บทที่ 2
แนวคิดทฤษฎีเครื่องมือและวรรณกรรมที่เกี่ยวข้อง

โครงงานเรื่องการพัฒนาเว็บแอพพลิเคชั่นเพื่อวิเคราะห์ข้อมูล ปริมาณการใช้งานอินเทอร์เน็ตไร้สาย 6 เขตพื้นที่ มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา โดยนำเทคโนโลยีมาประยุกต์ใช้ในการวิเคราะห์ข้อมูล ข้อมูลการปริมาณการเข้าใช้งานอินเทอร์เน็ตไร้สาย เพื่อสนับสนุนการตัดสินใจในการวางแผนปรับปรุงพัฒนาจุดให้บริการอินเทอร์เน็ตภายใน 6 เขตพื้นที่มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา ด้วยการทำเหมืองข้อมูล (Data Mining) แบบกฎความสัมพันธ์ของข้อมูล (Association analysis) และนำหลักการพยากรณ์ข้อมูลรูปแบบอนุกรมเวลา (Time-series models) และนำรูปแบบที่ได้มาวางแผน ประกอบการตัดสินใจในการพัฒนาจุดให้บริการอินเทอร์เน็ตในอนาคต

ในบทนี้เป็นการนำเสนอเกี่ยวกับ แนวคิด ทฤษฎี เครื่องมือและวรรณกรรมที่ เกี่ยวข้องของการวิเคราะห์ข้อมูลปริมาณการใช้งานอินเทอร์เน็ตไร้สาย นำไปสู่การวางแผนพัฒนาจุดให้บริการอินเทอร์เน็ตและการแสดงผลข้อมูลบนเว็บไซต์ ซึ่งได้รวบรวมการศึกษาเอกสารงานวิจัยที่เกี่ยวข้องกับการวิเคราะห์ข้อมูล เพื่อใช้เป็นแนวทางการศึกษาประกอบด้วยรายละเอียดตามลำดับดังนี้

2.1 แนวคิด

- 2.1.1 แนวคิดเกี่ยวกับการวิเคราะห์ข้อมูล (Data analytic)
- 2.1.2 แนวคิดเกี่ยวกับการพยากรณ์ข้อมูล (Forecasting data)
- 2.1.3 แนวคิดเกี่ยวกับการแสดงผลข้อมูล (Data visualization)

2.2 ทฤษฎี

- 2.2.1 ทฤษฎีเกี่ยวกับการจัดการข้อมูลขนาดใหญ่
- 2.2.2 ทฤษฎีเกี่ยวกับการทำเหมืองข้อมูล
- 2.2.3 ทฤษฎีเกี่ยวกับการออกแบบเว็บไซต์
- 2.2.4 ทฤษฎีเกี่ยวกับชุดคำสั่ง CSS และ Bootstrap
- 2.2.5 ทฤษฎีเกี่ยวกับการ Visualization
- 2.2.6 ทฤษฎีเกี่ยวกับอนุกรมเวลา (Time Series)

2.2.7 ทฤษฎีเกี่ยวกับ FP-Growth
2.3 เครื่องมือในการออกแบบและวิเคราะห์ข้อมูล
2.3.1 แบบจำลองอนุกรมเวลา (Time series analysis)
2.3.2 กระบวนการวิเคราะห์ข้อมูลด้วย (CRISP-DM)
2.3.3 ซอฟต์แวร์ Google Data Studio
2.3.4 โปรแกรม Minitap
2.3.5 โปรแกรม RapidMiner Studio
2.4 วรรณกรรมที่เกี่ยวข้อง
2.5 บทสรุป

2.1 แนวคิด

2.1.1 แนวคิดเกี่ยวกับการวิเคราะห์ข้อมูล (Data analytic) Data analytics

เป็นศาสตร์ของการวิเคราะห์ ข้อมูลต่าง ๆ จาก Big Data เพื่อช่วยในด้านธุรกิจ หรือตามวัตถุประสงค์ อื่นๆ ที่ต้องการซึ่งเริ่มต้น จากความสามารถของเราในการนำข้อมูลเหล่านั้น มาให้อยู่ในรูปแบบที่พร้อมจะประมวลผลได้โดย เทคโนโลยีหรือชุดคำ สั่งและแบบจำลองที่สร้างขึ้นเพื่อนำข้อมูลที่ได้ผ่านการวิเคราะห์แล้วมาใช้ หรือ แปลความหมายโดยบุคคลที่ได้รับการฝึกอบรมการใช้เทคโนโลยีเหล่านั้น รูปแบบของการวิเคราะห์ ข้อมูลสามารถแบ่งได้ดังนี้

1. การวิเคราะห์ข้อมูลแบบพื้นฐาน (Descriptive analytics) เป็นการวิเคราะห์เพื่อแสดงผลของรายการทางธุรกิจเหตุการณ์หรือกิจกรรมต่างๆที่ได้เกิดขึ้นหรืออาจกำลังเกิดขึ้นในลักษณะที่ง่ายต่อการเข้าใจ หรือต่อการตัดสินใจ ตัวอย่างเช่น รายงานการขาย รายงานผลการดำเนินงาน
2. การวิเคราะห์แบบเชิงวินิจฉัย (Diagnostic analytics) เป็นการอธิบายถึงสาเหตุของสิ่งที่เกิดขึ้นปัจจัยต่างๆ และความสัมพันธ์ของปัจจัยหรือตัวแปรต่างๆ ที่มีความสัมพันธ์ต่อกันของสิ่งที่เกิดขึ้น ตัวอย่างเช่นความสัมพันธ์ระหว่างยอดขายต่อกิจกรรมทางการตลาดแต่ละประเภทซึ่งเป็นก้าวใหม่ที่ช่วยเสริมให้ ตัดสินใจไปในทางที่ถูกต้อง
3. การวิเคราะห์แบบพยากรณ์ (Predictive analytics) เป็นการวิเคราะห์เพื่อพยากรณ์สิ่งที่กำลังจะเกิดขึ้นหรือน่าจะเกิดขึ้น โดยใช้ข้อมูลที่ได้เกิดขึ้นแล้วกับแบบจำลอง

ทางสถิติ หรือเทคโนโลยีปัญญาประดิษฐ์ต่างๆ (Artificial intelligence) ตัวอย่างเช่น การพยากรณ์ยอดขาย การพยากรณ์ผลประชามติ

4. การวิเคราะห์แบบให้คำแนะนำ (Prescriptive analytics) เป็นการวิเคราะห์ข้อมูลที่มีความ ซับซ้อนที่สุด เป็นทั้งการพยากรณ์สิ่งต่างๆ ที่ จะเกิดขึ้น ข้อดี ข้อเสีย สาเหตุ และระยะเวลา ของสิ่งที่จะเกิดขึ้น และการให้คำ แนะนำ ทาง เลือกต่างๆที่มีอยู่ และผลของแต่ละทางเลือก

2.1.2 แนวคิดเกี่ยวกับการพยากรณ์ข้อมูล (Forecasting data)

การพยากรณ์เป็นทั้งศาสตร์และศิลป์ในการทำนายเหตุการณ์ในอนาคต ซึ่งอาจนำหลาย ๆ วิธีมาใช้แล้วแต่สถาณการณ์ เช่น นำข้อมูลในอดีตมาพยากรณ์อนาคตโดยอาศัยหลักการทางคณิตศาสตร์เข้าช่วย หรือใช้ดุลยพินิจของผู้พยากรณ์เพียงอย่างเดียว หรืออาจใช้หลาย ๆ วิธีร่วมกันเพื่อให้การพยากรณ์มีความแม่นยำมากขึ้น

2.1.2.1 ความหมายของการพยากรณ์ข้อมูล (Forecasting data)

การพยากรณ์ หมายถึงการคาดการณ์ (Predict) เกี่ยวกับลักษณะหรือแนวโน้ม ของสิ่งที่สนใจที่จะเกิดขึ้นในอนาคต เพื่อใช้เป็นสารสนเทศ (Information) ประกอบการตัดสินใจซึ่ง การพยากรณ์จะต้องดำเนินการเป็นส่วนแรกสุดก่อนการวางแผน หรือการเตรียมการที่จะเริ่ม กระทำการใด ๆ เพื่อความถูกต้องและแม่นยำในการตัดสินใจ ดังนั้นในการดำเนินธุรกิจภายใต้ ความไม่แน่นอนและการเปลี่ยนแปลงที่รวดเร็วเช่นปัจจุบันจึงจำเป็นที่จะต้องทราบถึงความเป็นไป ในอนาคต โดยอาศัยเทคนิคการพยากรณ์ต่าง ๆ เพื่อเพิ่มประสิทธิภาพในการดำเนินงานทางธุรกิจ สามารถตอบสนองต่อความต้องการของลูกค้า เพิ่มผลตอบแทนทางธุรกิจ ลดต้นทุนและความสูญเสียต่าง ๆ ที่อาจเกิดขึ้นเป็นต้น

2.1.2.2 การเลือกเทคนิคการพยากรณ์

1) ระยะของการพยากรณ์ ผู้วิเคราะห์จะต้องกำหนดว่าต้องการพยากรณ์ ระยะสั้น ระยะปานกลาง หรือระยะยาว

2) ลักษณะของข้อมูล การเลือกวิธีการพยากรณ์ควรเลือกจากลักษณะหรือ รูปแบบของข้อมูล ข้อมูลบางชุดอาจจะไม่ค่อยเปลี่ยนแปลงมากนัก บางชุดมีแต่ความผันแปรไม่แน่นอน หรืออาจมีปัจจัย หรือตัวแปรประเภทอื่น ๆ ที่มีอิทธิพลต่อข้อมูลที่ต้องการศึกษา ซึ่งอาจ ต้องใช้วิธีการวิเคราะห์ความถดถอย และสหสัมพันธ์

3) ค่าใช้จ่าย การพิจารณาเลือกเทคนิคการพยากรณ์จะต้องคำนึงถึง ค่าใช้จ่ายด้วย เนื่องจากแต่ละวิธีจะต้องมีจำนวนข้อมูลที่ใช้วิเคราะห์แตกต่างกัน ถ้าใช้ข้อมูลมากจะ ยิ่งต้องเสียค่าใช้จ่ายมาก นอกจากนั้น ถ้าเลือกใช้การวิเคราะห์ความถดถอย และสหสัมพันธ์ จะต้องหาข้อมูลของตัวแปรอิสระที่มีอิทธิพลต่อตัวแปรที่จะพยากรณ์ จะต้องเสียค่าใช้จ่ายในการ หาข้อมูลตัวแปรอิสระเหล่านั้น โดยเฉพาะอย่างยิ่งถ้าตัวแปรอิสระเป็นข้อมูลภายนอกองค์กร ก็จะ ยิ่งเสียค่าใช้จ่ายมากขึ้น

4) ความถูกต้อง สิ่งที่สำคัญที่สุดในการพิจารณาเลือกเทคนิคพยากรณ์ คือ ความถูกต้องของค่าพยากรณ์ นั้นต้องการค่าคลาดเคลื่อนในการพยากรณ์ต่ำสุด

2.1.2.3 ช่วงเวลาของการพยากรณ์ สามารถแบ่งตามระยะเวลาของการพยากรณ์ได้ 3 ประเภทคือ

1) การพยากรณ์ในระยะสั้น เป็นการพยากรณ์เหตุการณ์ที่ไม่เกิน 1 ปี โดยทั่วไปมักจะอยู่ในช่วงไม่เกิน 3 เดือน เช่น การพยากรณ์การวางแผนจัดซื้อ การจัดตารางการ ทำงาน การมอบหมายงาน การพยากรณ์ยอดขาย และการพยากรณ์ระดับการผลิต

2) การพยากรณ์ระยะกลาง เป็นการพยากรณ์เหตุการณ์ที่อยู่ในช่วง 3 เดือน ถึง 3 ปี จะใช้มากในการพยากรณ์การวางแผนการขาย การวางแผนการผลิต การวางแผนด้านงบประมาณเงินสด และการวิเคราะห์การวางแผนการดำเนินงานต่าง ๆ

3) การพยากรณ์ระยะยาว เป็นการพยากรณ์เหตุการณ์ที่มากกว่า 3 ปีขึ้นไป มักใช้สำหรับการวางแผนผลิตภัณฑ์ใหม่ ค่าใช้จ่ายในการลงทุน การขยายทำเลที่ตั้ง และการวิจัย พัฒนา

2.1.2.4 เทคนิคการพยากรณ์ (Forecasting Techniques)

การพยากรณ์เป็นเรื่องการคาดคะเนในอนาคต ดังนั้นการพยากรณ์กับความ จริงนั้นอาจจะเหมือนหรือไม่เหมือนกัน แต่การพยากรณ์ที่ดีควรจะต้องให้ใกล้เคียงกับความเป็น จริงมากที่สุด ดังนั้นเทคนิคและวิธีการพยากรณ์จึงเป็นเรื่องที่จะต้องทำความเข้าใจร่วมกัน เทคนิค ของการพยากรณ์นั้นจะแบ่งออกเป็น 2 ประเภท ตามลักษณะวิธีการพยากรณ์ที่อาศัยข้อมูลในอดีตประกอบการสร้างรูปแบบในการพยากรณ์ได้ดังนี้

(1) เทคนิคการพยากรณ์แบบไม่เป็นทางการ (Informal Forecasting Technique) วิธีพยากรณ์แบบนี้จะใช้ประสบการณ์และดุลพินิจของ ผู้พยากรณ์ โดยผู้พยากรณ์จะเป็นผู้ที่มีความรู้ความเชี่ยวชาญในเรื่องนั้น ๆ อย่างดีจึงสามารถ คาดคะเนสิ่งต่าง ๆ ที่คาดว่าจะเกิดขึ้นได้

(2) เทคนิคการพยากรณ์แบบเป็นทางการ (Formal Forecasting Technique) การพยากรณ์วิธีนี้จะต้องอาศัยข้อมูลมาสนับสนุนและใช้ความรู้ทางสถิติ และคณิตศาสตร์เข้ามาช่วยในการวิเคราะห์ข้อมูล ดังนั้นผู้พยากรณ์จะต้องเข้าใจวิธีการและ ขั้นตอนในการนำข้อมูลมาวิเคราะห์ เทคนิควิธีการพยากรณ์แบบเป็นทางการนี้ยังแบ่งออกได้เป็น 2 จำพวก คือ การพยากรณ์เชิงคุณภาพ (Qualitative Methods) และการพยากรณ์เชิงปริมาณ (Quantitative Methods)

ก) เทคนิคเชิงคุณภาพ (Qualitative Techniques) เทคนิคการพยากรณ์เชิง คุณภาพเป็นเทคนิคที่อาศัยประสบการณ์ผู้พยากรณ์เป็นส่วนใหญ่ ซึ่งอาจจะไม่มีการใช้ข้อมูลในอดีต เนื่องจากไม่ได้มีการเก็บรวบรวมข้อมูลในอดีตไว้หรือมีแต่มีไม่พอเพียงต่อการนำมาสร้าง รูปแบบในการพยากรณ์ ดังนั้นความถูกต้องของการพยากรณ์เชิงคุณภาพจะขึ้นอยู่กับ ประสบการณ์และความสามารถของผู้พยากรณ์เป็นหลัก ซึ่งได้แก่ ทัศนคติของผู้จัดการ การพยากรณ์โดยกลุ่มผู้บริหาร (A Jury of Executive Opinion) พนักงานขาย ทำการพยากรณ์ (Sale forecast Estimate) สำรวจตลาด (Market Research) เทคนิคเดลฟี (Delphi Technique)

ข) เทคนิคเชิงปริมาณ (Quantitative Techniques) เทคนิคการพยากรณ์เชิงปริมาณ จะเป็นเทคนิคที่ต้องใช้ข้อมูลในอดีตมาสร้างรูปแบบการพยากรณ์ในรูปของสมการ คณิตศาสตร์ ดังนั้นความถูกต้องของการพยากรณ์นี้จะขึ้นอยู่กับความแม่นยำของข้อมูลที่มีอยู่ และวิธีการในรูปแบบของสมการทางคณิตศาสตร์ ซึ่งแบ่งได้เป็น 2 ประเภท ดังนี้

(1) รูปแบบปัจจัยสาเหตุ หรือรูปแบบเชิงเหตุผล (Associative Models) เป็นการพยากรณ์ด้วยการวิเคราะห์ปัจจัยต่าง ๆ ที่จะมีผลกระทบต่อสิ่งที่จะพยากรณ์เช่น การ วิเคราะห์การถดถอย (Regression Analysis) หรือ ตัวแบบเศรษฐมิติ (Econometric Model)

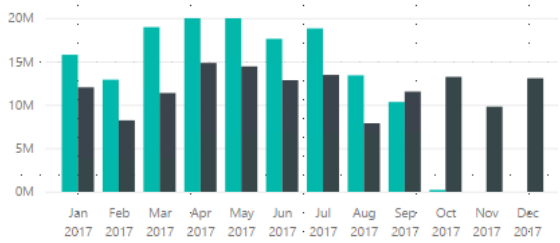
(2) รูปแบบอนุกรมเวลา (Time Series Models) ซึ่งได้แก่ วิธีการหา ค่าเฉลี่ยแบบตรงตัว (Naive Approach) วิธีหาค่าเฉลี่ยเคลื่อนที่ (Moving Average) วิธีการปรับเรียบ แบบเอ็กซ์โปเนนเชียล (Exponential Smoothing) วิธีแยกส่วน (Classical Decomposition) และวิธีการคาดคะเนแนวโน้ม (Trend Projection)

2.1.3 แนวคิดเกี่ยวกับการแสดงผลข้อมูล (Data visualization)

การเลือกรูปแบบ Visualization ให้เหมาะสมกับข้อมูล ในปัจจุบันเป็นยุคเทคโนโลยี เข้าถึงทุกคน ทำให้การรับรู้ข่าวสาร ข้อมูลต่าง ๆ เป็นไปได้ง่าย และรวดเร็วมากขึ้น คนที่ นำเสนอข้อมูล

จึงต้องนำเสนอข้อมูลให้น่าสนใจ เข้าใจง่าย และรวดเร็ว จึงเกิดการสร้าง Data Visualization ขึ้นมา Data Visualization เป็นการใช้ภาพเพื่อแสดงข้อมูลในเชิงปริมาณที่วัดได้ ซึ่ง อาจนำเสนอออกมาในรูปแบบ แผนภูมิ กราฟ กราฟิก และอื่น ๆ อีกมากมาย เพื่อให้เข้าใจได้

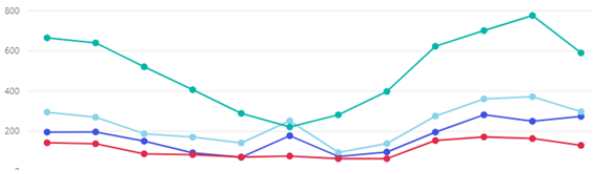
1) แผนภูมิแท่ง (Bar chart) คือกราฟที่ประกอบด้วยรูปสี่เหลี่ยมผืนผ้า (เรียกว่า แท่ง) จำนวนหนึ่ง โดยมีลักษณะเป็นแท่งสูงหรือยาวที่เปลี่ยนแปลงตามขนาด แต่มีความกว้างเท่ากันหมด เราอาจเรียงแท่งเหล่านี้ในทางตั้งหรือทางนอนก็ได้ โดยเว้นระยะช่องว่างตามสมควรและจะต้องเขียนโครงเรื่องจำแนกแต่ละแท่งให้ชัดเจนด้วย แผนภูมิแท่งอาจจะมีการระบายสีหรือแรเงาเพื่อให้ดูเด่น และในกรณีที่มีการเปรียบเทียบกันหลายแท่ง เช่น แผนภูมิแท่งซับซ้อนหรือเชิงประกอบจำเป็นจะต้องระบายสีหรือแรเงา เพื่อจำแนกความแตกต่างของแผนภูมิแต่ละชุดที่นำมาเปรียบเทียบกันนั้น วัตถุประสงค์ของการนำเสนอข้อมูลในรูปแผนภูมิแท่งเพื่อแสดงความเปลี่ยนแปลงและการเปรียบเทียบข้อมูล



ภาพที่ 2.1 แผนภูมิแท่ง (Bar Chart)

ที่มา : blog.1moby.com(ม.ป.ป.)

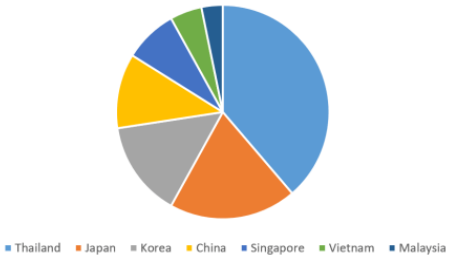
2) แผนภูมิเส้น (Line Charts) แผนภูมิแบบเส้น มีลักษณะคล้ายแผนภูมิแบบแท่ง ซึ่งประกอบด้วยแกนตั้งและนอน เพียงแต่เปลี่ยนจากแท่งข้อมูลเป็นจุดบนแผนภูมินั่นเองแผนภูมิประเภทนี้เหมาะกับการนำเสนอข้อมูลตัวเลขที่มีความสัมพันธ์กับข้อมูลที่มีลักษณะเป็นช่วงใช้แสดงให้เห็นการเปลี่ยนแปลงที่เกิดขึ้นตามเวลาเพื่อดูแนวโน้ม รวมถึงสามารถใช้พยากรณ์แนวโน้มในอนาคตได้เช่น ข้อมูลของยอดขายในแต่ละปี หรือไตรมาส และนำมาวิเคราะห์เพื่อดูแนวโน้ม เป็นต้น



ภาพที่ 2.2 แผนภูมิเส้น (Line Charts)

ที่มา : blog.1moby.com(ม.ป.ป.)

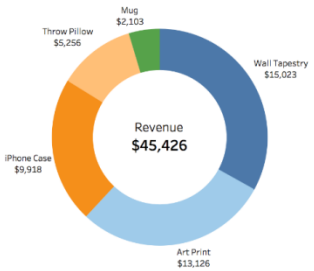
3) แผนภูมิวงกลม (Pie Charts) แผนภูมิวงกลมเหมาะกับการนำเสนอข้อมูลที่มีส่วนประกอบย่อยที่รวมกันเป็นส่วนใหญ่ มีการแบ่งส่วนให้ดูง่าย และสวยงามแต่ในทางกลับกันอาจจะดูยากในเรื่องของการประมาณขนาดของแต่ละชิ้น ยิ่งถ้ามีจำนวนชิ้นมาก จะยิ่งแยกยากเพราะต้องใช้หลายสี ในการนำเสนอข้อมูล เช่น ส่วนแบ่งทางการตลาด (Market Share), ข้อมูลแสดงส่วนผสมต่าง ๆ เป็นต้น



ภาพที่ 2.3 แผนภูมิวงกลม (Pie Charts)

ที่มา : blog.1moby.com(ม.ป.ป.)

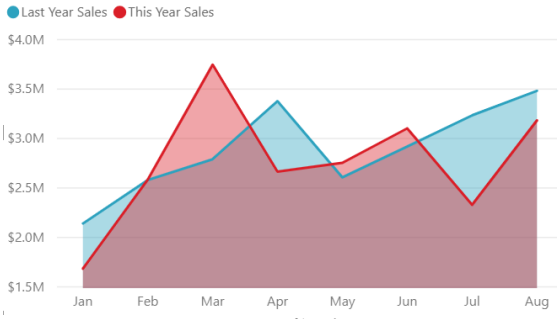
4) แผนภูมิโดนัท (Doughnut Charts) แผนภูมิโดนัทมีหลักการออกแบบเช่นเดียวกับแผนภูมิวงกลมแต่สามารถแสดงชุดข้อมูลได้มากกว่า 1 ชุด โดยนำเสนอข้อมูลเป็นวงกลมซ้อนกันหลายๆ ชั้น นั่นเอง



ภาพที่ 2.4 แผนภูมิโดนัท (Doughnut Charts)

ที่มา : blog.1moby.com(ม.ป.ป.)

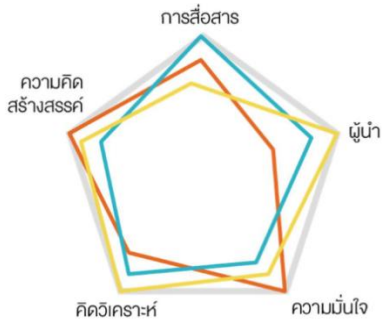
5) แผนภูมิพื้นที่ (Area Charts) มีหน้าตาคล้ายแผนภูมิเส้น แต่มีการแรเงาพื้นที่ใต้ เส้นข้อมูล หรือระหว่าง 2 เส้นเพื่อแสดงให้เห็นปริมาณความแตกต่างระหว่างเส้น เหมาะ สำหรับเน้นความสำคัญของการเปลี่ยนแปลงในช่วงเวลาแสดงให้เห็นผลรวมของความแตกต่าง ระหว่างข้อมูล เช่น ข้อมูลของการซื้อสินค้าในห้างสรรพสินค้าในแต่เดือน ตามหมวดหมู่ต่าง ๆ ไล่ไปเครื่องสำอางค์ เสื้อผ้าแฟชั่น อาหาร ตามลำดับ



ภาพที่ 2.5 แผนภูมิพื้นที่ (Area Charts)

ที่มา : blog.1moby.com(ม.ป.ป.)

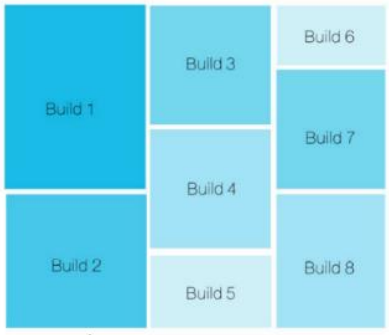
6) แผนภูมิเรดาร์ (Radar Charts) มีลักษณะคล้ายแผนภูมิเส้นที่มีการแสดงผลแบบ วงกลม จำนวนเหลี่ยมของเรดาร์เท่ากับจำนวนหัวข้อของข้อมูล แผนภูมินี้ไม่ได้บอกถึงความ ต่อเนื่องของข้อมูลแต่เหมาะสำหรับการนำเสนอข้อมูลเป็นหัวข้อ แล้วนำมาวิเคราะห์หาจุดอ่อน จุดแข็งของข้อมูลเช่น นำเสนอข้อมูลเกี่ยวกับคุณสมบัติของการรับพนักงานใหม่ เพื่อดูจุดอ่อน จุดแข็งของแต่ละคน เป็นต้น



ภาพที่ 2.6 แผนภูมิเรดาร์ (Radar Charts)

ที่มา : blog.1moby.com(ม.ป.ป.)

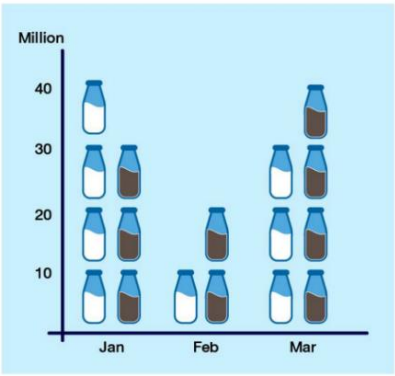
7) แผนภูมิต้นไม้ (Tree Maps) คือการนำเสนอข้อมูลแบบแสดงให้เห็นพื้นที่ แสดง ผลได้ในแบบลำดับชั้น เหมือนแบบโครงสร้างต้นไม้ อาจจะนำเสนอข้อมูลที่ต้องการให้เห็นถึง เขตพื้นที่แสดงพื้นที่สีที่แตกต่างกันได้



ภาพที่ 2.7 แผนภูมิต้นไม้ (Tree Maps)

ที่มา : blog.1moby.com(ม.ป.ป.)

8) แผนภูมิรูปภาพ (Picture Graph) เป็นแผนภูมิที่ประกอบไปด้วยแกนนอน และ แกนตั้ง แต่เลือกใช้รูปภาพ หรือไอคอนแทนจำนวนของสิ่งของนั้น ๆ เช่น การแสดงผลจำนวน ของนมที่ขายได้ในแต่ละเดือน โดยนำเสนอทั้งนมรสจืด รสช็อกโกแลต เปรียบเทียบในแต่ละ เดือนซึ่งมีการนำเสนอลักษณะคล้ายกับกราฟแท่ง แต่เปลี่ยนจากแท่งเป็นรูปกราฟิกของนม 2 รสชาติแทน ก็ทำให้การนำเสนอข้อมูลน่าสนใจมากยิ่งขึ้นซึ่งแนวทางการนำเสนอข้อมูลลักษณะ นี้ต้องอาศัยความคุ้นชินของคนดู เพื่อแทนสัญลักษณ์ภาพลงไป เช่น เมื่อพูดถึงจำนวนคน อาจจะแทนด้วยภาพไอคอนคนหรือเมื่อพูดถึงจำนวนเงิน ควรแทนภาพเป็นเหรียญเงิน หรือ แบงค์แทน ก็จะทำให้คนดูเข้าใจง่ายจากสัญลักษณ์ภาพที่คุ้นเคยอยู่แล้วและยังดึงดูดความ สนใจได้มากกว่าการใช้กราฟแท่งสี่เหลี่ยมอีกด้วย



ภาพที่ 2.8 แผนภูมิรูปภาพ (Picture Graph)

ที่มา : blog.1moby.com(ม.ป.ป.)

2.2 ทฤษฎี

2.2.1 ทฤษฎีเกี่ยวกับการจัดการข้อมูลขนาดใหญ่

2.2.1.1 ข้อมูลขนาดใหญ่

ข้อมูลขนาดใหญ่ หมายถึงข้อมูลที่มีขนาดใหญ่ เร็ว หรือซับซ้อนจนยาก หรือเป็นไปไม่ได้ที่จะประมวลผลโดยใช้วิธีการแบบเดิม การเข้าถึงและจัดเก็บข้อมูลจำนวนมาก เพื่อทำการวิเคราะห์มีมานานแล้ว แต่แนวคิดเกี่ยวกับข้อมูลขนาดใหญ่เป็นที่แพร่หลายในช่วง ต้นปี ค.ศ. 2000 เมื่อดัก ลานีย์ นักวิเคราะห์อุตสาหกรรมได้ให้คำจำกัดความที่เป็นที่เข้าใจกันใน ขณะนี้ ว่า ข้อมูลขนาดใหญ่ประกอบด้วยSVs ดังนี้

Volume (ปริมาณ):องค์กรต่าง ๆ รวบรวมข้อมูลจากหลากหลายแหล่ง ซึ่งรวมถึงธุรกรรมของธุรกิจ อุปกรณ์อัจฉริยะ (IoT) อุปกรณ์อุตสาหกรรม วิดีโอ โซเชียลมีเดีย และอื่น ๆ ในอดีต การจัดเก็บข้อมูลมือเป็นปัญหาใหญ่ – แต่เมื่อค่าใช้จ่ายในการจัดเก็บบนแพลตฟอร์มต่าง ๆ เช่น พื้นที่จัดเก็บข้อมูลส่วนกลาง (Data Lake) และ Hadoop ลดลง ภาระนี้ จึงบรรเทาลง

Velocity (ความเร็ว):ด้วยการเติบโตของ Internet of Things ข้อมูลจะถูกส่งไปยังธุรกิจต่าง ๆ ด้วยความเร็วที่ไม่เคยมีมาก่อนและต้องได้รับการจัดการในเวลาที่ เหมาะสม แท็ก RFID, เซ็นเซอร์ และสมาร์ทมิเตอร์ช่วยผลักดันความต้องการในการจัดการกับ กระแสข้อมูลเหล่านี้ในแบบเรียลไทม์

Variety (ความหลากหลาย):ข้อมูลมีในทุกรูปแบบ นับตั้งแต่ข้อมูลที่มีโครงสร้าง ตัวเลขในฐานข้อมูลแบบดั้งเดิม ไปจนถึงเอกสารข้อความ อีเมล วิดีโอ เสียง ข้อมูล หุ้น และธุรกรรมทางการเงิน

การวิเคราะห์ Big Data ช่วยให้องค์กรควบคุมข้อมูลของพวกเขาและใช้เพื่อ ระบุโอกาสใหม่ ๆ ในทางกลับกันนำไปสู่การเคลื่อนไหวทางธุรกิจที่ชาญฉลาดเพื่อการดำเนินงานที่มีประสิทธิภาพมากขึ้น ผลกำไรที่สูงขึ้นและลูกค้าที่มีความสุขมากขึ้น ใน รายงาน Big Data ใน บริษัท ขนาดใหญ่โดยผู้อำนวยการฝ่ายวิจัยของ Tom Davenport ให้ สัมภาษณ์ว่ามากกว่า 50 ธุรกิจใช้และทำความเข้าใจว่าพวกเขาใช้ Big Data อย่างไร และ พบว่าสามารถช่วยเหลือธุรกิจได้ดังต่อไปนี้

ลดต้นทุน:เทคโนโลยีข้อมูลขนาดใหญ่เช่น Hadoop และการวิเคราะห์ บนคลาวด์นำมาซึ่งความได้เปรียบด้านต้นทุนอย่างมีนัยสำคัญเมื่อพูดถึงการจัดเก็บข้อมูล จำนวนมาก รวมทั้งสามารถระบุวิธีการทำธุรกิจที่มีประสิทธิภาพมากขึ้น

เร็วกว่าและตัดสินใจดีกว่า:ด้วยความเร็วของ Hadoop และการ วิเคราะห์ในหน่วยความจำรวมกับความสามารถในการวิเคราะห์แหล่งข้อมูลใหม่ๆ ของธุรกิจจะ สามารถสร้างข้อมูลได้ทันทีและสามารถดำเนินการต่อได้ทันทีจากการวิเคราะห์นั้น ๆ

ผลิตภัณฑ์และบริการใหม่:ด้วยความสามารถในการวัดความต้องการและความพึงพอใจของลูกค้าผ่านการวิเคราะห์นำมาซึ่งสิ่งที่ลูกค้าต้องการ ซึ่งยกตัวอย่างโดย ดาเวนพอร์ท สามารถชี้ให้เห็นว่าด้วยการวิเคราะห์ข้อมูลขนาดใหญ่ของบริษัทต่าง ๆ จะ สามารถสร้างผลิตภัณฑ์ใหม่ ๆ เพื่อตอบสนองความต้องการของลูกค้าได้เป็นอย่างดี

2.2.1.2 การจัดการข้อมูลขนาดใหญ่

การรวบรวมข้อมูลของ Big Data เป็นการรวบรวมข้อมูลของจากหลากหลายทั้งที่มาและการใช้งานที่แตกต่างกันอย่างมากมาย ซึ่งกลไกและเทคโนโลยีแบบ ดั้งเดิม ETL (extract, transform, and load) ไม่สามารถทำได้ ซึ่ง Big Data หรือ ข้อมูลขนาด ใหญ่ต้องการเทคนิค วิธีการ และเทคโนโลยีใหม่ในการรวบรวมข้อมูลขนาด เทราไบต์ และ อาจจะเป็นระดับเพธาไบต์เลยก็มี ในการรวบรวมข้อมูลนั้นต้องมีการประมวลผล จัดรูปแบบ ให้ เหมาะสำหรับการใช้ในการวิเคราะห์หรือใช้งานสำหรับธุรกิจหรือวัตถุประสงค์นั้น ๆ

การจัดการข้อมูลข้อมูลขนาดใหญ่ หรือ Big Data นั้นมีความต้องการสถานที่จัดเก็บขนาด ใหญ่ การจัดเก็บข้อมูลมูลขนาดใหญ่จะเป็นชนิดใดก็ได้ไม่ว่าจะเป็นแบบ on premises หรือ แบบ cloud ขึ้นกับความต้องการหรือความสะดวกในการใช้ ซึ่งเราสามารถใช้และประเมินผลได้ เช่นเดียวกัน บางครั้งก็มีความจำเป็นที่ต้องจัดเก็บไว้ใกล้กับแหล่งข้อมูล หรือข้อมูลบางอัน ต้องการความยืดหยุ่นสูงและไม่ต้องการบริหารจัดการก็ใช้เป็นแบบ Cloud ซึ่งกำลังเป็นที่นิยม กันเป็นอย่างมาก

การวิเคราะห์การลงทุนสร้างข้อมูลขนาดใหญ่ หรือ Big data จะมีประโยชน์หรือคุ้มค่า ก็ต่อเมื่อคุณใช้และวิเคราะห์ข้อมูล การวิเคราะห์ข้อมูลทำให้เกิดความกระจ่างและชัดเจนในชุด ข้อมูลที่คุณมีอยู่ การสำรวจข้อมูลยังทำให้เราค้นพบสิ่งใหม่ แชร์สิ่งที่ค้นพบใหม่ๆ ต่อคนอื่น สร้างรูปแบบจำลองข้อมูล ด้วยการเรียนรู้ของเครื่องจักรและปัญญาประดิษฐ์ AI และนำข้อมูล เหล่านั้นไปใช้งาน

2.2.1.3 เทคโนโลยีการประมวลผลข้อมูล

การประมวลผลข้อมูลที่เป็น Big Data จะมีทั้งการวิเคราะห์ข้อมูลที่เป็น business intelligence (BI) เพื่อที่จะดึงข้อมูลมานำเสนอ หรือการทำ Predictive Analytics โดยใช้หลักการของ Data Science ความยากของการประมวลผลคือต้องการความเร็วในการ ประมวลผลข้อมูลที่นอกจากมีขนาดใหญ่แล้วบางครั้งยังเป็นข้อมูลที่ไม่มีโครงสร้าง ดังนั้นต้อง จึงมีการนำเทคโนโลยีหรือภาษาต่าง ๆ มาเพื่อให้สามารถประมวลผลข้อมูลได้ ซึ่งในบางครั้ง หน่วยงานอาจต้องพิจารณาต้องเลือกใช้

2.2.2 ทฤษฎีเกี่ยวกับการทำเหมืองข้อมูล

เหมืองข้อมูล คือกระบวนการที่กระทำกับข้อมูลจำนวนมากเพื่อค้นหารูปแบบและความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลนั้น ในปัจจุบันการทำเหมืองข้อมูลได้ถูกนำไประยุกต์ใช้ใน งานหลายประเภท ทั้งในด้านธุรกิจที่ช่วยในการตัดสินใจของผู้บริหาร ในด้านวิทยาศาสตร์และการแพทย์รวมทั้งในด้านเศรษฐกิจและสังคม

2.2.2.1 การทำเหมืองข้อมูล (Data Mining) เปรียบเสมือนวิวัฒนาการหนึ่งในการจัดเก็บและตีความหมายข้อมูล จากเดิมที่มีการจัดเก็บข้อมูลอย่างง่ายๆ มาสู่การจัดเก็บใน รูปแบบฐานข้อมูลที่สามารถดึงข้อมูลสารสนเทศมาใช้จนถึงการทำเหมืองข้อมูลที่สามารถ ค้นพบความรู้ที่ซ่อนอยู่ในข้อมูล หรือจะแยกๆ เป็นข้อๆ ได้ดังนี้

1) กระบวนการหรือการเรียงลำดับของการค้นข้อมูลจำนวนมากและ เก็บข้อมูลที่เกี่ยวข้อง

2) การนำมาใช้โดยหน่วยงานทางธุรกิจและนักวิเคราะห์ทางการเงินหรือการนำมาใช้งานในด้านวิทยาศาสตร์เพื่อเอาข้อมูลขนาดใหญ่ที่สร้างโดยวิธีการทดลองและการสังเกตการณ์ที่ทันสมัย

3) การสกัดหรือแยกข้อมูลที่เป็นประโยชน์จากข้อมูลขนาดใหญ่หรือฐานข้อมูล

4) การวางแผนทรัพยากรขององค์กรโดยสามารถวิเคราะห์ทางสถิติและตรรกะของข้อมูลขนาดใหญ่เป็นการมองหารูปแบบที่สามารถช่วยการตัดสินใจได้

2.2.2.2 ขั้นตอนการทำเหมืองข้อมูล

ประกอบด้วยขั้นตอนการทำงานย่อยที่จะเปลี่ยนข้อมูลดิบให้กลายเป็นความรู้ ประกอบด้วยขั้นตอนดังนี้

1) Business Understanding เน้นไปที่การทำความเข้าใจในงาน ระบุโอกาสและหาปัญหาที่จะเกิดขึ้น กำหนดขอบเขตของข้อมูลที่จะนำวิเคราะห์ ซึ่งต้องสามารถระบุ ผลลัพธ์ที่มีได้

2) Data Understanding ทำความเข้าใจข้อมูลโดยการรวบรวมข้อมูลที่เกี่ยวข้อง คัดเลือกให้เหลือเพียงข้อมูลที่มีความถูกต้องและสำคัญต่องานมาทำการวิเคราะห์

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

4) Modeling การสร้างแบบจำลองเพื่อวิเคราะห์ข้อมูลที่ได้จากขั้นตอนที่ 3 พร้อมทดสอบผลลัพธ์แบบจำลองเพื่อให้ได้คำตอบที่ดีที่สุด บางครั้งอาจมีการย้อนกลับไปปรับการเตรียมข้อมูลเพื่อให้ได้แบบจำลองที่เหมาะสมที่สุด

5) Evaluation การประเมินผลลัพธ์ที่ได้ก่อนที่จะนำไปใช้จริง ว่าตรงกับวัตถุประสงค์หรือเป้าหมายที่ได้ตั้งไว้หรือมีความน่าเชื่อถือมากน้อยเพียงใด หากไม่ได้ผลลัพธ์ ตามวัตถุประสงค์ต้องย้อนกลับไปปรับปรุงแก้ไขการดำเนินงานในขั้นตอนก่อนหน้า

6) Deployment การนำเอาข้อมูลที่เป็นผลลัพธ์จากทั้งหมด มาใช้ประโยชน์ตามวัตถุประสงค์ และทำการประเมินผลลัพธ์ที่ได้ว่ามีประสิทธิภาพตรงตามวัตถุประสงค์ที่ตั้งไว้หรือไม่

2.2.3 ทฤษฎีเกี่ยวกับการออกแบบเว็บไซต์

2.2.3.1 หลักในการออกแบบเว็บไซต์

หน้าเว็บเป็นสิ่งแรกที่ผู้ใช้จะได้เห็นขณะที่เปิดเข้าสู่เว็บไซต์ และยังเป็นสิ่งแรกที่แสดงถึงประสิทธิภาพในการออกแบบเว็บไซต์อีกด้วย หน้าเว็บจึงเป็นสิ่งสำคัญมาก เพราะเป็นสื่อกลางให้ผู้ชมสามารถใช้ประโยชน์จากข้อมูลของระบบงานของเว็บไซต์นั้นได้ โดย ปกติหน้าเว็บจะประกอบด้วย รูปภาพ ตัวอักษร สีพื้น ระบบเนวิเกชั่น และองค์ประกอบอื่น ๆ ที่ ช่วยสื่อความหมายของเนื้อหาและอำนวยความสะดวกต่อการใช้งานหลักสำคัญในการ ออกแบบหน้าเว็บก็

คือ การใช้รูปภาพและองค์ประกอบต่าง ๆ ร่วมกันเพื่อสื่อความหมาย เกี่ยวกับเนื้อหาหรือลักษณะสำคัญของเว็บไซต์ โดยมีเป้าหมายสำคัญเพื่อการสื่อความหมายที่ ชัดเจนและน่าสนใจ บนพื้นฐานของความเรียบง่ายและความสะดวกของผู้ใช้ การออกแบบ เว็บไซต์ ต้องคำนึงถึง

1) ความเรียบง่าย ได้แก่ มีรูปแบบที่เรียบง่าย ไม่ซับซ้อน และใช้งานได้สะดวก ไม่มีกราฟิกหรือตัวอักษรที่เคลื่อนไหวอยู่ตลอดเวลา ชนิดและสีของตัวอักษรไม่มากจนเกินไปทำให้วุ่นวาย

2) ความสม่ำเสมอ ได้แก่ ใช้รูปแบบเดียวกันตลอดทั้งเว็บไซต์ เช่น รูปแบบของหน้า สไตล์ของกราฟิก ระบบเนวิเกชันและโทนสี ควรมีความคล้ายคลึงกันตลอดทั้ง เว็บไซต์

3) ความเป็นเอกลักษณ์ การออกแบบเว็บไซต์ควรคำนึงถึงลักษณะขององค์กร เพราะรูปแบบของเว็บไซต์จะสะท้อนถึงเอกลักษณ์และลักษณะขององค์กรนั้น ๆ เช่น ถ้าเป็นเว็บไซต์ของทาง ราชการ จะต้องดูน่าเชื่อถือไม่เหมือนสวนสนุก ฯลฯ

4) เนื้อหาที่มีประโยชน์ เนื้อหาเป็นสิ่งที่สำคัญที่สุดในเว็บไซต์ ดังนั้นควรจัดเตรียมเนื้อหาและข้อมูลที่ผู้ใช้ต้องการให้ถูกต้อง และสมบูรณ์ มีการปรับปรุงและเพิ่มเติมให้ ทันเหตุการณ์อยู่เสมอ เนื้อหาไม่ควรซ้ำกับเว็บไซต์อื่น จึงจะดึงดูดความสนใจ

5) ระบบเนวิเกชันที่ใช้งานง่าย ต้องออกแบบให้ผู้ใช้เข้าใจง่ายและใช้งานสะดวก ใช้กราฟิกที่สื่อความหมายร่วมกับคำอธิบายที่ชัดเจน มีรูปแบบและลำดับของรายการที่สม่ำเสมอ เช่น วางไว้ ตำแหน่งเดียวกันของทุกหน้า

6) ลักษณะที่น่าสนใจ หน้าตาของเว็บไซต์จะต้องมีความสัมพันธ์กับคุณภาพขององค์ประกอบต่าง ๆ เช่น คุณภาพของกราฟิกที่จะต้องสมบูรณ์ การใช้สี การใช้ตัวอักษรที่อ่านง่าย สบายตา การใช้โทนสีที่เข้ากันลักษณะหน้าตาที่น่าสนใจนั้นขึ้นอยู่กับ ความชอบของแต่ละบุคคล

7) การใช้งานอย่างไม่จำกัด ผู้ใช้ส่วนใหญ่สามารถเข้าถึงได้มากที่สุด เลือกใช้บราวเซอร์ชนิดใดก็ได้ในการเข้าถึงเนื้อหาสามารถแสดงผลได้ทุกระบบปฏิบัติการและความละเอียดหน้าจอต่าง ๆ กันอย่างไม่มีปัญหาเป็นลักษณะสำคัญสำหรับผู้ใช้ที่มีจำนวนมาก

8) คุณภาพในการออกแบบ การออกแบบและเรียบเรียงเนื้อหาอย่างรอบคอบ สร้างความรู้สึกว่าเว็บไซต์มีคุณภาพ ถูกต้อง และเชื่อถือได้

9) ลิงค์ต่าง ๆ จะต้องเชื่อมโยงไปหน้าที่มีอยู่จริงและถูกต้อง ระบบการทำงานต่าง ๆ ในเว็บไซต์จะต้องมีความแน่นอนและทำหน้าที่ได้อย่างถูก

2.2.3.3 โครงสร้างเว็บไซต์

1) เว็บที่มีโครงสร้างแบบเรียงลำดับ (Sequential Structure) เป็นโครงสร้างแบบธรรมดาที่ใช้กันมากที่สุดเนื่องจากง่ายต่อการจัดระบบข้อมูล ข้อมูลที่นิยม จัด ด้วยโครงสร้างแบบนี้มักเป็นข้อมูลที่มีลักษณะเป็นเรื่องราวตามลำดับของเวลา เช่น การ เรียงลำดับตามตัวอักษร ดรรชนี สารานุกรม หรืออภิธานศัพท์ โครงสร้างแบบนี้ เหมาะกับ เว็บไซต์ที่มีขนาดเล็ก เนื้อหาไม่ซับซ้อนใช้การลิงก์ (Link) ไปทีละหน้า ทิศทางของการเข้าสู่ เนื้อหา (Navigation) ภายในเว็บจะเป็นการดำเนินเรื่องในลักษณะเส้นตรง โดยมี ปุ่มเดินหน้าถอยหลังเป็นเครื่องมือหลักในการกำหนดทิศทาง ข้อเสียของโครงสร้างระบบนี้คือ ผู้ใช้ไม่ สามารถกำหนดทิศทางการเข้าสู่เนื้อหาของตนเองได้ ทำให้เสียเวลาเข้าสู่เนื้อ

ภาพที่ 2.9 เว็บที่มีโครงสร้างแบบเรียงลำดับ

ที่มา : sites.google.com(ม.ป.ป.)

2) เว็บที่มีโครงสร้างแบบลำดับขั้น (Hierarchical Structure) เป็นวิธีที่ดีที่สุดวิธีหนึ่งในการจัดระบบโครงสร้างที่มีความซับซ้อนของข้อมูล โดยแบ่งเนื้อหา ออกเป็นส่วนต่างๆ และมีรายละเอียดย่อยๆ ในแต่ละส่วนลดหลั่นกันมาในลักษณะแนวคิดเดียวกับ แผนภูมิองค์กร จึงเป็นการง่ายต่อการทำความเข้าใจกับโครงสร้างของเนื้อหาในเว็บลักษณะนี้ ลักษณะ เด่นเฉพาะของ เว็บประเภทนี้คือการมีจุดเริ่มต้นที่จุดร่วมจุดเดียว นั่นคือ โฮมเพจ (Homepage) และเชื่อมโยงไปสู่เนื้อหา ในลักษณะเป็นลำดับจากบนลงล่าง

ภาพที่ 2.10 เว็บที่มีโครงสร้างแบบลำดับขั้น

ที่มา : sites.google.com(ม.ป.ป.)

3) เว็บที่มีโครงสร้างแบบตาราง (Grid Structure) โครงสร้างรูปแบบนี้มีความซับซ้อนมากกว่ารูปแบบที่ผ่านมา การออกแบบเพิ่มความยืดหยุ่น ให้แก่การเข้าสู่เนื้อหาของผู้ใช้ โดยเพิ่มการเชื่อมโยงซึ่งกันและกันระหว่างเนื้อหาแต่ละส่วน เหมาะแก่ การแสดงให้เห็นความสัมพันธ์กันของเนื้อหา การเข้าสู่เนื้อหาของผู้ใช้จะไม่เป็นลักษณะเชิงเส้นตรงเนื่องจากผู้ใช้สามารถเปลี่ยนทิศทางการเข้าสู่เนื้อหาของตนเองได้

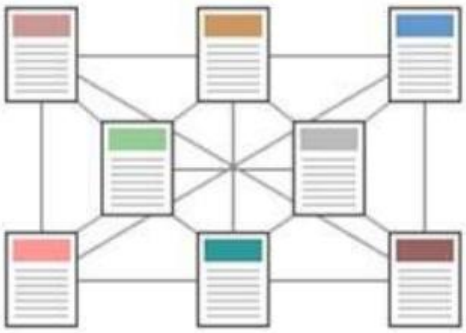
ภาพที่ 2.11 เว็บที่มีโครงสร้างแบบตาราง

ที่มา : sites.google.com(ม.ป.ป.)

ในการจัดระบบโครงสร้างแบบนี้ เนื้อหาที่นำมาใช้แต่ละส่วนควรมีลักษณะที่เหมือนกัน และ สามารถใช้รูปแบบร่วมกัน หลักการออกแบบคือนำหัวข้อทั้งหมดมาบรรจุลงในที่เดียวกันซึ่งโดยทั่วไป จะเป็นหน้าแผนภาพ (Map Page) ที่แสดงในลักษณะเดียวกับโครงสร้างของเว็บ เมื่อผู้ใช้คลิกเลือก หัวข้อใด ก็จะเข้าไปสู่หน้าเนื้อหา (Topic Page) ที่แสดงรายละเอียดของหัวข้อนั้นๆ และภายในหน้านั้น ก็จะมีการเชื่อมโยงไปยังหน้ารายละเอียดของ หัวข้ออื่นที่เป็นเรื่องเดียวกัน นอกจากนี้ยังสามารถนำ โครงสร้างแบบเรียงลำดับและแบบลำดับ ขั้นมาใช้ร่วมกันได้อีกด้วย ถึงแม้โครงสร้างแบบนี้ อาจจะสร้างความยุ่งยากในการเข้าใจได้ และ อาจเกิดปัญหาการคงค้าง ของหัวข้อ (Cognitive Overhead) ได้ แต่จะเป็นประโยชน์ที่สุดเมื่อผู้ ใช้ได้เข้าใจถึงความสัมพันธ์ ระหว่างเนื้อหา ในส่วนของการออกแบบจำเป็นจะต้องมีการ วางแผนที่ดีเนื่องจากมีการเชื่อมโยงที่เกิดขึ้น ได้หลายทิศทาง นอกจากนี้การปรับปรุงแก้ไขอาจ เกิดความยุ่งยากเมื่อต้องเพิ่มเนื้อหาในภายหลัง

4) เว็บที่มีโครงสร้างแบบใยแมงมุม (Web Structure) โครงสร้างประเภท นี้จะมีความยืดหยุ่นมากที่สุด ทุกหน้าในเว็บสามารถจะเชื่อมโยงไปถึงกัน ได้หมด เป็นการสร้างรูปแบบการเข้าสู่เนื้อหาที่เป็นอิสระ ผู้ใช้สามารถกำหนดวิธีการเข้าสู่เนื้อหาได้ด้วย ตนเอง การเชื่อมโยงเนื้อหาแต่ละหน้าอาศัยการโยงใยข้อความที่มีมโนทัศน์ (Concept) เหมือนกัน ของแต่ ละหน้าในลักษณะของไฮเปอร์เท็กซ์หรือไฮเปอร์มีเดีย โครงสร้างลักษณะนี้จัดเป็นรูปแบบที่ ไม่มี

โครงสร้างที่แน่นนอนตายตัว (Unstructured) นอกจากนี้การเชื่อมโยงไม่ได้จำกัดเฉพาะเนื้อหาภายในเว็บนั้น ๆ แต่สามารถเชื่อมโยงออกไปสู่เนื้อหาจากเว็บภายนอกได้



ภาพที่ 2.12 เว็บที่มีโครงสร้างแบบใยแมงมุม

ที่มา : sites.google.com(ม.ป.ป.)

ลักษณะการเชื่อมโยงในเว็บนั้น นอกเหนือจากการใช้ไฮเปอร์เท็กซ์หรือไฮเปอร์มีเดีย กับข้อความที่มีมโนทัศน์ (Concept) เหมือนกันของแต่ละหน้าแล้ว ยังสามารถใช้ลักษณะการเชื่อมโยง จากรายการที่รวบรวมชื่อหรือหัวข้อของเนื้อหาแต่ละหน้าไว้ ซึ่งรายการนี้ จะปรากฏอยู่บริเวณใด บริเวณหนึ่งในหน้าจอ ผู้ใช้สามารถคลิกที่หัวข้อใดหัวข้อหนึ่งในรายการ เพื่อเลือกที่จะเข้าไปสู่หน้าใด ๆ ก็ได้ตามความต้องการ ข้อดีของรูปแบบนี้คือง่ายต่อผู้ใช้ในการท่องเที่ยวบนเว็บ โดยผู้ใช้สามารถกำหนดทิศทาง การเข้าสู่เนื้อหาได้ด้วยตนเอง แต่ข้อเสียคือ ถ้ามีการเพิ่มเนื้อหาใหม่ๆ อยู่เสมอจะเป็นการยากในการ ปรับปรุง นอกจากนี้การเชื่อมโยง ระหว่างข้อมูลที่มีมากมายนั้นอาจทำให้ผู้ใช้เกิดการสับสนและ เกิดปัญหาการคงค้างของหัวข้อ (Cognitive Overhead) ได้

2.2.3.3 การใช้สีในการออกแบบเว็บไซต์

การสร้างสีสันบนหน้าเว็บเป็นสิ่งที่สื่อความหมายของเว็บไซต์ได้อย่างชัดเจน การเลือกใช้สีให้เหมาะสม กลมกลืน ไม่เพียงแต่จะสร้างความพึงพอใจให้กับผู้ใช้ แต่ยังสามารถทำให้เห็นถึงความแตกต่างระหว่างเว็บไซต์ได้ สีเป็นองค์ประกอบหลักสำหรับการ ตกแต่งเว็บ จึงจำเป็นอย่างยิ่งที่จะต้องทำความเข้าใจเกี่ยวกับการใช้สีระบบสีที่แสดงบน จอคอมพิวเตอร์ มีระบบการแสดงผลผ่านหลอดลำแสงที่เรียกว่า CRT (Cathode ray tube) โดย มีลักษณะระบบสีแบบบวก อาศัยการผสมของของแสงสีแดง สีเขียว และสีน้ำเงิน หรือระบบสี RGB สามารถกำหนดค่าสีจาก 0 ถึง 255 ได้ จากการรวมสีของแม่สีหลักจะทำให้เกิดแสงสีขาว มีลักษณะเป็นจุดเล็ก ๆ บนหน้าจอไม่สามารถมองเห็นด้วยตาเปล่าได้ จะมองเห็นเป็นสีที่ถูก ผสมเป็นเนื้อสีเดียวกัน

แล้ว จุดแต่ละจุดหรือพิกเซล (Pixel) เป็นส่วนประกอบของภาพบนหน้า จอคอมพิวเตอร์ โดยจำนวนบิตที่ใช้ในการกำหนดความสามารถของการแสดงสีต่าง ๆ เพื่อ สร้างภาพบนจอนั้นเรียกว่า บิตเด็ป (Bit depth) ในภาษา HTML มีการกำหนดสีด้วยระบบ เลขฐานสิบหก ซึ่งมีเครื่องหมาย (#) อยู่ด้านหน้าและตามด้วยเลขฐานสิบหกจำนวนอักษรอีก 6หลัก โดยแต่ละไบต์ (byte) จะมีตัวอักษรสองตัว แบ่งออกเป็น 3 กลุ่ม เช่น #FF12AC การใช้ ตัวอักษรแต่ละไบต์นี้เพื่อกำหนดระดับความเข้มของแม่สีแต่ละสีของชุดสี RGB โดย 2 หลักแรก แสดงถึงความเข้มของสีแดง 2 หลักต่อมา แสดงถึงความเข้มของสีเขียว 2 หลักสุดท้ายแสดงถึง ความเข้มของสีน้ำเงิน

สีมีอิทธิพลในเรื่องของอารมณ์การสื่อความหมายที่เด่นชัด กระตุ้นการรับรู้ทางด้านจิตใจมนุษย์ สีแต่ละสีให้ความรู้สึก อารมณ์ที่ไม่เหมือนกัน สีบางสีให้ความรู้สึก สงบ บางสีให้ความรู้สึกตื่นเต้นรุนแรง สีจึงเป็นปัจจัยสำคัญอย่างยิ่งต่อการออกแบบเว็บไซต์ ดังนั้นการเลือกใช้โทนสีภายในเว็บไซต์เป็นการแสดงถึงความแตกต่างของสีที่แสดงออกทาง อารมณ์ มีชีวิตชีวาหรือเศร้าโศก รูปแบบของสีที่สายตาของมนุษย์มองเห็น สามารถแบ่ง ออกเป็น 3 กลุ่ม คือ

1) สีโทนร้อน (Warm Colors) เป็นกลุ่มสีที่แสดงถึงความสุข ความปลอบโยน ความอบอุ่น และดึงดูดใจ สีกลุ่มนี้เป็นกลุ่มสีที่ช่วยให้หายจากความเฉื่อยชา มี ชีวิตชีวามากยิ่งขึ้น

2) สีโทนเย็น (Cool Colors) แสดงถึงความที่ดูสุภาพ อ่อนโยน เรียบร้อย เป็นกลุ่มสีที่มีคนชอบมากที่สุด สามารถโน้มนาวในระยะไกลได้

3) สีโทนกลาง (Neutral Colors) สีที่เป็นกลาง ประกอบด้วย สีดำ สีขาว สีเทา และสีน้ำตาล กลุ่มสีเหล่านี้คือ สีกลางที่สามารถนำไปผสมกับสีอื่น ๆ เพื่อให้เกิดสีกลาง ขึ้นมา

2.2.4 ทฤษฎีเกี่ยวกับชุดคำสั่งภาษาของการออกแบบเว็บไซต์

2.2.4.1 CSS (ย่อมาจาก Cascading Style Sheet) มักเรียกโดยย่อว่า "สไตล์ชีต" คือภาษาที่ใช้เป็นส่วนของการจัดรูปแบบการแสดงผลเอกสาร HTML โดยที่ CSS กำหนด กฎเกณฑ์ในการระบุรูปแบบ (หรือ "Style") ของเนื้อหาในเอกสาร อันได้แก่ สีของข้อความ สีพื้น หลัง ประเภทตัวอักษร และการจัดวางข้อความ ซึ่งการกำหนดรูปแบบ หรือ Style นี้ใช้หลักการ ของการแยกเนื้อหาเอกสาร HTML ออกจากคำสั่งที่ใช้ในการจัดรูปแบบการแสดงผล กำหนดให้ รูปแบบของการแสดงผลเอกสาร ไม่ขึ้นอยู่กับเนื้อหาของเอกสาร เพื่อให้ง่ายต่อการจัดรูปแบบ การแสดงผลลัพธ์ของเอกสาร HTML โดยเฉพาะในกรณีที่มีการเปลี่ยนแปลงเนื้อหาเอกสาร บ่อยครั้ง หรือต้องการควบคุมให้รูปแบบการแสดงผลเอกสาร HTML มีลักษณะของความ สม่ำเสมอทั่วกันทุกหน้าเอกสาร

ภายในเว็บไซต์เดียวกัน โดยกฎเกณฑ์ในการกำหนดรูปแบบ (Style) เอกสาร HTML ถูกเพิ่มเข้ามาครั้งแรกใน HTML 4.0 เมื่อปีพ.ศ. 2539 ในรูปแบบของ CSS level 1 Recommendations ที่กำหนดโดย องค์กร World Wide Web Consortium หรือ W3C

2.2.4.2 Bootstrap คือชุดคำสั่งที่ประกอบด้วยภาษา CSS, HTML และ Javascript เป็นชุดคำสั่งที่ถูกพัฒนาขึ้นมาเพื่อกำหนดกรอบหรือรูปแบบการพัฒนาเว็บไซต์ในส่วนของการปฏิสัมพันธ์กับผู้ใช้งานเว็บไซต์ (User Interface) เราจึงสามารถเรียก Bootstrapว่าเป็น Front-end framework คือใช้สำหรับ พัฒนาเว็บไซต์ส่วนการแสดงผล ซึ่งแตกต่างจากภาษาประเภท Server Side Script อย่าง PHP, Python หรือภาษาอื่น ๆ

2.2.5 ทฤษฎีเกี่ยวกับการ visualization

Data Visualization หรือ Information Visualization คือ การนำข้อมูลในเชิงปริมาณทั้งที่จัดเก็บไว้ในรูปแบบของข้อมูลจำนวนน้อย และข้อมูลจำนวนมหาศาล (Big Data) มาประมวลผล จากนั้นจึงนำมาแสดงผลในรูปแบบของกราฟ แผนภูมิอินฟอร์เมชันกราฟิก หรือแม้กระทั่งอินเทอร์แอคทีฟกราฟิก ที่ผู้บริโภคสามารถคลิกหรือมีปฏิสัมพันธ์กับกราฟิกนั้น ๆ ได้ ซึ่งปัจจุบันสื่อหลายสำนักในประเทศไทย เริ่มมีการนำเสนอข่าวโดยการนำข้อมูลแบบประยุกต์ ด้วยภาพ (Data Visualization) การถ่ายทอดข้อมูลในเชิงปริมาณที่มีความซับซ้อนได้อย่างมี ประสิทธิผล ซึ่งคำว่า “ประสิทธิผล” ในที่นี้หมายถึงมีความชัดเจน (Clarity), มีความแม่นยำ (Precision), และมีประสิทธิภาพ (Efficiency) หากไม่มีการทำ Data Visualization แล้ว อาจทำให้ เราไม่สามารถค้นพบนัยยะของข้อมูลในแง่ของแนวโน้ม, รูปแบบพฤติกรรม, และความสัมพันธ์ เชื่อมโยงได้

Visualization คือ การจินตนาการ หรือสร้างภาพขึ้นในความคิด ซึ่งเป็นกระบวนการ ทำให้เกิดความรู้ความเข้าใจสำหรับเรื่องที่จินตนาการยาก เข้าใจยาก วิธีการที่เป็นทางลัดก็คือ การสร้าง ภาพ ให้เป็น บันไดความคิด ไปสู่ การใช้ความคิดอีกระดับ บันไดนี้จะช่วยตัดปริมาณ ข้อมูล ช่วยลดภาระการคำนวณหรือการนำไปผ่านหลากหลายกระบวนการความคิด เพื่อนำไปสู่คำตอบที่ต้องการได้โดยเร็วและถูกต้อง

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

Data Visualization เป็นการนำข้อมูลมาผสมผสานกับจินตนาการ เพื่อสร้างภาพในความคิดขึ้นมา ซึ่งมีกระบวนการนำเสนอข้อมูลที่มีความซับซ้อนหรือข้อมูลเชิงปริมาณ ให้ สามารถเข้าใจได้ง่าย ในแบบของ กราฟ แผนภูมิ

2.2.5.1 โปรแกรมสำหรับการสร้างแดชบอร์ด (Data Visualization)

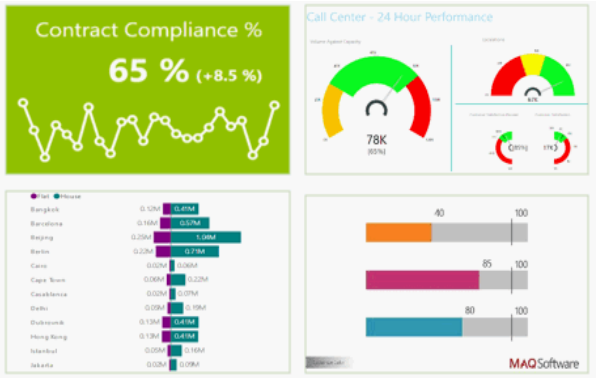
Tools ที่นิยมใช้ในตอนนี้ได้แก่ Tableau, Microsoft Power BI, Qlik View, Google Charts, Fusion Charts,Data wrapper และอื่น ๆ อีกมากมาย

ภาพที่ 2.13 Tools Data Visualization

ที่มา : autosoft.in.th(ม.ป.ป.)

2.2.5.2 รูปแบบในการใช้ Data Visualization

การนำเสนอแบบทิศทางหรือแนวโน้ม (Trending) เราใช้กราฟที่แสดงผลแบบทิศทางหรือแนวโน้ม เพื่อนำเสนอข้อมูลให้เห็นจำนวนข้อมูลที่เกิดขึ้นในแต่ละช่วงเวลา(period) รวมถึงเน้นข้อมูลที่ต้องการนำเสนอ เช่น Line Chart, Bar Chart, Radar Chart, Area Chart เป็นต้น



ภาพที่ 2.14 การนำเสนอแบบทิศทางหรือแนวโน้ม (Trending)

ที่มา : autosoft.in.th(ม.ป.ป.)

2.2.6 ทฤษฎีเกี่ยวกับอนุกรมเวลา (Time Series)

อนุกรมเวลาคือเซตของข้อมูลเชิงปริมาณที่มีการจัดเก็บในช่วงเวลาที่ติดต่อกัน ส่วนข้อมูลอนุกรมเวลา (Time Series Data) คือ ชุดของข้อมูลที่มีการเก็บรวบรวมตามระยะเวลาที่ ติดต่อกันอย่างเป็นระบบโดยทั่ว ไปข้อมูลอนุกรมเวลาประกอบด้วยองค์ประกอบ 4 ส่วน คือ แนวโน้ม (Trend : T), ฤดูกาล (Seasonal : S), วัฏจักร (Cycle : C) และเหตุการณ์ที่ผิดปกติหรือ เหตุการณ์ความไม่แน่นอน (Irregular: I) สำหรับรูปแบบของอนุกรมเวลาโดยทั่วไปนั้น มีอยู่ 2 รูปแบบคือ

1) รูปแบบบวก $Y = T + S + C + I$

2) รูปแบบคูณ $Y = T \times S \times C \times I$

โดยที่ Y คือ อนุกรมเวลา

T คือ อิทธิพลของแนวโน้ม

S คือ อิทธิพลของฤดูกาล

C คือ อิทธิพลของวัฏจักร

I คือ อิทธิพลของความไม่แน่นอน

การวิเคราะห์อนุกรมเวลาเป็นวิธีที่ใช้ในการวิเคราะห์ข้อมูลหรือค่าสังเกตที่มีการเปลี่ยนแปลงไปตามลำดับ เวลาที่เกิดขึ้น หรือการเปลี่ยนแปลงของตัวแปรในช่วงเวลาที่ผ่านไป ลักษณะของการเปลี่ยนแปลงอาจมีหรือไม่มีรูปแบบก็ได้ แต่ถ้าอนุกรมเวลาแสดงให้เห็นรูปแบบการเปลี่ยนแปลงในช่วงเวลาที่ผ่านมาในอดีตก็จะทำให้สามารถคาดการณ์ได้ว่าในอนาคต ลักษณะการเปลี่ยนแปลงควรอยู่ในรูปแบบใด และสามารถพยากรณ์การเปลี่ยนแปลงข้อมูลในอนาคตได้ การวิเคราะห์ข้อมูลอนุกรมเวลานี้ จะขึ้นอยู่กับการเปลี่ยนแปลงของเวลาในอดีตเป็นพื้นฐาน

2.2.6.1 องค์ประกอบของอนุกรมเวลา

ในการวิเคราะห์อนุกรมเวลา ผู้วิเคราะห์จะแยกองค์ประกอบต่าง ๆ ที่ ประกอบกันขึ้นเป็นอนุกรมเวลา โดยจะมีการเปลี่ยนแปลงไปตามอิทธิพลต่าง ๆ เช่น การ เปลี่ยนแปลงการผลิต เทคโนโลยี สภาวะอากาศ เป็นต้น ในการหาคุณลักษณะของอนุกรมเวลาเราสามารถใช้แบบจำลองได้หลายแบบ แบบจำลองที่ใช้โดยนักเศรษฐศาสตร์แบบหนึ่ง คือ แบบจำลอง แบบคลาสสิก (classical model) เป็นการอธิบายถึงองค์ประกอบของการแปรผันของอนุกรมเวลา 4 ส่วน ดังนี้

1) ค่าแนวโน้ม (Secular trend) แทนด้วย Tt เป็นการเปลี่ยนแปลงข้อมูล มีลักษณะราบเรียบ แนวโน้ม อาจมีลักษณะเป็นเส้นตรงหรือเส้นโค้งในทางเพิ่มขึ้นหรือลดลง ค่า

แนวโน้มของข้อมูลเป็นการเคลื่อนไหวในช่วงระยะเวลาที่ค่อนข้างนานพอสมควร ควรเป็นข้อมูล รายปี และควรมีข้อมูลอย่างน้อย 15 ปี ซึ่งจะแสดงทิศทางของอนุกรมเวลาตัวอย่างกราฟแสดงค่า แนวโน้มของผลผลิตเคมีภัณฑ์ชนิดหนึ่ง

2) การเปลี่ยนแปลงหรือความแปรผันตามฤดูกาล (Seasonal Variation) แทนด้วย St เป็นการเปลี่ยนแปลงข้อมูลมีลักษณะการเพิ่มขึ้น หรือลดลงในลักษณะเดียวกันของรอบระยะเวลาหนึ่งที่แน่นอน เรียกว่า การเปลี่ยนแปลงตามฤดูกาล หน่วยของระยะเวลาสำหรับ ข้อมูลอาจเป็นรายชั่วโมง รายวัน รายสัปดาห์ รายเดือน รายไตรมาส สำหรับข้อมูลรายปีไม่มีการ แปรผันตามฤดูกาล การเปลี่ยนแปลงตามฤดูกาลนั้นกำหนดระยะเวลาการเกิดซ้ำในรอบหนึ่ง ๆ ได้ ค่อนข้างแน่นอน ตัวอย่างเช่น ยอดขายรายเดือนของห้างสรรพสินค้าแห่งหนึ่ง

3) การเปลี่ยนแปลงหรือความแปรผันตามวัฏจักร (Cyclical Variation) แทนด้วย Ct การเปลี่ยนแปลงตามวัฏจักร มีการเปลี่ยนแปลงเคลื่อนไหวในลักษณะซ้ำ ๆ กันและจะ มีลักษณะคล้ายคลึงกับการเปลี่ยนแปลงตามฤดูกาล จะต่างกันก็ตรงที่การเปลี่ยนแปลงตามวัฏ จักรแต่ละรอบจะใช้ระยะเวลาที่นานกว่า คือ ตั้งแต่ 5 ปีขึ้นไป ข้อมูลที่มีการเปลี่ยนแปลงตามวัฏ จักรในทางธุรกิจ เรียกว่า "วัฏจักรธุรกิจ" (Business Cyclical) โดยทั่วไปประกอบด้วย ระยะ เจริญรุ่งเรือง (prosperity) ระยะฝืดเคือง (recession) ระยะตกต่ำ (depression) และระยะขยายตัว (recovery)

การเปลี่ยนแปลงหรือความแปรผัน เนื่องจากเหตุการณ์ผิดปกติ (Irregularly Variation) แทนด้วย It เป็นการเปลี่ยนแปลงของข้อมูลอนุกรมเวลาที่เกิดจากเหตุการณ์ที่เราไม่ สามารถคาดการณ์ได้ล่วงหน้า เช่น การเกิดไฟไหม้ในโรงงาน การเกิดอุทกภัย การนัดหยุดงานของคนงาน แผ่นดินไหว เป็นต้น ซึ่งเหตุการณ์เหล่านี้เป็นสิ่งที่เกิดขึ้นโดยบังเอิญไม่คาดคิดมาก่อน เป็นการเปลี่ยนแปลงที่เป็นเชิงสุ่ม (random variation) เพราะไม่ได้อยู่ภายใต้เงื่อนไขที่เรากำหนด จากองค์ประกอบของอนุกรมเวลาทั้ง 4 อย่าง คือ T S C และ I ในข้อมูลอนุกรมชุดหนึ่ง ไม่ จำเป็นต้องครบองค์ประกอบข้างต้นก็ได้ ทั้งนี้ ขึ้นอยู่กับชนิดของมูลของเรา

2.2.6.2 รูปแบบของอนุกรมเวลา จากปัจจัยทั้ง 4 ข้างต้น ถ้า Y แทนข้อมูลอนุกรมเวลาชุดหนึ่ง ๆ เราสามารถ กำหนดแบบจำลองได้ 2 แบบ ดังนี้ 1) แบบจำลองผลบวก (Additive model) ถือว่าข้อมูลในแต่ละอนุกรมเวลา ประกอบด้วยผลบวกขององค์ประกอบ

ทั้ง 4 อย่าง Yt = Tt + St + Ct + It 2) แบบจำลองผลคูณ (Multiplicative model) ถือว่าข้อมูลในแต่ละอนุกรม เวลาประกอบด้วยผลคูณขององค์ประกอบทั้ง 4 อย่าง Y t = Tt * St * Ct * It โดยทั่วไปข้อมูลอนุกรมเวลา ในทางธุรกิจจะมีความสัมพันธ์ในรูป แบบจำลองผลคูณ เนื่องจากเป็นการพิจารณาการเปลี่ยนแปลงในรูปอัตราร้อยละ ซึ่งจะทำให้ผล การวิเคราะห์ใกล้เคียงความเป็นจริงมากกว่าการใช้แบบจำลองผลบวก

2.2.7 ทฤษฎีเกี่ยวกับ FP–Growth

หลักการทำงานของขั้นตอนวิธี FP–Growth มีลักษณะการคนหากลุมขอมูลที่ปรากฏบอยแบบการ เติบโตอยางเปนรูปแบบ (Pattern Growth) โดยการทำงานของขั้นตอนวิธี FP–Growth สามารถ อธิบายหลักการทำงานไดดังนี้

1) อานขอมูลจากฐานขอมูลครั้งแรกเพื่อนับคาความถี่ของแตละชิ้นขอมูล แลวนำชิ้นขอมูลที่ไมนอยกวาคาสนับสนุนขั้นต่ำ (L1) มาเรียงลำดับตามคาความถี่ ของแตละชิ้นขอมูลจากมากไปหานอยแลวนำมาสรางตาราง Header (Header Table)

2) อานขอมูลจากฐานขอมูลครั้งที่สองเพื่อสรางตนไม FP Tree โดยอานขอมูลจากฐานขอมูลทีละรายการขอมูล จากนั้นตัดชิ้นขอมูลในรายการขอมูลนั้นที่ ไมปรากฏอยูในตาราง Header ทิ้งไป แลวเรียงชิ้นขอมูลที่เหลือตามลำดับในตาราง Header แลวนำชิ้นขอมูลดังกลาวไปสรางโหนด (Node Tree) เพิ่มเขาไปในตนไม FP–Tree แลวเชื่อมแต ละโหนดที่เปนชิ้นขอมูลเดียวกันเพิ่มเขาไปกับตาราง Header

3) สราง Conditional pattern base และสราง Conditional FP–Tree ของแตละชิ้นขอมูล เพื่อใชในขั้นตอนการคนหากลุมขอมูลที่ปรากฏรวมกันบอย โดย การพิจารณาจะเริ่มจากชิ้นขอมูลลางสุดจนถึงชิ้นขอมูลที่อยูบนสุดในตาราง Header ซึ่ง Conditional pattern base หมายถึงเซตของชิ้นขอมูลที่เกิดขึ้นพรอมกับชิ้นขอมูลที่กำลัง พิจารณาในแตละเสนทาง (Path Tree) และกำหนดใหทุกชิ้นขอมูลมีคาความถี่เทากับคาความถี่ ของชิ้นขอมูลที่กำลังพิจารณาจากตนไม FP–Tree หลังจากนั้นสรางตนไม FP–Tree บน Conditional pattern base นี้เรียกวา Conditional FP–Tree ซึ่งเกิดจากการนำคาความถี่ของ แตละชิ้นขอมูลในทุกเสนทางมารวมกัน และเลือกเฉพาะชิ้นขอมูลที่ผานคาสนับสนุนขั้นต่ำจาก Conditional FP–Tree เพื่อนำไปสรางกลุมขอมูลที่ปรากฏรวมกันบอยตอไป

4) คนหากลุมขอมูลที่ปรากฏรวมกันบอยจากการสราง Conditional pattern base และสราง Conditional FP–Tree ของแตละชิ้นขอมูล โดยใชหลักการ ทำงานแบบแบงแยกแลวเอาชนะ (Divide and Conquer)

2.3. เครื่องมือในการออกแบบและวิเคราะห์ข้อมูล

2.3.1 แบบจำลองอนุกรมเวลา (Time series analysis)

2.3.1.1 ส่วนประกอบของอนุกรมเวลา

1) ค่าแนวโน้ม (Long Term Trend: T) ค่าแนวโน้มเป็นการแสดงถึงการ เคลื่อนไหวหรือเปลี่ยนแปลงของข้อมูลในระยะยาว เช่น ปริมาณการใช้ไฟฟ้าของประเทศไทย , ปริมาณการนำเข้าน้ำมันดิบ เป็นต้น

2) ค่าการผันแปรตามฤดูกาล (Seasonal Variation: S) หมายถึงการ เปลี่ยนแปลงตามฤดูกาล โดยเกิดขึ้นซ้ำ ๆ กันในรอบ 1 ปี จนกลายเป็นแบบแผนเดียวกัน เช่น ผลผลิตข้าวจะสูงในช่วงไตรมาสแรกของปี ยอดขายของห้างสรรพสินค้าจะสูงในช่วงปลายปี เป็นต้น ในการวิเคราะห์การผันแปรตามฤดูกาลนี้จะวัดออกมาในรูปของดัชนีฤดูกาล (Seasonal Index)

3) ค่าการผันแปรตามวัฏจักร (Cyclical Variation: C) หมายถึงการเคลื่อนไหวที่เป็นไปตามวัฏจักร ซึ่งการเคลื่อนไหวตามวัฏจักรนี้จะมีลักษณะคล้ายกับการผันแปรตามฤดูกาล แต่จะมีระยะเวลาที่ยาวนานกว่า

4) การผันแปรเนื่องจากเหตุการณ์ไม่ปกติ (Irregular Variation: I) การผันแปร ชนิดนี้ไม่แน่นอน ไม่สามารถคาดการณ์ได้ล่วงหน้า เช่น ภัยธรรมชาติ สงคราม การนัดหยุดงาน เป็นต้น

ข้อมูลอนุกรมเวลา อาจได้รับอิทธิพลของปัจจัยที่เป็นส่วนประกอบของอนุกรมเวลาทั้ง 4 ปัจจัยหรือเพียงปัจจัยใดปัจจัยหนึ่งเท่านั้น การวิเคราะห์จึงควรแยกวิเคราะห์ทีละปัจจัย ซึ่งในที่นี้จะกล่าวถึงการวิเคราะห์ ปัจจัยค่าแนวโน้ม และค่าผันแปรตามฤดูกาลเท่านั้น

2.3.1.2 วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล (Exponential smoothing)

วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล (Exponential smoothing) เป็นวิธีพยากรณ์ที่เหมาะสมกับการพยากรณ์ในระยะสั้นและปานกลางวิธีการนี้ได้ให้ความสำคัญ

กับข้อมูล ล่าสุด และความสำคัญของข้อมูลที่ห่างออกไปจะลดลง วิธีการ Exponential smoothing ที่นิยมใน ปัจจุบันมีอยู่ 3 วิธีดังนี้

1) Single Exponential Smoothing (SES) เป็นวิธีการทำให้เรียบอย่างง่าย โดยใช้การหาค่าเฉลี่ยแบบถ่วงน้ำหนัก และสมมติให้ค่าน้ำหนักหรือค่าความสำคัญของข้อมูล คือ α (Alpha) วิธีการนี้มีเงื่อนไขว่า ข้อมูล อนุกรมเวลาที่ใช้ในการวิเคราะห์จะต้องไม่มีแนวโน้ม (Trend) และอิทธิพลของฤดูกาล (Seasonality) หมายความว่าข้อมูลอนุกรมเวลาที่นำมาใช้การวิเคราะห์ต้องมีลักษณะคงที่ สำหรับสมการที่ใช้ใน การพยากรณ์

2) Holt's Two-Parameter Method

วิธี Holt's Two-Parameter Method เป็นวิธีที่ Holt (1957) ได้ปรับปรุงวิธี Single Exponential Smoothing (SES) -7hoใหม่เพื่อให้สามารถใช้ข้อมูลอนุกรมเวลาที่มีแนวโน้ม (trend) ของเวลา เรียกวิธีการนี้ว่า "Holt's Two-Parameter Method" วิธีการนี้ให้ความสำคัญกับ ข้อมูลล่าสุดและแนวโน้มเวลา ดังนั้นจึงมีค่าคงที่ในการทำให้เรียบ 2 ค่า คือ α (Alpha) และ β (Beta)

3) Holt-Winters-Trend and Seasonal (Three-Parameter)

วิธีนี้ถูกพัฒนา เพิ่ม ขึ้นจากวิธีการของ Holtโดย Winters (1960) ได้พัฒนาให้วิธีการนี้สามารถวิเคราะห์ข้อมูล อนุกรมเวลาที่มีทั้งแนวโน้ม (Trend) และฤดูกาล (Seasonality) วิธีการนี้ให้ความสำคัญกับข้อมูล ล่าสุดแนวโน้มเวลาและฤดูกาล ดังนั้นจึงมีค่าคงที่ในการทำให้เรียบ 3 ค่าคือ (Alpha), (Beta) และ (Gamma)

2.3.1.3 การประเมินค่าการพยากรณ์

การศึกษาครั้งนี้ใช้ค่าสถิติ RMSE (Root Mean Squared Error) และ MAPE (Mean Absolute Percentage Error) ในการประเมินค่าการพยากรณ์เพื่อทดสอบความแม่นยำของผลการพยากรณ์

2.3.2 กระบวนการวิเคราะห์ข้อมูลด้วย (CRISP-DM)

งานวิทยาการด้านข้อมูล (Data Science) ซึ่งกำลังมีบทบาทอย่างมากในโลกยุคปัจจุบัน และทวีความสำคัญยิ่งขึ้นในอนาคต ก็มี CRISP-DM เป็นกระบวนการหลักในการจัดทำเหมือง ข้อมูลเพื่อการวิเคราะห์และใช้ประโยชน์ในทางธุรกิจ กระบวนการวิเคราะห์ข้อมูล ด้วย CRISP-DM หรือ Cross Industry Standard Process for Data Mining พัฒนาขึ้นในปี ค.ศ. 1996

โดยความ ร่วมมือของ 3 บริษัทคือ Daimler Chrysler, SPSS และ NCR ประกอบด้วย 6 ขั้นตอนหลัก ได้แก่

1) ทำความเข้าใจระบุโอกาสหรือหาปัญหาที่จะเกิดขึ้นกับธุรกิจ (Business Understanding)กำหนดขอบเขตของข้อมูลที่จะนำวิเคราะห์เพื่อหาความได้เปรียบทางการตลาดเพื่อนำมาแก้ไขปัญหาองค์กร และต้องสามารถระบุผลลัพธ์ที่มีได้

2) ทำความเข้าใจข้อมูล (Data Understander) โดยการรวบรวมข้อมูลที่เกี่ยวข้องให้อยู่ รวมกลุ่มกัน คัดเลือกให้เหลือเพียงข้อมูลที่มีความสำคัญต่อการทำงาน

3) เตรียมข้อมูล (Data Preparation) ขั้นตอนนี้จะเป็นขั้นตอนที่ใช้เวลามากที่สุดใน 6 ขั้นตอน เพราะคุณภาพของงานที่ได้นั้นจะดีเพียงใด ต้องขึ้นอยู่กับคุณภาพข้อมูลที่ได้จัดเตรียมในขั้นตอนนี้ การเตรียมข้อมูลประกอบด้วย การคัดเลือกข้อมูล การกลั่นกรองข้อมูล และแปลงรูปแบบของข้อมูล

4) สร้างแบบจำลอง (Modeling) นำผลที่ได้มาทดลองทำแบบจำลอง เพื่อใช้ถ่ายทอด ข้อมูลหรือทำการเปลี่ยนแปลงข้อมูลให้อยู่ในรูปสารสนเทศเป็นข้อมูลที่นำไปใช้ประโยชน์ได้จริง ในทางธุรกิจ

5) การประเมินผล (Evaluation) เพื่อเป็นการทดสอบผลที่ได้โดยการวัดประสิทธิภาพ จากผลลัพธ์ที่ได้มาจัดทำงานจริงและติดตามตรวจสอบผลที่ได้ใหม่ ซึ่งสามารถตรวจสอบได้หลาย ทางเช่น วัดจากส่วนแบ่งของตลาด วัดจากปริมาณลูกค่า หรือวัดจากกำไรสุทธิ เป็นต้น

6) การปฏิบัติตามผลเสนอแนะ (Deployment) คือ การนำเอาข้อมูลที่เป็นผลลัพธ์มา ลองปฏิบัติจริงกับธุรกิจ โดยแปลงแนวคิดที่มีให้เกิดเป็นการนำไปใช้ประโยชน์ และติดตามรวบรวม ผลที่ได้เพื่อการปรับปรุงต่อไป จากขั้นตอนที่กล่าวมาคือการทำเหมืองข้อมูลในงานระบบทางธุรกิจเป็นกระบวนการ ทางสถิติที่เน้นการทำกับข้อมูลที่มีจำนวนมากในหลากหลายรูปแบบ คัดเลือกเฉพาะข้อมูลที่สำคัญ และจำเป็นต้องนำมาใช้งาน จัดทำการกำหนดรูปแบบจัดแบ่งกลุ่มลำดับความสำคัญ จากนั้นจึงจะ เริ่มค้นหารูปแบบแนวทางและความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลนั้น โดยแต่ละขั้นตอนจะอาศัย ผลลัพธ์จากอีกขั้นตอนหนึ่งกลายเป็นข้อมูลให้ขั้นตอนต่อไป การทำเหมืองข้อมูลจะช่วยเปลี่ยน ข้อมูลดิบให้เป็นสารสนเทศที่มีประโยชน์ การระบุแหล่งข้อมูลที่ถูกต้องจึงเป็นสิ่งที่สำคัญต่อผลลัพธ์ ที่ได้จากการวิเคราะห์

2.3.3 โปรแกรม Google Data Studio

Google Data Studio คือ โปรแกรมที่ใช้ในการแสดงผลข้อมูลในรูปแบบของภาพ (Data Visualization) เพื่อทำให้มีความเข้าใจในข้อมูลมากขึ้น ทำให้สามารถนำข้อมูลจำนวนมากมาวิเคราะห์และมองออกได้อย่างง่ายขึ้น ซึ่งตัวโปรแกรม Google Data Studio จะสามารถใช้งานได้ผ่านเว็บไซต์ Google Data Studio Online เป็นต้น

1. Google Data Studio ในมุมมองด้าน Visualization ในความเห็นของผู้เขียน ถ้าจะให้คำนิยามว่า Google Data Studio คือซอฟต์แวร์ประเภท ไหนที่ตรงประเด็นที่สุดแล้ว ผู้เขียนเห็นว่า Data Visualization น่าจะเป็นคำนิยามที่ตรงประเด็นมาก ที่สุด กล่าวคือ Google Data Studio ทำหน้าที่แปลงข้อมูล (Data) ให้เป็นภาพ (Visualization) ตัวอย่างของการ แปลงข้อมูลเป็นภาพเป็นสิ่งที่เราคุ้นเคย และมีมาเนิ่นนานแล้วก็คือการสร้างกราฟนั่นเอง จะเห็นได้ ว่าเรามีการสร้าง ตาราง กราฟแท่ง กราฟเส้น กราฟวงกลม (pie) มาก่อนแล้ว กราฟที่กล่าวมาถือ เป็น Visualization อย่างหนึ่ง แต่ก็ยังมีกราฟรูปแบบอื่น ๆ อีกมากมายที่มีประสิทธิภาพและทรง พลัง ในการสื่อสารมากกว่า

2. ภาพรวมในการทำงานของ Google Data Studio

- เตรียมข้อมูลจากแหล่งต่าง ๆ (Data) เนื่องจาก Google Data Studio เป็นเครื่องมือที่ช่วยรายงานผล ดังนั้นการที่เราจะนำข้อมูลมากมายมาวิเคราะห์ ตีความ ได้นั้นต้องมีการเก็บข้อมูลที่ดีและแม่นยำก่อน เพราะการแสดงผลจะอ้างอิงจากข้อมูลที่มีอยู่
- เชื่อมโยงข้อมูล (Connect) คือการดึงข้อมูลจากแหล่งข้อมูล (Data Source) ที่เราต้องการนำมาใช้ในการแสดงผล รวมถึงสามารถเลือกเขตข้อมูล (Field) ให้แสดงผลตามที่ต้องการได้อีกด้วย
- เลือกรูปแบบ (Templates) Google Data Studio มี Templates ที่ใช้ในการนำเสนอเป็นภาพที่เข้าใจง่าย ให้เลือกหลากหลายรูปแบบและสามารถตกแต่งครีเอทได้อย่างสวยงาม
- การส่งต่อข้อมูล (Share) สามารถส่งต่อข้อมูลนำเสนอให้เพื่อนร่วมงานหรือลูกค้าได้อย่างรวดเร็ว โดยไม่มีค่าใช้จ่าย ทำข้อมูลให้ออกมาเป็นภาพที่เข้าใจง่าย สามารถวิเคราะห์ข้อมูลเชิงสถิติได้อย่างง่ายดายไม่ซับซ้อน

2.3.4 โปรแกรม Minitab

Minitab เป็นโปรแกรมสำเร็จรูปใช้ประมวลผลข้อมูลทางด้านสถิติ โดยพัฒนาจากกลุ่มนักวิชาการทางด้านสถิติมากว่า 30 ปีแล้ว โดยปัจจุบันได้พัฒนาปรับปรุงมาจนถึงเวอร์ชัน 16 ซึ่งสามารถใช้กับระบบปฏิบัติการวินโดวส์ โดย Minitab เข้ามามีบทบาทสำหรับผู้ใช้สถิติในส่วนของการประมวลผลและการแสดงผลข้อมูลในลักษณะของตัวเลขและผลในลักษณะของกราฟ ประกอบกับเทคโนโลยีทางด้านคอมพิวเตอร์ได้พัฒนาและมีบทบาทต่อชีวิตประจำวัน

Minitab จึงถูกเลือกใช้ด้วยเหตุผลหลัก 3 ประการ คือ

1. ความซับซ้อนในการประมวลผล (Complexity)
2. ความเที่ยงตรงและแม่นยำในการประมวลผลข้อมูล (Accuracy)
3. ความรวดเร็วและความสามารถในการทำซ้ำ (Repeatability)

Minitab เป็นโปรแกรมที่มีความโดดเด่นในด้านการใช้งานที่ง่ายและมีการพัฒนาปรับปรุงฟังก์ชันต่าง ๆ ให้สอดคล้องกับความรู้และทฤษฎีใหม่ ๆ รวมถึงการประยุกต์ทางด้านสถิติโดยเฉพาะในงานด้านคุณภาพอย่างต่อเนื่อง จึงเป็นที่รู้จักกันเป็นอย่างดีสำหรับกลุ่มผู้ที่พัฒนาปรับปรุงคุณภาพด้วยหลักการ “ซิกซ์ ซิกม่า” เนื่องจาก Minitab เป็นเครื่องมือที่สำคัญสำหรับกลุ่มผู้ใช้กลุ่มนี้ แต่ไม่ได้หมายความว่าจะจำเพาะกลุ่มผู้ใช้กลุ่มนี้เท่านั้น Minitab ยังเป็นโปรแกรมที่นิยมใช้กันอย่างแพร่หลายในกลุ่มนักวิชาการ นักวิจัย นักวิทยาศาสตร์ และผู้ใช้สถิติทั่วไปถึงแม้ว่าโปรแกรม Minitab จะเป็นโปรแกรมที่ช่วยให้การทำงานของเราสะดวกขึ้นมาก แต่ความรู้ความเข้าใจในการทำงานของโปรแกรมก็ยังเป็นสิ่งสำคัญสำหรับผู้ใช้เสมอ

2.3.5 โปรแกรม RapidMiner Studio

RapidMiner Studio เป็นโปรแกรมที่ออกแบบมาสำหรับการวิเคราะห์ข้อมูล การทำเหมืองข้อมูล (Data mining) เป็นกระบวนการ (Process) ที่กระทำกับข้อมูลขนาดใหญ่เพื่อค้นหารูปแบบ แนวทาง และความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลนั้นโดยอาศัยหลักสถิติ การรู้จำ การเรียนรู้ของเครื่อง และหลักคณิตศาสตร์เพื่อให้ได้สารสนเทศที่เราไม่รู้ออกมา โดยสารสนเทศที่ได้จะมีเหตุผลและสามารถนำไปใช้ประโยชน์ได้ โปรแกรม RapidMiner สามารถช่วยให้คุณเรียกดูข้อมูลและสร้างแบบจำลองเพื่อระบุแนวโน้มได้อย่างง่ายดาย เมื่อคุณจัดการกับฐานข้อมูลขนาดใหญ่การระบุการเชื่อมต่อระหว่างสองเหตุการณ์อาจเป็นเรื่องยากหรือเป็นไปไม่ได้ เนื่องจากมีหลายธุรกิจที่ต้องพึ่งพาข้อมูลที่มีอยู่เพื่อทำการตัดสินใจที่สำคัญ นักวิเคราะห์ข้อมูลจึงใช้แอปพลิเคชันพิเศษเพื่อให้เห็นภาพและเข้าใจข้อมูล โดยโปรแกรม RapidMiner Studio มีวัตถุประสงค์เพื่อมอบ

เครื่องมือที่ใช้งานง่ายและมีประสิทธิภาพสำหรับการวิเคราะห์ข้อมูลจากแหล่งต่าง ๆ ช่วยให้คุณสามารถโหลดข้อมูลที่ต้องการจากไฟล์ข้อความธรรมดา, เอกสาร Office หรือแม้แต่เซิร์ฟเวอร์ฐานข้อมูลเช่น Oracle, MySQL หรือ PostgreSQL

2.4 วรรณกรรมที่เกี่ยวข้อง

2.4.1 การศึกษาพฤติกรรมในการใช้อินเทอร์เน็ตและศึกษาปัจจัยที่มีผลต่อการใช้อินเทอร์เน็ตในสถานการณ์โรคไวรัสโคโรนา มหาวิทยาลัยธรรมศาสตร์

สิริมา สุโพธ (2562) ได้ศึกษาพฤติกรรมในการใช้อินเทอร์เน็ตและศึกษาปัจจัยที่มีผลต่อการใช้อินเทอร์เน็ตในสถานการณ์โรคไวรัสโคโรนา 2019 กลุ่มตัวอย่างในการศึกษาคือ นักศึกษาคณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์ จำนวน 400 วิจัยนำข้อมูลที่เก็บรวบรวมได้จากกลุ่มตัวอย่าง จำนวน 400 คน คิดเป็น 100 % มาประมวลผลตามระเบียบวิธีวิจัยทางสถิติโดยใช้โปรแกรมสำเร็จรูปเพื่อวิเคราะห์ข้อมูล หาค่าร้อยละ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน ในส่วนของสมมติฐานเพื่อหาความแตกต่าง และหาความสัมพันธ์ของตัวแปรอิสระและตัวแปรตาม โดยใช้การวิเคราะห์ความแปรปรวนทางเดียว, ทดสอบความแตกต่างของค่ากลางของสองประชากรไม่อิสระ และหาความสัมพันธ์โดยใช้สถิติวิเคราะห์ความถดถอยเชิง พหุคูณโดยกำหนดระดับนัยสำคัญ

2.4.2 การวิเคราะห์ข้อมูลปริมาณน้ำฝนรายเดือนของจังหวัดตาก โดยใช้เทคนิคการวิเคราะห์อนุกรมเวลาแบบคลาสสิค

รุ่งใจ แสงทอง (2547) ใช้เทคนิคการวิเคราะห์อนุกรมเวลาแบบคลาสสิค โดยนำเทคนิคการทำให้เรียบแบบการเฉลี่ยเคลื่อนที่และเทคนิคการทำให้เรียบแบบเอ็กซ์โพแนนเชียล โดยนำข้อมูลที่มาศึกษาคือ ข้อมูลปริมาณน้ำฝนรายเดือนของจังหวัดตาก ตั้งแต่เดือน มกราคม 2514 ถึงเดือน 2544 โดยการวิเคราะห์ข้อมูลแบ่งเป็น 2 รูปแบบ คือ รูปแบบคงที่ และ รูปแบบเชิงเส้น

2.4.3 การศึกษาพฤติกรรมและปัญหาการใช้อินเทอร์เน็ตของนักศึกษาระดับปริญญาตรี มหาวิทยาลัยราชภัฏบ้านสมเด็จเจ้าพระยา

ชเนตตี สยนานนท์ (2555) การวิจัยนี้มีความมุ่งหมายเพื่อ ศึกษาพฤติกรรมและปัญหาการใช้อินเทอร์เน็ตของนักศึกษาระดับปริญญาตรีและเปรียบเทียบพฤติกรรมการใช้อินเทอร์เน็ตของนักศึกระดับปริญญาตรี มหาวิทยาลัยราชภัฏบ้านสมเด็จเจ้าพระยา โดยใช้วิธีการจำแนกตามตัวแปรเพศ ระดับชั้นปี คณะวิชา และผลสัมฤทธิ์ทางการเรียน ใช้กลุ่มตัวอย่างประกอบด้วย

นักศึกษาที่กำลังศึกษาอยู่ในภาคเรียนที่ 2 ปีการศึกษา 2554 จำนวน 400 คน เครื่องมือที่ใช้ในการเก็บรวบรวมได้แก่ แบบสอบพฤติกรรมและปัญหาการใช้อินเทอร์เน็ตของนักศึกษา จำนวน 30 ข้อ แบบสอบถามมีลักษณะเป็นแบบมาตราส่วนประมาณค่า 5 ระดับ ตามแบบไลเคิร์ท มีความเชื่อมั่นเท่ากับ 0.90 สถิติที่ใช้ในการวิเคราะห์ข้อมูล คือ ค่าความถี่ ค่าร้อยละ ค่าเฉลี่ยและความเบี่ยงเบนมาตรฐาน การทดสอบที และการวิเคราะห์ความแปรปรวน

2.4.4 การศึกษาการหารูปแบบการพยากรณเพื่อนำไปใช้ในการวางแผนการใช้พลังงานไฟฟ้าในอนาคตของแต่ละโรงงานอุตสาหกรรม

อัครัช บรรจงศิลป์ (2550) ศึกษาการหารูปแบบการพยากรณ์ (Forecasting Model) ที่เหมาะสมกับโรงงานอุตสาหกรรมแต่ละประเภท เพื่อนำไปใช้ในการวางแผนการใช้พลังงานไฟฟ้าในอนาคตของแต่ละโรงงานอุตสาหกรรม โดยใช้ข้อมูลการใช้พลังงานไฟฟ้าของผู้ใช้รายใหญ่ในระบบ 115KV. ในเขตพื้นที่รับผิดชอบของการไฟฟ้าส่วนภูมิภาคเขต 3 ซึ่งข้อมูลการใช้พลังงานไฟฟ้าสามารถแบ่งออกได้เป็น 2 ประเภท คือ ข้อมูลค่าโหลดการใช้พลังงานไฟฟ้าสูงสุด (KW.) และข้อมูลค่าหน่วยการใช้พลังงานไฟฟ้า (KWH.) ซึ่งจากการศึกษาพบว่าการใช้ค่าเฉลี่ยเปอร์เซ็นต์ของความคลาดเคลื่อนสมบูรณ์ (Mean Absolute Percent Error , MAPE) และการให้ระดับความสำคัญของข้อมูลมาใช้พิจารณาเลือกรูปแบบการพยากรณ์ สามารถหารูปแบบการพยากรณ์(Forecasting Model) ที่เหมาะสมกับโรงงานอุตสาหกรรมแต่ละประเภทได้

2.5 บทสรุป

จากแนวคิด ทฤษฎี เครื่องมือและวรรณกรรมที่เกี่ยวข้องที่ได้กล่าวมาในข้างต้น ทั้งหมดนั้น คณะผู้จัดทำได้เลือกใช้ขั้นตอนกระบวนการวิเคราะห์ข้อมูลด้วย กระบวนการวิเคราะห์ CRISP-DM และใช้เทคนิคการทำเหมืองข้อมูลแบบ Association analysis ด้วยการสร้างโมเดล ทำการทดสอบโมเดล เพื่อดูแนวโน้มของโมเดลที่สร้างขึ้น ด้วยโปรแกรม KNIME จากนั้นนำข้อมูลสารสนเทศมาทำการ แสดงผลแบบ Visualization ในรูปแบบของภาพโดยใช้โปรแกรม Power BI เผยแพร่บนเว็บไซต์ที่เป็นที่นิยมในยุคอินเตอร์เน็ตคือการเผยแพร่ทางสื่อออนไลน์ โดยใช้ภาษา HTML และ CSS ในการเขียนเว็บไซต์ขึ้นมา