

## บทที่ 2

### แนวคิด ทฤษฎี เครื่องมือและวรรณกรรมที่เกี่ยวข้อง

โครงการเรื่อง การวิเคราะห์ข้อมูลการแข่งขันกีฬาโอลิมปิกเพื่อใช้สำหรับเผยแพร่ข้อมูลบนเว็บไซต์ ในบทนี้จะเป็นการนำเสนอเกี่ยวกับ แนวคิด ทฤษฎี เครื่องมือและวรรณกรรมที่เกี่ยวข้องของการวิเคราะห์ข้อมูลการแข่งขันกีฬาโอลิมปิกเพื่อใช้สำหรับเผยแพร่ข้อมูลบนเว็บไซต์ ซึ่งได้รวบรวมการศึกษาเอกสารงานวิจัยที่เกี่ยวข้องกับการวิเคราะห์ เพื่อใช้เป็นแนวทางการศึกษาประกอบและรายละเอียดตามลำดับ ดังนี้

#### 2.1 แนวคิด

- 2.1.1 แนวคิดเกี่ยวกับการวิเคราะห์ข้อมูล ( Data analytic )
- 2.1.2 แนวคิดเกี่ยวกับการแสดงผลข้อมูล ( Data visualization )
- 2.1.3 แนวคิดเกี่ยวกับการพยากรณ์ข้อมูล ( Forecasting )
- 2.1.4 แนวคิดเกี่ยวกับข้อมูลอนุกรมเวลา ( Data Time Series )
- 2.1.5 การวิเคราะห์อนุกรมเวลา ( Time Series Analysis )

#### 2.2 ทฤษฎี

- 2.2.1 ทฤษฎีเกี่ยวกับการออกแบบเว็บไซต์
- 2.2.2 ทฤษฎีเกี่ยวกับการทำเหมืองข้อมูล
- 2.2.3 ทฤษฎีเกี่ยวกับการ visualization
- 2.2.4 ทฤษฎีเกี่ยวกับการต้นไม้ตัดสินใจ

## 2.3 เครื่องมือในการออกแบบและวิเคราะห์ข้อมูล

2.3.1) การแสดงผลแบบ visualization

2.3.2) แบบจำลองอนุกรมเวลา ( Time Series Analysis )

2.3.3) กระบวนการทำเหมืองข้อมูล ( CRISP-DM)

## 2.4 วรรณกรรมที่เกี่ยวข้อง

### 2.1 แนวคิด

#### 2.1.1 แนวคิดเกี่ยวกับการวิเคราะห์ข้อมูล ( Data analytic )

การวิเคราะห์ข้อมูล หมายถึง การนำข้อมูลที่เก็บรวบรวมได้ จากกลุ่มตัวอย่างหรือจากประชากรการวิจัยจำนวนหนึ่ง มาจำแนกเพื่อตอบประเด็นปัญหาการวิจัย หรือทดสอบสมมุติฐานการวิจัยให้ครบทุกข้อ ถ้าข้อมูลเชิงปริมาณหรือเป็นตัวเลข ผู้วิจัยจะใช้วิธีการทางสถิติสรุปรวมข้อมูล แต่ถ้าเป็นข้อมูลเชิงคุณภาพ ผู้วิจัยจะใช้วิธีการสรุปความ หรือสังเคราะห์ข้อความ

การวิเคราะห์ข้อมูลในการวิจัยเชิงคุณภาพ เป็นกระบวนการประกอบด้วย การจำแนกและจัดระบบข้อมูล และการหาความสัมพันธ์ของข้อมูล มีจุดมุ่งหมายที่จะแยกและอธิบายองค์ประกอบความหมาย และความสัมพันธ์ ภายใต้เงื่อนไขและสภาพแวดล้อมทางสังคมและวัฒนธรรม ซึ่งการใช้หลักการของการวิจัยเชิงคุณภาพในการวิเคราะห์ข้อมูล นั้นคือการมองภาพรวม การอธิบายเงื่อนไขและสภาพแวดล้อมทางสังคมและวัฒนธรรมของเหตุการณ์ การนำเอาทัศนคติของผู้ให้ข้อมูลมาอธิบายเหตุการณ์ โดยให้ความสำคัญต่อมุมมองของผู้ให้ข้อมูล ซึ่งถือเป็นผู้อยู่ในเหตุการณ์หรืออยู่ในเหตุการณ์นั้น ๆ วิธีการวิเคราะห์ข้อมูลในการวิจัยเชิงคุณภาพ อาจใช้การวิเคราะห์ข้อมูลแบบสร้างข้อสรุป หรือการวิเคราะห์เนื้อหา ความน่าเชื่อถือของงานวิจัยเชิงคุณภาพ พิจารณาได้จากความเชื่อถือได้ การถ่ายโอนผลการวิจัย การพึ่งพากับเกณฑ์อื่น และการยืนยันผล

การวิเคราะห์ข้อมูลเชิงปริมาณเป็นวิธีการวิเคราะห์ที่จะต้องใช่วิธีการทางสถิติช่วยสรุปรวมข้อมูลเพื่อ ตอบประเด็นปัญหาการวิจัยต่างๆ วิธีการทางสถิติแบ่งได้เป็น 2 ประเภทคือ สถิติบรรยาย และสถิติอ้างอิง ก่อนที่จะกล่าวถึงรายละเอียดของสถิติแต่ละประเภท จะขอกล่าวถึงการใช่วิธีการทางสถิติในการวิจัยก่อนดังนี้ การหาค่าสถิติต่าง ๆ ในปัจจุบันผู้วิจัยไม่จำเป็นต้องคำนวณหาค่าโดยการแทนค่าลงในสูตร เพราะเรามี โปรแกรมคอมพิวเตอร์สำเร็จรูปสำหรับคำนวณหาค่าสถิติต่าง ๆ ที่ผู้วิจัยต้องการได้โดยที่ผู้วิจัยจะต้องมีโมโน ทัศน์ (Concept) ดังนี้

1) ผู้วิจัยต้องเลือกใช่วิธีการทางสถิติให้เหมาะสมกับลักษณะของข้อมูล หรือ สมมุติฐานการวิจัย เช่น ผู้วิจัยต้องมีความรู้ว่าจะข้อมูลแบบต่อเนื่อง หรือไม่ต่อเนื่องควรใช้สถิติอะไรที่เหมาะสม หรือสมมุติฐานการวิจัย อย่างนี้ควรใช้สถิติอะไร เป็นต้น

2) ผู้วิจัยต้องอ่านค่าสถิติหรือแปลความหมายค่าสถิติที่โปรแกรมคอมพิวเตอร์ คำนวณมาให้ได้ว่า หมายความว่าอย่างไร เช่น ค่าร้อยละ ค่าเฉลี่ย ที่ได้หมายความว่าอย่างไร หรือค่าสถิติทดสอบที่ได้ผู้วิจัยจะ ตัดสินใจปฏิเสธ หรือไม่ปฏิเสธสมมุติฐานการวิจัย เป็นต้น

การวิเคราะห์ข้อมูลโดยไม่ใช้สถิติ มีดังนี้

2.1) การวิเคราะห์เอกสาร ( Documentary Analysis ) ใช้ในการวิจัยเอกสารและวิเคราะห์เกี่ยวกับ เนื้อหา ( content ) ในเอกสาร

2.2) การวิเคราะห์โดยสังเกต โดยสังเกตพฤติกรรม / เหตุการณ์ที่อยู่ในสังคมแล้วผู้วิจัยแปลความหมาย เอง สรุปความเอง วิเคราะห์เอง การวิเคราะห์ข้อมูลของการวิจัยเชิงคุณภาพทางด้านสังคมศาสตร์นั้นต้องมีการเชื่อมโยงความรู้อื่น ๆ ดังนี้

2.2.1) ความรู้ทางด้านประวัติศาสตร์

2.2.2) ความรู้ด้านชุมชน สังคม วัฒนธรรม

2.2.3) ความรู้ทางปรัชญาและศาสนาเพื่ออธิบายสรุปนามธรรมจากข้อมูลที่ปรากฏในรูปธรรมได้ด้วย

### 2.1.2 แนวคิดเกี่ยวกับการแสดงผลข้อมูล ( Data visualization )

การแสดงผลข้อมูล (Data Visualization) เป็นส่วนประกอบสำคัญใน Cognitive System ซึ่งเป็นส่วนในการแสดงข้อมูลหรือผลลัพธ์ต่าง ๆ ในระหว่างคอมพิวเตอร์และผู้ใช้งานในรูปแบบของภาพ โดยผู้ใช้สามารถเรียนรู้และจดจำข้อมูลผ่านการมองได้มากกว่าการใช้ประสาทสัมผัสอื่น ๆ หรือจะกล่าวได้ว่า Visualization ก็คือ การสร้างมโนภาพของสิ่งต่าง ๆ ที่เราสนใจขึ้นมาในใจ ซึ่งต่อมาได้กลายเป็นการนำภาพมาใช้กับการนำเสนอหรือนามามาเป็นกรอบความคิด ซึ่งได้นำไปใช้ในการสนับสนุนการตัดสินใจ หากไม่มีการทำ Data Visualization แล้ว อาจทำให้เราไม่สามารถค้นพบนัยยะของข้อมูลในแง่ของแนวโน้ม, รูปแบบพฤติกรรม, และความสัมพันธ์เชื่อมโยงได้

การเลือกรูปแบบ Visualization ให้เหมาะสมกับข้อมูล ในปัจจุบันเป็นยุคที่เทคโนโลยีเข้าถึงทุกคน ทำให้การรับรู้ข่าวสาร ข้อมูลต่างๆ เป็นไปได้ง่าย และรวดเร็วมากขึ้น คนที่นำเสนอข้อมูลจึงต้องนำเสนอข้อมูลให้น่าสนใจ เข้าใจง่าย และรวดเร็ว จึงเกิดการสร้าง Data Visualization ขึ้นมา Data Visualization เป็นการใช้ภาพเพื่อแสดงข้อมูลในเชิงปริมาณที่วัดได้ ซึ่งอาจนำเสนอออกมาในรูปแบบ แผนภูมิ กราฟ กราฟิก และอื่นๆ อีกมากมาย เพื่อให้เข้าใจได้ง่าย และรวดเร็ว

### 2.1.3 แนวคิดเกี่ยวกับการพยากรณ์ข้อมูล ( Forecasting )

การพยากรณ์ หมายถึง การคาดคะเนหรือประมาณการเหตุการณ์หรือสิ่งหนึ่งที่จะเกิดขึ้นในช่วงเวลาในอนาคต โดยอาศัยข้อมูลในอดีตหรือปัจจุบัน ตลอดจนพิจารณาความรู้ ประสบการณ์ของบุคคล เพื่อประโยชน์ในการช่วยให้การตัดสินใจมีความถูกต้อง โดยเริ่มจากการใช้ข้อมูลที่เกี่ยวข้องในอดีตมาเข้ากระบวนการพยากรณ์ด้วยเทคนิคต่างๆ ที่เหมาะสม จะทำให้ได้ผลจากการพยากรณ์ที่น่าเชื่อถือ สามารถใช้ในการวางแผนต่างๆ

การพยากรณ์ที่ดีต้องเป็นการพยากรณ์ที่รวดเร็ว แม่นยำ และสามารถทราบทิศทางการเปลี่ยนแปลงได้ จะทำให้การดำเนินการต่างๆ สามารถแก้ไขสถานการณ์ หรือ เพิ่มโอกาสที่ดีในอนาคตได้ จึงควรพิจารณาองค์ประกอบที่ดีส่งผลให้การพยากรณ์มีความแม่นยำ ประกอบไปด้วย

- 1) ระบุวัตถุประสงค์ในการนำผลการพยากรณ์ไปใช้อย่างชัดเจน และช่วงเวลาที่ต้องการในการพยากรณ์ เพื่อประโยชน์ในการเลือกใช้วิธีที่เหมาะสมในการพยากรณ์
- 2) การรวบรวมข้อมูลควรดำเนินการอย่างมีระบบ สามารถจำแนกชนิดและลักษณะของข้อมูลได้อย่างชัดเจน เพียงพอ และ ถูกต้องตามความเป็นจริงอย่างต่อเนื่อง เพราะข้อมูลเปรียบเสมือนปัจจัยหลักในการวิเคราะห์ตามวิธีที่เหมาะสมกับลักษณะข้อมูลนั้น
- 3) การกำหนดนิยามเฉพาะของตัวแปรที่ใช้หรือเกี่ยวข้องในการพยากรณ์ รวมถึงข้อจำกัดผลกระทบที่มีต่อการพยากรณ์
- 4) การตรวจสอบความแม่นยำในการพยากรณ์อย่างต่อเนื่อง เนื่องจากสถานการณ์ในช่วงเวลาต่างๆ อาจเกิดการเปลี่ยนแปลง การตรวจสอบความแม่นยำในการพยากรณ์เป็นระยะจะเป็นตัวชี้วัดถึงวิธีการที่เหมาะสม
- 5) การพยากรณ์ที่ดีจะต้องง่ายต่อการนำไปประยุกต์ใช้ สร้างความเข้าใจที่ตรงกันทั่วทั้งองค์กร การพยากรณ์มีเป้าหมายที่แตกต่างกันตามลักษณะและการนำไปใช้

#### 2.1.4 แนวคิดเกี่ยวกับข้อมูลอนุกรมเวลา ( Data Time Series )

ข้อมูลอนุกรมเวลา (time series data) คือ ข้อมูลที่ได้จากการจัดเก็บตามลำดับเวลา ต่อเนื่องกันเป็นช่วงๆ เช่น ทุก 7 วัน ทุก 1 เดือน ทุก 1 ปี ข้อมูลประเภทนี้ช่วยแสดงให้เห็นการเปลี่ยนแปลงของข้อมูลที่นักเรียนสนใจในช่วงเวลาหนึ่งได้อย่างชัดเจน เช่น การเพิ่มขึ้นของข้อมูล การลดลงของข้อมูล ค่าสูงสุด ค่าต่ำสุด ข้อมูลดังกล่าวจะเป็นประโยชน์ในการทำโครงการ หรือ ทำวิจัยในระยะยาว ตัวอย่างข้อมูลอนุกรมเวลา เช่นจำนวนนักเรียนที่ใช้บริการห้องสมุดของโรงเรียนแห่งหนึ่ง ตั้งแต่เดือน มกราคม พ.ศ.2563 ถึงเดือนธันวาคม พ.ศ.2564 ราคาน้ำมันแก๊สโซฮอล์ 95 ในช่วงเดือน เมษายน พ.ศ.2562 ถึงเดือน เมษายน พ.ศ.2564

### 2.1.5 การวิเคราะห์อนุกรมเวลา ( Time Series Analysis )

ในการวิเคราะห์อนุกรมเวลา ผู้วิเคราะห์จะแยกองค์ประกอบต่าง ๆ ที่ประกอบกันขึ้นเป็นอนุกรมเวลา โดยจะมีการเปลี่ยนแปลงไปตามอิทธิพลต่าง ๆ เช่น การเปลี่ยนแปลงการผลิต เทคโนโลยี สภาพอากาศ เป็นต้น ในการหาคุณลักษณะของอนุกรมเวลาเราสามารถที่ใช้แบบจำลองได้หลายแบบ แบบจำลองที่ใช้โดยนักเศรษฐศาสตร์แบบหนึ่ง คือ แบบจำลองแบบคลาสสิก (classical model) เป็นการอธิบายถึงองค์ประกอบของการแปรผันของอนุกรมเวลา 4 ส่วน ดังนี้

#### 1) ค่าแนวโน้ม (Secular trend)

เป็นการเปลี่ยนแปลงข้อมูลมีลักษณะราบเรียบ แนวโน้ม อาจมีลักษณะเป็นเส้นตรงหรือเส้นโค้งในทางเพิ่มขึ้นหรือลดลง ค่าแนวโน้มของข้อมูลเป็นการเคลื่อนไหวในช่วงระยะเวลาที่ค่อนข้างนานพอสมควร ควรเป็นข้อมูลรายปี และควรมีข้อมูลอย่างน้อย 15 ปี ซึ่งจะแสดงทิศทางของอนุกรมเวลา

#### 2) การเปลี่ยนแปลงหรือความแปรผันตามฤดูกาล (Seasonal Variation)

เป็นการเปลี่ยนแปลงข้อมูลมีลักษณะการเพิ่มขึ้น หรือลดลงในลักษณะเดียวกันของรอบระยะเวลาหนึ่งที่แน่นอน เรียกว่า การเปลี่ยนแปลงตามฤดูกาล หน่วยของระยะเวลาสำหรับข้อมูลอาจเป็นรายชั่วโมง รายวัน รายสัปดาห์ รายเดือน รายไตรมาส สำหรับข้อมูลรายปีไม่มีการแปรผันตามฤดูกาล การเปลี่ยนแปลงตามฤดูกาลนั้นกำหนดระยะเวลาการเกิดซ้ำในรอบหนึ่ง ๆ ได้ค่อนข้างแน่นอน

#### 3) การเปลี่ยนแปลงหรือความผันแปรตามวัฏจักร (Cyclical Variation)

เป็นการเกิดของเหตุการณ์ที่คล้ายกับส่วนฤดูกาล คือการเกิดรูปแบบซ้ำ ๆ กันในอนุกรมเวลา แต่จะเกิดในช่วงระยะเวลาที่ยาวกว่า ซึ่งจำเป็นที่จะต้องสังเกตจากอนุกรมเวลาที่มีระยะเวลานาน

#### 4) การเปลี่ยนแปลงหรือความผันแปรเนื่องจากเหตุการณ์ผิดปกติ (Irregular Variation)

## 2.2 ทฤษฎี

### 2.2.1 ทฤษฎีเกี่ยวกับการออกแบบเว็บไซต์

#### 2.2.1.1 หลักการออกแบบเว็บไซต์

เว็บไซต์เป็นสื่อที่ได้รับความนิยมอย่างมากบนอินเทอร์เน็ต ซึ่งเว็บไซต์เป็นสื่อที่อยู่ในความควบคุมของผู้ใช้โดยสมบูรณ์ กล่าวคือ ผู้ใช้สามารถตัดสินใจเลือกได้ว่า จะดูเว็บไซต์ใดและจะไม่เลือกดูเว็บไซต์ใด ได้ตามต้องการ จึงทำให้ผู้ใช้ไม่มีความอดทนต่ออุปสรรคและปัญหาที่เกิดจากการออกแบบเว็บไซต์ผิดพลาดถ้าผู้ใช้เห็นว่าเว็บที่กำลังดูอยู่นั้นไม่มีประโยชน์ต่อตัวเขา หรือไม่เข้าใจว่าเว็บไซต์นี้จะใช้งานอย่างไร เขาก็สามารถที่จะเปลี่ยนไปดูเว็บไซต์อื่น ๆ ได้อย่างรวดเร็ว เนื่องจากในปัจจุบันมีเว็บไซต์อยู่มากมายและยังมีเว็บไซต์ที่เกิดขึ้นใหม่ ๆ ทุกวัน ผู้ใช้จึงมีทางเลือกมากขึ้น และสามารถเปรียบเทียบคุณภาพของเว็บไซต์ต่าง ๆ ได้เองเว็บไซต์ที่ได้รับการออกแบบอย่างสวยงาม มีการใช้งานที่สะดวก ย่อมได้รับความสนใจจากผู้ใช้งานมากกว่าเว็บไซต์ที่ดูสับสน วุ่นวาย มีข้อมูลมากมายแต่หาอะไรไม่เจอนอกจากนี้ยังใช้เวลาในการแสดงผลแต่ละหน้านานเกินไป ซึ่งปัญหาเหล่านี้ล้วนเป็นผลมาจากการออกแบบเว็บไซต์ไม่ดีทั้งสิ้น ดังนั้น การออกแบบเว็บไซต์จึงเป็นกระบวนการสำคัญในการสร้างเว็บไซต์ ให้ประทับใจผู้ใช้ ทำให้เขาอยากกลับมาเข้าเว็บไซต์เดิมอีกในอนาคต ซึ่งนอกจากต้องพัฒนาเว็บไซต์ที่ดีมีประโยชน์แล้ว ยังต้องคำนึงถึงการแข่งขันกับเว็บไซต์อื่น ๆ ดังนั้นการออกแบบเว็บไซต์ ต้องคำนึงถึง

1) ความเรียบง่าย เว็บไซต์มีรูปแบบที่เรียบง่าย ไม่ซับซ้อน และใช้งานได้สะดวก ไม่มีกราฟิกหรือตัวอักษรที่เคลื่อนไหวอยู่ตลอดเวลา ชนิดและสีของตัวอักษรไม่มากจนเกินไปทำให้วุ่นวาย

2) ความสม่ำเสมอ เว็บไซต์ควรใช้รูปแบบเดียวกันตลอดทั้งเว็บไซต์ เช่น รูปแบบของหน้าสไตล์ของกราฟิก ระบบเมนูเก็ชชั่นและโทนสี ควรมีความคล้ายคลึงกันตลอดทั้งเว็บไซต์

3) ความเป็นเอกลักษณ์ การออกแบบเว็บไซต์ควรคำนึงถึงลักษณะขององค์กร เพราะรูปแบบของเว็บไซต์จะสะท้อนถึงเอกลักษณ์และลักษณะขององค์กรนั้น ๆ เช่น ถ้าเป็นเว็บไซต์ของทางราชการ จะต้องดูน่าเชื่อถือไม่เหมือนสวนสนุก

4) เนื้อหาที่มีประโยชน์ เนื้อหาเป็นสิ่งที่สำคัญที่สุดในเว็บไซต์ ดังนั้นควรจัดเตรียมเนื้อหาและข้อมูลที่ใช้ต้องการให้ถูกต้อง และสมบูรณ์ มีการปรับปรุงและเพิ่มเติมให้ทันเหตุการณ์อยู่เสมอ เนื้อหาไม่ควรซ้ำกับเว็บไซต์อื่น จึงจะดึงดูดความสนใจ

5) ระบบเนวิเกชันที่ใช้งานง่าย ต้องออกแบบให้ผู้ใช้เข้าใจง่ายและใช้งานสะดวก ใช้กราฟิกที่สื่อความหมายร่วมกับคำอธิบายที่ชัดเจน มีรูปแบบและลำดับของรายการที่สม่ำเสมอ เช่น วางไว้ตำแหน่งเดียวกันของทุกหน้า

6) ลักษณะที่น่าสนใจ หน้าตาของเว็บไซต์จะต้องมีความสัมพันธ์ กับคุณภาพขององค์ประกอบต่าง ๆ เช่น คุณภาพของกราฟิกที่จะต้องสมบูรณ์ การใช้สี การใช้ตัวอักษรที่อ่านง่าย สบายตา การใช้โทนสีที่เข้ากันลักษณะหน้าตาที่น่าสนใจนั้นขึ้นอยู่กับความชอบของแต่ละบุคคล

7) การใช้งานอย่างไม่จำกัด ผู้ใช้ส่วนใหญ่สามารถเข้าถึงได้มากที่สุดเลือกใช้เบราว์เซอร์ชนิดใดก็ได้ในการเข้าถึงเนื้อหาสามารถแสดงผลได้ทุกระบบปฏิบัติการ และความละเอียดหน้าจอต่าง ๆ กันอย่างไม่มีปัญหาเป็นลักษณะสำคัญสำหรับผู้ใช้ที่มีจำนวนมาก

8. คุณภาพในการออกแบบ การออกแบบและเรียบเรียงเนื้อหาอย่างรอบคอบ สร้างความรู้สึกว่าเว็บไซต์มีคุณภาพ ถูกต้อง และเชื่อถือได้

9. ลิงค์ต่าง ๆ จะต้องเชื่อมโยงไปหน้าที่มีอยู่จริงและถูกต้อง ระบบการทำงานต่าง ๆ ในเว็บไซต์จะต้องมีความแน่นอนและทำหน้าที่ได้อย่างถูกต้อง



### 2.2.1.2 การออกแบบเว็บไซต์

ในการออกแบบเว็บไซต์นั้นประกอบด้วยกระบวนการต่าง ๆ มากมาย เช่น การออกแบบโครงสร้าง ลักษณะหน้าตา หรือการเขียนโปรแกรม แต่มีหลายคนที่พัฒนาเว็บไซต์ โดยขาดการวางแผนและทำงานไม่เป็นระบบ ตัวอย่างเช่น การลงมือออกแบบโดยใช้โปรแกรมช่วยสร้างเว็บไซต์ เนื้อหาและรูปแบบก็เป็นไปตามที่นึกขึ้นได้ขณะนั้น และเมื่อเห็นว่าดูดีแล้วก็เปิดตัวเลย ทำให้เว็บนั้นมีเป้าหมายและแนวทางที่ไม่แน่นอน ผลลัพธ์ที่ได้จึงเสี่ยงกับความล้มเหลวค่อนข้างมาก ความล้มเหลวที่พบเห็นได้ทั่วไป ได้แก่ เว็บไซต์ที่แสดงข้อความว่าอยู่ระหว่างการก่อสร้าง (Under Construction หรือ Coming soon) ซึ่งแสดงให้เห็นถึงการขาดการวางแผนที่ดีบางเว็บถือได้ว่าตายไปแล้ว เนื่องจากข้อมูลไม่ทันสมัย ขาดการพัฒนาปรับปรุงเทคโนโลยีล้ำสมัย ลิงค์ผิดพลาด สิ่งเหล่านี้แสดงให้เห็นถึงการขาดการดูแล ตรวจสอบและพัฒนาให้ทันสมัยอยู่เสมอการออกแบบเว็บไซต์อย่างถูกต้องจะช่วยลดความผิดพลาดเหล่านี้ และช่วยลดความเสี่ยงที่จะทำให้เว็บประสบความล้มเหลว การออกแบบเว็บไซต์ที่ดีต้องอาศัยการออกแบบและจัดระบบข้อมูลอย่างเหมาะสม กระบวนการแรกของการออกแบบเว็บไซต์คือการกำหนดเป้าหมายของเว็บไซต์กำหนดกลุ่มผู้ใช้ ซึ่งการจะให้ได้มาซึ่งข้อมูล ผู้พัฒนาต้องเรียนรู้ผู้ใช้ หรือจำลองสถานการณ์ สิ่งเหล่านี้จะช่วยให้เราสามารถออกแบบเนื้อหาและการใช้งานเว็บไซต์ได้อย่างเหมาะสม ตรงกับความต้องการของผู้ใช้อย่างแท้จริง

1) กำหนดเป้าหมายของเว็บไซต์ขั้นตอนแรกของการออกแบบเว็บไซต์ คือการกำหนดเป้าหมายของเว็บไซต์ให้แน่ชัดเสียก่อน เพื่อจะได้ออกแบบการใช้งานได้ตรงกับเป้าหมายที่ได้ตั้งเอาไว้ โดยทั่วไปมักจะเข้าใจว่าการทำเว็บไซต์มีจุดมุ่งหมายเพื่อบริการข้อมูลของหน่วยงานหรือองค์กรเท่านั้น แต่ในความเป็นจริงแล้ว เว็บไซต์แต่ละแห่งก็จะมีเป้าหมายของตนเองแตกต่างกันออกไป

2) กำหนดกลุ่มผู้ใช้เป้าหมาย ผู้ออกแบบเว็บไซต์จำเป็นต้องทราบกลุ่มผู้ใช้เป้าหมายที่เข้ามาใช้บริการเว็บไซต์ เพื่อที่จะได้ตอบสนองความต้องการของผู้ใช้ได้อย่างชัดเจน ตัวอย่างเช่น เว็บไซต์ที่มีกลุ่มผู้ใช้หลากหลาย เช่น เสิร์ชเอนจิน เว็บท่า และเว็บไดเรกทอรี แต่เว็บไซต์ส่วนใหญ่นั้นจะตอบสนองความต้องการเฉพาะกลุ่มเท่านั้น ไม่สำหรับทุกคน เพราะคุณไม่สามารถตอบสนองความต้องการของคนที่หลากหลายได้ในเว็บไซต์เดียว

### 2.2.1.3 ออกแบบหน้าเว็บไซต์ (Page Design)

หลักสำคัญในการออกแบบหน้าเว็บก็คือ การใช้รูปภาพและองค์ประกอบต่าง ๆ ร่วมกันเพื่อสื่อความหมาย เกี่ยวกับเนื้อหาหรือลักษณะสำคัญของเว็บไซต์ โดยมีเป้าหมายสำคัญเพื่อการสื่อความหมายที่ชัดเจนและน่าสนใจ บนพื้นฐานของความเรียบง่ายและความสะดวกของผู้ใช้ ดังนี้

1) ความเรียบง่าย ได้แก่ มีรูปแบบที่เรียบง่าย ไม่ซับซ้อน และใช้งานได้สะดวก ไม่มีกราฟิกหรือตัวอักษรที่เคลื่อนไหวอยู่ตลอดเวลา ชนิดและสีของตัวอักษรไม่มากจนเกินไปทำให้วุ่นวาย

2) ความสม่ำเสมอ ได้แก่ ใช้รูปแบบเดียวกันตลอดทั้งเว็บไซต์ เช่น รูปแบบของหน้า สไตล์ของกราฟิก ระบบเมนูเช็คนและโทนสี ควรมีความคล้ายคลึงกันตลอดทั้งเว็บไซต์

3) ความเป็นเอกลักษณ์ การออกแบบเว็บไซต์ควรคำนึงถึงลักษณะขององค์กร เพราะรูปแบบของเว็บไซต์จะสะท้อนถึงเอกลักษณ์และลักษณะขององค์กรนั้น ๆ เช่น ถ้าเป็นเว็บไซต์ของทางราชการ จะต้องดูน่าเชื่อถือไม่เหมือนสวนสนุก ฯลฯ

4) เนื้อหาที่มีประโยชน์ เนื้อหาเป็นสิ่งที่สำคัญที่สุดในเว็บไซต์ ดังนั้นควรจัดเตรียมเนื้อหา และข้อมูลที่ใช้ต้องการให้ถูกต้อง และสมบูรณ์ มีการปรับปรุงและเพิ่มเติมให้ทันเหตุการณ์อยู่เสมอ เนื้อหาไม่ควรซ้ำกับเว็บไซต์อื่น จึงจะดึงดูดความสนใจ

5) ระบบเนวิเกชันที่ใช้งานง่าย ต้องออกแบบให้ผู้ใช้เข้าใจง่ายและใช้งานสะดวก ใช้กราฟิกที่สื่อความหมายร่วมกับคำอธิบายที่ชัดเจน มีรูปแบบและลำดับของรายการที่สม่ำเสมอ เช่น วางไว้ตำแหน่งเดียวกันของทุกหน้า

6) ลักษณะที่น่าสนใจ หน้าตาของเว็บไซต์จะต้องมีความสัมพันธ์กับคุณภาพขององค์ประกอบต่าง ๆ เช่น คุณภาพของกราฟิกที่จะต้องสมบูรณ์ การใช้สี การใช้ตัวอักษรที่อ่านง่าย สบายตา การใช้โทนสีที่เข้ากันลักษณะหน้าตาที่น่าสนใจนั้นขึ้นอยู่กับความชอบของแต่ละบุคคล

7) การใช้งานอย่างไม่จำกัด ผู้ใช้ส่วนใหญ่สามารถเข้าถึงได้มากที่สุดเลือกใช้เบราว์เซอร์ชนิดใดก็ได้ในการเข้าถึงเนื้อหาสามารถแสดงผลได้ทุกระบบปฏิบัติการและความละเอียดหน้าจอต่าง ๆ กันอย่างไม่มีปัญหาเป็นลักษณะสำคัญสำหรับผู้ที่มีจำนวนมาก

8) คุณภาพในการออกแบบ การออกแบบและเรียบเรียงเนื้อหาอย่างรอบคอบ สร้างความรู้สึกว่าเว็บไซต์มีคุณภาพ ถูกต้อง และเชื่อถือได้

9) ลิงค์ต่าง ๆ จะต้องเชื่อมโยงไปหน้าที่มีอยู่จริงและถูกต้อง ระบบการทำงานต่าง ๆ ในเว็บไซต์จะต้องมีความแน่นอนและทำหน้าที่ได้อย่างถูกต้อง

#### 2.2.1.4 การออกแบบโครงสร้างเว็บไซต์ (Site Structure Design)

โครงสร้างเว็บไซต์ ( Site Structure ) เป็นแผนผังของการลำดับเนื้อหาหรือการจัดวางตำแหน่งเว็บเพจทั้งหมด ซึ่งจะทำให้เรารู้ว่าทั้งเว็บไซต์ประกอบไปด้วยเนื้อหาอะไรบ้าง และมีเว็บเพจหน้าไหนที่เกี่ยวข้องเชื่อมโยงถึงกัน ดังนั้นการออกแบบโครงสร้างเว็บไซต์จึงเป็นเรื่องสำคัญเปรียบเสมือนกับการเขียนแบบอาคารก่อนที่จะลงมือสร้างเพราะจะทำให้เรามองเห็นหน้าตาของเว็บไซต์เป็นรูปธรรมมากขึ้นสามารถออกแบบระบบเนวิเกชันได้เหมาะสมและเป็นแนวทางการทำงานที่ชัดเจนสำหรับขั้นตอนต่อ ๆ ไปนอกจากนี้โครงสร้างเว็บไซต์ที่ดียังช่วยให้ผู้ชมไม่สับสนและค้นหาข้อมูลที่ต้องการได้อย่างรวดเร็ว วิธีการจัดโครงสร้างเว็บไซต์สามารถทำได้หลายแบบ แต่แนวคิดหลัก ๆ ที่นิยมใช้กันมีอยู่ 2 แบบคือ จัดตามกลุ่มเนื้อหา (Content-based Structure) จัดตามกลุ่มผู้ชม (User-based Structure)

#### 2.2.2.5 รูปแบบของโครงสร้างเว็บไซต์

เราสามารถวางรูปแบบโครงสร้างเว็บไซต์ได้หลายแบบตามความเหมาะสม

- 1) แบบเรียงลำดับ ( Sequence ) เหมาะสำหรับเว็บไซต์ที่มีจำนวนเว็บเพจไม่มากนัก หรือเว็บไซต์ที่มีการนำเสนอข้อมูลแบบทีละขั้นตอน
- 2) แบบระดับชั้น ( Hierarchy ) เหมาะสำหรับเว็บไซต์ที่มีจำนวนเว็บเพจมากขึ้น เป็นรูปแบบที่เราจะพบได้ทั่วไป
- 3) แบบผสม ( Combination ) เหมาะสำหรับเว็บไซต์ที่ซับซ้อน เป็นการนำข้อดีของรูปแบบทั้ง 2 ข้างต้นมาผสมกัน

### 2.2.2.6 การใช้สีในการออกแบบเว็บไซต์

ระบบสีที่แสดงบนจอคอมพิวเตอร์ มีระบบการแสดงผลผ่านหลอดลำแสงที่เรียกว่า CRT (Cathode ray tube) โดยมีลักษณะระบบสีแบบบวก อาศัยการผสมของของแสงสีแดง สีเขียว และสีน้ำเงิน หรือระบบสี RGB สามารถกำหนดค่าสีจาก 0 ถึง 255 ได้ จากการรวมสีของแม่สีหลักจะทำให้เกิดแสงสีขาว มีลักษณะเป็นจุดเล็ก ๆ บนหน้าจอไม่สามารถมองเห็นด้วยตาเปล่าได้ จะมองเห็นเป็นสีที่ถูกผสมเป็นเนื้อสีเดียวกันแล้ว จุดแต่ละจุดหรือพิกเซล (Pixel) เป็นส่วนประกอบของภาพบนหน้าจอคอมพิวเตอร์ โดยจำนวนบิตที่ใช้ในการกำหนดความสามารถของการแสดงสีต่าง ๆ เพื่อสร้างภาพบนจอขึ้นเรียกว่า บิตเด็ป (Bitdepth) ในภาษา HTML มีการกำหนดสีด้วยระบบเลขฐานสิบหก ซึ่งมีเครื่องหมาย (#) อยู่ด้านหน้าและตามด้วยเลขฐานสิบหกจำนวนอักษรอีก 6 หลัก โดยแต่ละไบต์ (byte) จะมีตัวอักษรสองตัว แบ่งออกเป็น 3 กลุ่ม เช่น #FF12AC การใช้ตัวอักษรแต่ละไบต์นี้เพื่อกำหนดระดับความเข้มของแม่สีแต่ละสีของชุดสี RGB โดย 2 หลักแรก แสดงถึงความเข้มของสีแดง 2 หลักต่อมา แสดงถึงความเข้มของสีเขียว 2 หลักสุดท้ายแสดงถึงความเข้มของสีน้ำเงินสีมีอิทธิพลในเรื่องของอารมณ์การสื่อความหมายที่เด่นชัด กระตุ้นการรับรู้ทางด้านจิตใจมนุษย์ สีแต่ละสีให้ความรู้สึก อารมณ์ที่ไม่เหมือนกัน สีบางสีให้ความรู้สึกสงบ บางสีให้ความรู้สึกตื่นตัวรุนแรง สีจึงเป็นปัจจัยสำคัญอย่างยิ่งต่อการออกแบบเว็บไซต์ ดังนั้นการเลือกใช้โทนสีภายในเว็บไซต์เป็นการแสดงถึงความแตกต่างของสีที่แสดงออกทางอารมณ์ มีชีวิตชีวาหรือเศร้าโศก รูปแบบของสีที่สายตาของมนุษย์มองเห็น สามารถแบ่งออกเป็น 3 กลุ่ม คือ

- 1) สีโทนร้อน (Warm Colors) เป็นกลุ่มสีที่แสดงถึงความสุข ความปลอบโยน ความอบอุ่น และดึงดูดใจ สีกลุ่มนี้เป็นกลุ่มสีที่ช่วยให้หายจากความเฉื่อยชา มีชีวิตชีวามากยิ่งขึ้น
- 2) สีโทนเย็น (Cool Colors) แสดงถึงความที่ดูสุภาพ อ่อนโยน เรียบร้อย เป็นกลุ่มสีที่มีคนชอบมากที่สุด สามารถโน้มน้าวในระยะไกลได้
- 3) สีโทนกลาง (Neutral Colors) สีที่เป็นกลาง ประกอบด้วย สีดำ สีขาว สีเทา และสีน้ำตาล กลุ่มสีเหล่านี้คือ สีกลางที่สามารถนำไปผสมกับสีอื่น ๆ เพื่อให้เกิดสีกลางขึ้นมา

## 2.2.2 ทฤษฎีเกี่ยวกับการทำเหมืองข้อมูล

การทำเหมืองข้อมูล (Data Mining) คือกระบวนการค้นหาความสัมพันธ์และรูปแบบทั้งหมด ซึ่ง มีอยู่จริงในฐานข้อมูลแต่ได้ถูกซ่อนไว้ภายในข้อมูลจำนวนมาก ในปัจจุบันการทำเหมืองข้อมูลได้ถูกนำไปประยุกต์ใช้ในงานหลายประเภท โดยการทำเหมืองข้อมูลจะเหมาะกับการแก้ปัญหาบางชนิดเท่านั้น มีเทคนิคต่าง ๆ ที่ใช้ในการแก้ไขปัญหาย่อยหลายเทคนิค ซึ่งไม่มีเทคนิคใดสามารถแก้ปัญหาได้ทุกปัญหา ดังนั้นความหลากหลายของเทคนิคเป็นสิ่งที่จำเป็นที่จะนำไปสู่วิธีการแก้ปัญหาที่ดีที่สุดของการทำเหมืองข้อมูล และการค้นพบสิ่งที่ซ่อนอยู่ในข้อมูลนั้น จะแบ่งออกเป็นข้อต่าง ๆ ได้ดังนี้

- 1) การระบุปัญหาที่เกิดขึ้นกับธุรกิจเป็นการระบุขอบเขตของข้อมูลที่จะนำมาทำการวิเคราะห์เพื่อหาความได้เปรียบทางการตลาดหรือเพื่อนำมาแก้ไขปัญหา
- 2) การทำเหมืองข้อมูล เป็นการนำเทคนิคของการทำเหมืองข้อมูลไปใช้ถ่ายทอดหรือทำการเปลี่ยนแปลงข้อมูลดิบให้อยู่ในรูปของข้อมูลจะนำไปใช้ได้จริงในทางธุรกิจ
- 3) การนำเอาข้อมูลที่เป็นผลลัพธ์ของส่วนการทำเหมืองข้อมูลมาทดลองปฏิบัติจริงกับธุรกิจ
- 4) การวัดประสิทธิภาพของเทคนิคการทำเหมืองข้อมูลที่จะนำมาใช้จากผลลัพธ์ เช่น วัดจากส่วนแบ่งของตลาด วัดจากปริมาณลูกค้า หรือ วัดจากกำไรสุทธิ

จาก 4 ขั้นตอนที่กล่าวมาข้างต้นคือการนำเอาเหมืองข้อมูล (Data Mining) ไปใช้กับระบบทางธุรกิจ โดยแต่ละขั้นตอนจะพึ่งพาอาศัยกันผลลัพธ์จากขั้นตอนหนึ่งจะกลายมาเป็นอินพุต (input) จากอีกขั้นตอนต่อไป ซึ่งเหมืองข้อมูล (Data Mining) จะเปลี่ยนข้อมูลดิบให้เป็นข้อมูลประยุกต์ ดังนั้นการระบุแหล่งข้อมูลที่ต้องการจึงเป็นสิ่งสำคัญอย่างยิ่ง

### ขั้นตอนการทำเหมืองข้อมูล

เป็นขั้นตอนของการทำงานที่จะเปลี่ยนข้อมูลดิบให้ไปเป็นความรู้ ประกอบด้วยขั้นตอนดังนี้

1) Business Understanding เป็นการทำความเข้าใจในงาน ระบุโอกาส และหาปัญหาที่จะเกิดขึ้น กำหนดขอบเขตของข้อมูลที่จะนำวิเคราะห์ ซึ่งต้องสามารถระบุผลลัพธ์ที่มีได้

2) Data Understanding ทำความเข้าใจข้อมูลโดยการรวบรวมข้อมูลที่เกี่ยวข้อง เลือกให้เหลือเพียงข้อมูลที่มีความถูกต้องและสำคัญต่องานมาทำการวิเคราะห์

3) Data Preparation ทำการแปลงข้อมูล (Raw Data) ให้เป็นข้อมูลที่สามารถนำมาช่วยในการวิเคราะห์ต่อไปได้ ขั้นตอนนี้ต้องจะใช้เวลามากกว่าขั้นตอนอื่น เพราะคุณภาพของงานที่ได้จะดีเพียงใดขึ้นอยู่กับคุณภาพข้อมูลที่จัดเตรียมในขั้นนี้ การเตรียมข้อมูลประกอบด้วย การคัดเลือกข้อมูล การกลั่นกรองข้อมูล และแปลงรูปแบบของข้อมูล

4) Modeling การสร้างแบบจำลองเพื่อวิเคราะห์ข้อมูลที่ได้จากขั้นตอนที่ 3 พร้อมทดสอบผลลัพธ์แบบจำลองเพื่อให้ได้คำตอบที่ดีที่สุด บางครั้งอาจมีการย้อนกลับไปปรับการเตรียมข้อมูลเพื่อให้ได้แบบจำลองที่เหมาะสมที่สุด

5) Evaluation การประเมินผลลัพธ์ที่ได้ก่อนที่จะนำไปใช้จริง ว่าตรงกับวัตถุประสงค์หรือไม่ ตามที่ได้ตั้งไว้หรือมีความน่าเชื่อถือมากน้อยเพียงใด หากไม่ได้ผลลัพธ์ตามวัตถุประสงค์ต้องย้อนกลับไปปรับปรุงแก้ไขการดำเนินงานในขั้นตอนก่อนหน้า

6) Deployment การนำเอาข้อมูลที่เป็นผลลัพธ์จากทั้งหมด มาใช้ประโยชน์ตามวัตถุประสงค์ และทำการประเมินผลลัพธ์ที่ได้ว่ามีประสิทธิภาพตรงตามวัตถุประสงค์ที่ตั้งไว้หรือไม่

### 2.2.3 ทฤษฎีเกี่ยวกับการ visualization

Visualization คือ การจินตนาการ หรือสร้างภาพขึ้นในความคิด ซึ่งเป็นกระบวนการทำให้เกิดความรู้ความเข้าใจสำหรับเรื่องที่จินตนาการยาก เข้าใจยากวิธีการที่เป็นทางลัดก็คือการสร้างภาพ ให้เป็น บันไดความคิด ไปสู่ การใช้ความคิดอีกระดับ บันไดนี้จะช่วยลดปริมาณข้อมูล ช่วยลดภาระการคำนวณหรือการนำไปผ่านหลากหลายกระบวนการความคิด เพื่อนำไปสู่คำตอบที่ต้องการได้โดยเร็วและถูกต้อง

Visualization System คือระบบ ซึ่งประกอบด้วยกระบวนการต่าง ๆ ที่ออกแบบมาเพื่อสร้าง รักษา นำไปใช้ และปรับปรุงทัศนสนเทศ เพื่อทำให้เกิดความรู้ความเข้าใจ อย่างถูกต้อง รวดเร็ว และได้ผลเป็นอย่างสูง

Data Visualization เป็นการนำข้อมูลมาผสมผสานกับจินตนาการ เพื่อสร้างภาพในความคิดขึ้นมา ซึ่งมีกระบวนการนำเสนอข้อมูลที่มีความซับซ้อนหรือข้อมูลเชิงปริมาณ ให้สามารถเข้าใจได้ง่าย ในแบบของ กราฟ แผนภูมิ



### 2.2.3.1 เครื่องมือในการทำ Data Visualization

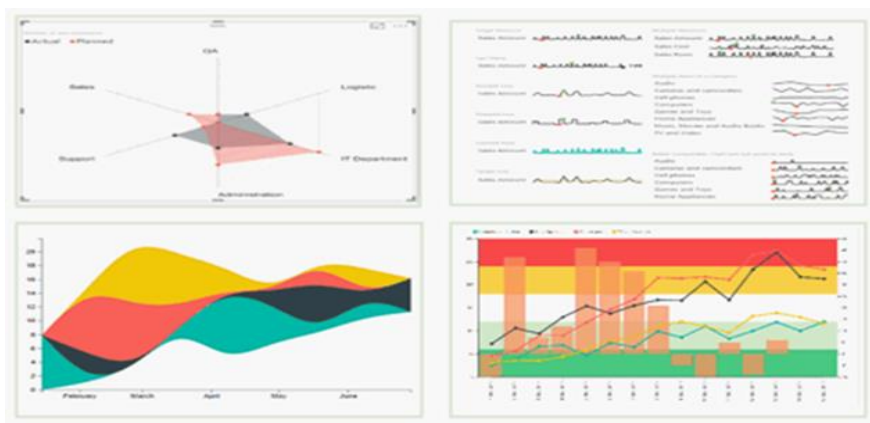
เครื่องมือที่นิยมนำมาใช้ในกระบวนการ Data Visualization ตอนนี้ คือ Tableau , Microsoft Power BI , Qlik View , Google Charts , Fusion Charts , Data wrapper



ภาพที่ 2.1 เครื่องมือที่นิยมนำมาใช้ในกระบวนการ Data Visualization

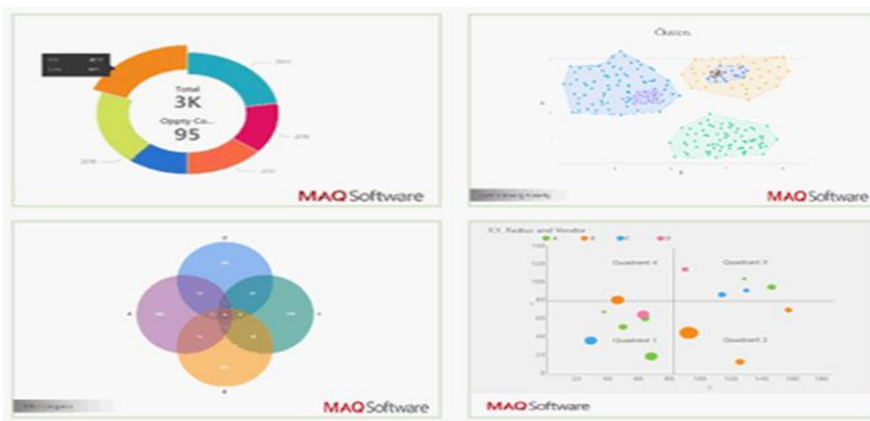
### 2.2.3.2 รูปแบบในการใช้ Data Visualization

1) การนำเสนอแบบทิศทางหรือแนวโน้ม (Trending) เราใช้กราฟที่แสดงผล แบบทิศทางหรือแนวโน้ม เพื่อนำเสนอข้อมูลให้เห็นจำนวนข้อมูลที่เกิดขึ้นในแต่ละช่วงเวลา (period) รวมถึงเน้นข้อมูลที่ต้องการนำเสนอ เช่น Line Chart, Bar Chart , Radar Chart, Area Char



ภาพที่ 2.2 กราฟการนำเสนอแบบทิศทางหรือแนวโน้ม

2) การนำเสนอแบบกลุ่มข้อมูล (Classification) เป็นการนำเสนอโดยนำข้อมูลมาจัดเป็นกลุ่มๆ เช่น Donut Chart, Ring Chart, Pie Chart,



ภาพที่ 2.3 การนำเสนอแบบกลุ่มข้อมูล

3) การนำเสนอเชิงเปรียบเทียบข้อมูล (Comparison) เหมาะสำหรับการนำเสนอที่ต้องการเปรียบเทียบผลการดำเนินงาน เช่นเทียบกับปีที่แล้ว(YoY) เปรียบเทียบกับเป้าที่ตั้งไว้ (Target) ซึ่งกราฟที่เหมาะสมและมักนำมาใช้ เช่น KPI Indicator, Bullet Chart, Power BI Card with state เป็นต้น



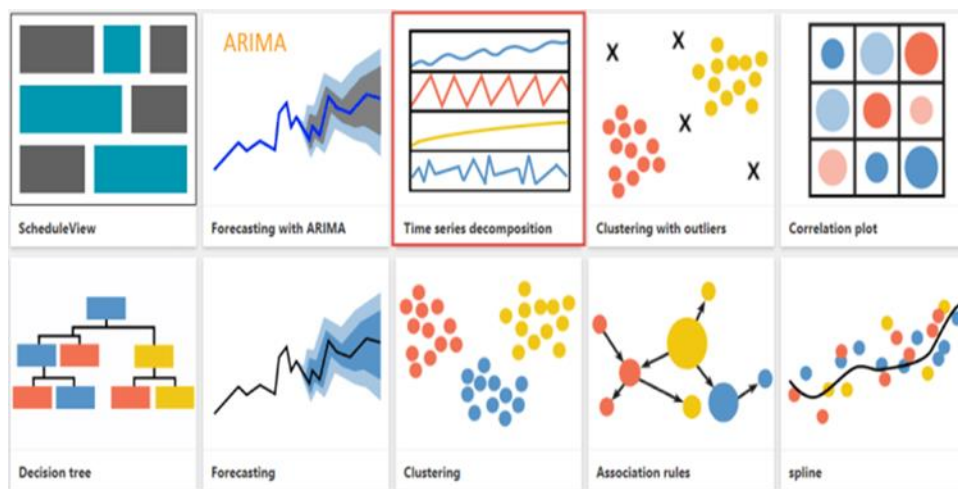
ภาพที่ 2.4 การนำเสนอเชิงเปรียบเทียบข้อมูล

4) การนำเสนอรูปแบบแผนที่ (Geographical) เหมาะสำหรับนำเสนอข้อมูลบนแผนที่ โดยสามารถที่จะนำยอดขาย, รายได้, ความหนาแน่นของประชากร เพื่อ Focus กลุ่มลูกค้าในแต่ละพื้นที่ที่เราสนใจ เช่น Globe Map, Google Map, Flow Map เป็นต้น



ภาพที่ 2.5 การนำเสนอรูปแบบแผนที่

5) กลุ่มที่ต้องพยากรณ์ล่วงหน้าและการวิเคราะห์ข้อมูลที่ซับซ้อน (Analytics) เราสามารถใช้ภาษา R หรือ Python ดึงข้อมูลในอดีตมาเพื่อวิเคราะห์เพื่อพยากรณ์อนาคต และนำเสนอข้อมูลที่ได้ในรูปของกราฟ เช่น Association Rules, Clustering, Forecasting Time series, Calculation plot เป็นต้น



ภาพที่ 2.6 กลุ่มการพยากรณ์ล่วงหน้าและการวิเคราะห์ข้อมูลที่ซับซ้อน

### 2.2.3.3 ทฤษฎีเกี่ยวกับการต้นไม้ตัดสินใจ

ต้นไม้การตัดสินใจจะทำการจัดกลุ่ม (classify) ชุดข้อมูลนำเข้าในแต่ละกรณี (Instance) แต่ละบัพ (node) ของต้นไม้การตัดสินใจคือตัวแปร (attribute) ต่าง ๆ ของชุดข้อมูล เช่นหากต้องการตัดสินใจว่าจะไปเล่นกีฬาหรือไม่ก็จะมีตัวแปรต้นที่จะต้องพิจารณาคือ ทัศนียภาพ ลม ความชื้น อุณหภูมิ เป็นต้น และมีตัวแปรตามซึ่งเป็นผลลัพธ์จากต้นไม้คือการตัดสินใจว่าจะไปเล่นกีฬาหรือไม่ ซึ่งแต่ละตัวแปรนั้นก็จะมีค่าของตัวเอง (value) เกิดเป็นชุดของตัวแปรค่าของตัวแปร (attribute-value pair) เช่น ทัศนียภาพเป็นตัวแปร ก็อาจมีค่าได้เป็น ฝนตก แดดออก หรือการตัดสินใจว่าจะไปเล่นกีฬาหรือไม่นั้นก็อาจมีค่าได้เป็นใช่ กับ ไม่ใช่ เป็นต้น การทำนายประเภทด้วยต้นไม้ตัดสินใจจะเริ่มจากบัพราก โดยทดสอบค่าตัวแปรของบัพ แล้วจึงตามกิ่งของต้นไม้ที่กำหนดค่า เพื่อไปยังบัพลูกถัดไป การทดสอบนี้จะกระทำไปจนกระทั่งเจอบัพใบซึ่งจะแสดงผลการทำนาย

ส่วนประกอบของต้นไม้ตัดสินใจ

- 1) โหนด คือ คุณสมบัติต่าง ๆ เป็นจุดที่แยกข้อมูลว่าจะให้ไปในทิศทางใด ซึ่งโหนดที่อยู่จุดสูงสุดเรียกว่า โหนดราก (Root Node)
- 2) กิ่ง คือ คุณสมบัติของคุณสมบัติในโหนดที่แตกออกมา โดยจำนวนของกิ่งจะเท่ากับคุณสมบัติของโหนด
- 3) ใบ คือ กลุ่มของผลลัพธ์ในการแยกแยะข้อมูล โดยสามารถแสดงส่วนประกอบของต้นไม้ตัดสินใจ

ในการเรียนรู้ของเครื่อง (machine learning) ต้นไม้ตัดสินใจ เป็นโมเดลทางคณิตศาสตร์ที่ใช้ทำนายประเภทของวัตถุโดยพิจารณาจากลักษณะของวัตถุ บัพภายใน (inner node) ของต้นไม้จะแสดงตัวแปร ส่วนกิ่งจะแสดงค่าที่เป็นไปได้ของตัวแปร ส่วนบัพใบ (leaf node) จะแสดงประเภทของวัตถุ

ต้นไม้ตัดสินใจที่บัพใบแสดงถึงข้อมูลที่เป็นข้อมูลไม่ต่อเนื่อง (discrete values) จะเรียกว่าต้นไม้ตัดสินใจแบบจำแนก (classification trees) และต้นไม้ตัดสินใจที่บัพใบเป็นข้อมูลต่อเนื่อง (continuous values) จะเรียกว่าต้นไม้ตัดสินใจแบบถดถอย (regression trees)

ต้นไม้การตัดสินใจในการบริหารธุรกิจ เป็นแผนผังต้นไม้ช่วยในการตัดสินใจ โดยแสดงถึงมูลค่าของทรัพยากรที่จะใช้ ความเสี่ยงในการลงทุนและผลลัพธ์ที่มีโอกาสเกิดขึ้น ต้นไม้ตัดสินใจสร้างขึ้นเพื่อช่วยการตัดสินใจเพื่อใช้ในการสร้างแผนงาน นิยมใช้มากในการบริหารความเสี่ยง (risk management) ต้นไม้ตัดสินใจเป็นส่วนหนึ่งของทฤษฎีการตัดสินใจ (decision theory) และ ทฤษฎีกราฟ ต้นไม้ตัดสินใจเป็นวิธีการพื้นฐานอย่างหนึ่งสำหรับการทำเหมืองข้อมูล

## 2.3 เครื่องมือในการออกแบบและวิเคราะห์ข้อมูล

### 2.3.1 การแสดงผลแบบ visualization

#### 2.3.1.1 การแสดงผลแบบ visualization ด้วย Tableau Public

ทาโบลิว (Tableau) Business Intelligence (BI) Software เป็นซอฟต์แวร์เพื่อการวิเคราะห์ข้อมูลและ Tableau ยังทำหน้าที่แปลงข้อมูล (Data) ให้เป็นภาพ (Visualization) สามารถนำข้อมูลออกมาเป็น Chart แบบต่าง ๆ ได้อย่างง่ายดายรวมถึงการนำหลาย Chart มาสร้างเป็น Dashboard และ Story Teller ก็ได้เช่นกัน รองรับการเข้าถึงข้อมูลจากหลายฐานข้อมูล และสามารถนำข้อมูลจากฐานข้อมูลในองค์กรนั้นมาใช้งาน เช่น Excel, Access, Firebird 2.0, IBM DB2, MS SQL, Oracle, MySQL เป็นต้น

ทาโบล์ว (Tableau) ถูกพัฒนาอย่างต่อเนื่องเพื่อให้แสดงผลได้บนทุก Platform อย่างง่ายดาย ไม่ว่าจะเป็น Desktop , Mobile หรือแม้แต่ Browser โดยองค์กรสามารถเลือกติดตั้งใช้งานภายในองค์กรเองก็ได้ หรือเช่าใช้บริการ Tableau Cloud ก็ได้เช่นกัน ทำให้การทำ Business Intelligence และ Data Analytics ไม่ใช่เรื่องยากอีกต่อไป และง่ายขึ้นจนถึงขนาดผู้บริหารก็สามารถใช้งานด้วยตัวเองได้

#### 2.3.1.2 ลักษณะเด่นของ Tableau Public

1) ใช้งานง่ายด้วยหลัก Drag-and-Drop Interface โดยสามารถสร้าง Visualization จากการ Drag ข้อมูล (Dimension หรือ Measure) ไป Drop บน Shelf ต่าง ๆ แล้วโปรแกรม Tableau จะตีความ และสร้างเป็นภาพให้เห็นทันที

2) รองรับการเข้าถึงข้อมูลจากหลายฐานข้อมูล และสามารถนำข้อมูลจากฐานข้อมูลในองค์กรนั้นมาใช้งาน เช่น Excel, Access, Firebird 2.0, IBM DB2, MS SQL Server, Microsoft Power pivot

3) รูปแบบการนำเสนอรายงานที่สวยงาม เข้าใจง่าย และง่ายในการนำเสนอต่อผู้บริหาร

4) การเข้าถึงหรือการใช้งาน สามารถทำได้โดยง่าย ตั้งแต่ผู้บริหารระดับสูง ไปจนถึง พนักงานระดับปฏิบัติการ ทั้งนี้ทั้งนั้น สามารถเพิ่มทักษะและแนวคิดให้กับพนักงานระดับปฏิบัติการ ให้มีแนวคิดเชิงสถิติและการประยุกต์ใช้งานมากขึ้น

### 2.3.2 แบบจำลองอนุกรมเวลา ( Time Series Analysis )

การพยากรณ์ ในการพยากรณ์อนุกรมเวลาแบบฉบับ ส่วนประกอบที่สำคัญที่มีอิทธิพลต่อการพยากรณ์ คือส่วนประกอบของการเปลี่ยนแปลงตามฤดูกาลแนวโน้ม และการเปลี่ยนแปลงตามวัฏจักร ส่วนการเปลี่ยนแปลงที่ผิดปกติที่เราไม่สามารถพยากรณ์ได้ เนื่องจากเป็นตัวแปรสุ่มที่มีรูปแบบที่ไม่แน่นอน ทำให้ไม่อาจคาดการณ์ได้ล่วงหน้าว่าจะเกิดอะไรขึ้น ณ เวลาใด และรุนแรงเพียงใด ดังนั้น ค่าพยากรณ์  $Y$  ในหน่วย เวลา  $t$  จะหาค่าได้ดังนี้

$$Y = T \times S \times C \times I$$

$Y$  คือ ข้อมูลอนุกรมเวลา

$T$  คือ ค่าประมาณของส่วนประกอบแนวโน้มของหน่วยเวลา (Trend)

$S$  คือค่าประมาณของส่วนประกอบการผันแปรตามฤดูกาล (Seasonal Variation)

$C$  คือ การผันแปรตามวัฏจักร (Cyclical Variation)

$I$  คือ การผันแปรผิดปกติ (Irregular Variation)

ในการพยากรณ์ค่าส่วนประกอบแนวโน้มและส่วนประกอบการเปลี่ยนแปลงตามฤดูกาลกระทำได้ไม่ยากนัก แต่การพยากรณ์ค่าของส่วนประกอบการเปลี่ยนแปลงตามวัฏจักรนั้นไม่ใช่สิ่งที่ยากนัก เพราะการเปลี่ยนแปลงตามวัฏจักรไม่มีรูปแบบ และความยาวของการเปลี่ยนแปลงตามวัฏจักรไม่แน่นอน ในการพยากรณ์การเปลี่ยนแปลงตามวัฏจักรในระยะสั้น หรืออาจใช้ความคิดเห็นของผู้เชี่ยวชาญในธุรกิจนั้น ๆ เช่น การประมาณอัตราการขยายตัวของธุรกิจโดยใช้ค่าเฉลี่ยของความสูง และความยาวของการเปลี่ยนแปลงตามวัฏจักรในอดีต เป็นต้น แต่ในทางปฏิบัติก็ไม่ใช่สิ่งที่จะทำได้โดยง่าย และการเปลี่ยนแปลงตามวัฏจักรในระยะสั้นก็มีความสำคัญต่อความแม่นยำของ



ค่าพยากรณ์มากพอสมควร ดังนั้นการประมาณการเปลี่ยนแปลงตามวัฏจักรโดยวิธีการนำเสนอไปในหัวข้อ 4 เป็นแนวทางหนึ่ง ที่อาจใช้ประมาณค่าการเปลี่ยนแปลงตามวัฏจักรได้อย่างเป็นระบบ นอกเหนือไปจากวิธีหาค่าเฉลี่ยแบบเคลื่อนที่ ค่าประมาณของส่วนประกอบการเปลี่ยนแปลงตามวัฏจักรที่จะใช้พยากรณ์ในฤดูกาลหน้าจะเป็นค่าในฤดูกาลปัจจุบัน

หากข้อมูลอนุกรมเวลาที่ใช้ในการศึกษาครั้งนี้ ไม่พบการผันแปรตามวัฏจักรและการผันแปรผิดปกติของข้อมูลที่ชัดเจน ดังนั้น การพยากรณ์จึงมีเฉพาะองค์ประกอบที่เป็นค่าแนวโน้มและองค์ประกอบที่แสดงค่า การผันแปรตามฤดูกาล ดังนี้

$$Y = T \times S$$

การวัดความแม่นยำของตัวแบบที่นำมาใช้พยากรณ์ ค่าเฉลี่ยเปอร์เซ็นต์ ความคลาดเคลื่อนสมบูรณ์ (Mean Absolute Percentage Error : MAPE) สามารถคำนวณได้จากสูตร ค่าเฉลี่ยเปอร์เซ็นต์ความคลาดเคลื่อนสมบูรณ์ (MAPE) ยิ่งน้อย หมายถึง การพยากรณ์ยิ่งมีความแม่นยำ

### 2.3.3 กระบวนการทำเหมืองข้อมูล ( CRISP-DM)

การทำเหมืองข้อมูล (Data Mining) คือ การค้นหาข้อมูลที่มีประโยชน์จากแหล่งข้อมูลที่มีเป็นจำนวนมากมายามหาศาล เพื่อดึงข้อมูลที่มีประโยชน์มาทำการวิเคราะห์ค้นหารูปแบบหรือความสัมพันธ์ที่เกิดในฐานข้อมูล และจัดทำเป็นสารสนเทศเพื่อใช้ในการวางแผนบริหารจัดการธุรกิจ โดยการแยกข้อมูลที่มีประโยชน์ออกมาใช้งานเปรียบเทียบกับการทำเหมืองแร่ ที่จะต้องทำการแยกเศษหินดินทรายที่ไม่มีค่าและมีปริมาณมากออกจากแร่ที่มีมูลค่ามากและมักจะมีปริมาณน้อย โดยจะมีขั้นตอนคือ

1) Business Understanding เป็นการทำความเข้าใจในงาน ระบุโอกาส และหา ปัญหา ที่จะเกิดขึ้นกับธุรกิจ กำหนดขอบเขตของข้อมูลที่จะนำวิเคราะห์หาความได้เปรียบทางการตลาด และแก้ไขปัญหาองค์กร ซึ่งต้องสามารถระบุผลลัพธ์ที่มีได้

2) Data Understanding ทำความเข้าใจข้อมูลโดยการรวบรวมข้อมูลที่เกี่ยวข้อง คัดเลือกให้ เหลือเพียงข้อมูลที่มีความถูกต้องและสำคัญต่องานมาทำการวิเคราะห์

3) Data Preparation ทำการแปลงข้อมูล (Raw Data) ให้กลายเป็นข้อมูลที่สามารถนำมา ช่วยในการวิเคราะห์ต่อไปได้ ขั้นตอนนี้จะใช้เวลามากที่สุดในทุกขั้นตอน เพราะคุณภาพของงานที่ได้ จะดีเพียงใดขึ้นอยู่กับคุณภาพข้อมูลที่จัดเตรียมในขั้นนี้ การเตรียมข้อมูลประกอบด้วย การคัดเลือก ข้อมูล การกลั่นกรองข้อมูล และแปลงรูปแบบของข้อมูล

4) Modeling การสร้างแบบจำลองเพื่อวิเคราะห์ข้อมูลที่ได้จากขั้นตอนที่ 3 พร้อมทดสอบ ผลลัพธ์แบบจำลองเพื่อให้ได้คำตอบที่ดีที่สุด บางครั้งอาจมีการย้อนกลับไปปรับการเตรียม ข้อมูล เพื่อให้ได้แบบจำลองที่เหมาะสมที่สุด

5) Evaluation การประเมินผลลัพธ์ที่ได้ก่อนที่จะนำไปใช้จริง ว่าตรงกับวัตถุประสงค์หรือ เป้าหมายที่ได้ตั้งไว้หรือมีความน่าเชื่อถือมากน้อยเพียงใด หากไม่ได้ผลลัพธ์ตามวัตถุประสงค์ต้อง ย้อนกลับไปปรับปรุงแก้ไขการดำเนินงานในขั้นตอนก่อนหน้า

6) Deployment การนำเอาข้อมูลที่เป็นผลลัพธ์จากทั้งหมด มาลองปฏิบัติจริงกับธุรกิจใ นองค์กร โดยแปลงแนวคิดที่มีให้เกิดขึ้นเป็นสารสนเทศเพื่อให้ผู้บริหารหรือนักการตลาดเข้าใจสามารถ นำไปใช้ประโยชน์ในทางธุรกิจได้จริง และติดตามประเมินผลที่ได้เพื่อนำกลับไปปรับปรุง Data Mining ต่อเนื่องต่อไป ซึ่งการประเมินผลสามารถทำได้หลายทางเช่น วัดจากส่วนแบ่งการตลาด วัด จากปริมาณลูกค้า หรือ วัดจากกำไรสุทธิ เป็นต้น

จากขั้นตอนที่ยกขึ้นมา นั้นคือการทำเหมืองข้อมูลในงานระบบทางธุรกิจ เป็นกระบวนการทางสถิติที่เน้นการจัดการข้อมูลที่มีจำนวนมากในหลากหลายรูปแบบ คัดเลือกเฉพาะข้อมูลที่สำคัญและจำเป็นต้องนำมาใช้งาน ทำการกำหนดรูปแบบจัดแบ่งกลุ่มลำดับความสำคัญ จากนั้นจึงจะเริ่มค้นหารูปแบบแนวทางและความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลนั้น โดยแต่ละขั้นตอนจะอาศัยผลลัพธ์จากอีกขั้นตอนหนึ่งกลายเป็นข้อมูลให้ขั้นตอนต่อไป การทำเหมืองข้อมูลจะช่วยเปลี่ยนข้อมูลดิบให้เป็นสารสนเทศที่มีประโยชน์ การระบุแหล่งข้อมูลที่ถูกต้องจึงเป็นสิ่งสำคัญต่อผลลัพธ์ที่ได้จากการวิเคราะห์

การทำเหมืองข้อมูลจำเป็นต้องอาศัยบุคลากรจากหลายฝ่ายและต้องอาศัยความรู้จำนวนมากถึงจะได้รับประโยชน์อย่างแท้จริง เพราะสิ่งที่ได้เป็นเพียงตัวเลขและข้อมูล ที่อาจจะนำไปใช้ประโยชน์ได้หรือใช้ประโยชน์อะไรไม่ได้เลยก็เป็นได้ ผู้ที่ศึกษาการทำเหมืองข้อมูลจึงควรมีความรู้รอบด้านและต้องติดต่อกับทุก ๆ ฝ่าย เพื่อให้เข้าใจถึงขอบเขตของปัญหาโดยแท้จริงก่อน เพื่อให้การทำเหมืองข้อมูลเกิดประโยชน์อย่างแท้จริง

## 2.4 วรรณกรรมที่เกี่ยวข้อง

การวิเคราะห์ข้อมูลการจำแนกเขตพืชมหรือเขตที่ได้รับประทานได้ ในอเมริกาเหนือ เพื่อเผยแพร่ข้อมูลบนเว็บไซต์ กุลธิดา เป็นแผ่น , ญัฐพงษ์ นันตะเสน , (2563) นำข้อมูลเขตในอเมริกาเหนือที่ได้รับประทานได้หรือมีพืชม เพื่อเผยแพร่ข้อมูลบนเว็บไซต์ ด้วยขั้นตอนกระบวนการวิเคราะห์ข้อมูลด้วย CRISP-DM จากเทคนิคการทำเหมืองข้อมูลแบบ Classification ด้วยการสร้างโมเดล Decision Tree เพื่อจัดกลุ่มข้อมูลที่มีความสัมพันธ์กันเพื่อทำการเปรียบเทียบ และเลือกผลลัพธ์ที่ดีที่สุดมานำเสนอ จากนั้นนำข้อมูลสารสนเทศมาทำการแสดงผลแบบ Visualization ในรูปแบบของภาพ โดยใช้โปรแกรม Tableau Public เผยแพร่บน Web Browser

การวิเคราะห์ข้อมูลปริมาณน้ำฝนจากกรมอุตุนิยมวิทยาเพื่อเผยแพร่ข้อมูลบนเว็บไซต์ นที กุลพะกา , วรรณกร กรุขบาง , (2563) นำข้อมูลปริมาณน้ำฝนจากกรม อุตุนิยมวิทยาเพื่อเผยแพร่ข้อมูลบนเว็บไซต์ มาวิเคราะห์และพัฒนาเป็นข้อมูลสารสนเทศ การรายงานผลการแนะนำแนวทางในการแก้ไขปัญหาในพื้นที่ได้อย่างถูกต้อง แม่นยำ และรวดเร็วเพื่อให้เกิดการแก้ไขปัญหาอย่างทันเหตุการณ์ โดยการนำเอากระบวนการหลัก ในการจัดทำ เหมืองข้อมูลเพื่อวิเคราะห์ CRISP-DM หรือ (Cross Industry Standard Process for Data Mining) ดำเนินการตั้งแต่ การวิเคราะห์ปัญหาของข้อมูล จัดการกับการรวบรวมข้อมูล การแปลงข้อมูลให้อยู่ในรูปแบบที่เหมาะสมแก่การวิเคราะห์ข้อมูล การเลือกเทคนิคทางเหมืองข้อมูล การวัดประสิทธิภาพของผลลัพธ์ที่ได้ และเทคโนโลยีสารสนเทศเข้ามาช่วยสนับสนุนกระบวนการใน การนำผลการวิเคราะห์ข้อมูลที่ได้สร้างเป็นสารสนเทศนำไปสู่การรายงานผลทางเว็บไซต์ทำให้ สามารถเข้าถึงข้อมูลได้อย่างเหมาะสมและรวดเร็ว รวมถึงการแนะนำแนวทางในการแก้ไขปัญหา ในพื้นที่อย่างเหมาะสมและนำไปสู่การแก้ไขปัญหาอย่างทันเหตุการณ์

การวิเคราะห์ข้อมูลการค้าชายแดนไทยกับประเทศเพื่อนบ้านจากกรมการค้าต่างประเทศ เพื่อเผยแพร่ข้อมูลบนเว็บไซต์ วรณันท์ภักดิ์ ใจอุดม , เมธิณี ลอจาก , (2563) นำข้อมูลจากกรมการค้าต่างประเทศมา วิเคราะห์และพัฒนาเป็นข้อมูลสารสนเทศด้านการวิเคราะห์ข้อมูลจากกรมการค้าต่างประเทศ เพื่อนำเสนอข้อมูลสารสนเทศบนเว็บไซต์ การรายงานผล การแนะแนวทางในการปฏิบัติและ แนวทางในการแก้ไขปัญหาได้อย่างถูกต้อง แม่นยำ เหมาะสมและรวดเร็ว เพื่อให้เกิดการแก้ไข ปัญหาอย่างทันเหตุการณ์ โดยการนำเอากระบวนการหลักในการจัดทำเหมืองข้อมูลเพื่อการ วิเคราะห์ CRISP-DM หรือ (Cross Industry Standard Process for Data Mining) ดำเนินการตั้งแต่การวิเคราะห์ปัญหาของข้อมูล จัดการรวบรวมข้อมูล การแปลงข้อมูลให้อยู่ในรูปแบบที่เหมาะสมแก่การวิเคราะห์ข้อมูล การเลือกเทคนิคทางเหมืองข้อมูล การวัดประสิทธิภาพของ ผลลัพธ์ที่ได้และนำเทคโนโลยีสารสนเทศเข้ามาช่วยสนับสนุนกระบวนการในการนำผลการ วิเคราะห์ ข้อมูลที่ได้สร้างเป็นสารสนเทศนำไปสู่การรายงานผลทางเว็บไซต์ ทำให้สามารถ เข้าถึงข้อมูลได้อย่างเหมาะสมและรวดเร็วความเห็นของผู้จัดทำคือ เป็นการนำข้อมูลมาวิเคราะห์เพื่อให้สามารถ นำข้อมูลไปช่วยตัดสินใจและเพิ่มความสะดวกระบายในการนำข้อมูลไปใช้

การวิเคราะห์ข้อมูลจากศูนย์ข้อมูลหมอกควันแห่งประเทศไทย เพื่อนำเสนอข้อมูลและสารสนเทศบนเว็บไซต์ วรวิช วิยะกิจ , อนุชา รินหาวุธ , (2562) นำข้อมูลจากศูนย์ข้อมูลหมอกควันแห่งประเทศไทยมาวิเคราะห์และพัฒนาเป็นข้อมูลสารสนเทศด้าน การวิเคราะห์ข้อมูลจากศูนย์ข้อมูลหมอกควันแห่งประเทศไทยเพื่อนำเสนอข้อมูลสารสนเทศบนเว็บไซต์ การรายงานผล การแนะแนวทางในการปฏิบัติ และแนวทางในการแก้ไขปัญหาในพื้นที่ได้อย่างถูกต้อง แม่นยำ เหมาะสมและรวดเร็วเพื่อให้เกิดการแก้ไขปัญหาอย่างทันเหตุการณ์ โดย การนำเอา กระบวนการหลักในการจัดทำเหมืองข้อมูลเพื่อการวิเคราะห์ CRISP-DM หรือ (Cross Industry Standard Process for Data Mining) ดำเนินการตั้งแต่ การวิเคราะห์ปัญหาของข้อมูล จัดการ กับรวบรวมข้อมูล การแปลงข้อมูลให้อยู่ในรูปแบบที่เหมาะสมแก่การวิเคราะห์ข้อมูล การเลือกเทคนิค ทางเหมืองข้อมูล การวัดประสิทธิภาพของผลลัพธ์ที่ได้ และนำเทคโนโลยีสารสนเทศเข้ามาช่วย สนับสนุนกระบวนการในการนำผลการวิเคราะห์ข้อมูลที่ได้สร้างเป็นสารสนเทศนำไปสู่การ รายงานผลทางเว็บไซต์ทำให้สามารถเข้าถึงข้อมูลได้อย่างเหมาะสมและรวดเร็ว เพื่อให้ข้อมูลสารสนเทศเหล่านี้

สามารถนำไปเผยแพร่ต่อสาธารณะชนสร้างความรู้ตระหนักรู้ถึงปัญหา หมอกควันรวมถึงการประชาสัมพันธ์แนะแนวทางในการปฏิบัติ และแนวทางในการแก้ไขปัญหา ในพื้นที่ได้อย่างเหมาะสม นำไปสู่การแก้ไขปัญหาอย่างทันเหตุการณ์ความเห็นของผู้จัดทำคือ เป็นการนำข้อมูลมาวิเคราะห์ เพื่อให้สามารถนำข้อมูลไปช่วยตัดสินใจและช่วยให้คำแนะนำแนวทางการป้องกันการเกิดเหตุการณ์และยังลดปัญหาในการวิเคราะห์ปัญหาควัน

การวิเคราะห์ข้อมูลรายงานการฆาตกรรมปี 2011-2013 เพื่อใช้สำหรับเผยแพร่ข้อมูลบนเว็บไซต์ วรรณพร นันตะเสน , (2563) วิเคราะห์ข้อมูลรายงานการฆาตกรรมปี 2011-2013 เพื่อใช้สำหรับเผยแพร่ข้อมูลบนเว็บไซต์โดยผู้จัดทำได้เลือกใช้ขั้นตอนกระบวนการวิเคราะห์ข้อมูลด้วย CRISP-DM หรือ (Cross Industry Standard Process for Data Mining) มาดำเนินการตั้งแต่ การวิเคราะห์ปัญหา จัดการรวบรวมข้อมูล และแปลงข้อมูลให้อยู่ในรูปแบบที่เหมาะสมแก่การนำไปวิเคราะห์ โดยเลือกใช้เหมือนข้อมูลแบบ Classification ที่สร้างโมเดล Decision Tree โดยใช้โปรแกรม RapidMiner ในการทำเหมืองข้อมูลและข้อมูลที่ได้จากการวิเคราะห์ข้อมูลและแสดงผลในรูปแบบของแผนภาพโดยใช้โปรแกรม Tableau Public ผู้ใช้งานสามารถกรองข้อมูลผ่านเว็บไซต์ โดยระบบ Tableau Public ได้ ซึ่งผู้จัดทำได้เว็บไซต์สำหรับเผยแพร่ข้อมูลสารสนเทศนี้บน Web browser โดยเว็บไซต์ที่สร้างขึ้น ได้ใช้ภาษา HTML CSS PHP และ JavaScript ในการสร้างเว็บไซต์ขึ้นมา และมีส่วนประกอบต่าง ๆ ดังนี้

1. ส่วนเนื้อหาปัญหาอาชญากรรม
2. ส่วนการวิเคราะห์ข้อมูลด้วย Decision Tree
3. ส่วนแสดงผลข้อมูลอาชญากรรม
4. ส่วนแบบฟอร์มแบบทดสอบ
5. ส่วนแบบฟอร์มแบบสอบถาม
6. ส่วนดาวน์โหลดข้อมูล
7. ส่วนอัปโหลดเพิ่ม ลบ แก้ไขข้อมูล

ความเห็นของผู้จัดทำคือ การวิเคราะห์ข้อมูลโดยใช้ HTML CSS PHP Java และการจัดการข้อมูลด้วย Decision tree น่าจะตอบสนองความต้องการได้