มีทั้งหมด 6 องค์ประกอบ ดังนี้

1) ส่วนหัว (Header) สิ่งสำคัญที่ช่วยสร้างการจดจำให้กับธุรกิจ คือ โลโก้ (Logo) ซึ่งทำหน้าที่เป็นสัญลักษณ์ที่สะท้อนถึงตัวตน และลักษณะเฉพาะของธุรกิจอย่างชัดเจน การออกแบบโลโก้ที่เหมาะสมจะช่วยให้ผู้เยี่ยมชมเว็บไซต์สามารถเข้าใจได้ทันทีว่าเว็บไซต์นั้นมีเนื้อหาหรือบริการเกี่ยวข้องกับอะไร โลโก้จึงถือเป็นองค์ประกอบที่สำคัญในการสื่อสารภาพลักษณ์ของธุรกิจ ควรจัดวางโลโก้ให้อยู่บริเวณด้านบนสุดของหน้าแรกของเว็บไซต์ เพื่อให้ผู้เข้าชมสามารถมองเห็นได้ทันทีที่เข้าสู่หน้าเว็บไซต์ และสร้างการจดจำธุรกิจของคุณอย่างมีประสิทธิภาพ

2) เมนู (Navigator) แถบเมนูเป็นองค์ประกอบสำคัญที่ช่วยอำนวยความสะดวกในการใช้งานเว็บไซต์ให้กับผู้เยี่ยมชม ควรออกแบบแถบเมนูให้สอดคล้องกับโครงสร้าง และการใช้งานของเว็บไซต์ โดยตำแหน่งของเมนูสามารถอยู่ในตำแหน่งเดียวกับโลโก้ด้านบนของหน้าเว็บไซต์ หรือวางไว้ที่ด้านซ้ายของหน้าจอก็ได้ ทั้งนี้ เมนูควรถูกจัดเรียงตามลำดับความสำคัญ เพื่อให้ผู้เยี่ยมชมสามารถเข้าถึงข้อมูลและบริการต่าง ๆ ของเว็บไซต์ได้อย่างรวดเร็ว และสะดวก อย่างไรก็ตาม ควรหลีกเลี่ยงการใส่เมนูมากเกินไปจนเกิดความจำเป็น เพราะอาจก่อให้เกิดความสับสนในการใช้งาน และส่งผลให้การนำทางภายในเว็บไซต์ดูซับซ้อน และไม่เป็นระเบียบ

3) เนื้อหา (Contents) ในหน้าแรกของเว็บไซต์ธุรกิจ ควรนำเสนอเนื้อหาที่ชัดเจน และครอบคลุมรายละเอียดเกี่ยวกับธุรกิจ หรือบริการของคุณ เพื่อให้ผู้เยี่ยมชมได้รู้จักธุรกิจของคุณมากยิ่งขึ้น ข้อมูลที่ควรนำเสนอ ได้แก่ การแนะนำว่าสินค้าหรือบริการของคุณคืออะไร คุณสมบัติเด่นของสินค้า หรือบริการ ประโยชน์ที่ผู้ใช้จะได้รับ ข่าวสารหรือกิจกรรมสำคัญ รวมถึงโปรโมชั่นที่เกี่ยวข้อง ช่องทางการติดต่อที่สะดวกและเข้าถึงง่าย คำรับรองจากลูกค้าที่เคยใช้บริการ รวมถึงรางวัล หรือการรับรองคุณภาพบริการที่ธุรกิจของคุณได้รับ

นอกจากนี้ หากต้องการเพิ่มโอกาสให้ธุรกิจของคุณถูกค้นพบได้ง่ายขึ้น ควรใส่คำสำคัญ (Keywords) ที่เกี่ยวข้องกับสินค้า หรือบริการในเนื้อหาของเว็บไซต์ ซึ่งจะช่วยให้กลุ่มเป้าหมายสามารถเข้าถึงเว็บไซต์ของคุณได้สะดวก และตรงกับความต้องการมากยิ่งขึ้น

4) ฟังก์ชันการค้นหา (Search Function) การจัดทำฟังก์ชันการค้นหาในเว็บไซต์เป็นสิ่งสำคัญที่ช่วยให้ผู้เยี่ยมชมสามารถค้นหาสินค้าหรือบริการได้อย่างรวดเร็วและสะดวกยิ่งขึ้น โดยเฉพาะในกรณีที่เว็บไซต์ของคุณมีสินค้าหรือบริการจำนวนมาก ฟังก์ชันนี้จะช่วยให้ผู้ใช้สามารถค้นหาสิ่งที่ต้องการได้อย่างง่ายดาย การออกแบบ และพัฒนาฟังก์ชันการค้นหาที่มีประสิทธิภาพจึงเป็นสิ่งที่คุณควรให้ความสำคัญ เพื่อให้ผู้เข้าชมสามารถเข้าถึงข้อมูลที่ต้องการได้โดยไม่มีอุปสรรค และเพิ่มประสบการณ์การใช้งานที่ดีขึ้นบนเว็บไซต์

5) ส่วนท้าย (Footer) ส่วนท้ายของเว็บไซต์ควรถูกออกแบบให้ช่วยอำนวยความสะดวกแก่ผู้เข้าชมในการนำไปยังหน้าต่าง ๆ ภายในเว็บไซต์ พร้อมทั้งเชื่อมโยงไปยังแพลตฟอร์มโซเชียลมีเดีย (Social Media) ข้อมูลที่ควรแสดงในส่วนนี้ประกอบด้วย ชื่อของธุรกิจ ช่องทางการติดต่อ เช่น หมายเลขโทรศัพท์และอีเมล ที่ตั้งของบริษัท รวมถึงลิงก์ไปยังโซเชียลมีเดียต่าง ๆ และลิงก์ที่เชื่อมโยงไปยังหน้าหลักอื่น ๆ ของเว็บไซต์ ทั้งนี้ การจัดวางข้อมูลในส่วนท้ายของ

เว็บไซต์ควรคำนึงถึงการใช้งานที่ง่ายและการเข้าถึงข้อมูลได้อย่างรวดเร็ว เคทีเอ็น ดีเวลลอป, (2565)

2.2.4.5 ทฤษฎีสีในการออกแบบติดต่อประสานงานผู้ใช้

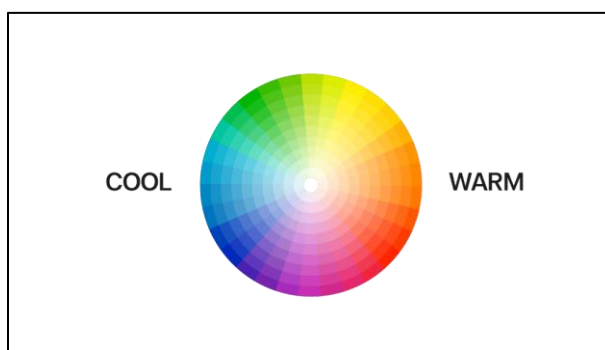
วิชาญ ทุมทอง, (2557) ได้กล่าวว่าจากสีที่แสดงในวงล้อสีมาตรฐานสามารถนำมาจัดเป็นกลุ่มของชุดสีเพื่อใช้ในการออกแบบ ส่วนติดต่อกับผู้ใช้ได้ดังนี้

1) ชุดสีโทนร้อน (Warm Color Scheme)

เป็นชุดสีที่มีความสัมพันธ์กับความอบอุ่น ความสุข ความรู้สึกสดชื่น การออกแบบที่มีสีสันทัดกัน อบอุ่น เหมาะสำหรับเด็ก วัยรุ่น สามารถดึงดูดความสนใจได้ดี ชุดสีโทนร้อน ประกอบไปด้วยสีเหลือง สีส้มเหลือง สีส้มแดง สีแดง และสีม่วงแดง

2) ชุดสีโทนเย็น (Cool Color Scheme)

เป็นชุดสีที่ให้ความรู้สึกเย็นสบาย กลมกลืนกับธรรมชาติ เป็นโทนสีที่เรียบง่าย มีความเป็นผู้ใหญ่ สุภาพเรียบร้อย ไม่มีสีสันทัดกันหรือสีตัดกันมากนัก บางครั้งดู ไม่ค่อยน่าสนใจ โทนสีชุดเย็น ประกอบไปด้วย สีม่วง สีม่วงน้ำเงิน สีน้ำเงิน สีเขียวน้ำเงิน สีเขียว และสีเขียวเหลือง



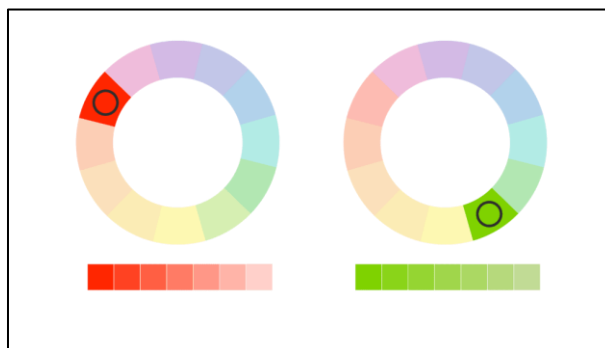
ภาพที่ 2.14 ชุดสีโทนเย็น-ร้อน

(ที่มา: <https://www.jallearnmedia.com/blog/color-theory/>)

3) ชุดสีแบบเดียว (Monochromatic Color Scheme)

เป็นชุดสีที่แสดงถึงความเรียบง่าย บริสุทธิ์ มีคุณค่า แต่ในบางครั้งอาจดู ไม่น่าสนใจลักษณะการใช้สีจะเป็นการไล่เฉดสีในระดับความเข้มมากไปจนถึงความเข้มน้อย

หรืออ่อน เช่น ถ้าใช้สีเขียวจะมีการไล่เฉดสีเป็น สีเขียวเข้ม สีเขียวแก่ สีเขียวปานกลาง และสีเขียวอ่อน

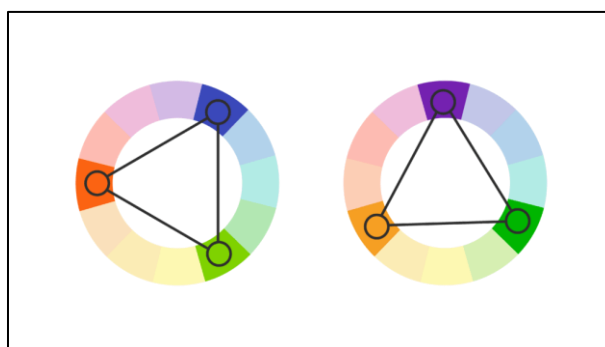


ภาพที่ 2.15 ตัวอย่างชุดสีแบบเดี่ยว

(ที่มา: <https://www.jlearnmedia.com/blog/color-theory/>)

4) ชุดสีสามสีตรงข้าม (Triadic Color Scheme)

เป็นสีที่อยู่มุมสามเหลี่ยมของวงล้อสี เป็นสีที่เข้าชุดกันอย่างลงตัว มีความสมดุลภายในชุดสีที่มุมสามเหลี่ยมเป็นสีที่ตัดกันทำให้เกิดความน่าสนใจ มีชีวิตชีวา ฉูดฉาด และสะดุดตา เหมาะสำหรับเด็ก หรือวัยรุ่น แต่ไม่เหมาะสำหรับผู้สูงอายุ แต่สีที่ฉูดฉาดเกินไปอาจจะรบกวนการสื่อความหมายได้

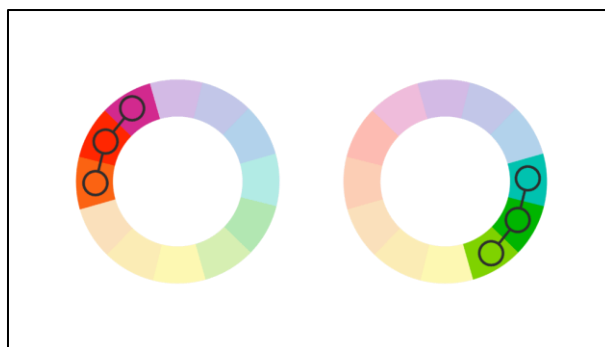


ภาพที่ 2.16 ตัวอย่างชุดสีสามสีตรงข้าม

(ที่มา: <https://www.jlearnmedia.com/blog/color-theory/>)

5) ชุดสีคล้ายคลึงกัน (Analogous Color Scheme)

เป็นชุดสีที่อยู่ติดกันในวงล้อสี อาจจะมี 2 สี 3 สี หรือมากกว่านั้นชุดสีที่อยู่ติดกันในวงล้อสีจะมีความสัมพันธ์กัน ให้อารมณ์ความรู้สึกที่ใกล้เคียงกัน และเมื่อนำมาใช้จะมีความกลมกลืนเข้ากันได้ ให้ความรู้สึกเรียบง่าย

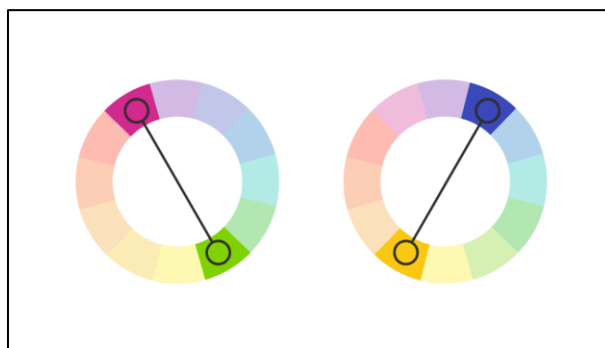


ภาพที่ 2.17 ตัวอย่างชุดสีคล้ายคลึงกัน

(ที่มา: <https://www.jlearnmedia.com/blog/color-theory/>)

6) ชุดสีตรงข้าม (Complementary Color Scheme)

เป็นชุดสีที่อยู่ตรงข้ามกันของวงล้อสี เช่น สีแดงกับสีเขียว สีเหลืองกับสีม่วง สีน้ำเงินเขียว กับสีส้มแดง เป็นต้น เมื่อนำมาใช้จะให้ความรู้สึกสดใส สะดุดตา โดดเด่น และมีความน่าสนใจ

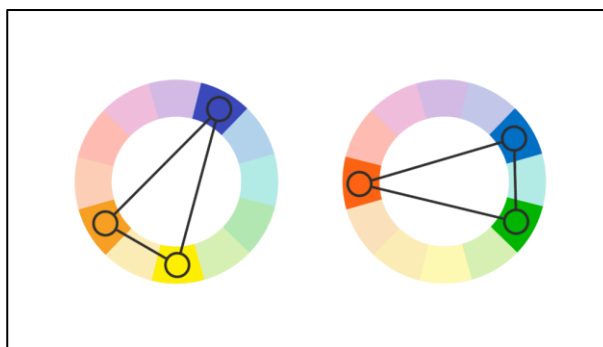


ภาพที่ 2.18 ตัวอย่างชุดสีตรงข้าม

(ที่มา: <https://www.jlearnmedia.com/blog/color-theory/>)

7) ชุดสีตรงข้ามข้างเคียง (Split Complementary Color Scheme)

เป็นชุดสีที่อยู่ด้านข้างของสีฝั่งตรงข้ามทั้งสองด้าน โดยสีฝั่งตรงข้ามจะถูกแทนที่ด้วยสีทั้งสอง การนำมาใช้จะทำให้มีความหลากหลายในการเลือกใช้สี แต่จะทำให้ความสดใส หรือความสะดุด ตาลดลง และทำให้ความเข้ากันของสีลดลงไปด้วย

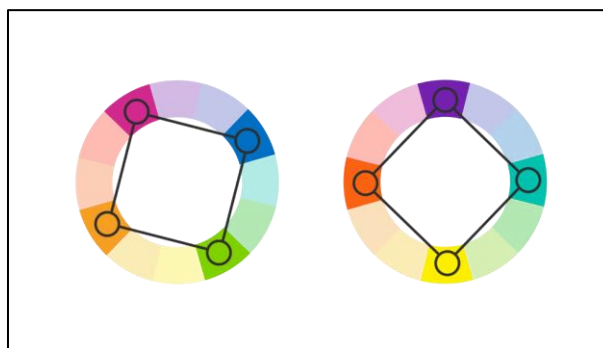


ภาพที่ 2.19 ตัวอย่างชุดสีตรงข้ามข้างเคียง

(ที่มา: <https://www.jlearnmedia.com/blog/color-theory/>)

8) ชุดสีสี่สีตรงข้าม (Tetradic)

การใช้สีสามสีตรงข้าม คือ การใช้สีที่อยู่ตรงข้ามกันอย่างชัดเจนจำนวน 4 สีบนวงล้อสี เป็นชุดสีที่สามารถใช้ได้แต่ได้รับความนิยมน้อยเนื่องจากมีสีหลักมากเกินไป แต่ถ้าใช้ในอัตราส่วนที่พอเหมาะก็ทำให้งานออกมาโดดเด่นได้มากขึ้น



ภาพที่ 2.20 ตัวอย่างชุดสีสี่สีตรงข้าม

(ที่มา: <https://www.jlearnmedia.com/blog/color-theory/>)

2.3 วรรณกรรมที่เกี่ยวข้อง

กิตติเทพ ไชยเทพ และไชยรัช เมฆแก้ว (2566) ได้ศึกษาการพยากรณ์ปริมาณการส่งออกตู้คอนเทนเนอร์ไปประเทศอินโดนีเซีย ช่วงสถานการณ์โควิด - 19 กรณีศึกษา : บริษัทเฟรทพอเว็ดเดอร์แห่งหนึ่ง ได้รวบรวมข้อมูลความต้องการตู้คอนเทนเนอร์ ตั้งแต่เดือนมกราคม 2563 - เดือนมิถุนายน 2564 โดยเปรียบเทียบรูปแบบการพยากรณ์ทั้งหมด 9 วิธี ได้แก่ วิธีค่าเฉลี่ยเคลื่อนที่อย่างง่าย วิธีค่าเฉลี่ยเคลื่อนที่แบบถ่วงน้ำหนัก วิธีการปรับเรียบแบบเอ็กซ์โพเนนเชียลครั้งเดียว วิธีแบบจำลองเชิงฤดูกาลแบบบวก วิธีแบบจำลองฤดูกาลแบบคูณ วิธีค่าเฉลี่ยเคลื่อนที่ซ้ำสองครั้ง วิธีการปรับเรียบแบบเอ็กซ์โพเนนเชียลซ้ำสองครั้ง วิธีการปรับให้เรียบเอ็กซ์โพเนนเชียลแบบไฮลด์-วินเทอร์ ที่มีอิทธิพลของฤดูกาลแบบบวก วิธีการปรับให้เรียบเอ็กซ์โพเนนเชียลแบบไฮลด์-วินเทอร์ ที่มีอิทธิพลของฤดูกาลแบบคูณ พบว่าค่าพยากรณ์ตู้คอนเทนเนอร์ ได้จากวิธีการปรับให้เรียบเอ็กซ์โพเนนเชียลแบบไฮลด์-วินเทอร์ที่ให้ค่าคลาดเคลื่อน (MSE) น้อยที่สุดเท่ากับ 3,855.97

เฉลิมวุฒิ คำเมือง และวชิรารักษ์ โอรสรัมย์ (2565) ได้ศึกษาการพยากรณ์ข้าวเปลือกหอมมะลิ ด้วยเทคนิคการทำเหมืองข้อมูล ข้อมูลที่ใช้ในการศึกษาราคาข้าวเปลือกหอมมะลิเฉลี่ยรายเดือน ตั้งแต่ มกราคม พ.ศ. 2553 - ธันวาคม พ.ศ. 2563 จำนวน 132 ข้อมูลโดยแบ่งข้อมูลออกเป็น 2 ชุด ข้อมูล ชุดที่ 1 ตั้งแต่เดือน มกราคม พ.ศ. 2553 - ธันวาคม พ.ศ. 2562 จำนวน 120 ข้อมูล สำหรับศึกษาตัวแบบพยากรณ์ โดยวิธีบ็อกซ์เจนกินส์ (Box-Jenkins) วิธีแบ็กกิง (Bagging) และวิธีแรนดอมฟอเรสต์ (Random Forest) ข้อมูลชุดที่ 2 ตั้งแต่เดือนมกราคม พ.ศ. 2563 - ธันวาคม พ.ศ. 2563 จำนวน 12 ค่า สำหรับการเปรียบเทียบความแม่นยำของค่าพยากรณ์โดยใช้เกณฑ์รากที่สองของความคลาดเคลื่อนเฉลี่ย (Root Mean Square Error: RMSE) และเกณฑ์เปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percentage Error: MAPE) พบว่าวิธีแรนดอมฟอเรสต์ได้ทำนายราคาข้าวเปลือกหอมมะลิ ได้ใกล้เคียงกับค่าจริงที่สุด และเมื่อคำนวณค่า RMSE ได้ 1,699.64 และได้ค่า MAPE เป็น 9.80

อุดมโชค อาษาวิมลกิจ (2563) ได้ศึกษาข้อมูลเปิดภาครัฐ : มาตรฐานการพัฒนา และการใช้งาน กล่าวว่าช่วงทศวรรษที่ผ่านมา องค์กรภาครัฐในหลายประเทศได้เริ่ม และพัฒนา “ข้อมูลเปิด” (open data) เพื่อเผยแพร่ข้อมูลของหน่วยงานไปสู่สังคมภายนอกอย่างกว้างขวาง ศูนย์ข้อมูลเปิดของรัฐบาลไทยอันเป็นแหล่งรวบรวมข้อมูลจากองค์กรภาครัฐเพื่อให้บริการสังคมจำนวนมากที่สุด ได้แก่ domain name ชื่อ data.go.th (ต่อมาเปลี่ยนเป็น opendata.data.go.th) โดย

สำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน) หรือ สพร. บทความนี้นำเสนอพัฒนาการของระบบศูนย์กลางข้อมูลเปิดภาครัฐ นับตั้งแต่ปี พ.ศ. 2558 ถึง ปัจจุบัน ตลอดจนการให้บริการข้อมูลเปิดจากศูนย์กลางข้อมูลภาครัฐดังกล่าว ซึ่งยังมีโอกาสในการพัฒนา โดยเฉพาะอย่างยิ่งคุณภาพความเป็นข้อมูลเปิด และการใช้ประโยชน์จากข้อมูลเปิด วิธีการศึกษาอาศัยการวิเคราะห์เนื้อหาจากมาตรฐาน รูปแบบ และการใช้ข้อมูลเปิด ข้อค้นพบสำคัญ ได้แก่ การควบคุมมาตรฐานข้อมูลเปิด รวมถึงการส่งเสริมให้เกิด การใช้ประโยชน์เชิงนวัตกรรมทางธุรกิจ การวิจัย หรือการพัฒนาต่อยอดในด้านต่าง ๆ ตามความต้องการของผู้ใช้งานวงกว้าง การพัฒนาต่อยอดระบบศูนย์กลางข้อมูลเปิดภาครัฐแห่งนี้อาจก่อให้เกิดการใช้งานที่มีผลกระทบด้านคุณค่าเชิงวิชาการ เชิงธุรกิจ การพัฒนาองค์กร และนโยบายสาธารณะ ในวงกว้าง

วารางคณา เรียนสุทธิ (2563) ได้ศึกษาการพยากรณ์มูลค่าการส่งออกพริกไทย โดยใช้ข้อมูลค่าเฉลี่ยรายเดือนจากเว็บไซต์ของสำนักงานเศรษฐกิจการเกษตร ตั้งแต่เดือนมกราคม 2554 ถึงเดือนตุลาคม 2563 จำนวน 118 เดือน ผู้ศึกษาได้แบ่งข้อมูลออกเป็น 2 ชุด ชุดที่ 1 ตั้งแต่ เดือนมกราคม 2554 ถึงเดือนธันวาคม 2562 จำนวน 108 เดือน สำหรับการสร้างตัวแบบพยากรณ์ด้วยวิธีการทางสถิติทั้งหมด 7 วิธี ได้แก่ วิธีบ็อกซ์-เจนกินส์ วิธีการทำให้เรียบด้วยเลขชี้กำลังของไฮลด์ วิธีการทำให้เรียบด้วยเลขชี้กำลังของบราวน์ วิธีการทำให้เรียบด้วยเลขชี้กำลังที่มีแนวโน้มแบบแตรม วิธีการทำให้เรียบด้วยเลขชี้กำลังที่มีฤดูกาลอย่างง่าย วิธีการทำให้เรียบด้วยเลขชี้กำลังของวินเทอร์แบบบวก และวิธีการทำให้เรียบด้วยเลขชี้กำลังของวินเทอร์แบบคูณ ข้อมูลชุดที่ 2 ตั้งแต่เดือนมกราคมถึงเดือนตุลาคม 2563 จำนวน 10 เดือน นำมาใช้สำหรับการเปรียบเทียบความแม่นยำของตัวแบบพยากรณ์ ด้วยเกณฑ์รากของค่าคลาดเคลื่อนกำลังสองเฉลี่ยที่ต่ำที่สุดพบว่ามูลค่าการส่งออกพริกไทย ด้วยวิธีการทางสถิติ 7 วิธี ได้แก่ วิธีบ็อกซ์-เจนกินส์ วิธีการทำให้เรียบด้วยเลขชี้กำลังของไฮลด์ วิธีการทำให้เรียบด้วยเลขชี้กำลังของบราวน์ วิธีการทำให้เรียบด้วยเลขชี้กำลังที่มีแนวโน้มแบบแตรม วิธีการทำให้เรียบด้วยเลขชี้กำลังที่มีฤดูกาลอย่างง่าย วิธีการทำให้เรียบด้วยเลขชี้กำลังของวินเทอร์แบบบวก และวิธีการทำให้เรียบด้วยเลขชี้กำลังของวินเทอร์แบบคูณผลการศึกษาพบว่า วิธีบ็อกซ์-เจนกินส์มีความแม่นยำมากที่สุด

ทัตพงศ์ อภิโรธนานนท์ สัตยา ต้นจันทร์พงศ์ และณัฐวรราช จำลองราษฎร์ (2563) ได้ศึกษาปัจจัยที่มีผลกระทบต่อมูลค่าการส่งออกข้าวในตลาดประเทศจีน ปัจจัยที่ใช้ในการศึกษาประกอบด้วย ดัชนีผู้บริโภคอุณหภูมิ ปริมาณน้ำฝน อัตราแลกเปลี่ยน และมูลค่าการส่งออก

ยางพาราในตลาดประเทศจีน ซึ่งทำการศึกษาข้อมูลรายเดือนตั้งแต่เดือนมกราคม พ.ศ. 2551 ถึง เดือนกันยายน พ.ศ. 2562 โดยใช้การวิเคราะห์ความถดถอยเชิงพหุคูณ พบว่า ดัชนีผู้บริโภค และมูลค่าการส่งออกยางพาราในตลาดประเทศจีนมีผลกระทบต่อมูลค่าการส่งออกข้าวในตลาดประเทศจีนอย่างมีระดับนัยสำคัญทางสถิติ ณ ระดับความเชื่อมั่นที่ร้อยละ 95 ส่วนปริมาณน้ำฝน อุณหภูมิ และอัตราแลกเปลี่ยน ไม่มีผลกระทบต่อมูลค่าการส่งออกข้าวในตลาดประเทศจีน

เซาวลิต ประสิทธิ์ และกนกพร ชัยประสิทธิ์ (2562) ได้ศึกษาความสัมพันธ์ระหว่างปัจจัยภายนอกกับความสำเร็จของวิสาหกิจขนาดกลางและขนาดย่อมในจังหวัดฉะเชิงเทรา ในการศึกษาครั้งนี้มีกลุ่มตัวอย่างที่ศึกษา คือ ผู้ประกอบการวิสาหกิจขนาดกลางและขนาดย่อม ในจังหวัดฉะเชิงเทรา โดยใช้แบบสอบถามในการเก็บข้อมูลจำนวน 393 ชุด โดยสถิติที่ใช้ในการวิเคราะห์ข้อมูล คือ สถิติเชิงพรรณนา (Descriptive Statistics) ได้แก่ ความถี่ ร้อยละ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน สถิติเชิงอนุมาน พบว่าลักษณะกิจการและจำนวนพนักงานมีผลต่อความสำเร็จของกิจการที่แตกต่างกัน ปัจจัยภายนอก เช่น การเมือง เศรษฐกิจ สังคมและวัฒนธรรม และเทคโนโลยี มีความสัมพันธ์ในระดับปานกลางกับความสำเร็จของ SMEs โดยภาพรวม ระดับความสำคัญของปัจจัยภายนอก และความสำเร็จอยู่ในระดับปานกลาง ที่ระดับนัยสำคัญทางสถิติ .05

ได้ศึกษาเรื่องการพยากรณ์ราคาที่พักโดยใช้การวิเคราะห์แบ่งช่วง (Divided Regression Analysis) โดยการวิเคราะห์สหสัมพันธ์ (Correlation Analysis) ข้อมูลที่ใช้ในการศึกษานี้มาจากชุดข้อมูล KC House Sales จาก Kaggle เพื่อทำนายราคาขายของบ้าน โดยใช้การวิเคราะห์การถดถอย (Regression Analysis) เป็นเครื่องมือทางสถิติในการศึกษาความสัมพันธ์ระหว่างตัวแปรต้น (Independent Variables) และตัวแปรตาม (Dependent Variables) การศึกษานี้ได้แบ่งข้อมูลออกเป็นชุดย่อยจำนวน 5 ชุด และทำการ Yann Ling Goh Yeh Huann Goh Chun-Chieh Yip และ Kooi Huat Ng (2022)วิเคราะห์ถดถอยเชิงเส้น (Multiple Regression Analysis) ในแต่ละชุด จากนั้นจึงรวมผลลัพธ์เพื่อสร้างโมเดลการวิเคราะห์ถดถอยแบ่งช่วง (Divided Regression Analysis) และทดสอบความเหมาะสมของโมเดล รวมถึงตรวจสอบการซ้ำซ้อนของตัวแปร (Multicollinearity) ที่อาจส่งผลกระทบต่อผลลัพธ์ พบว่าการแบ่งข้อมูลเป็นชุดย่อยและใช้การวิเคราะห์ถดถอยเชิงเส้นสามารถสร้างแบบจำลองที่ตอบสนองต่อข้อมูลได้ดี นอกจากนี้ ยังใช้วิธี Box-Cox Method เป็นเครื่องมือในการปรับปรุงแบบจำลอง เพื่อเพิ่มความแม่นยำของการพยากรณ์ และตรวจสอบ

ความเพียงพอของแบบจำลอง ซึ่งเป็นขั้นตอนสำคัญในการยืนยันประสิทธิภาพของแบบจำลองที่สร้างขึ้น

2.4 เครื่องมือในการออกแบบและวิเคราะห์ข้อมูล

2.4.1 กระบวนการวิเคราะห์ข้อมูลด้วย (CRISP-DM)

ขั้นตอนการทำเหมืองข้อมูลที่นิยมใช้ในปัจจุบันมีหลายแบบ และรูปแบบที่เป็นที่นิยม คือ Cross-Industry Standard Process for Data Mining (CRISP-DM) ที่มีการพัฒนา เป็น Workflow มาตรฐานสำหรับการทำเหมืองข้อมูล ประกอบด้วย 6 ขั้นตอนคือ

1) Business Understanding คือ การทำความเข้าใจในงาน ระบุโอกาส และหาปัญหาที่จะเกิดขึ้นกับธุรกิจ กำหนดขอบเขตของข้อมูลที่จะนำวิเคราะห์หาความได้เปรียบทางการตลาด และแก้ไขปัญหาองค์กร ซึ่งต้องสามารถระบุผลลัพธ์ที่มีได้

2) Data Understanding ทำความเข้าใจข้อมูลโดยการรวบรวมข้อมูลที่เกี่ยวข้อง คัดเลือกให้เหลือเพียงข้อมูลที่มีความถูกต้อง และสำคัญต่องานมาทำการวิเคราะห์

3) Data Preparation ทำการแปลงข้อมูล (Raw Data) ให้กลายเป็นข้อมูลที่สามารถนำมาช่วยในการวิเคราะห์ต่อไปได้ ขั้นตอนนี้จะใช้เวลามากที่สุดในทุกขั้นตอน เพราะคุณภาพของงานที่ได้จะดีเพียงใดขึ้นอยู่กับคุณภาพข้อมูลที่จัดเตรียมในขั้นนี้ การเตรียมข้อมูลประกอบด้วย การคัดเลือกข้อมูล การกลั่นกรองข้อมูล และแปลงรูปแบบของข้อมูล

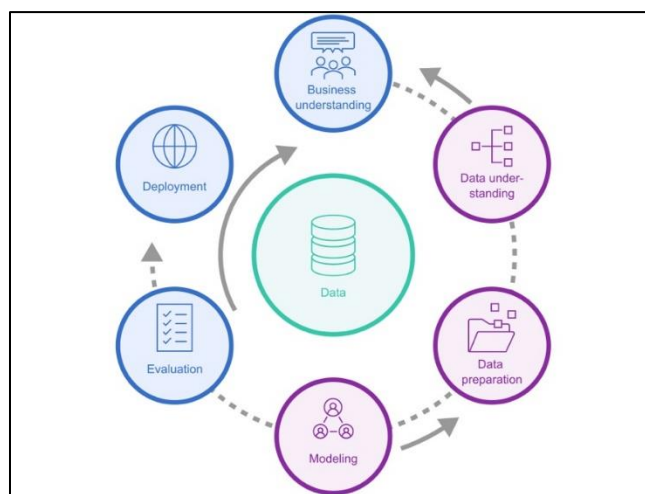
4) Modeling การสร้างแบบจำลองเพื่อวิเคราะห์ข้อมูลที่ได้จากขั้นตอนที่ 3 พร้อมทดสอบผลลัพธ์แบบจำลองเพื่อให้ได้คำตอบที่ดีที่สุด บางครั้งอาจมีการย้อนกลับไปปรับการเตรียมข้อมูลเพื่อให้ได้แบบจำลองที่เหมาะสมที่สุด

5) Evaluation การประเมินผลลัพธ์ที่ได้ก่อนที่จะนำไปใช้จริง ว่าตรงกับวัตถุประสงค์หรือเป้าหมายที่ได้ตั้งไว้ หรือมีความน่าเชื่อถือมากน้อยเพียงใด หากไม่ได้ผลลัพธ์ตามวัตถุประสงค์ ต้องย้อนกลับไปปรับปรุงแก้ไขการดำเนินงานในขั้นตอนก่อนหน้า

6) Deployment การนำเอาข้อมูลที่เป็นผลลัพธ์จากทั้งหมด มาลองปฏิบัติจริงกับธุรกิจ ในองค์กร โดยแปลงแนวคิดที่มีให้เกิดเป็นสารสนเทศเพื่อให้ผู้บริหาร หรือนักการตลาดเข้าใจสามารถนำไปใช้ประโยชน์ในทางธุรกิจได้จริง และติดตามประเมินผลที่ได้เพื่อนำกลับไปปรับปรุง Data Mining ต่อเนื่องต่อไป ซึ่งการประเมินผลสามารถทำได้หลายทาง เช่น วัดจากส่วนแบ่งของตลาด วัดจากปริมาณลูกค้า หรือ วัดจากกำไรสุทธิ เป็นต้น

ทั้งนี้การทำเหมืองข้อมูลในงานระบบทางธุรกิจ เป็นกระบวนการทางสถิติที่เน้นการจัดการข้อมูลที่มีจำนวนมากในหลากหลายรูปแบบ คัดเลือกเฉพาะข้อมูลที่สำคัญและจำเป็นต้องนำมาใช้งาน ทำการกำหนดรูปแบบจัดแบ่งกลุ่มลำดับความสำคัญ จากนั้นจึงจะเริ่มค้นหารูปแบบแนวทาง และความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลนั้นโดยแต่ละขั้นตอนจะอาศัยผลลัพธ์จากอีกขั้นตอนหนึ่งกลายเป็นข้อมูลให้ขั้นตอนต่อไป การทำเหมืองข้อมูลจะช่วยเปลี่ยนข้อมูลดิบให้เป็นสารสนเทศที่มีประโยชน์ การระบุแหล่งข้อมูลที่ต้องการจึงเป็นสิ่งสำคัญต่อผลลัพธ์ที่ได้จากการวิเคราะห์

การทำเหมืองข้อมูลจำเป็นต้องอาศัยบุคลากรจากหลายฝ่าย และต้องอาศัยความรู้จำนวนมากถึงจะได้รับประโยชน์อย่างแท้จริง เพราะสิ่งที่ได้เป็นเพียงตัวเลข และข้อมูล ที่อาจจะนำไปใช้ประโยชน์ได้หรือใช้ประโยชน์อะไรไม่ได้เลยก็เป็นได้ ผู้ที่ศึกษาการทำเหมืองข้อมูลจึงควรมีความรู้รอบด้าน และต้องติดต่อกับทุก ๆ ฝ่าย เพื่อให้เข้าใจถึงขอบเขตของปัญหาโดยแท้จริงก่อน เพื่อให้การทำเหมืองข้อมูลเกิดประโยชน์อย่างแท้จริง (กรมส่งเสริมอุตสาหกรรม, 2562)



ภาพที่ 2.21 แบบจำลอง CRISP-DM

(ที่มา: ist.fraunhofer.de)

2.4.2 การวิเคราะห์การถดถอยเชิงเส้น (Linear Regression)

การวิเคราะห์การถดถอย (Regression Analysis) เป็นวิธีการทางสถิติที่ใช้ในการศึกษาและตรวจสอบความสัมพันธ์ระหว่างตัวแปร โดยปกติจะมีตัวแปรที่ต้องการทำนาย หรืออธิบาย (ตัวแปรตาม) และตัวแปรที่ใช้ในการทำนาย (ตัวแปรอิสระ) ซึ่งตัวแปรอิสระอาจมีมากกว่าหนึ่งตัว

การวิเคราะห์การถดถอยมีวัตถุประสงค์หลัก ดังนี้

1) ศึกษาความสัมพันธ์ระหว่างตัวแปรอิสระ และตัวแปรตาม การวิเคราะห์จะช่วยให้ทราบถึงขนาดของความสัมพันธ์ระหว่างตัวแปรอิสระกับตัวแปรตาม รวมถึงการระบุว่าตัวแปรอิสระใดมีความสัมพันธ์สูง หรือต่ำกว่าตัวแปรตาม เพื่อให้สามารถคาดการณ์ได้ว่าตัวแปรอิสระใดมีอิทธิพลมากที่สุดต่อตัวแปรตาม เช่น การศึกษาความสัมพันธ์ระหว่างน้ำหนักทารกแรกเกิดกับอายุ น้ำหนัก และส่วนสูงของมารดา

2) สร้างแบบจำลองเพื่อทำนายตัวแปรตาม การสร้างแบบจำลองทางคณิตศาสตร์ที่สามารถทำนายค่าของตัวแปรตามจากค่าของตัวแปรอิสระ เช่น การศึกษาปริมาณการใช้ยาที่ส่งผลต่อความดันโลหิต

3) ทราบความสัมพันธ์ระหว่างตัวแปรอิสระบางตัวที่มีผลต่อตัวแปรตาม การวิเคราะห์จะช่วยในการควบคุมอิทธิพลของตัวแปรอิสระอื่น ๆ ที่อาจมีผลกระทบต่อความสัมพันธ์ที่สนใจ เช่น การศึกษาผลกระทบของความวิตกกังวลต่อประสิทธิภาพการทำงาน โดยควบคุมระยะเวลาในการทำงาน

4) สร้างแบบจำลองที่ดีที่สุดเพื่อการทำนาย การพัฒนา และเลือกแบบจำลองที่ดีที่สุดจากหลายแบบจำลองที่สร้างขึ้น เพื่อใช้ในการทำนายตัวแปรตามอย่างมีประสิทธิภาพ

5) ตรวจสอบประสิทธิภาพของแบบจำลองในการทำนาย การทดสอบว่าแบบจำลองที่พัฒนาขึ้นมีความเสถียรและมีประสิทธิภาพในการทำนายตัวแปรตามในกลุ่มเป้าหมายที่แตกต่างกันหรือไม่

ชนิดของการวิเคราะห์การถดถอย

การวิเคราะห์การถดถอย มีหลายชนิด ขึ้นกับลักษณะของตัวแปรตาม รูปแบบความสัมพันธ์ และการทำนายค่าตัวแปรอิสระ (ตัวแปรต้น) ซึ่งโดยทั่วไปแบ่งการวิเคราะห์การถดถอยได้เป็น 2 ประเภท คือ

1) การวิเคราะห์การถดถอยเชิงเส้น (Linear regression analysis) เป็นการวิเคราะห์การถดถอยที่ตัวแปรอิสระมีความสัมพันธ์ที่เป็นสัดส่วนกับตัวแปรตาม ต้นแบบที่เป็นเส้นตรงตัวแปรตาม สามารถแทนได้ด้วยสมการทางคณิตศาสตร์ที่เป็นเชิงเส้น (Linear model)

2) การวิเคราะห์การถดถอยแบบไม่เป็นเชิงเส้น (Non linear regression) เป็นการวิเคราะห์การถดถอย ที่รูปแบบของความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตามสามารถแทนได้ด้วยสมการทางคณิตศาสตร์ที่ไม่เป็นเชิงเส้น (non - Linear model)

การวิเคราะห์การถดถอยเชิงเส้นมี 2 แบบ คือ

1) การวิเคราะห์การถดถอยเชิงเส้นอย่างง่าย (Simple Linear Regression)

การวิเคราะห์การถดถอยเชิงเส้นอย่างง่าย จะประกอบด้วยตัวแปรตาม 1 ตัว และตัวแปรอิสระเพียง 1 ตัว การวิเคราะห์การถดถอยจะหาความสัมพันธ์เชิงเส้นตรงเพียงเส้นเดียว และมีรูปแบบสมการทางคณิตศาสตร์ที่ใช้ในการทำนายหาความสัมพันธ์ระหว่างตัวแปรทั้งสอง เช่น สมการทำนายปริมาณ Carbon Monoxide ในอุตสาหกรรมกับปริมาณก๊าซบางชนิดที่ปล่อยต่อวัน

2) การวิเคราะห์การถดถอยเชิงเส้นแบบพหุ (Multiple Linear Regression)

จะประกอบด้วยตัวแปรตาม 1 ตัว และ ตัวแปรอิสระตั้งแต่ 2 ตัวขึ้นไป การวิเคราะห์ที่เป็นการหาหาขอบของความสัมพันธ์ และสร้างรูปแบบสมการทางคณิตศาสตร์ที่เป็นการทำนายค่าของตัวแปรตาม โดยใช้ตัวแปรอิสระที่ศึกษา เช่น ศึกษาความสัมพันธ์ระหว่างอายุ เพศ ปัญหาในการทำงาน ความขัดแย้งในครอบครัว กับความรู้สึกเก็บกดของผู้ป่วยในโรงพยาบาลแห่งหนึ่ง เป็นต้น (Faculty of Nursing, Chiang Mai University. (n.d.))

2.5 บทสรุป

จากการทบทวนแนวคิด ทฤษฎี เครื่องมือ และวรรณกรรมที่เกี่ยวข้องกับการวิเคราะห์ข้อมูล เพื่อพยากรณ์มูลค่าการส่งออกสินค้าของวิสาหกิจขนาดกลางและขนาดย่อมโดยใช้เทคนิคการทำเหมืองข้อมูล ผู้ศึกษาได้เลือกใช้กระบวนการวิเคราะห์ข้อมูลตามขั้นตอนของ Cross-Industry Standard Process for Data Mining (CRISP-DM) ซึ่งเป็นเทคนิคที่ได้รับการยอมรับในงานวิจัยด้านการทำเหมืองข้อมูล โดยใช้โปรแกรมที่เหมาะสมสำหรับการพัฒนาแบบจำลองสมการถดถอยเชิงพหุคูณ (Multiple Linear Regression) เพื่อพยากรณ์มูลค่าการส่งออกสินค้าของวิสาหกิจขนาดกลางและขนาดย่อม นอกจากนี้ยังมีการนำเสนอข้อมูลสารสนเทศผ่านเว็บไซต์ที่พัฒนาขึ้นโดยใช้ภาษา PHP, HTML และ CSS ซึ่งช่วยให้สามารถเข้าถึงข้อมูลได้อย่างสะดวก และมีประสิทธิภาพ