

## บทที่ 2

### แนวคิด ทฤษฎี เครื่องมือและวรรณกรรมที่เกี่ยวข้อง

โครงการเรื่อง การวิเคราะห์ปัญหาการลงทะเบียนเรียนออนไลน์ของนักศึกษาในมหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เชียงใหม่ เป็นการศึกษาปัญหาที่เกิดขึ้นจริงจากการใช้ระบบลงทะเบียนเรียนออนไลน์ เพื่อนำข้อมูลเชิงประจักษ์มาวิเคราะห์หาสาเหตุและเสนอแนวทางปรับปรุงระบบให้มีประสิทธิภาพมากขึ้น บทนี้จะนำเสนอแนวคิดพื้นฐาน ทฤษฎี เครื่องมือ และงานวิจัยที่เกี่ยวข้อง เพื่อสร้างกรอบการศึกษาและวางแผนการวิจัยอย่างเป็นระบบ ประกอบด้วย รายละเอียดตามลำดับ ดังนี้

#### 2.1 แนวคิด

- 2.1.1 แนวคิดเกี่ยวกับการวิเคราะห์ข้อมูล (Data Analytics)
- 2.1.2 แนวคิดเกี่ยวกับการประยุกต์ใช้ข้อมูลทางสถิติ
- 2.1.3 แนวคิดเกี่ยวกับการแสดงผลข้อมูล (Data Visualization)
- 2.1.4 แนวคิดเกี่ยวกับการทำความสะอาดข้อมูล (Data Cleaning)
- 2.1.5 แนวคิดการจับกลุ่มข้อมูลด้วย K-Means

#### 2.2 ทฤษฎี

- 2.2.1 การทำเหมืองข้อมูล (Data Mining)
- 2.2.2 หลักในการทำเหมืองข้อมูล
- 2.2.3 ทฤษฎีเกี่ยวกับกฎความสัมพันธ์ (Association Rules)
- 2.2.4 การวิเคราะห์เชิงพรรณนา (Descriptive Analytics)
- 2.2.5 ทฤษฎีที่เกี่ยวข้องกับการสร้างเว็บไซต์

#### 2.3 วรรณกรรมที่เกี่ยวข้อง

#### 2.4 เครื่องมือในการออกแบบและวิเคราะห์ข้อมูล

- 2.4.1 กระบวนการวิเคราะห์ข้อมูลด้วย (CRISP-DM)
- 2.4.2 การวิเคราะห์เกี่ยวกับกฎความสัมพันธ์ (Association Rules)

#### 2.5 บทสรุป

## 2.1 แนวคิด

2.1.1 แนวคิดเกี่ยวกับการวิเคราะห์ข้อมูล (Data Analytics) ในการดำเนินงานเรื่องการวิเคราะห์ปัญหาการลงทะเบียนเรียนออนไลน์ของนักศึกษาในมหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี เชียงใหม่ 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 และวิเคราะห์มีความแม่นยำรวมถึงสามารถนำไปใช้งานได้อย่างมีประสิทธิภาพ แบ่งออกเป็น 4 ประเภท ดังนี้

1) Diagnostic Analytics คือ การวิเคราะห์เชิงวินิจฉัย ซึ่งเป็นการวิเคราะห์เชิงลึกที่ต่อยอดจาก Descriptive Analytics เพื่อตอบคำถามว่าเหตุการณ์ปัญหาการลงทะเบียนเรียนออนไลน์นั้น ๆ เกิดขึ้นได้อย่างไร โดยใช้เทคนิคต่าง ๆ เช่น Data Discovery และ Log File Mining เพื่อตรวจสอบและค้นหาปัจจัยหรือตัวแปรที่ส่งผลให้เกิดปัญหา การวิเคราะห์สาเหตุที่ระบบล่มในช่วง 10 นาทีแรกหลังเปิดลงทะเบียน การตรวจสอบปริมาณ Request/Minute ที่สูงเกินปกติจนทำให้เซิร์ฟเวอร์ตอบสนองช้า การค้นหาว่าขั้นตอนใดในกระบวนการลงทะเบียน (เช่น ยืนยันรายวิชา หรือการตรวจสอบสถานะรายวิชา) เป็นจุดที่เกิด Error บ่อยที่สุด ตลอดจนการวิเคราะห์พฤติกรรมของนักศึกษาที่เข้าใช้งานพร้อมกันในช่วงเวลา Peak Load ซึ่งอาจเป็นสาเหตุของการเกิดปัญหา ข้อมูลที่ได้จะช่วยให้นักศึกษาเข้าใจสาเหตุเชิงลึกของปัญหา สามารถปรับปรุงโครงสร้างของระบบ เช่น การเพิ่มขีดความสามารถของเซิร์ฟเวอร์ การใช้ระบบ Queue-based Access การทำ Load Balancing

2) การวิเคราะห์ข้อมูลในปัจจุบันเป็นสิ่งสำคัญอย่างยิ่งสำหรับสถาบันการศึกษาในการบริหารจัดการเชิงกลยุทธ์และตัดสินใจอย่างมีข้อมูล โดยเฉพาะเมื่อเกี่ยวข้องกับกระบวนการหลักของมหาวิทยาลัยอย่างการลงทะเบียนเรียนออนไลน์ ซึ่งต้องการข้อมูลเชิงลึกเพื่อวางแผนล่วงหน้าและป้องกันปัญหาที่อาจเกิดขึ้น Predictive Analytics คือ การวิเคราะห์ข้อมูลในอดีตและปัจจุบันเพื่อตรวจสอบแนวโน้มและทำนายเหตุการณ์ที่จะเกิดขึ้น โดยใช้แบบจำลองทางสถิติร่วมกับเทคโนโลยี Machine Learning เช่น Decision Tree Random Forest หรือ LSTM ซึ่งในบริบทของโครงการนี้นำมาใช้เพื่อคาดการณ์ปริมาณผู้ใช้งานในแต่ละช่วงเวลา ระบุโอกาสที่ระบบอาจล่ม และหาว่ากลุ่มนักศึกษาใดมีแนวโน้มประสบปัญหา มากที่สุด Forecast Analytics คือ การวิเคราะห์เชิงคาดการณ์ที่มุ่งเน้นการทำนายเหตุการณ์

เฉพาะเจาะจงในอนาคต โดยใช้ข้อมูลแบบอนุกรมเวลา (Time Series Data) เพื่อคาดการณ์แนวโน้มในระยะยาว เช่น การคาดการณ์ช่วง Peak Load ของการลงทะเบียนในแต่ละภาคเรียน การประเมินความเสี่ยงของระบบล่มในช่วงเวลาเปิดลงทะเบียน และการคาดการณ์จำนวนวิชาที่มีโอกาสเต็มอย่างรวดเร็ว ข้อมูลเหล่านี้สามารถนำไปวางแผนการเพิ่มขีดความสามารถของเซิร์ฟเวอร์และการบริหารจัดการทรัพยากรของมหาวิทยาลัยได้อย่างแม่นยำ การใช้ Predictive Analytics และ Forecast Analytics ในโครงการนี้ ช่วยให้การวิเคราะห์ปัญหาไม่เพียงหยุดอยู่ที่การอธิบายเหตุการณ์ที่เกิดขึ้น แต่สามารถ คาดการณ์ปัญหาที่อาจเกิดขึ้นในอนาคต ทำให้มหาวิทยาลัยสามารถเตรียมมาตรการเชิงรุก ลดโอกาสการเกิดปัญหาซ้ำ และเพิ่มความพึงพอใจของนักศึกษาได้อย่างมีประสิทธิภาพ

3) Predictive Analytics และ Forecast Analytics จะเป็นการใช้ข้อมูลในอดีตและปัจจุบันเพื่อนำมาคาดการณ์อนาคต แต่ทั้งสองวิธีมีวัตถุประสงค์และการใช้งานที่แตกต่างกันอย่างชัดเจน Predictive Analytics มักใช้เพื่อคาดการณ์พฤติกรรมของผู้ใช้งานหรือเหตุการณ์เฉพาะเจาะจง โดยเน้นการวิเคราะห์ในระดับจุลภาค (Micro-level) เช่น การคาดการณ์จำนวนนักศึกษาที่จะเข้าใช้งานระบบลงทะเบียนในแต่ละนาที่ การคาดการณ์ขั้นตอนที่มีโอกาสเกิด Error สูง หรือการระบุว่านักศึกษากลุ่มใดมีแนวโน้มจะประสบปัญหามากที่สุด Forecast Analytics มุ่งเน้นไปที่การมองภาพรวมแนวโน้มในอนาคตในระดับมหภาค (Macro-level) โดยใช้ข้อมูลแบบอนุกรมเวลา (Time-series data) เช่น การพยากรณ์จำนวนผู้ใช้งานสูงสุดของระบบในแต่ละภาคการศึกษา การคาดการณ์ความหนาแน่นของระบบในช่วงเปิดลงทะเบียน และการประมาณจำนวนวิชาที่คาดว่าจะเต็มอย่างรวดเร็ว เพื่อให้ผู้บริหารสามารถวางแผนทรัพยากรและเพิ่มขีดความสามารถของระบบได้อย่างแม่นยำ การใช้ Predictive และ Forecast Analytics ร่วมกันในโครงการนี้ จะช่วยให้มหาวิทยาลัยเข้าใจทั้ง ปัญหาเฉพาะจุด (เช่น สาเหตุการล่มในช่วงเวลาเร่งด่วน) และ ภาพรวมของระบบในอนาคต (เช่น แนวโน้มความต้องการลงทะเบียน) เพื่อปรับปรุงระบบลงทะเบียนให้มีความเสถียร ตอบสนองต่อจำนวนผู้ใช้งาน และลดผลกระทบต่อการเรียนของนักศึกษา

4) Prescriptive Analytics คือ การวิเคราะห์เพื่อให้คำแนะนำเชิงปฏิบัติที่เหมาะสมที่สุด เป็นขั้นตอนวิเคราะห์ที่ซับซ้อนที่สุด และเป็นการต่อยอดจาก Predictive Analytics เมื่อได้ข้อมูลแนวโน้มที่คาดว่าจะเกิดขึ้นแล้ว Prescriptive Analytics จะช่วยเสนอแนวทางปฏิบัติที่ดีที่สุด วิเคราะห์ผลลัพธ์ที่อาจเกิดขึ้นจากแต่ละทางเลือก และช่วยผู้มีส่วนเกี่ยวข้องตัดสินใจได้อย่างแม่นยำในบริบทของโครงการ การวิเคราะห์ปัญหาการลงทะเบียนเรียนออนไลน์ของนักศึกษาในมหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เชียงใหม่ Prescriptive

Analytics สามารถนำมาใช้เพื่อแนะนำช่วงเวลาที่เหมาะสมที่สุดให้นักศึกษาทำการลงทะเบียน เพื่อลดความหนาแน่นของระบบแนะนำให้เจ้าหน้าที่เพิ่มทรัพยากรเซิร์ฟเวอร์หรือทำ Load Balancing ก่อนช่วง Peak Load เสนอแนะการออกแบบระบบแจ้งเตือนอัตโนมัติเมื่อตรวจพบ ความเสี่ยงของการล่มประเมินผลลัพธ์ในแต่ละทางเลือก เช่น ถ้าเพิ่มเซิร์ฟเวอร์จะช่วยลดความ เสี่ยงได้กี่เปอร์เซ็นต์ Prescriptive Analytics ใช้การประมวลผลข้อมูลขนาดใหญ่ (Big Data) ร่วมกับอัลกอริทึม Machine Learning และเทคโนโลยี AI เพื่อจำลองผลลัพธ์และหาทางเลือกที่ เหมาะสมที่สุด ซึ่งเกินขีดความสามารถของการวิเคราะห์ด้วยมือหรือการตัดสินใจแบบใช้ สัญชาตญาณ ช่วยให้นักศึกษาลดตัดสินใจได้อย่างมีหลักฐานเชิงข้อมูลรองรับ และลดโอกาส ที่ปัญหาจะเกิดขึ้นในอนาคต

2.1.2 สถิติ (Statistic) คือ ศาสตร์ที่ศึกษาเกี่ยวกับข้อมูล การเก็บรวบรวม การวิเคราะห์ ข้อมูล การแปลความหมาย และการนำเสนอข้อมูลข้อสรุปของการวิเคราะห์ โดยสถิติจะแสดง ข้อมูล ข้อเท็จจริงเกี่ยวกับด้านในด้านหนึ่ง โดยประเภทของแบ่งตามวิธีการดำเนินการกับ ข้อมูล เป็น 2 ประเภท ดังนี้

1) สถิติเชิงพรรณนา (Descriptive Statistics) คือ การอธิบาย บรรยาย หรือ นำเสนอข้อมูลที่เก็บมาศึกษาในลักษณะตารางหรือแผนภูมิ เพื่อให้เข้าใจคุณลักษณะหรือ แนวโน้มของกลุ่มข้อมูลนั้น ๆ ตัวอย่างค่าทางสถิติที่ใช้ ได้แก่ ค่าเฉลี่ย (Mean) ใช้บอกค่ากลาง ของข้อมูลในเชิงปริมาณ ค่ามัธยฐาน (Median) ค่าที่อยู่กึ่งกลางของชุดข้อมูลเมื่อเรียงลำดับ แล้ว ค่าฐานนิยม (Mode) ค่าที่เกิดซ้ำมากที่สุดหรือมีความถี่สูงสุด ค่าการกระจาย (Standard Deviation) ใช้วัดความแปรปรวนของข้อมูล

2) สถิติเชิงอนุมาน (Inferential Statistics) การศึกษาลักษณะข้อมูลที่ได้จาก กลุ่มตัวอย่าง แล้วนำผลการวิเคราะห์ไปอธิบายหรือสรุปอ้างอิงถึงลักษณะของประชากร ทั้งหมด โดยมีจุดประสงค์เพื่อให้สามารถคาดการณ์และตัดสินใจได้อย่างมั่นใจบนพื้นฐานของ ข้อมูลที่มีอยู่ เนื่องจากในหลายกรณีไม่สามารถเก็บข้อมูลจากประชากรทั้งหมดได้จริง องค์ประกอบสำคัญของสถิติเชิงอนุมาน ได้แก่

1) การประมาณค่า (Estimation) การใช้สถิติจากกลุ่มตัวอย่างมาเป็น ค่าประมาณของพารามิเตอร์ประชากร เช่น การประมาณค่าเฉลี่ยของเวลาที่นักศึกษาใช้ ลงทะเบียนต่อครั้ง หรือการประมาณร้อยละของนักศึกษาที่พบ Error

2) การทดสอบสมมติฐาน (Hypothesis Testing) เป็นการทดสอบว่า ความแตกต่างหรือความสัมพันธ์ที่พบในกลุ่มตัวอย่างเป็นเพราะความบังเอิญหรือมีนัยสำคัญ

ทางสถิติจริง ตัวอย่างเช่น  $H_0$ : อัตราการเกิด Error ในช่วงเวลา Peak ไม่แตกต่างจาก Off-Peak  
 $H_1$ : อัตราการเกิด Error ในช่วงเวลา Peak สูงกว่าอย่างมีนัยสำคัญ

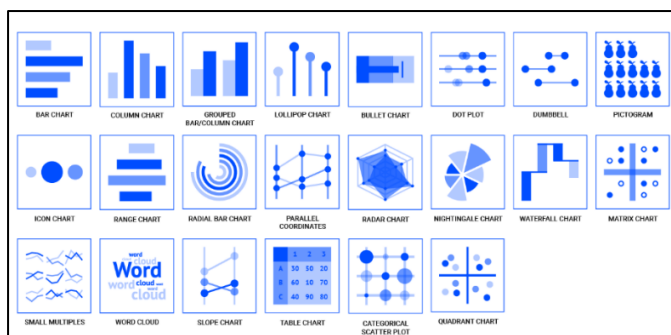
3) การวิเคราะห์ความสัมพันธ์ (Correlation Analysis) ใช้ตรวจสอบความสัมพันธ์ระหว่างตัวแปร เช่น จำนวนผู้ใช้งานพร้อมกัน (Concurrent Users) กับอัตราการล่มของระบบ (Downtime)

4) การวิเคราะห์การถดถอย (Regression Analysis) เพื่อสร้างสมการคาดการณ์ เช่น การพยากรณ์เวลาตอบสนองของระบบ (Response Time) จากปัจจัยจำนวน Request ต่อวินาที (blendata, 2567)

### 2.1.3 แนวคิดเกี่ยวกับการแสดงผลข้อมูล (Data Visualization)

Data Visualization เป็นการนำข้อมูลมาแสดงในรูปแบบของ ภาพกราฟิก แผนภูมิ กราฟ ไดอะแกรม หรืออื่น ๆ (Charts Graphs Diagrams หรือ Maps) เพื่อให้เราเห็นและเข้าใจลักษณะ แนวโน้ม และความสัมพันธ์ของข้อมูล (Patterns trends และ correlations) ได้ง่ายขึ้น ชัดเจนขึ้น เนื่องจากโดยทั่วไปแล้ว ข้อมูลที่เราสนใจมักจะอยู่ในรูปแบบข้อความอักขระ หรือตัวเลข ซึ่งก็มีอยู่ เป็นจำนวนมาก 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 เพื่อที่จะเลือกชนิดของกราฟได้เหมาะสม เช่น ต้องการดูจำนวน ความถี่ใช้กราฟแท่ง Histogram ต้องการดูการกระจาย ต้องการดูความสัมพันธ์ หรือแนวโน้ม ของข้อมูล ใช้ Scatter และยังต้องคำนึงถึงด้วยว่า ข้อมูลที่นำมาพล็อตเป็นข้อมูลชนิดใด ชนิดหนึ่ง ต่อเนื่อง หรือแบบ Category ฯลฯ (cheewin.j, 2568)

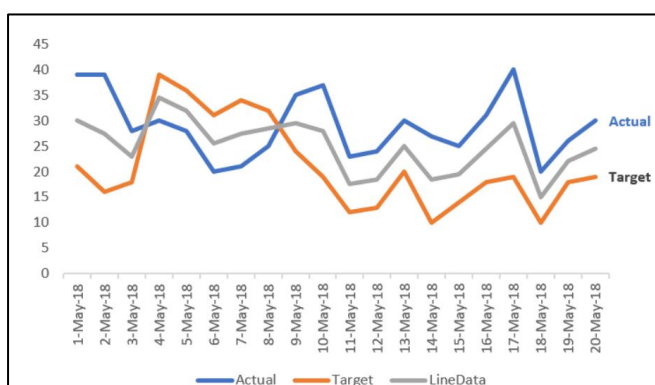
1) แผนภูมิ (Charts) เป็นรูปแบบการแสดงผลข้อมูลที่นิยมมากที่สุด เนื่องจากสามารถสื่อสารข้อมูลจำนวนมากให้เข้าใจได้อย่างรวดเร็ว และเหมาะสมกับการนำเสนอข้อมูลที่แตกต่างกันตามวัตถุประสงค์ โดยแผนภูมิมียุคหลายชนิด ซึ่งแต่ละชนิดเหมาะกับการใช้งานที่เฉพาะเจาะจง



ภาพที่ 2.1 ตัวอย่างแผนภูมิรูปแบบต่าง ๆ

(ที่มา: <https://www.datylon.com/blog/types-of-charts-graphs-examples-data-visualization#comparison-charts>)

2) กราฟ (Graphs) คือ Subset หรือประเภทหนึ่งของแผนภูมิ โดยมีหน้าที่หลักในการแสดงความสัมพันธ์ระหว่างข้อมูลสองตัวแปรผ่านแกนพิกัด ได้แก่ แกนแนวนอน (แกน X) และแกนแนวตั้ง (แกน Y) ทำให้สามารถสังเกตเห็นแนวโน้มของข้อมูล



ภาพที่ 2.2 ตัวอย่างกราฟเส้นตัวอย่าง

(ที่มา: <https://reportingengineer.com/2018/05/25/show-variance-between-2-line-chart-techniques/>)

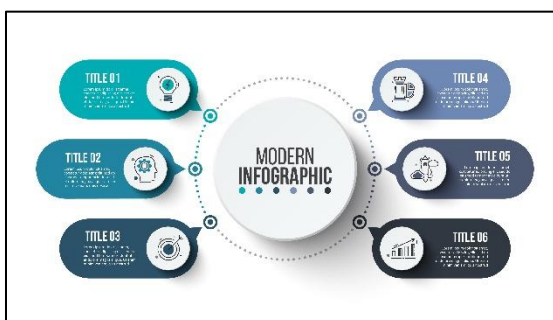
3) ตาราง (Tables) เป็นอีกรูปแบบที่ใช้กันมากเพื่อนำเสนอข้อมูลให้ออกมา ดูง่าย ตารางประกอบไปด้วย 2 ส่วน ได้แก่ คอลัมน์ และแถว ซึ่งช่วยจัดการข้อมูลให้เรียบร้อย ช่วยให้เห็นบริบท และความสัมพันธ์ของข้อมูลหลาย ๆ ชุดได้อย่างง่าย

Marks	Number of Students		Total
	Males	Females	
30 – 40	8	6	14
40 – 50	16	10	26
50 – 60	14	16	30
60 – 70	12	8	20
70 – 80	6	4	10
Total	56	44	100

ภาพที่ 2.3 ตัวอย่างตารางแสดงข้อมูล

(ที่มา: <https://1stcraft.com/what-is-data-visualization/#5-3-ตาราง-tables>)

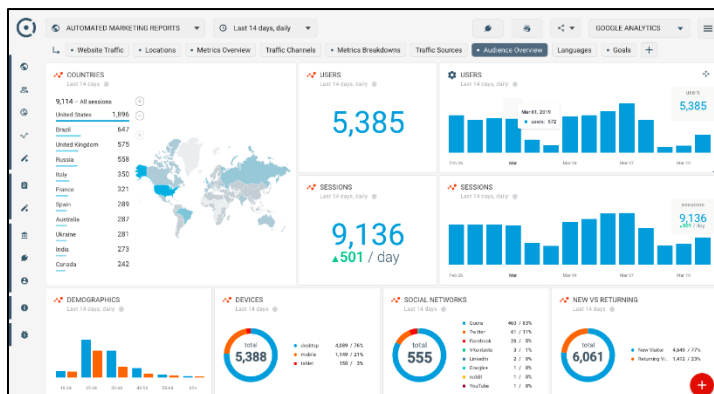
4) อินโฟกราฟิก (Infographic) คือ การนำเสนอสารสนเทศ (Info: information) ด้วยภาพกราฟิก (Graphic) เป็นรูปแบบการนำเสนอข้อมูลที่ใช้ภาพสื่อแทน ทำให้ผู้อ่านข้อมูลเข้าใจ ข้อมูลได้ง่ายหรือสามารถทำความเข้าใจผ่านภาพแทนที่คุ้นเคย นอกจากนี้ อินโฟกราฟิกยังเป็น 16 รูปแบบการนำเสนอข้อมูลที่น่าสนใจ มีการนำเทคนิคการเล่าเรื่อง (Storytelling) มาใช้ทำให้ข้อมูล น่าสนใจ น่าดึงดูด จึงมักจะใช้เพื่อนำเสนอเนื้อหา ความรู้ หรือ เป็นสื่อการเรียนการสอน



ภาพที่ 2.4 ตัวอย่างอินโฟกราฟิก

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

5) แดชบอร์ด (Dashboards) คือ การนำข้อมูลต่างๆ มาเรียบเรียงและสรุปเป็นภาพ โดยใช้แผนภูมิและกราฟต่างๆ มาใช้นำเสนอ ปัจจุบันแดชบอร์ดเป็น Data Visualization ที่นิยมใช้กับการนำเสนอข้อมูลแบบ Real-time ผ่านซอฟต์แวร์หรือเครื่องมือจัดการและวิเคราะห์ข้อมูลต่างๆ เช่น เครื่องมือการตลาด เครื่องมือบริหารจัดการข้อมูล เครื่องมือติดตามและดูแลเว็บไซต์ ฯลฯ



ภาพที่ 2.5 แสดงตัวอย่างหน้าแดชบอร์ด

(ที่มา: <https://1stcraft.com/what-is-data-visualization/#8-6-แดชบอร์ด-dashboards>)

2.1.4 แนวคิดเกี่ยวกับการทำความสะอาดข้อมูล (Data Cleaning) การทำความสะอาดข้อมูล (Data Cleansing หรือ Data Cleaning) คือ กระบวนการตรวจสอบ แก้ไข หรือลบข้อมูลที่ไม่สมบูรณ์ ซ้ำซ้อน ผิดพลาด หรืออยู่ในรูปแบบที่ไม่เหมาะสม เพื่อให้ข้อมูลมีความถูกต้อง เป็นระเบียบและสามารถนำไปใช้วิเคราะห์ได้อย่างมีประสิทธิภาพ (ditto, 2567)

#### 2.1.4.1 ความสำคัญ

1) ช่วยให้ผลลัพธ์ของการวิเคราะห์แม่นยำขึ้น ลดความคลาดเคลื่อน (errors) จากข้อมูลที่ผิดหรือขาด

2) ทำให้ได้รายงาน (Insight / Report) ที่สามารถนำไปใช้สนับสนุนการตัดสินใจได้อย่างรวดเร็วและเชื่อถือได้

3) ป้องกันผลกระทบทางลบจากการตัดสินใจที่อาศัยข้อมูลที่ไม่สมบูรณ์

#### 2.1.4.2 ลักษณะของข้อมูลที่ต้องผ่านการทำความสะอาด

1) ข้อมูลที่อยู่ในหลายรูปแบบ (หลุดหลายไฟล์หรือหลายแหล่ง) ซึ่งอาจมีนามสกุลไฟล์ต่างกัน เช่น .pdf .doc .xls ฯลฯ

2) ข้อมูลที่ไม่ได้อยู่ในรูปแบบที่ต้องการ เช่น เป็นรูปภาพ (.jpg, .png) แต่ต้องการไฟล์ที่อ่านได้เป็นข้อความหรือไฟล์ structured (.csv, .json)

3) ข้อมูลที่ไม่ถูกต้อง เช่น มีการกรอกผิดสะกดผิด ข้อมูลที่เกินความเป็นจริง หรือมีความผิดพลาดจากการป้อนข้อมูลคน

### 2.1.3.3 ขั้นตอนการทำ Data Cleansing

- 1) ลบข้อมูลที่ซ้ำซ้อน (Duplicate Removal)
- 2) ลบข้อมูลที่ไม่เกี่ยวข้องกับวัตถุประสงค์การวิเคราะห์ (Irrelevant Data Removal)
- 3) แทนที่ข้อมูลที่หายไปหรือผิดพลาด (Data Imputation / Replacement)
- 4) บำรุงรักษาข้อมูลให้คงอยู่ในสภาพดีอยู่เสมอ (Data Maintenance)
- 5) ตรวจสอบความถูกต้องของข้อมูลหลังทำความสะอาด (Validation)

### 2.1.5 แนวคิดการจับกลุ่มข้อมูลด้วย K-Means

เทคนิค K-Means Clustering เป็นหนึ่งในวิธีการทำเหมืองข้อมูล (Data Mining) อยู่ในกลุ่มของ Unsupervised Learning ใช้สำหรับแบ่งข้อมูลออกเป็นกลุ่มย่อย (Clusters) โดยอาศัยหลักการหาจุดศูนย์กลาง (Centroid) ของแต่ละกลุ่ม และจัดกลุ่มข้อมูลให้มีความคล้ายคลึงกันมากที่สุดในแต่ละรอบของการทำงาน

#### 2.1.5.1 หลักการทำงานของ K-Means มีขั้นตอนดังนี้

- 1) กำหนดจำนวนกลุ่ม (K) ที่ต้องการแบ่ง และเลือกจุดศูนย์กลางเริ่มต้นจำนวน K จุด
- 2) นำข้อมูลทุกตัวมาเปรียบเทียบกับจุดศูนย์กลาง และจัดข้อมูลให้อยู่ในกลุ่มที่มีศูนย์กลางใกล้ที่สุด
- 3) คำนวณหาค่าเฉลี่ยใหม่ของข้อมูลในแต่ละกลุ่ม และใช้เป็นศูนย์กลางใหม่
- 4) ทำซ้ำขั้นตอนการจัดกลุ่มและการคำนวณศูนย์กลาง จนกว่าตำแหน่งของจุดศูนย์กลางจะไม่เปลี่ยนแปลงอีก

2.1.5.2 การประเมินจำนวนกลุ่มที่เหมาะสม ในการใช้งานจริง มักใช้ Elbow Method เพื่อหาจำนวน K ที่เหมาะสม โดยสังเกตจากกราฟค่าความคลาดเคลื่อน (Within-Cluster Sum of Squares: WCSS) ว่าจุดใดเป็น “ข้อศอก” ของกราฟ ซึ่งบ่งบอกถึงจำนวนกลุ่มที่เหมาะสมที่สุด (St, 2567)

## 2.2 ทฤษฎี

### 2.2.1 การทำเหมืองข้อมูล (Data Mining)

การทำเหมืองข้อมูล (Data Mining) คือ กระบวนการสกัดหาความรู้ จากฐานข้อมูลขนาดใหญ่ ด้วยการใช้เทคนิคทางสถิติ คณิตศาสตร์ และอัลกอริทึมของวิทยาการคอมพิวเตอร์ เพื่อค้นหารูปแบบ ความสัมพันธ์ และแนวโน้ม ที่ซ่อนอยู่ในข้อมูล ซึ่งไม่สามารถมองเห็นได้ด้วยการสังเกตทั่วไป

### 2.2.2 หลักในการทำเหมืองข้อมูล

เหมืองข้อมูลเป็นการวิเคราะห์ความสัมพันธ์ของข้อมูลจากฐานข้อมูลขนาดใหญ่ในอดีต เพื่อการค้นหากฎความสัมพันธ์ของข้อมูล การจัดกลุ่มข้อมูล และจำแนกประเภทข้อมูล เพื่อพยากรณ์สิ่งที่จะเกิดขึ้นในอนาคตโดยมีรายละเอียด ดังต่อไปนี้

1) การค้นหากฎความสัมพันธ์ของข้อมูล (Association Rule) เป็นเทคนิคที่ใช้ค้นหารูปแบบของเหตุการณ์ที่เกิดขึ้นร่วมกันบ่อยในข้อมูล เช่น การพบความสัมพันธ์ระหว่างช่วงเวลาเปิดระบบ (08.00–09.00 น.) กับจำนวน Error สูง หรือการพบว่าการลงทะเบียนรายวิชาเฉพาะกลุ่มทำให้ระบบมีภาระสูงกว่าปกติ การวิเคราะห์นี้ใช้ค่าสนับสนุน (Support) และค่าความเชื่อมั่น (Confidence) เพื่อคัดเลือกกฎที่มีนัยสำคัญ ช่วยให้มหาวิทยาลัยสามารถวางแผนแบ่งช่วงเวลาลงทะเบียน ลดความหนาแน่น และป้องกันปัญหาซ้ำซ้อน

2) การจัดกลุ่มข้อมูล (Clustering) ใช้จัดกลุ่มข้อมูลที่มีลักษณะคล้ายกัน เช่น กลุ่มนักศึกษาที่มีกลางทะเบียนวิชาหมวดเดียวกันพร้อมกัน หรือกลุ่มเวลาที่มีรูปแบบการใช้งานคล้ายกัน (Peak/Off-Peak) เทคนิคนี้ช่วยลดขนาดข้อมูลและทำให้เห็นโครงสร้างของปัญหาอย่างชัดเจน ช่วยให้สามารถวางแผนกระจายโหนดระบบ และบริหารทรัพยากรเซิร์ฟเวอร์ให้เหมาะสมกับแต่ละกลุ่ม

3) การจำแนกประเภทข้อมูล (Classification) เป็นเทคนิคที่ใช้จำแนกประเภทของปัญหาที่เกิดขึ้นตามคุณลักษณะที่กำหนด เช่น จำแนกว่า Error เกิดจากฝั่งเซิร์ฟเวอร์ (Server Error) ฝั่งเครือข่าย (Network Error) หรือจากฝั่งผู้ใช้ (User Error) ข้อมูลที่ได้จะช่วยให้ทีมเทคนิควางแผนแก้ไขตรงจุด เช่น เพิ่มความสามารถเซิร์ฟเวอร์ ปรับปรุง UX ของระบบ หรือเสริมคู่มือการใช้งานสำหรับนักศึกษา (TNIC, 2567)

2.2.3 กฎความสัมพันธ์ (Association Rule Mining) เป็นเทคนิคหนึ่งในกระบวนการทำเหมืองข้อมูล (Data Mining) ที่ใช้ค้นหารูปแบบความสัมพันธ์ระหว่างรายการข้อมูลในฐานข้อมูลขนาดใหญ่ โดยมีวัตถุประสงค์เพื่อค้นหาความสัมพันธ์ที่ซ่อนอยู่ภายในข้อมูล ซึ่งไม่สามารถมองเห็นได้จากการวิเคราะห์เชิงพรรณนาเพียงอย่างเดียว กฎที่ได้จะแสดงในรูปแบบเงื่อนไขในการประเมินคุณภาพของกฎความสัมพันธ์ จะใช้ตัวชี้วัดสำคัญ 3 ค่า ได้แก่

Support หมายถึง สัดส่วนของรายการข้อมูลทั้งหมดที่มี X และ Y เกิดร่วมกัน

Confidence หมายถึง ความน่าจะเป็นที่ Y จะเกิดขึ้นเมื่อ X เกิดขึ้นแล้ว

Lift หมายถึง ระดับความสัมพันธ์ที่แท้จริงระหว่าง X และ Y โดยหากมีค่ามากกว่า 1 แสดงว่าทั้งสองรายการมีความสัมพันธ์เชิงบวกต่อกัน เทคนิคกฎความสัมพันธ์ถูกนำไปประยุกต์ใช้ในหลากหลายด้าน เช่น การวิเคราะห์พฤติกรรมการซื้อสินค้า (Market Basket Analysis) การวิเคราะห์ข้อมูลด้านสุขภาพ และการวิเคราะห์ข้อมูลทางการศึกษา ในบริบทของการศึกษา แนวคิดดังกล่าวเป็นส่วนหนึ่งของสาขา Educational Data Mining ซึ่งมุ่งเน้นการค้นหารูปแบบพฤติกรรมของผู้เรียนจากข้อมูลขนาดใหญ่ เพื่อสนับสนุนการตัดสินใจเชิงนโยบายและการปรับปรุงระบบสารสนเทศทางการศึกษา (Mindphp, 2562)

สำหรับโครงการการวิเคราะห์ปัญหาการลงทะเบียนเรียนออนไลน์ กฎความสัมพันธ์สามารถประยุกต์ใช้เพื่อค้นหารูปแบบการเลือกลงทะเบียนรายวิชาของนักศึกษา เช่น ความสัมพันธ์ระหว่างรายวิชาที่มักถูกลงทะเบียนพร้อมกัน ความสัมพันธ์ระหว่างช่วงเวลาเข้าใช้งานระบบกับโอกาสเกิดข้อผิดพลาด หรือความสัมพันธ์ระหว่างจำนวนหน่วยกิตที่ลงทะเบียนกับพฤติกรรมการยกเลิกรายวิชา ข้อมูลดังกล่าวสามารถนำไปใช้ในการวางแผนการเปิดรายวิชา การบริหารจัดการที่นั้งเรียน และการปรับปรุงประสิทธิภาพของระบบลงทะเบียนให้รองรับปริมาณผู้ใช้งานได้อย่างเหมาะสม

2.2.4 การวิเคราะห์เชิงพรรณนา (Descriptive Analytics) วิธีการทางสถิติที่ใช้ในการสรุปผล จัดระเบียบ และอธิบายคุณลักษณะของชุดข้อมูลที่เก็บรวบรวมมาได้ โดยมุ่งเน้นการนำเสนอข้อเท็จจริงที่ปรากฏในข้อมูลโดยมิได้มีการอ้างอิงไปถึงประชากรส่วนใหญ่ในลักษณะการทดสอบสมมติฐาน (Inference) แต่เป็นการทำให้ข้อมูลตีความหมายและเข้าใจง่ายขึ้น

#### 2.2.4.1 รูปแบบของสถิติเชิงพรรณนาที่ใช้ในการวิจัย

การแจกแจงความถี่และร้อยละ (Frequency and Percentage) ใช้สำหรับการอธิบายข้อมูลทั่วไปของกลุ่มตัวอย่าง (Demographic Data) เช่น เพศ ระดับการศึกษา หรือสถานภาพ เพื่อเปรียบเทียบสัดส่วนของข้อมูลแต่ละกลุ่มเทียบกับภาพรวมทั้งหมด การวัดค่ากลางของข้อมูล

(Measures of Central Tendency) ค่าเฉลี่ยเลขคณิต ( $\bar{x}$ ) ใช้เป็นตัวแทนของข้อมูลเชิงปริมาณเพื่อแสดงระดับความคิดเห็นหรือค่าเฉลี่ยของตัวแปรที่ศึกษา

1) ฐานนิยม (Mode) และมัธยฐาน (Median): ใช้ในกรณีที่ต้องการทราบค่าที่มีความถี่สูงสุดหรือค่าที่อยู่ตำแหน่งกลางของชุดข้อมูล (หากจำเป็น)

2) การวัดการกระจายของข้อมูล (Measures of Dispersion) ส่วนเบี่ยงเบนมาตรฐาน (S.D.): ใช้เพื่ออธิบายการกระจายตัวของข้อมูลว่ามีความสอดคล้องหรือแตกต่างกันเพียงใด หากค่า S.D. ต่ำ แสดงว่ากลุ่มตัวอย่างมีความคิดเห็นไปในทิศทางเดียวกัน

## 2.2.5 ทฤษฎีที่เกี่ยวข้องกับการสร้างเว็บไซต์

### 2.2.5.1 การสร้างเว็บไซต์สิ่งสำคัญอยู่ที่การออกแบบเว็บเพราะเว็บไซต์ที่มี

รูปแบบสวยงามจะสามารถดึงดูดความสนใจจากผู้ คนได้ดีกว่าทำให้ผู้ คนเกิดความรู้สึกประทับใจอยากกลับมาใช้งานเว็บไซต์อีกครั้งในอนาคต ดังนั้นเริ่มแรกก่อนทำเว็บไซต์จึงจำเป็นต้องทำความเข้าใจกับหลักการออกแบบและรูปแบบโครงสร้างของเว็บก่อน การออกแบบเว็บไซต์เพื่อให้มีประสิทธิภาพและสามารถดึงดูดความสนใจของผู้ คนได้ดีจะต้องมีองค์ประกอบของเว็บไซต์อย่างครบถ้วน ซึ่งได้แก่ต่าง ๆ ได้เองเว็บไซต์ที่ได้รับการออกแบบอย่างสวยงาม มีการใช้งานที่สะดวกย่อม ได้รับความสนใจจากผู้ ใช้ มากกว่าเว็บไซต์ที่ดูสับสนวุ่นวาย มีข้อมูลมากมายแต่หาอะไรไม่เจอ นอกจากนี้ยังใช้เวลาในการแสดงผลแต่ละหน้านานเกินไป ซึ่งปัญหาเหล่านี้ล้วนเป็นผลมาจาก การออกแบบเว็บไซต์ไม่ดีทั้งสิ้น ดังนั้น การออกแบบเว็บไซต์จึงเป็นกระบวนการสำคัญในการสร้างเว็บไซต์ ให้ประทับใจผู้ ใช้ ทำให้เขาอยากกลับมาเข้ามาเว็บไซต์เดิมอีกในอนาคต ซึ่งนอกจากต้องพัฒนาเว็บไซต์ที่ดีมีประโยชน์แล้ว ยังต้องคำนึงถึงการแข่งขันกับเว็บไซต์อื่น ๆ อีกด้วย องค์ประกอบของ การออกแบบเว็บไซต์ การออกแบบเว็บไซต์ที่มีประสิทธิภาพ นั้นต้องคำนึงถึง องค์ประกอบ สำคัญดังต่อไปนี้

1. ความเรียบง่าย (Simplicity) การออกแบบเว็บไซต์ในรูปแบบที่เรียบง่ายไม่ซับซ้อน และใช้งานอย่างสะดวก

2. ความสม่ำเสมอ (Consistency) การสร้างความสม่ำเสมอให้เกิดขึ้นตลอดทั้งเว็บไซต์ โดยอาจเลือกใช้รูปแบบเดียวกันตลอดทั้งเว็บไซต์

3. ความเป็นเอกลักษณ์ (Identity) เว็บไซต์จะสะท้อนถึงเอกลักษณ์และลักษณะขององค์กร การเลือกใช้ตัวอักษร ชุดสี รูปภาพหรือกราฟิกจะมีผลต่อรูปแบบของเว็บไซต์

4. เนื้อหา (Useful Content) เนื้อหาในเว็บไซต์ต้องสมบูรณ์และได้รับการปรับปรุงพัฒนาให้ทันสมัยอยู่เสมอ

5. ระบบเนวิเกชัน (User-Friendly Navigation) ระบบเนวิเกชันช่วยไม่ให้ผู้ใช้เกิด ความสับสนระหว่างดูเว็บไซต์ จึงเปรียบเสมือนป้ายบอกทาง ควรทำให้เข้าใจง่ายและใช้งานได้สะดวก

6. คุณภาพของสิ่งที่ปรากฏให้เห็นในเว็บไซต์ (Visual Appeal) ลักษณะที่น่าสนใจ ของเว็บไซต์นั้น ควรมีคุณภาพ เช่น กราฟิกควรสมบูรณ์ไม่มีรอยหรือขอบขั้นบันไดให้เห็น ชนิดตัวอักษรอ่านง่ายสบายตา มีการเลือกใช้โทนสีที่เข้ากันอย่างสวยงาม เป็นต้น

7. ความสะดวกของการใช้ในสภาพต่าง ๆ (Compatibility) สามารถใช้งานได้ดีใน สภาพแวดล้อมที่หลากหลาย ไม่มีการบังคับให้ผู้ใช้ต้องติดตั้งโปรแกรมอื่นใดเพิ่มเติม นอกเหนือจากเว็บเบราว์เซอร์

8. ความคงที่ในการออกแบบ (Design Stability) ให้ความสำคัญกับการ ออกแบบเว็บไซต์ ต้องออกแบบวางแผนและเรียบเรียงเนื้อหาอย่างรอบคอบ

9. ความคงที่ของการทำงาน (Function Stability) ระบบการทำงานต่าง ๆ ใน เว็บไซต์ควรมีความถูกต้องแน่นอน และตรวจสอบอยู่เสมอ

2.2.5.2 รูปแบบโครงสร้างของเว็บไซต์ การออกแบบโครงสร้างของเว็บไซต์ สามารถทำได้หลากหลายแบบ ซึ่งก็ขึ้นอยู่กับความชอบและความถนัดของแต่ละบุคคล นอกจากนี้ยังขึ้นอยู่กับกลุ่มเป้าหมายที่ต้องการนำเสนอเพราะจะต้องออกแบบให้เหมาะกับการใช้งานของกลุ่มเป้าหมายมากที่สุด โดยโครงสร้างของเว็บไซต์ส่วนใหญ่ก็จะประกอบไปด้วย 4 รูปแบบดังนี้

## 1. โครงสร้างแบบเรียงลำดับ



ภาพที่ 2.6 แสดงโครงสร้างเว็บไซต์แบบเรียงลำดับ

(ที่มา: <https://jacktnp.github.io/webtech/category/basic/3.html>)

โครงสร้างเว็บไซต์แบบเรียงลำดับจะเป็นโครงสร้างแบบธรรมดาที่นิยมใช้งานกันมากที่สุด เนื่องจากมีความง่ายต่อการจัดระบบข้อมูลและสามารถนำเสนอเรื่องราวตามลำดับได้ เป็นอย่างดี เหมาะกับเว็บไซต์ที่มีขนาดเล็กมีเนื้อหาที่ไม่ซับซ้อน ส่วนใหญ่ก็จะเป็นพวกเว็บไซต์ ที่ให้ความรู้ หรือเว็บไซต์องค์กรขนาดย่อม โดยลักษณะการลิ้งค์เนื้อหา ก็จะลิ้งค์ไปที่ละหน้า มีทิศทาง การเข้าสู่เนื้อหาต่างๆในแบบเส้นตรงใช้ปุ่มเดินหน้า-ถอยหลังในการกำหนดทิศทางจึง ทำให้การใช้งานเป็นไปอย่างง่าย แต่โครงสร้างเว็บไซต์แบบเรียงลำดับก็มีข้อเสียคือจะทำให้ ผู้ใช้งานต้องเสียเวลาในการเข้าสู่เนื้อหาเพราะไม่สามารถกำหนดทิศทาง การเข้าสู่เนื้อหาด้วยตัวเองได้

## 2. โครงสร้างแบบลำดับชั้น



ภาพที่ 2.7 แสดงโครงสร้างแบบลำดับชั้น

(ที่มา: <https://jacktnp.github.io/webtech/category/basic/3.html>)

เป็นวิธีที่ดีที่สุดวิธีหนึ่งในการจัดระบบโครงสร้างที่มีความซับซ้อนของข้อมูล โดยแบ่งเนื้อหา ออกเป็นส่วนต่างๆ และมีรายละเอียดย่อยๆ ในแต่ละส่วนลดหลั่นกันมาในลักษณะแนวคิด

เดียวกับ แผนภูมิองค์กร จึงเป็นการถ่ายทอดการทำหน้าที่กับโครงสร้างของเนื้อหาในเว็บไซต์  
ลักษณะนี้ ลักษณะเด่นเฉพาะของ เว็บประเภทนี้คือการมีจุดเริ่มต้นที่จุดร่วมจุดเดียว นั่นคือ  
โฮมเพจ

### 3. โครงสร้างแบบตาราง



ภาพที่ 2.8 แสดงโครงสร้างแบบตาราง

(ที่มา: <https://jacktnp.github.io/webtech/category/basic/3.html>)

โครงสร้างรูปแบบนี้มีความซับซ้อนมากกว่ารูปแบบที่ผ่านมา การออกแบบเพิ่มความยืดหยุ่น  
ให้แก่การเข้าสู่เนื้อหาของผู้ใช้ โดยเพิ่มการเชื่อมโยงซึ่งกันและกันระหว่างเนื้อหาแต่ละส่วน  
เหมาะแก่ การแสดงให้เห็นความสัมพันธ์กันของเนื้อหา การเข้าสู่เนื้อหาของผู้ใช้จะไม่เป็น  
ลักษณะเชิงเส้นตรง เนื่องจากผู้ใช้สามารถเปลี่ยนทิศทางการเข้าสู่เนื้อหาของตนเองได้

### 4. โครงสร้างแบบใยแมงมุม



ภาพที่ 2.9 แสดงโครงสร้างแบบใยแมงมุม

(ที่มา: <https://jacktnp.github.io/webtech/category/basic/3.html>)

โครงสร้างแบบใยแมงมุม เป็นโครงสร้างที่มีความยืดหยุ่นมากที่สุด เพราะทุกหน้าบนเว็บไซต์  
สามารถเชื่อมโยงหากันได้ทั้งหมดทั้งภายในเว็บไซต์หรือแม้กระทั่งเชื่อมโยงไปสู่เว็บไซต์  
ภายนอกได้ ทำให้ผู้ใช้สามารถเลือกที่จะเข้าสู่เนื้อหาได้อย่างอิสระด้วยตนเอง

2.2.5.3 ส่วนประกอบสำคัญของหน้าเว็บเพจ บนหน้าเว็บเพจ จะมี ส่วนประกอบ สำคัญ ที่จำเป็นต้องมีอยู่ 3 ส่วน

1) ส่วนหัวของหน้า (Header) อยู่ตอนบนสุดของหน้าและเป็นส่วนที่สำคัญที่สุด โดยจะต้องทำให้สามารถดึงดูดผู้ชมให้รู้สึกอยากติดตามเนื้อหาในเว็บไซด์ต่อไป ซึ่งส่วนใหญ่มักจะมีการใส่ภาพกราฟิกให้ดูสวยงาม สิ่งสำคัญหลักๆเลยก็คือ โลโก้ ชื่อเว็บไซด์และเมนู หลัก ที่สามารถลิงค์ไปยังเนื้อหาในหน้าเว็บเพจต่างๆได้

2) ส่วนของเนื้อหา (Body) อยู่บริเวณตอนกลางของหน้าเว็บ โดยจะแสดงข้อมูลเกี่ยวกับเนื้อหาบนเว็บแบบคร่าวๆ ซึ่งก็จะมีข้อความกราฟิก ตารางข้อมูลหรือวิดีโอ ประกอบ อยู่และหากมีเมนูแบบเฉพาะกลุ่มก็จะถูกจัดไว้ในหน้านี้เช่นกันและที่สำคัญเนื้อหาในส่วนนี้ควร จะมีความกระชับ เข้าใจง่าย มีการใช้รูปแบบตัวอักษรแบบเรียบง่ายและเป็นระเบียบ

3) ส่วนท้ายของหน้า (Footer) อยู่ล่างสุดของหน้าเว็บซึ่งจะมีหรือไม่มีก็ได้ ส่วนนี้จะ แสดงถึงข้อมูลต่างๆเพิ่มเติมเข้าไป เช่น ข้อความที่แสดงถึงการเป็นลิขสิทธิ์ข้อมูลเจ้าของ เว็บไซด์ วิธีการติดต่อและคำแนะนำต่างๆเกี่ยวกับการใช้งานเว็บไซด์อย่างถูกต้อง

2.2.5.4 วิธีการเลือกใช้สีสำหรับการออกแบบเว็บไซด์การเลือกใช้สีในการ ออกแบบ เว็บไซด์มีความสำคัญเป็นอย่างมาก เพราะสีสามารถกำหนดอารมณ์ ความรู้สึก และ กระตุ้น การรับรู้ทางด้านจิตใจของมนุษย์ได้ดี ดังนั้นสีที่ใช้จึงต้องมีความสอดคล้องกับ เนื้อหา และจุดประสงค์ของเว็บว่าต้องการให้ผู้เข้าชมรู้สึกอย่างไรต่อเนื้อหาที่ได้อ่าน โดย รูปแบบของสี ที่สายตาของมนุษย์สามารถมองเห็นได้ก็แบ่งออกเป็น 3 กลุ่ม

1) สีโทนร้อน (Warm Colors) เป็นสีแห่งความอบอุ่น ปลอดภัยและกระตุ้นความสุขได้ดี ซึ่งจะทำให้ผู้เข้าชมรู้สึกมีชีวิตชีวาและมีแรงผลักดันมากขึ้น อีกทั้งยังช่วยดึงดูดให้ผู้ชมรู้สึกอยากติดตามเนื้อหามากขึ้น

2) สีโทนเย็น (Cool Colors) เป็นสีแห่งความสุภาพและความอ่อนโยน ทำให้ผู้ชมรู้สึกผ่อนคลายและเพลิดเพลินมากขึ้นและยังสามารถใช้โน้มน้าวจากในระยะไกลได้อีกด้วย

3) สีโทนกลาง (Neutral Colors) สีเหล่านี้ มักจะถูกนำไปผสมกับสีอื่นๆ เพื่อให้เกิดสีที่เป็นกลางมากขึ้น และให้ความรู้สึกที่เป็นธรรมชาติ

### 2.3 วรรณกรรมที่เกี่ยวข้อง

เตชะประสิทธิ์วานิช พ. (2020) ศึกษาเรื่อง *กระบวนการสร้างความรู้ ความเข้าใจในการใช้ระบบทะเบียนออนไลน์ในระดับอุดมศึกษา* ตีพิมพ์ในวารสารนวัตกรรมการศึกษาและการวิจัย ปีที่ 3 ฉบับที่ 2 หน้า 121-130 โดยมีวัตถุประสงค์เพื่อวิเคราะห์กระบวนการส่งเสริมความเข้าใจและพฤติกรรมการใช้งานระบบทะเบียนออนไลน์ของนักศึกษาในมหาวิทยาลัย งานวิจัยพบว่าปัจจัยด้านความเข้าใจระบบ ความชัดเจนของขั้นตอนการใช้งาน และการสื่อสารข้อมูลจากมหาวิทยาลัย มีผลต่อประสิทธิภาพการใช้งานระบบอย่างมีนัยสำคัญ อีกทั้งยังพบปัญหาที่เกิดขึ้นบ่อยในช่วงเปิดระบบ เช่น ความล่าช้าในการประมวลผล และความสับสนในขั้นตอนการลงทะเบียน ผลการศึกษาดังกล่าวสะท้อนให้เห็นว่าปัญหาของระบบไม่ได้เกิดจากปัจจัยทางเทคนิคเพียงอย่างเดียว แต่รวมถึงมิติด้านการรับรู้และประสบการณ์ของผู้ใช้ ซึ่งสอดคล้องกับแนวทางของโครงการนี้ที่มุ่งวิเคราะห์ปัญหาการลงทะเบียนออนไลน์ทั้งเชิงเทคนิคและเชิงพฤติกรรมผู้ใช้งาน

จิตรลดา นาคประเสริฐ (2562) ศึกษาเรื่อง *ปัญหาและแนวทางการพัฒนาระบบสารสนเทศเพื่อการศึกษาในมหาวิทยาลัย* โดยมุ่งวิเคราะห์ข้อจำกัดของระบบสารสนเทศด้านโครงสร้างพื้นฐาน เทคโนโลยี และการบริหารจัดการข้อมูล ผลการศึกษาพบว่าปัญหาหลักของระบบสารสนเทศในมหาวิทยาลัยมักเกิดจากการออกแบบฐานข้อมูลที่ไม่รองรับปริมาณผู้ใช้งานจำนวนมาก และการขาดแผนรองรับการใช้งานในช่วงเวลาที่มีความต้องการสูง งานวิจัยนี้ช่วยสนับสนุนกรอบแนวคิดเชิงเทคนิคของโครงการ โดยเฉพาะในประเด็นความเสถียรของระบบและการบริหารจัดการทรัพยากรสารสนเทศ

ศุภชัย ทองดี (2563) ศึกษาเรื่อง *การวิเคราะห์และออกแบบระบบลงทะเบียนเรียนออนไลน์สำหรับสถานศึกษาขนาดกลาง* โดยใช้แนวทางวิเคราะห์ระบบ (System Analysis and Design) เพื่อศึกษาปัญหาที่เกิดขึ้นในระบบเดิม และออกแบบโครงสร้างระบบใหม่ให้รองรับการใช้งานได้อย่างมีประสิทธิภาพ ผลการศึกษาชี้ว่าปัญหาที่พบส่วนใหญ่เกี่ยวข้องกับการประมวลผลคำร้องขอพร้อมกันจำนวนมาก และการจัดลำดับความสำคัญของข้อมูล งานวิจัยนี้ให้กรอบเชิงเทคนิคที่ชัดเจนเกี่ยวกับโครงสร้างระบบลงทะเบียน ซึ่งสามารถนำมาใช้เป็นพื้นฐานในการวิเคราะห์องค์ประกอบของระบบในโครงการนี้

วัชรพงษ์ เจริญกิจ และ บุญญาเนตร ศรีสุข (2561) ศึกษาการวิเคราะห์ข้อความ ร้องเรียนของนักศึกษาในระบบ e-Learning โดยใช้การวิเคราะห์เนื้อหา (Content Analysis) เพื่อ จำแนกประเภทของปัญหาที่ผู้ใช้งานสะท้อนผ่านระบบออนไลน์ ผลการวิจัยพบว่าปัญหาส่วนใหญ่เกี่ยวข้องกับความปลอดภัยในการเข้าถึงระบบ ความไม่ชัดเจนของข้อมูล และการสื่อสารที่ไม่ ครบถ้วน งานวิจัยนี้สะท้อนมิติด้านประสบการณ์ผู้ใช้ ซึ่งมีความสำคัญต่อการทำความเข้าใจ ปัญหาการลงทะเบียนออนไลน์ในเชิงคุณภาพ และสามารถใช้เป็นแนวทางในการออกแบบ คำถามหรือการวิเคราะห์ข้อมูลเชิงข้อความในโครงการงานนี้

ณัฐพล สุขสวัสดิ์ และ ภัคพร เทพวงศ์ (2565) ศึกษาการวิเคราะห์พฤติกรรมการใช้งานระบบสารสนเทศนักศึกษาโดยใช้ข้อมูล Log โดยเก็บข้อมูลการเข้าใช้งานระบบตลอดภาค การศึกษา และนำมาวิเคราะห์เชิงพรรณนาเพื่อหาช่วงเวลาที่มีการใช้งานสูงสุดและรูปแบบ พฤติกรรมการใช้งาน ผลการศึกษาพบว่าช่วงเปิดระบบลงทะเบียนเป็นช่วงเวลาที่มียอดการใช้งาน เข้าใช้งานหนาแน่นสูงสุด และมีการร้องขอข้อมูลซ้ำจำนวนมาก งานวิจัยนี้มีความสอดคล้อง โดยตรงกับแนวทางของโครงการงานที่ใช้ข้อมูล Log ในการวิเคราะห์ปัญหาเชิงพฤติกรรมของ ผู้ใช้งาน

วรพจน์ สมสุข และ สุวิมล วิริยะ (2564) ศึกษาการใช้เทคนิค Data Mining วิเคราะห์ พฤติกรรมการใช้งานระบบสารสนเทศของนักศึกษา โดยประยุกต์ใช้เทคนิคเหมืองข้อมูลเพื่อ ค้นหารูปแบบและความสัมพันธ์ของพฤติกรรมผู้ใช้ ผลการศึกษาชี้ให้เห็นว่าเทคนิคการทำ เหมืองข้อมูลสามารถช่วยระบุรูปแบบการใช้งานที่ซ่อนอยู่ และสนับสนุนการตัดสินใจเชิง บริหารในการพัฒนาระบบ งานวิจัยนี้สนับสนุนแนวทางการประยุกต์ใช้เทคนิควิเคราะห์ข้อมูล เช่น การค้นหากฎความสัมพันธ์ (Association Rules) ซึ่งเป็นเครื่องมือสำคัญในโครงการฉบับนี้

## 2.4 เครื่องมือในการออกแบบและวิเคราะห์ข้อมูล

### 2.4.1 กระบวนการวิเคราะห์ข้อมูลด้วย (CRISP-DM)

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

1) Business Understanding คือ การทำความเข้าใจในงาน ระบุโอกาส และหาปัญหาที่จะเกิดขึ้นกับธุรกิจ กำหนดขอบเขตของข้อมูลที่จะนำวิเคราะห์หาความได้เปรียบทางการตลาด และแก้ไขปัญหาคอร์ป ซึ่งต้องสามารถระบุผลลัพธ์ที่มีได้

2) Data Understanding ทำความเข้าใจข้อมูลโดยการรวบรวมข้อมูลที่เกี่ยวข้อง คัดเลือกให้เหลือเพียงข้อมูลที่มีความถูกต้อง และสำคัญต่องานมาทำการวิเคราะห์

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

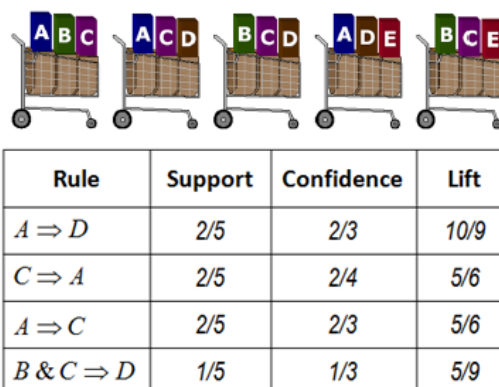
4) Modeling การสร้างแบบจำลองเพื่อวิเคราะห์ข้อมูลที่ได้จากขั้นตอนที่ 3 พร้อม ทดสอบผลลัพธ์แบบจำลองเพื่อให้ได้คำตอบที่ดีที่สุด บางครั้งอาจมีการย้อนกลับไปปรับการเตรียมข้อมูลเพื่อให้ได้แบบจำลองที่เหมาะสมที่สุด

5) Evaluation การประเมินผลลัพธ์ที่ได้ก่อนที่จะนำไปใช้จริง ว่าตรงกับวัตถุประสงค์ หรือเป้าหมายที่ได้ตั้งไว้ หรือมีความน่าเชื่อถือมากน้อยเพียงใด หากไม่ได้ผลลัพธ์ตามวัตถุประสงค์ ต้องย้อนกลับไปปรับปรุงแก้ไขการดำเนินงานในขั้นตอนก่อนหน้า

6) Deployment การนำเอาข้อมูลที่เป็นผลลัพธ์จากทั้งหมด มาลงปฏิบัติจริงกับธุรกิจ ในองค์กร โดยแปลงแนวคิดที่มีให้เกิดเป็นสารสนเทศเพื่อให้ผู้บริหารสามารถนำไปใช้ประโยชน์ในทางธุรกิจได้จริง และติดตามประเมินผลที่ได้เพื่อนำกลับไปปรับปรุง Data Mining ต่อเนื่องต่อไป ซึ่งการประเมินผลสามารถทำได้หลายทาง เช่น การสร้าง Dashboard แสดงจำนวนผู้ใช้งานแบบเรียลไทม์ การแจ้งเตือนช่วงเวลาเสี่ยง การปรับตารางเวลาการเปิดลงทะเบียน และการเพิ่มขีดความสามารถของเซิร์ฟเวอร์

2.4.2 เทคนิคกฎความสัมพันธ์ (Association Rule Analysis) เพื่อค้นหารูปแบบความสัมพันธ์ระหว่างเหตุการณ์หรือพฤติกรรมที่เกิดขึ้นภายในระบบ เช่น การเลือกลงทะเบียนรายวิชาพร้อมกัน ช่วงเวลาเข้าใช้งานที่สัมพันธ์กับปัญหาระบบ หรือรูปแบบการเปลี่ยนแปลงรายวิชา การวิเคราะห์ดำเนินการโดยใช้หลักการของอัลกอริทึม Apriori ซึ่งเป็นวิธีที่ใช้ค้นหาชุดข้อมูลที่เกิดร่วมกันบ่อย (Frequent Itemsets) จากนั้นจึงสร้างกฎความสัมพันธ์ในรูปแบบ  $X \rightarrow$

Y เพื่ออธิบายแนวโน้มของเหตุการณ์ที่เกิดขึ้นร่วมกันในกระบวนการวิเคราะห์ มีการกำหนดค่าพารามิเตอร์สำคัญ ได้แก่ Minimum Support เพื่อคัดกรองความถี่ขั้นต่ำของเหตุการณ์ที่เกิดขึ้นร่วมกัน Minimum Confidence เพื่อกำหนดระดับความน่าเชื่อถือของกฎ การพิจารณาค่า Lift เพื่อประเมินความสัมพันธ์ที่แท้จริงระหว่างตัวแปร ข้อมูลที่ใช้วิเคราะห์ประกอบด้วยข้อมูล Log File การใช้งานระบบ และข้อมูลการลงทะเบียนรายวิชาของนักศึกษา ซึ่งถูกจัดเตรียมให้อยู่ในรูปแบบ Transaction Data ก่อนนำเข้าสู่กระบวนการคำนวณกฎความสัมพันธ์



ภาพที่ 2.10 แสดงข้อมูลของกฎความสัมพันธ์

ผลลัพธ์จากการวิเคราะห์จะได้ชุดกฎที่แสดงรูปแบบพฤติกรรม เช่น รายวิชาที่มักถูกเลือกลงทะเบียนพร้อมกัน ช่วงเวลาที่มีแนวโน้มเกิดข้อผิดพลาดสูง รูปแบบที่สัมพันธ์กับการยกเลิกรายวิชา ผลการวิเคราะห์ดังกล่าวถูกนำไปใช้ในการเสนอแนวทางปรับปรุงระบบลงทะเบียน เช่น การวางแผนเปิดรายวิชา การเพิ่มทรัพยากรระบบในช่วงเวลาที่มีความหนาแน่นสูง และการปรับโครงสร้างกระบวนการลงทะเบียนเพื่อลดปัญหาที่เกิดขึ้นซ้ำ

## 2.5 บทสรุป

จากการทบทวนแนวคิด ทฤษฎี เครื่องมือ และวรรณกรรมที่เกี่ยวข้องกับการวิเคราะห์ปัญหาการลงทะเบียนเรียนออนไลน์ของนักศึกษา พบว่าการประยุกต์ใช้เทคนิคการทำเหมืองข้อมูล (Data Mining) โดยเฉพาะการวิเคราะห์กฎความสัมพันธ์ (Association Rules) ร่วมกับการวิเคราะห์เชิงสถิติ เป็นแนวทางที่เหมาะสมในการค้นหารูปแบบ ความสัมพันธ์ และปัจจัยที่เกี่ยวข้องกับการเกิดปัญหาในระบบ ซึ่งช่วยให้สามารถอธิบายสถานการณ์ที่เกิดขึ้นได้อย่างมีหลักฐานเชิงข้อมูลรองรับการดำเนินงานวิจัยกำหนดกรอบกระบวนการตามแนวทาง CRISP-DM เพื่อให้การวิเคราะห์เป็นระบบและตรวจสอบได้ ครอบคลุมตั้งแต่การทำความเข้าใจปัญหา และวัตถุประสงค์ การรวบรวมและเตรียมข้อมูลจากระบบลงทะเบียน การสร้างและทดสอบแบบจำลอง การประเมินผลความเหมาะสมของแบบจำลอง ตลอดจนการนำผลลัพธ์ไปประยุกต์ใช้จริงในเชิงปฏิบัติ ผลการวิเคราะห์จะนำเสนอในรูปแบบกราฟ แผนภูมิ และแดชบอร์ด เพื่อสนับสนุนการตัดสินใจของผู้ที่เกี่ยวข้อง ช่วยให้สามารถมองเห็นรูปแบบการใช้งาน และปัจจัยเสี่ยงได้ชัดเจน อันนำไปสู่ข้อเสนอแนะแนวทางปรับปรุงระบบให้มีประสิทธิภาพ เสถียรภาพ และตอบสนองต่อความต้องการของนักศึกษาได้ดียิ่งขึ้น