บทที่ 2
แนวคิดทฤษฎีเอกสารที่เกี่ยวข้อง

ในบทนี้เป็นการนำเสนอเกี่ยวกับ แนวคิด ทฤษฎี เครื่องมือและวรรณกรรมที่เกี่ยวข้องของการพัฒนาเว็บไซต์สำหรับวิเคราะห์ข้อมูลปริมาณการใช้งานเน็ตประชารัฐปี ซึ่งได้รวบรวมการศึกษาเอกสารงานวิจัยที่เกี่ยวข้องกับระบบ เพื่อใช้เป็นแนวทางการศึกษา ประกอบด้วยรายละเอียดตามลำดับ ดังนี้

2.1 แนวคิด
- 2.1.1 แนวคิดเกี่ยวกับการวิเคราะห์ข้อมูล (Data Analytic)
- 2.1.2 แนวคิดเกี่ยบกับการพยากรณ์ข้อมูล (Forecasting Data)
- 2.1.3 แนวคิดเกี่ยวกับการแสดงผลข้อมูล (Data Visualization)

2.2 ทฤษฎี
- 2.2.1 ทฤษฎีเกี่ยวกับระบบจัดการฐานข้อมูล
- 2.2.2 ทฤษฎีเกี่ยวกับการออกแบบเว็บไซต์
- 2.2.3 ทฤษฎีเกี่ยวกับการทำนอร์มัลไลเซชัน (Normalization)

2.3 เครื่องมือในการออกแบบ และวิเคราะห์ข้อมูล
- 2.3.1 แบบจำลองการพยากรณ์ (Forecasting)
- 2.3.2 การแสดงผลแบบ (Visualization)
- 2.3.3 กระบวนการวิเคราะห์ข้อมูลด้วย (CRISP–DM)
- 2.3.4 แผนภาพกระแสข้อมูล (Data Flow Diagram : DFD)

2.4 วรรณกรรมที่เกี่ยวข้อง

2.5 บทสรุป

2.1 แนวคิด

2.1.1 แนวคิดเกี่ยวกับการวิเคราะห์ข้อมูล (Data Analytic)

การวิเคราะห์ข้อมูลเป็นศาสตร์ในการวิเคราะห์และประมวลผลข้อมูลที่มีอยู่ในปัจจุบัน อดีต และทำนายอนาคต เพื่อนำผลนั้นมาปรับใช้ให้เกิดประโยชน์ ความรู้เรื่องดังกล่าวสำคัญยิ่งขึ้น เมื่อเทคโนโลยีในปัจจุบันมีการเก็บข้อมูลมหาศาลในระดับ Big Data มากเกินกว่าคนทั่วไปจะสามารถใช้งานได้เต็มที่ Data Analytics จึงเหมาะที่จะเป็นเครื่องมือสำหรับธุรกิจ (Business Intelligence) และ Data Analytics นี้ไม่จำเป็นต้องเป็นธุรกิจขนาดใหญ่ ธุรกิจขนาดเล็กก็สามารถทำได้เหมือนกัน เราสามารถแบ่งเป็น 3 แบบคือ

1) Descriptive Analytics เป็นรูปแบบการใช้ข้อมูลแบบพื้นฐานที่สุด โดยเน้นการอธิบายว่ากำลังเกิดขึ้นหรืออาจจะเกิดอะไรขึ้น สามารถอธิบายถึงสาเหตุการเกิดต่าง ๆ ได้ว่าทำไม ซึ่ง Descriptive Analytics ตัวอย่างคือรายงานธุรกิจ รายงานด้านการทำ Campaign หรือโฆษณา หรือรายงานผลดำเนินงานที่ผ่านมาเป็นข้อมูลพื้นฐานที่แสดงผลในการทำกิจกรรมต่าง ๆ ให้เราได้รับทราบ

2) Predictive Analytics เป็นรูปแบบการใช้ข้อมูลที่มีความซับซ้อนขึ้นมา โดยจะเป็นการ “พยากรณ์” หรือ “ทำนาย” สิ่งที่กำลังเกิดขึ้น โดยใช้ข้อมูลในอดีต ร่วมกับโมเดลทางคณิตศาสตร์ต่าง ๆ หรือร่วมกับการทำ Data Mining นอกจากนี้ Predictive Analytics ยังทำให้เราสามารถวิเคราะห์หาโอกาสและความเสี่ยงต่าง ๆ ที่จะเกิดขึ้นในอนาคตได้ด้วย เช่นการรู้เทรนด์ทางการตลาด การพยากรณ์ยอดการขายหรือการทำ Campaign ว่าจะมีคนร่วมเท่าไหร่

3) Prescriptive Analytics เป็นรูปแบบการวิเคราะห์ข้อมูลที่มีความซับซ้อนและยากที่สุด เพราะไม่เพียงพยากรณ์หรือทำนายว่าอะไรจะเกิดขึ้น แต่ยังให้คำแนะนำในทางเลือกต่าง ๆ และผลแต่ละทางเลือกว่าจะมี Pros & Cons อย่างไร โมเดลของ Prescriptive Analytics นั้นจะสามารถปรับเปลี่ยนได้ตามข้อมูลที่เพิ่มเติมเข้ามามากขึ้นและ Prescriptive Analytics นี้ยังเป็นการใช้ข้อมูลที่มากที่สุดและเกี่ยวพันกับเรื่อง Big Data เป็นอย่างมาก

2.1.2 แนวคิดเกี่ยบกับการพยากรณ์ข้อมูล (Forecasting Data)

การพยากรณ์ หมายถึงการคาดการณ์ (Predict) เกี่ยวกับลักษณะหรือแนวโน้มของสิ่งที่สนใจที่จะเกิดขึ้นในอนาคตเพื่อใช้เป็นสารสนเทศ (Information) ประกอบการตัดสินใจซึ่งการพยากรณ์จะต้องดำเนินการเป็นส่วนแรกสุดก่อนการวางแผนหรือการเตรียมการที่จะเริ่มกระทำการใด ๆ เพื่อความถูกต้องและแม่นยำในการตัดสินใจ ดังนั้นในการดำเนินธุรกิจภายใต้ความไม่

แน่นอนและการเปลี่ยนแปลงที่รวดเร็วเช่นปัจจุบันจึงจำเป็นที่จะต้องทราบถึงความเป็นไปในอนาคต โดยอาศัยเทคนิคการพยากรณ์ต่าง ๆ เพื่อเพิ่มประสิทธิภาพในการดำเนินงานทางธุรกิจ สามารถตอบสนองต่อความต้องการของลูกค้า เพิ่มผลตอบแทนทางธุรกิจ ลดต้นทุนและความสูญเสียต่าง ๆ ที่อาจเกิดขึ้นเป็นต้น การพยากรณ์จึงเป็นทั้งศาสตร์และศิลป์ในการทำนายเหตุการณ์ในอนาคต ซึ่งอาจนำหลายๆวิธีมาใช้แล้วแต่สถาณการณ์ เช่น นำข้อมูลในอดีตมาพยากรณ์อนาคตโดยอาศัยหลักการทางคณิตศาสตร์เข้าช่วย หรือใช้ดุลยพินิจของผู้พยากรณ์เพียงอย่างเดียว หรืออาจใช้หลาย ๆ วิธีร่วมกันเพื่อให้การพยากรณ์มีความแม่นยำมากขึ้น

2.1.2.1 ช่วงเวลาของการพยากรณ์

แบ่งตามระยะเวลาของการพยากรณ์ได้ 3 ประเภทคือ

1) การพยากรณ์ในระยะสั้น (Short-Range Forecasting) เป็นการพยากรณ์เหตุการณ์ที่ไม่เกิน 1 ปี โดยทั่วไปมักจะอยู่ในช่วงไม่เกิน 3 เดือน เช่นการพยากรณ์การวางแผนจัดซื้อ การจัดตารางการทำงาน การมอบหมายงาน การพยากรณ์ยอดขาย และการพยากรณ์ระดับการผลิต

2) การพยากรณ์ระยะกลาง (Medium-Range Forecasting) เป็นการพยากรณ์เหตุการณ์ที่อยู่ในช่วง 3 เดือน ถึง 3 ปี จะใช้มากในการพยากรณ์การวางแผนการขาย การวางแผนการผลิต การวางแผนด้านงบประมาณเงินสด และการวิเคราะห์การวางแผนการดำเนินงานต่าง ๆ

3) การพยากรณ์ระยะยาว (Long-Range Forecasting) เป็นการพยากรณ์เหตุการณ์ที่มากกว่า 3 ปีขึ้นไป มักใช้สำหรับการวางแผนผลิตภัณฑ์ใหม่ ค่าใช้จ่ายในการลงทุน การขยายทำเลที่ตั้ง และการวิจัยพัฒนา

2.1.2.2 เทคนิคการพยากรณ์

เทคนิคการพยากรณ์แบบมีหลักเกณฑ์ (Formal Forecasting Techniques) การพยากรณ์วิธีนี้ต้องอาศัยข้อมูลมาสนับสนุนและใช้ความรู้ทางสถิติ คณิตศาสตร์เข้ามาช่วยในการวิเคราะห์ข้อมูล ดังนั้น ผู้จะพยากรณ์เป็นใครก็ได้ที่เข้าใจวิธีการ ขั้นตอนในการนำข้อมูลมาและวิเคราะห์ เทคนิคการพยากรณ์แบบมีหลักเกณฑ์นี้ยังแบ่งได้อีก 2 พวกคือ

1) การพยากรณ์เชิงคุณภาพ (Qualitative Forecasting) การพยากรณ์แบบนี้ไม่เน้นวิธีการทางสถิติมากนัก แต่เน้นการสอบถามความเห็น ความรู้สึก ความคิดเห็นจากนักบริหาร จากฝ่ายขาย จากฝ่ายซื้อ แล้วนำข้อมูลเหล่านั้นรวบรวมมาสรุปเป็นคำพยากรณ์ณ์สำหรับวิธีการให้ได้มา ซึ่งคำพยากรณ์เชิงคุณภาพนี้มีวิธีการปฏิบัติหลายอย่าง คือ

1.1 วิธีเดลฟาย (Delphi Method) เป็นวิธีที่อาศัยผู้เชียวชาญใช้แสดงความคิดเห็นโดยการสอบถามนั้นจะทำอยางต่อเนื่องอยางน้อย 2-3 ครั้ง เพื่อให้ได้ข้อสรุปที่ดีที่สุดแล้วนำข้อมูลที่สอดคล้องและมีความถี่มากมาเป็นค่าพยากรณ์

1.2 วิธีวิจัยตลาด (Market Research) เป็นการสำรวจข้อมูลโดยตรงจากผู้ใช้หรือกลุ่มเป้าหมาย ดูว่าแนวโน้มหรือสภาพการใช้สินค้าต่าง ๆ เป็นอย่างไร ทำได้โดยการแจกแบบสอบถาม/สัมภาษณ์ แล้วนำข้อมูลมาหาความถี่แล้วสรุปมาเป็นค่าพยากรณ์

1.3 วิธีการอภิปราย (Panel Discussion) โดยน าผู้รู้ผู้เชี่ยวชาญมาสัก 4-5 คน อภิปรายร่วมกน และหาข้อสรุปเป็นค่าพยากรณ์

1.4 วิธีถามบุคคลใกล้ชิด (Grass-Roots Forecasting) วิธีการแบบนี้คือสอบถามจากผู้ใกล้ชิดกบปัญหานั้น ๆ แล้วนำมาเป็นข้อสรุป

1.5 วิธีการยึดอดีตเป็นหลัก (Historical Analogy) เป็นแนวการพยากรณ์ที่ยึดข้อมูลของเหตุการณ์หนึ่งมาพยากรณ์อีกเหตุการณ์หนึ่ง หมายถึงเหตุการณ์ที่เกิดขึ้นในอดีตและเมื่อเหตุการณ์อยางเดียวกันเกิดขึ้นอีกก็จะมีผลคล้ายกับที่เกิดขึ้นมาแล้วเช่นกันได้ เป็นต้น

2) การพยากรณ์เชิงปริมาณ (Quantitative Forecasting) เทคนิคการพยากรณ์แบบนี้นั้นอาศัยตัวเลขเป็นข้อมูลนำมาคำนวณหาตัวเลข หรือแนวโน้มที่จะเกิดขึ้นในอนาคต ซึ่งจะแบ่งจำแนกออกได้เป็น 2 ประเภท คือ

2.1 การพยากรณ์แบบอนุกรมเวลา (Time Series Forecasting) เป็นวิธีการคาดคะเนเหตุการณ์ โดยใช้ข้อมูลต่าง ๆ ที่มีอยู่ในอดีตมาพยากรณ์ข้อมูลในอนาคต

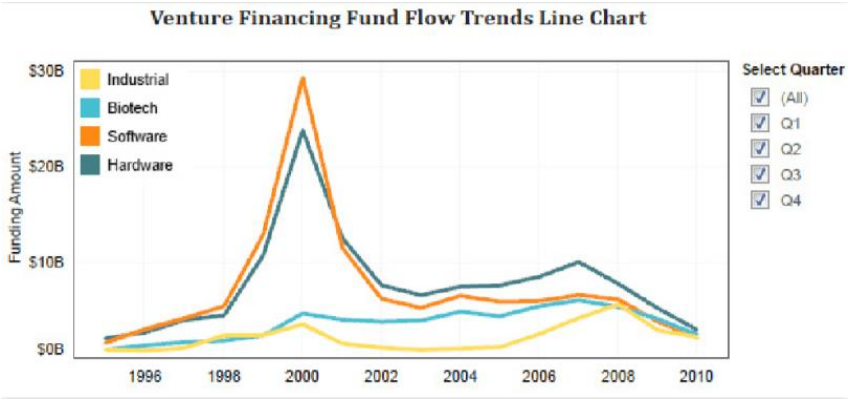
2.2 การพยากรณ์เชิงสหสัมพันธ์ (Correlation Forecasting) และการวิเคราะห์การถดถอย (Regression analysis) เป็นวิธีการพยากรณ์อนาคตโดยการถูกกำหนดโดยข้อมูลอื่นที่มีความสัมพันธ์กนกับสิ่งที่เป็นตัวพยากรณ์

2.1.3 แนวคิดเกี่ยวกับการแสดงผลข้อมูล (Data Visualization)

Data Visualization เป็นการนำ Data เชิงลึกที่ได้มาจากช่องทางต่าง ๆ มาวิเคราะห์ข้อมูล, จัดการข้อมูล และนำข้อมูลมาแสดงผลในรูปแบบของแผนภูมิ กราฟในรูปแบบที่หลากหลาย แสดงให้เห็นถึงการเปลี่ยนแปลงของข้อมูลเชิงปริมาณ ซึ่งจะทำให้ผู้อ่านเข้าใจได้ง่ายขึ้น ดูน่าสนใจมากกว่าการอ่านข้อมูลแบบตารางทั่วไป ซึ่งข้อมูลที่นำมาประมวลผลนั้น สามารถแปลงเป็นรูปภาพกราฟต่าง ๆ หรือแม้กระทั่งวิดีโอ เพื่อนำเสนอผลงานในรูปแบบต่าง ๆ ทำให้เกิดการจดจำได้ง่าย

2.1.3.2 ประเภทของ Data Visualization

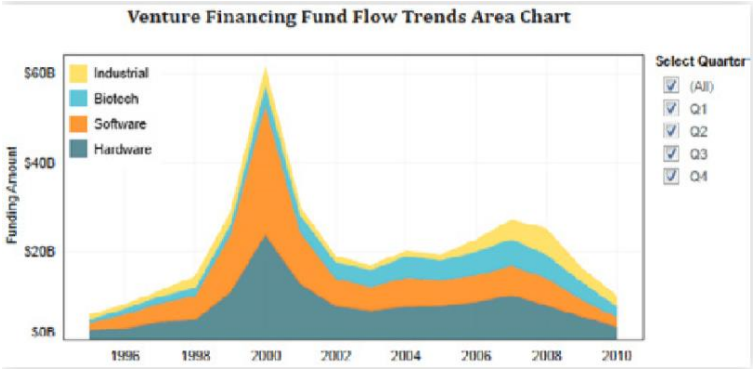
1. Trends over Time การวิเคราะห์แนวโน้มตามช่วงเวลา : Trends over Time มักจะถูกจำมาใช้เพื่อดู Trend หรือแนวโน้ม ตามช่วงเวลา โดยในตัวอย่างที่เขายกมา จะเป็นเรื่องของเงินลงทุนรายปีในแต่ละ Sector ซึ่ง Chart ที่เหมาะสมกับการดูแนวโน้มที่สุดก็มักจะเป็น Line Chart โดยให้แกนตั้ง หรือ แกน Y เป็น Sector และแกนนอน หรือ แกน X เป็นช่วงเวลา และอีกสิ่งที่ช่วยให้เห็นภาพของแต่ละ Sector ได้ชัดเจนขึ้นก็คือ สีของแต่ละเส้น



ภาพที่ 2.1 Trends over Time การวิเคราะห์แนวโน้มตามช่วงเวลา

ที่มา : https://www.d4biz.com/data-and-analytics/select-chart-for-data-visualization-and-analysis/

หากเราต้องการดูภาพรวมหรือสัดส่วนให้ชัดเจนขึ้น การใช้ Chart แบบอื่นอาจจะเหมาะสมกว่า เช่น

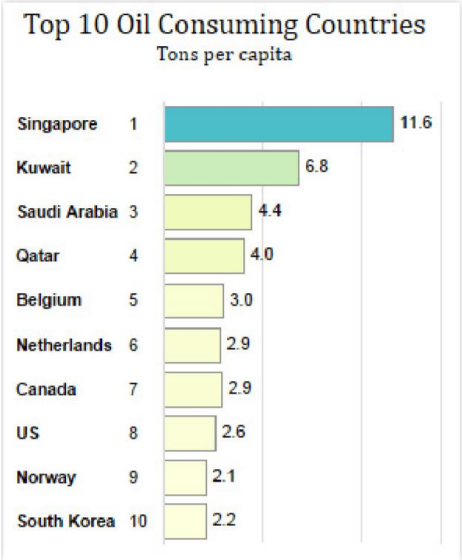


ภาพที่ 2.2 Area Chart

ที่มา : https://www.d4biz.com/data-and-analytics/select-chart-for-data-visualization-and-analysis/

หนึ่งใน Chart ที่ควรใช้ก็คือ Area Chart เพราะ Area Chart นั่นบอกทั้งแนวโน้มและสัดส่วนได้พร้อมกัน เพียงแต่อาจจะเปรียบเทียบความต่างของแต่ละ Sector ได้ไม่เท่า Line Chart

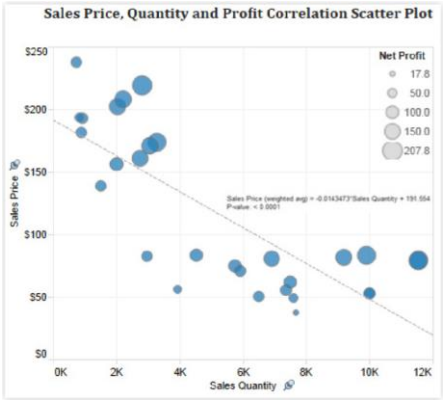
2. Comparison and Ranking การเปรียบเทียบและจัดอันดับ : Comparison and Ranking หรือ การเปรียบเทียบและจัดอันดับตามหมวดหมู่ต่าง ๆ ไม่ว่าจะเป็นแบ่งตามประเทศ หมวดหมู่ธุรกิจ และอื่น ๆ ซึ่ง Chart ที่เหมาะสมที่สุดก็คือ Bar Chart



ภาพที่ 2.3 Comparison and Ranking การเปรียบเทียบและจัดอันดับ

ที่มา : https://www.d4biz.com/data-and-analytics/select-chart-for-data-visualization-and-analysis/

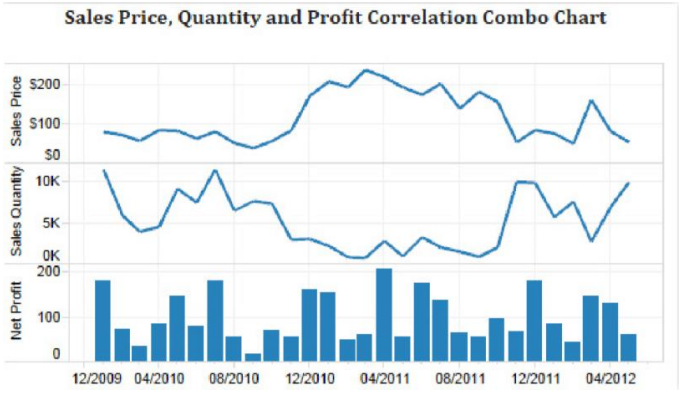
3. Correlation ความสัมพันธ์ระหว่างตัวแปร : เมื่อมีข้อสงสัยว่า สิ่งที่เกิดขึ้น มีความสัมพันธ์หรือส่งผลอะไรกับอีกสิ่งนึงรึเปล่า เช่น ราคาส่งผลกับยอดขายหรือเปล่า เราก็อาจจะเอาลองตัวเลขของราคากับยอดขายมาเปรียบเทียบดูว่า มันสอดคล้องหรือล้อกันไปหรือเปล่า ซึ่ง Chart ที่เหมาะสมกับการดู Correlation ก็คือ Scatter Plot หรือ Bubble Chart



ภาพที่ 2.4 Bubble Chart

ที่มา : https://www.d4biz.com/data-and-analytics/select-chart-for-data-visualization-and-analysis/

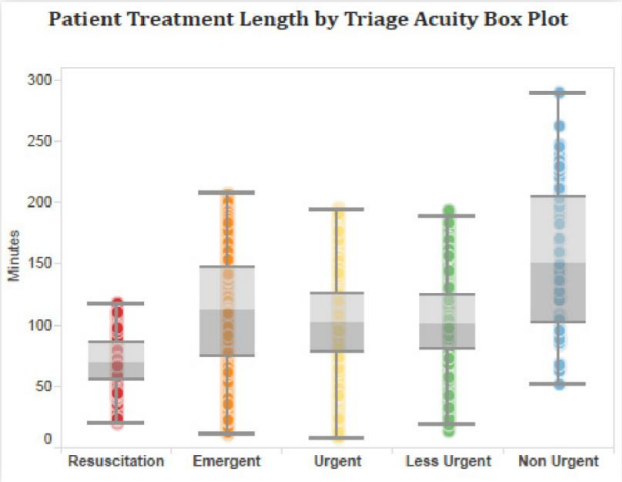
หรือหากเราต้องการดูเทรนด์ประกอบไปด้วย ก็อาจเปลี่ยนไปใช้ Chart หลายแบบประกอบกันได้ เช่น



ภาพที่ 2.5 Combo Chart

ที่มา : https://www.d4biz.com/data-and-analytics/select-chart-for-data-visualization-and-analysis/

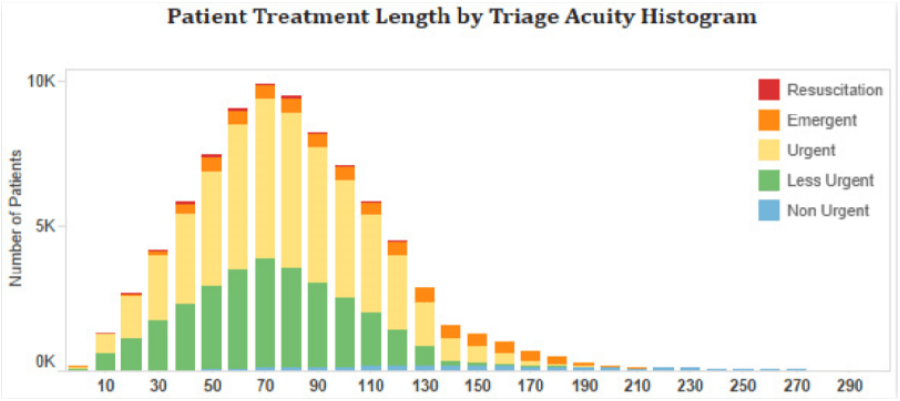
4. Distribution การกระจายตัวของข้อมูล : การทำ Distribution Analysis นั้น ช่วยให้เห็นภาพรายละเอียดที่ชัดเจนของข้อมูลยิ่งขึ้น เช่น ในแต่ละกลุ่มนั้น มีค่าสูงสุด ต่ำสุด ที่เท่าไหร่ หรือ ความหนาแน่นที่ค่าไหนมากที่สุด รวมถึงการ detect outlier ในข้อมูลได้ชัดเจน ซึ่ง Chart ที่สามารถบอกการกระจายตัวได้ดีก็คือ Box Plot



ภาพที่ 2.6 Box Plot Chart

ที่มา : https://www.d4biz.com/data-and-analytics/select-chart-for-data-visualization-and-analysis/

อีก Chart ที่ช่วยให้เราเห็นการกระจายตัวของข้อมูลได้ดีเช่นกัน ก็คือ Histogram



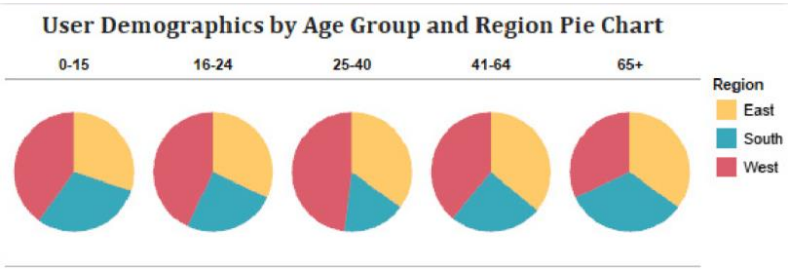
ภาพที่ 2.7 Histogram Chart

ที่มา : https://www.d4biz.com/data-and-analytics/select-chart-for-data-visualization-and-analysis/

5. Part to Whole การเปรียบเทียบสัดส่วน : ในการเปรียบเทียบสัดส่วนนั้นส่วนมาก Chart ที่ถูกนำมาใช้กันก็คือ Pie Chart ซึ่งหลายๆ คนที่ทำ Data Analysis มักแนะนำให้หลีกเลี่ยงด้วยสองเหตุผล ก็คือ

1.คนเรามักประเมินพื้นที่ด้วยสายตาได้ยาก

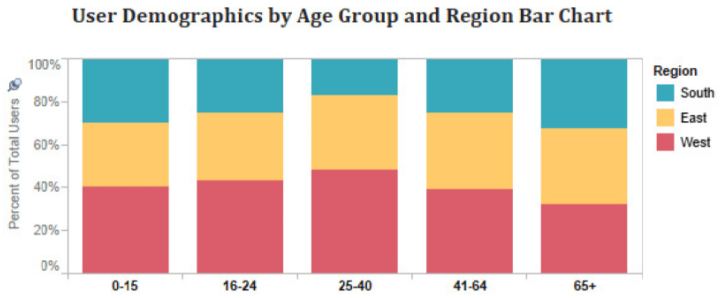
2.เรามักจะเปรียบเทียบสัดส่วนได้ภายใน Chart เดียวกัน แต่เปรียบเทียบกับ Chart อื่นได้ยาก เช่น



ภาพที่ 2.8 Pie Chart

ที่มา : https://www.d4biz.com/data-and-analytics/select-chart-for-data-visualization-and-analysis/

หากเราต้องการดูว่า สัดส่วนของ West ใน Chart ในช่วงอายุไหนสูงสุด ก็คงไม่ง่ายเท่าไหร่ แต่หากเราเปลี่ยน Chart เป็นแบบนี้



ภาพที่ 2.9 Stacked Bar Chart

ที่มา : https://www.d4biz.com/data-and-analytics/select-chart-for-data-visualization-and-analysis/

ซึ่งเป็น Stacked Bar Chart แบบเปรียบเทียบสัดส่วนเป็นเปอร์เซ็นต์ ก็น่าจะช่วยให้เห็นชัดและตอบคำถามว่า สัดส่วนของ West ใน Chart ในช่วงอายุ 25-40 นั่นสูงสุดในทุกช่วงอายุ

6. Geographical Data ข้อมูลเชิงพื้นที่ : หากต้องการดูข้อมูลรายพื้นที่หรือ Location แล้ว Chart ที่ชัดเจนที่สุดก็คือ Map แต่ Map Chart จะวิเคราะห์ได้ดีนั้น ควรใช้ Chart อื่นประกอบไปด้วยเพื่อให้เห็นภาพและวิเคราะห์ได้ดียิ่งขึ้น เช่น การใช้ Pie Chart หรือ Cross-tab Table

	Blog	Community	Examples	User-generated
United States	10K	16K	62K	67K
United Kingdom	2K	4K	12K	14K
Canada	1K	1K	6K	9K
India	1K	1K	3K	10K
Australia	1K	1K	4K	6K
China	0K	0K	1K	7K
France	1K	1K	2K	4K

ภาพที่ 2.10 Map Chart

ที่มา : https://www.d4biz.com/data-and-analytics/select-chart-for-data-visualization-and-analysis/

แม้ว่า Pie Chart มักจะถูกแนะนำว่า ไม่ควรนำมาใช้ในการวิเคราะห์สัดส่วน แต่หากนำมาใช้ประกอบกับ Map แล้ว เพื่อให้เราประเมินข้อมูลในแต่ละพื้นที่ได้ดียิ่งขึ้น

2.2 ทฤษฎี

2.2.1 ทฤษฎีเกี่ยวกับระบบจัดการฐานข้อมูล

2.2.1.1 ความหมายของระบบจัดการฐานข้อมูล

การจัดการฐานข้อมูล(Database Management) คือ การบริหารแหล่งข้อมูลที่ถูกเก็บรวบรวมไว้ที่ศูนย์กลาง เพื่อตอบสนองต่อการใช้ของโปรแกรมประยุกต์อย่างมีประสิทธิภาพ และลดการซ้ำซ้อนของข้อมูล รวมทั้งความขัดแย้งของข้อมูลที่เกิดขึ้นภายในองค์การ ในอดีตการเก็บข้อมูลมักจะเป็นอิสระต่อกันไม่มีการเชื่อมโยงของข้อมูลเกิดการ สิ้นเปลืองพื้นที่ในการเก็บข้อมูล เช่น องค์การหนึ่งจะมีแฟ้มบุคคล (Personnel) แฟ้มเงินเดือน (Payroll) และแฟ้ม สวัสดิการ (Benefits) อยู่แยกจากกัน เวลาผู้บริหารต้องการข้อมูลของพนักงานท่านใดจำเป็นจะต้องเรียกดูแฟ้มข้อมูลทั้ง 3 แฟ้ม ซึ่งเป็นการไม่สะดวก จงทำให้เกิดแนวความคิดในการรวมแฟ้มข้อมูลทั้ง 3

เข้าด้วยกันแล้วเก็บไว้ที่ ศูนย์กลางในลักษณะฐานข้อมูล (Database) จึงทำให้เกิดระบบการจัดการฐานข้อมูล (Database Management system (DBMS) ซึ่งจะต้องอาศัยโปรแกรมเฉพาะในการสร้างและบำรุงรักษา (Create and Maintenance) ฐานข้อมูลและสามารถที่จะให้ผู้ใช้ประยุกต์ใช้กับธุรกิจส่วนตัวได้โดยการดึงข้อมูล (Retrieve) ขึ้นมาแล้วใช้โปรแกรมสำเร็จรูปอื่นสร้างงานขึ้นมาโดยใช้ข้อมูลที่มีอยู่ในฐานข้อมูลแสดงการรวมแฟ้มข้อมูล 3 แฟ้มเข้าด้วยกัน

2.2.1.2 การออกแบบฐานข้อมูล

การออกแบบฐานข้อมูล (Designing Databases) มีความสำคัญต่อการจัดการระบบฐานข้อมูล (DBMS) ทั้งนี้เนื่องจากข้อมูลที่อยู่ภายในฐานข้อมูลจะต้องศึกษาถึงความสัมพันธ์ของข้อมูล โครงสร้างของข้อมูลการเข้าถึงข้อมูลและกระบวนการที่โปรแกรมประยุกต์จะเรียกใช้ฐานข้อมูล ดังนั้น เราจึงสามารถแบ่งวิธีการสร้างฐานข้อมูลได้ 3 ประเภท

1. ฐานข้อมูลเชิงสัมพันธ์ (Relational Database) : เป็นการเก็บข้อมูลในรูปแบบที่เป็นตาราง (Table) หรือเรียกว่า รีเลชั่น (Relation) มีลักษณะเป็น 2 มิติ คือเป็นแถว (row) และเป็นคอลัมน์ (column) การเชื่อมโยงข้อมูลระหว่างตาราง จะเชื่อมโยงโดยใช้แอททริบิวต์ (attribute) หรือคอลัมน์ที่เหมือนกันทั้งสองตารางเป็นตัวเชื่อมโยงข้อมูล ฐานข้อมูลเชิงสัมพันธ์นี้จะเป็นรูปแบบของฐานข้อมูลที่นิยมใช้ในปัจจุบัน ดังตัวอย่าง

รหัสพนักงาน	ชื่อพนักงาน	ที่อยู่	เงินเดือน	รหัสแผนก
12501535	นายสมพงศ์	กรุงเทพ	12000	VO
12534568	นายมนตรี	นครปฐม	12500	VN
12503452	นายเอก	กรุงเทพ	13500	VO
12356892	นายบรรทัด	นนทบุรี	11500	VD
15689730	นายราชัน	สมุทรปราการ	12000	VA

ภาพที่ 2.11 ภาพแสดงตารางพนักงาน

ที่มา : https://sites.google.com/site/pond1619/home

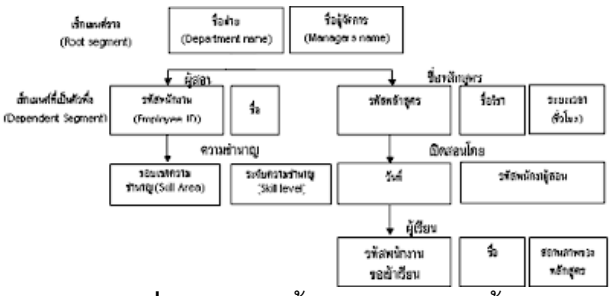
2. ฐานข้อมูลแบบเครือข่าย (Network Database) : ฐานข้อมูลแบบเครือข่ายจะเป็นการรวมระเบียนต่าง ๆ และความสัมพันธ์ระหว่างระเบียนแต่จะต่างกับฐานข้อมูลเชิงสัมพันธ์คือ ในฐานข้อมูลเชิงสัมพันธ์จะแฝงความสัมพันธ์เอาไว้ โดยระเบียนที่มีความสัมพันธ์กันจะต้องมี

ค่าของข้อมูลในแอททริบิวต์ใดแอททริบิวต์หนึ่งเหมือนกัน แต่ฐานข้อมูลแบบเครือข่าย จะแสดงความสัมพันธ์อย่างชัดเจน ตัวอย่างเช่น

ภาพที่ 2.12 ฐานข้อมูลแบบเครือข่าย

ที่มา : https://sites.google.com/site/pond1619/home

3. ฐานข้อมูลแบบลำดับชั้น (Hierarchical Database) : ฐานข้อมูลแบบลำดับชั้นเป็นโครงสร้างที่จัดเก็บข้อมูลในลักษณะความสัมพันธ์แบบพ่อ–ลูก (Parent–Child Relationship Type : PCR Type) หรือเป็นโครงสร้างรูปแบบต้นไม้ (Tree) ข้อมูลที่จัดเก็บในที่นี่ คือ ระเบียน (Record) ซึ่งประกอบด้วยค่าของเขตข้อมูล (Field) ของเอนทิตี้หนึ่ง ๆ



ภาพที่ 2.13 ฐานข้อมูลแบบลำดับชั้น

ที่มา : https://sites.google.com/site/pond1619/home

2.2.1.3 ประเภทของระบบฐานข้อมูล

การแบ่งประเภทของระบบฐานข้อมูลมีการแบ่งออกหลายประเภท ขึ้นอยู่กับชนิดและประเภทที่นำมาจำแนกในบทเรียนนี้จะแบ่งประเภทของระบบฐานข้อมูลออกเป็น 4 ประเภทใหญ่ ตามชนิดต่าง ๆ ดังนี้

1. แบ่งตามจำนวนของผู้ใช้ : การแบ่งโดยใช้จำนวนผู้ใช้เป็นหลัก สามารถแบ่งออกเป็น 2 ประเภทได้แก่

1.1 ผู้ใช้คนเดียวเป็นระบบฐานข้อมูลที่ใช้ภายในองค์กรขนาดเล็ก เช่น ระบบ Point of sale ของร้านสะดวกซื้อ หรือระบบบัญชีของร้านเล็ก ๆ ทั่วไป เป็นต้น มีเครื่องคอมพิวเตอร์

เพียงเครื่องเดียวและผู้ใช้เพียงคนเดียว ไม่มีการแบ่งฐานข้อมูลร่วมกันใช้กับผู้อื่น ถ้าผู้ใช้คนอื่นต้องการใช้ระบบนี้จะต้องรอให้ผู้ใช้คนแรกเลิกใช้ก่อนจึงจะใช้ได้

1.2 ผู้ใช้หลายคน แบ่งออกเป็น 2 ประเภทย่อย ๆ ได้แก่ ผู้ใช้เป็นกลุ่ม หรือ Workgroup database และประเภทฐานข้อมูลขององค์กรขนาดใหญ่หรือ Enterprise database เป็นฐานข้อมูลที่มีผู้ใช้หลายกลุ่มหรือหลายแผนก และแต่ละกลุ่มอาจมีผู้ใช้หลายคน มีการแลกเปลี่ยนข้อมูลซึ่งกันหรืออาจจะใช้ฐานข้อมูลเดียวกันก็ได้ แต่จะอยู่ในองค์กรเดียวกันเท่านั้น องค์การขนาดใหญ่ เป็นระบบฐานข้อมูลที่ใช้กับองค์กรขนาดใหญ่ที่มีสาขาหลายสาขา ทั้งในประเทศหรือมีสาขาในต่างประเทศ จะใช้ฐานข้อมูลขนาดใหญ่ มีระบบสำรอง การรักษาความปลอดภัยเป็นอย่างดี

2. แบ่งโดยใช้ขอบเขตของงาน : การแบ่งโดยใช้ขอบเขตของงาน แบ่งออกเป็น 3 ประเภทใหญ่ ๆ ได้แก่ ประเภทผู้ใช้คนเดียว ประเภทผู้ใช้เป็นกลุ่มและประเภทองค์การขนาดใหญ่ ดังได้กล่าวรายละเอียดในตอนต้นแล้ว

3. แบ่งตามสถานที่ตั้ง : การแบ่งตามสถานที่ตั้ง แบ่งออกเป็น 2 ประเภทใหญ่ ๆ ได้แก่ ประเภท ศูนย์กลาง และประเภทกระจาย ทั้งสองประเภทมีรายละเอียดดังนี้

3.1 ประเภทศูนย์กลาง เป็นระบบฐานข้อมูลที่นำเอามาเก็บไว้ในตำแหน่งศูนย์กลาง ผู้ใช้ทุกแผนก ทุกคนจะต้องมาใช้ข้อมูลร่วมกัน ตามสิทธิ์ของผู้ใช้แต่ละกลุ่มหรือแต่ละคน

3.2 ประเภทกระจาย เป็นระบบฐานข้อมูลที่เก็บฐานข้อมูลไว้ ณ ตำแหน่งใด ๆ ของแผนก และแต่ละแผนกใช้ฐานข้อมูลร่วมกันโดยผู้มีสิทธิ์ใช้ตามสิทธิ์ที่ได้กำหนดจากผู้มีอำนาจ การเข้าถึงข้อมูล เช่น ฐานข้อมูลของฝ่ายบุคคลเก็บไว้ที่แผนกทรัพยากรบุคคล ยอมให้ฝ่ายบัญชีนำรายชื่อของพนักงานไปใช้ร่วมกับฐานข้อมูลการจ่ายโบนัส และในขณะเดียวกันฝ่ายบัญชีมีฐานข้อมูลเก็บเงินเดือน สวัสดิการและรายจ่ายต่าง ๆ ของพนักงานเพื่อให้แผนกอื่น ๆ เข้ามาใช้ได้เช่นกัน

4. แบ่งตามการใช้งาน : การแบ่งตามการใช้งานแบ่งออกเป็น 3 ประเภทใหญ่ ๆ ได้แก่ ฐานข้อมูลสำหรับงานประจำวัน ฐานข้อมูลเพื่อใช้ในการตัดสินใจ และเพื่อเป็นคลังข้อมูล

4.1 ฐานข้อมูลสำหรับงานประจำวัน เป็นระบบฐานข้อมูลที่ใช้ในงานประจำวันของพนักงานระดับปฏิบัติการป้อนข้อมูลเข้าสู่ระบบ เช่น งานสินค้าคงคลัง งานระบบซื้อมาขายไป สำหรับร้านสะดวกซื้อ หรือระบบงานขายของร้านค้าทั่วไป เป็นต้น ฐานข้อมูลประเภทนี้มีการนำข้อมูลเข้า เปลี่ยนแปลงและลบออกตลอดทั้งวัน จึงทำให้ข้อมูลเปลี่ยนแปลงตลอดเวลา

4.2 ฐานข้อมูลเพื่อการตัดสินใจ ระบบฐานข้อมูลประเภทนี้มีไว้เพื่อใช้ในการสนับสนุนการตัดสินใจของผู้ใช้ระดับผู้บริหารระดับกลางขึ้นไป ข้อมูลที่นำเข้ามาในระบบได้จากการป้อนข้อมูลงานประจำวันของฐานข้อมูลสำหรับงานประจำวัน ส่วนใหญ่ฐานข้อมูลประเภทนี้นำไปใช้ในงานวางแผนกลยุทธ์ในองค์กร

4.3 ฐานข้อมูลเพื่อเป็นคลังข้อมูล ฐานข้อมูลประเภทนี้เกิดจากการนำข้อมูลเข้ามาในระบบทุก ๆ วันจึงทำให้เกิดมีข้อมูลขนาดใหญ่ จึงนำเอาข้อมูลที่มีประโยชน์มาสร้างฟังก์ชันหรือสมการต่างเพื่อประมวลผลหาผลลัพธ์ต่าง ๆ ให้เป็นประโยชน์กับองค์กร

2.2.1.4 หน้าที่ของระบบการจัดการฐานข้อมูล

ซอฟต์แวร์ระบบฐานการจัดการฐานข้อมูลที่ดีจะต้องทำหน้าที่แก้ปัญหาความไม่สมบูรณ์ ไม่คงเส้นคงวาของข้อมูลและทำให้ข้อมูลมีความถูกต้องไม่ขัดแย้งกันได้ จึงต้องมีหน้าที่ให้ครอบคลุมหลาย ๆ ด้าน ดังนี้

1. หน้าที่จัดการพจนานุกรมข้อมูล : ในการออกแบบฐานข้อมูลโดยปกติผู้ออกแบบได้เขียนพจนานุกรมข้อมูลในรูปของเอกสารให้กับโปรแกรมเมอร์ โปรแกรมเมอร์จะใช้ซอฟต์แวร์ระบบการจัดการฐานข้อมูลสร้างพจนานุกรมข้อมูลต่อไป และสามารถกำหนดความสัมพันธ์ระหว่างตาราง เมื่อมีการเปลี่ยนแปลงโครงสร้างข้อมูล จำเป็นต้องเปลี่ยนที่พจนานุกรมข้อมูลด้วย โปรแกรมเมอร์สามารถเปลี่ยนแปลงโครงสร้างข้อมูลได้ทันที ต่อจากนั้นจึงให้พจนานุกรมข้อมูลพิมพ์รายงาน พจนานุกรมข้อมูลที่เปลี่ยนแปลงไปแล้วเป็นเอกสารได้เลยทันทีโดยไม่ต้องแก้ไขที่เอกสาร

2. หน้าที่จัดการแหล่งจัดเก็บข้อมูล : ระบบการจัดการฐานข้อมูลที่ทันสมัยจะไม่ทำหน้าที่เพียงจัดการแหล่งจัดเก็บข้อมูลเท่านั้น แต่ยังเพิ่มหน้าที่ที่เกี่ยวกับการสร้างฟอร์มป้อนข้อมูลเข้าหรือกำหนดแบบจอภาพ แบบรายงาน หรือแม้แต่การตรวจสอบข้อมูลนำเข้าว่าถูกต้องหรือไม่ และจัดการเรื่องอื่น ๆ อีกหลายอย่าง

3. การเปลี่ยนรูปแบบและการแสดงผลข้อมูล : การเปลี่ยนรูปแบบและการแสดงผลข้อมูล เป็นหน้าที่สำหรับเปลี่ยนข้อมูลที่ถูกป้อนเข้าไปเป็นโครงสร้างข้อมูลจะจัดเก็บ ซึ่งอยู่ในมุมมองทางกายภาพ หรืออาจจะกล่าวได้ว่า ระบบจัดการฐานข้อมูลทำข้อมูลให้เป็นอิสระจากโปรแกรมประยุกต์ได้

4. จัดการด้านความปลอดภัยของข้อมูล : ระบบจัดการฐานข้อมูลทำหน้าที่รักษาความมั่นคง ความปลอดภัยของข้อมูล การไม่ยินยอมเข้าถึงข้อมูลจากผู้ใช้ที่ไม่มีสิทธิ์เข้าไปใช้ฐานข้อมูล โดยเฉพาะอย่างยิ่งฐานข้อมูลประเภทผู้ใช้หลายคน นอกจากนี้ยังสามารถกำหนดสิทธิ์ให้ผู้ใช้แต่ละคนใช้คำสั่ง เพิ่ม หรือลบ ปรับปรุงข้อมูลได้เป็นรายคนหรือรายกลุ่ม

5. ควบคุมการเข้าถึงข้อมูลของผู้ใช้ : การควบคุมการเข้าถึงข้อมูล เป็นการทำหน้าที่ให้ผู้ใช้เข้าใช้ได้หลาย ๆ คนในเวลาเดียวกันโดยไม่ทำให้เกิดขัดข้องของข้อมูล ซึ่งจะเน้นกฎความสมบูรณ์ของข้อมูลและการใช้ข้อมูลพร้อมกัน

6. สำรองข้อมูลและการกู้คืนข้อมูล : การสำรองข้อมูลและการกู้คืนข้อมูล เป็นหน้าที่ที่จำเป็นอย่างยิ่งเพื่อให้ผู้ใช้ระบบฐานข้อมูลมั่นใจว่าข้อมูลที่จัดเก็บอยู่ในเครื่องคอมพิวเตอร์ไม่ได้เสียหาย ยังมีความสมบูรณ์อยู่ตลอดเวลา ผู้ใช้ที่เป็นผู้บริหารฐานข้อมูลสามารถใช้คำสั่งสำรองข้อมูลและคำสั่งกู้คืนข้อมูลได้

7. จัดการด้านบูรณภาพของข้อมูล : เป็นข้อกำหนดให้มีกฎความสมบูรณ์เป็นบูรณภาพ โดยจะให้มีข้อมูลที่ซ้ำซ้อนกันให้น้อยที่สุด แต่ให้มีความถูกต้องตรงกันให้มากที่สุด เพราะในระบบฐานข้อมูลเชิงสัมพันธ์จะมีหลาย ๆ ตารางที่สัมพันธ์กันตารางที่เกี่ยวข้องกันจะขัดแย้งกันไม่ได้

8. เป็นภาษาสำหรับจัดการข้อมูลและจัดสร้างส่วนประสานกับผู้ใช้ : ระบบจัดการฐานข้อมูลจัดให้มีภาษาสำหรับสอบถาม เป็นภาษาที่เขียนเข้าใจง่ายไม่เหมือนภาษาชั้นสูงประเภท Procedural ทั่วไป ทำให้ผู้เขียนโปรแกรมภาษาระดับสูงเขียนคำสั่งเข้าไปสอบถามข้อมูลหรือประมวลผลสารสนเทศได้ตามต้องการ

9.เป็นส่วนประสานกับผู้ใช้ในด้านการสื่อสารฐานข้อมูล : ระบบการจัดการฐานข้อมูลสมัยใหม่จะสนับสนุนการทำงานแบบเครือข่ายอินเทอร์เน็ต เพื่อให้ผู้ใช้สามารถเขียนคำสั่งด้วยโปรแกรมที่ทำงานบน www เช่น browser ของ Internet Explorer หรือ Netscape เป็นต้น

2.2.2 ทฤษฎีเกี่ยวกับการออกแบบเว็บไซต์

การสร้างเว็บไซต์สิ่งสำคัญอยู่ที่การ ออกแบบเว็บ เพราะเว็บไซต์ที่มีรูปแบบสวยงาม จะสามารถดึงดูดความสนใจจากผู้คนได้ดีกว่า ทำให้ผู้คนเกิดความรู้สึกประทับใจ อยากกลับมาใช้งานเว็บไซต์อีกครั้งในอนาคต ดังนั้นเริ่มแรกก่อนทำเว็บไซต์ จึงจำเป็นต้องทำความเข้าใจ กับหลักการออกแบบ และรูปแบบโครงสร้างของเว็บก่อน

2.2.2.1 องค์ประกอบในการออกแบบเว็บไซต์

การออกแบบเว็บไซต์เพื่อให้มีประสิทธิภาพ และสามารถดึงดูดความสนใจของผู้คนได้ดี จะต้องมีองค์ประกอบของเว็บไซต์อย่างครบถ้วน ซึ่งได้แก่

1. ความเรียบง่าย เข้าใจง่าย : การออกแบบเว็บไซต์ที่ดี จะต้องเน้นที่ความเรียบง่ายเป็นหลัก โดยเลือกนำเสนอเฉพาะสิ่งที่ต้องการนำเสนอจริง ๆ ในรูปแบบที่หลากหลาย โดยอาจจะเป็นสีสัน กราฟิก ภาพเคลื่อนไหวหรือตัวอักษร ที่สำคัญจะต้องมีการนำเสนอที่ไม่ดูรกหน้าเว็บจนเกินไป เพื่อไม่ให้เกิดความรู้สึกรกสายตา หรือสร้างความเบื่อหน่าย น่ารำคาญให้กับผู้ที่เข้าชมเว็บไซต์ มีตัวอย่างเว็บไซต์ที่มีการออกแบบโดยเน้นความเรียบง่ายได้ดี คือ Apple, Nokia และ Microsoft เป็นต้น

2. ความสม่ำเสมอ ไม่สับสน : ควรออกแบบเว็บไซต์ด้วยความสม่ำเสมอ คือจะต้องมีรูปแบบ กราฟิก โทนสีและการตกแต่งต่าง ๆ ให้แต่ละหน้าบนเว็บไซต์มีความคล้ายคลึงกันและเป็นแนวเดียวกันไปตลอดทั้งเว็บไซต์ ดังตัวอย่างเว็บไซต์ทั่ว ๆ ไปที่จะสังเกตเห็นได้ว่าทุกหน้าของเว็บไซต์นั้น จะเน้นการตกแต่งในรูปแบบเดียวกันทั้งหมด ต่างก็แค่การนำเสนอของแต่ละหน้าเท่านั้น

3. สร้างความโดดเด่น เป็นเอกลักษณ์ : การออกแบบเว็บไซต์เพื่อให้สามารถสื่อถึงจุดประสงค์ในการนำเสนอเว็บได้ดี จะต้องมีการสร้างความเป็นเอกลักษณ์และจุดเด่นให้กับเว็บไซต์ เพื่อให้สามารถสะท้อนถึงลักษณะขององค์กรได้มากที่สุด โดยการสร้างเอกลักษณ์ดังกล่าวนั้น อาจใช้ชุดสี รูปภาพ ตัวอักษรหรือกราฟิก นอกจากนี้ก็ต้องขึ้นอยู่กับว่า เป็นเว็บไซต์แบบบทางการหรือไม่ เพื่อจะได้ออกแบบได้อย่างเหมาะสมที่สุด

4. เนื้อหาต้องดี ครบถ้วน : เนื้อหาเป็นสิ่งที่สำคัญที่สุดของการสร้างเว็บไซต์ เพราะสิ่งที่ทำให้ผู้คนเกิดความสนใจ และหมั่นติดตามเว็บไซต์เหล่านั้นอยู่เสมอ ก็คือเนื้อหาที่มีความสมบูรณ์และน่าสนใจ นอกจากนี้จะต้องมีการปรับปรุง พัฒนาเนื้อหาบนเว็บให้มีความทันสมัยอยู่เสมอ รวมถึงข้อมูลต้องมีความถูกต้องที่สุด

5. ระบบเนวิเกชัน ใช้ง่าย : ระบบเนวิเกชัน เป็นเสมือนป้ายบอกทางเพื่อให้ผู้ใช้งาน ไม่เกิดความสับสนในขณะใช้งานเว็บไซต์ ซึ่งการออกแบบเนวิเกชันก็จะต้องเน้นที่ความเรียบง่าย ใช้งานสะดวก และมีความเข้าใจได้ง่าย ที่สำคัญจะต้องมีตำแหน่งการวางที่สม่ำเสมอเพื่อให้ดูเป็นแนวทางเดียวกัน ทำให้ผู้ใช้งานหรือผู้ชมรู้สึกประทับใจ และจดจำเว็บไซต์ได้ง่ายขึ้น

ส่วนใครที่มีการนำกราฟิกมาใช้ในระบบเนวิเกชั่น ก็จะต้องเลือกกราฟิกที่สามารถสื่อความหมายได้ดีเช่นกัน

6. คุณภาพของเว็บไซต์ : เว็บไซต์ที่ดีจะต้องมีคุณภาพ ทั้งสิ่งที่ปรากฏให้เห็นบนเว็บไซต์ ไม่ว่าจะเป็นกราฟิก ชนิดตัวอักษร รูปภาพหรือสีสันที่ใช้ เนื้อหาที่นำมาแสดงผล ซึ่งหากเว็บไซต์มีคุณภาพก็จะสร้างความน่าเชื่อถือ และเป็นจุดเด่นที่ทำให้ผู้คนส่วนใหญ่เกิดความสนใจได้ดี เพราะฉะนั้นห้ามละเลยในส่วนของคุณภาพเด็ดขาด

7. ความสะดวกในการเข้าใช้งาน : เว็บไซต์ควรให้ความสะดวกสบายแก่ผู้ใช้งานได้ดี คือจะต้องมีการแสดงผลได้ในทุกระบบปฏิบัติการ ไม่ว่าจะเป็นเว็บเบาว์เซอร์ คอมพิวเตอร์โน๊ตบุ๊คหรือบนโทรศัพท์มือถือ ที่สำคัญจะต้องมีความละเอียดของการแสดงผลและสามารถใช้งานได้โดยไม่มีปัญหาด้วย

8. ความคงที่ของการออกแบบ : การออกแบบเว็บไซต์ควรจะมีความคงที่ในการออกแบบ ด้วยการสร้างเว็บไซต์ด้วยแบบแผนเดียวกัน และมีการเรียบเรียงเนื้อหาอย่างรอบคอบ ทำให้เว็บมีความน่าเชื่อถือ และดูมีคุณภาพ ช่วยสร้างความประทับใจให้กับผู้ใช้งานได้เป็นอย่างดี

9. ความคงที่ของการทำงาน : ระบบการทำงานบนเว็บไซต์จะต้องมีความคงที่และสามารถใช้งานได้ดี ซึ่งนอกจากการออกแบบระบบการทำงานให้มีความทันสมัยและสร้างสรรค์แล้ว ก็จะต้องหมั่นตรวจสอบอยู่เสมอ เพราะหากระบบการใช้งานมีความผิดปกติก็จะได้แก้ปัญหาได้ทัน นอกจากนี้อาจมีการอัพเดตดีไซน์ให้ทันสมัยขึ้นบ่อย ๆ เพื่อให้ผู้ใช้งานรู้สึกสนุกไปกับการใช้งานเว็บไซต์

2.2.2.2 รูปแบบโครงสร้างของเว็บไซต์

การออกแบบโครงสร้างของเว็บไซต์ สามารถทำได้หลากหลายแบบ ซึ่งก็ขึ้นอยู่กับความชอบและความถนัดของแต่ละบุคคล นอกจากนี้ยังขึ้นอยู่กับกลุ่มเป้าหมายที่ต้องการนำเสนอ เพราะจะต้องออกแบบให้เหมาะกับการใช้งานของกลุ่มเป้าหมายมากที่สุด โดยโครงสร้างของเว็บไซต์ส่วนใหญ่ก็จะประกอบไปด้วย 4 รูปแบบดังนี้

1. โครงสร้างแบบเรียงลำดับ : จะเป็นโครงสร้างแบบธรรมดาที่นิยมใช้งานกันมากที่สุด เนื่องจากมีความง่ายต่อการจัดระบบข้อมูล และสามารถนำเสนอเรื่องราวตามลำดับได้เป็นอย่างดี เหมาะกับเว็บไซต์ที่มีขนาดเล็ก มีเนื้อหาที่ไม่ซับซ้อน ส่วนใหญ่ก็จะเป็นพวกเว็บไซต์ที่ให้ความรู้ หรือเว็บไซต์องค์กรขนาดย่อม โดยลักษณะการลิ้งค์เนื้อหา ก็จะลิ้งค์ไปทีละหน้า มีทิศ

ทางการเข้าสู่เนื้อหาต่าง ๆ ในแบบเส้นตรง ใช้ปุ่มเดินหน้า-ถอยหลังในการกำหนดทิศทาง จึงทำให้การใช้งานเป็นไปอย่างง่าย แต่โครงสร้างเว็บไซต์แบบเรียงลำดับก็มีข้อเสีย คือจะทำให้ผู้ใช้งานต้องเสียเวลาในการเข้าสู่เนื้อหาเพราะไม่สามารถกำหนดทิศทางการเข้าสู่เนื้อหาด้วยตัวเองได้

ภาพที่ 2.14 โครงสร้างแบบเรียงลำดับ

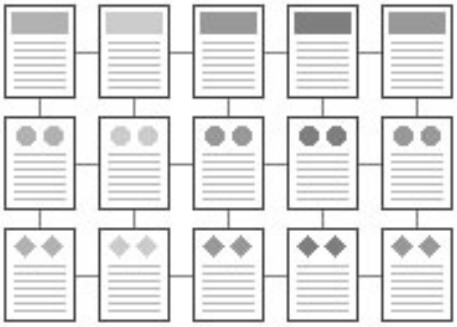
ที่มา : https://www.1belief.com/article/website-design/

2. โครงสร้างแบบลำดับขั้น : นิยมใช้กับเว็บที่มีความซับซ้อนของข้อมูล เพื่อให้สามารถเข้าถึงข้อมูลต่าง ๆ ได้ง่ายขึ้น โดยจะมีการแบ่งเนื้อหาออกเป็นส่วนๆ และมีการนำเสนอรายละเอียดย่อย ๆ ที่ลดหลั่นกันมา ทำให้สามารถทำความข้าใจกับโครงสร้างเนื้อหาได้ง่ายขึ้น โดยจะมีโฮมเพจเป็นจุดเริ่มต้น และจุดร่วมจุดเดียวที่จะนำไปสู่การเชื่อมโยงเนื้อหาเป็นลำดับจากบนลงล่าง

ภาพที่ 2.15 โครงสร้างแบบลำดับขั้น

ที่มา : https://www.1belief.com/article/website-design/

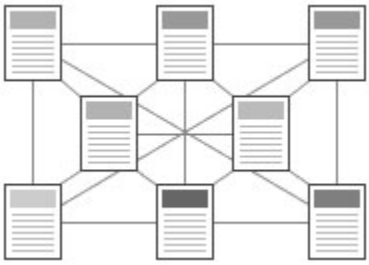
3. โครงสร้างแบบตาราง : โครงสร้างแบบตาราง เป็นโครงสร้างการออกแบบเว็บไซต์ที่มีความซับซ้อน แต่ก็มีความยืดหยุ่นในระดับหนึ่ง เพื่อให้ผู้ใช้งานสามารถเข้าสู่เนื้อหาต่าง ๆ ได้ง่ายขึ้น การออกแบบในลักษณะนี้จะมีการเชื่อมโยงเนื้อหาในแต่ละส่วนซึ่งกันและกัน ทำให้ผู้ใช้งานสามารถเปลี่ยนทิศทาง หรือกำหนดทิศทางในการเข้าสู่เนื้อหาด้วยตัวเองได้ จึงไม่ทำให้เสียเวลา แถมยังทำให้เว็บไซต์มีความทันสมัยขึ้น



ภาพที่ 2.16 โครงสร้างแบบตาราง

ที่มา : https://www.1belief.com/article/website-design/

4. โครงสร้างแบบใยแมงมุม : เป็นโครงสร้างที่ได้รับความนิยมเป็นอย่างมาก เพราะมีความยืดหยุ่นมากที่สุด โดยทุกหน้าเว็บจะมีการเชื่อมโยงถึงกันหมด ทำให้สามารถเข้าถึงหน้าเว็บเพจต่าง ๆ ที่ต้องการได้อย่างง่าย และมีความอิสระมากขึ้น นอกจากนี้ก็สามารถเชื่อมโยงไปสู่เว็บไซต์ภายนอกได้ดี



ภาพที่ 2.17 โครงสร้างแบบใยแมงมุม

ที่มา : https://www.1belief.com/article/website-design/

2.2.2.3 การออกแบบโครงสร้างของเว็บไซต์

โครงสร้างเว็บไซต์ ก็คือการจัดลำดับของเนื้อหาบนเว็บไซต์ออกเป็นแผนผังที่เข้าใจง่าย ว่าต้องการให้เว็บไซต์มีเนื้อหาอะไรบ้าง มีเว็บเพจอยู่ตรงไหน หน้าไหนบ้างที่จะนำมาเชื่อมโยงถึงกัน หรือกล่าวง่ายๆ ก็คือเหมือนการวางโครงเรื่องก่อนจะเขียนเนื้อหาให้ออกมาอย่างสมบูรณ์แบบนั่นเอง ดังนั้นการออกแบบโครงสร้างเว็บไซต์จึงมีความสำคัญมาก ซึ่งก็สามารถทำได้หลากหลายรูปแบบด้วยกัน แต่มีแนวคิดหลักๆ ที่ได้รับความนิยมมากที่สุดอยู่ 2 รูปแบบ คือ

1.จัดโครงสร้างตามกลุ่มเนื้อหา (Content-based Structure)

2.จัดโครงสร้างตามกลุ่มผู้ชม (User-based Structure)

2.2.2.4 ส่วนประกอบสำคัญบนหน้าเว็บไซต์

บนหน้าเว็บเพจ จะมีส่วนประกอบสำคัญที่จำเป็นต้องมีอยู่ 3 ส่วน ได้แก่

1. ส่วนหัวของหน้า (Header) : อยู่ตอนบนสุดของหน้าและเป็นส่วนที่สำคัญที่สุด โดยจะต้องทำให้สามารถดึงดูดผู้ชมให้รู้สึกอยากติดตามเนื้อหาในเว็บไซต์ต่อไป ซึ่งส่วนใหญ่ก็มักจะมีการใส่ภาพกราฟฟิคให้ดูสวยงาม สิ่งสำคัญหลักๆ เลย ก็คือ โลโก้ ชื่อเว็บไซต์และเมนูหลักที่สามารถลิ้งค์ไปยังเนื้อหาในหน้าเว็บเพจต่าง ๆ ได้

2. ส่วนของเนื้อหา (Body) : อยู่บริเวณตอนกลางของหน้าเว็บ โดยจะแสดงข้อมูลเกี่ยวกับเนื้อหาบนเว็บแบบคร่าวๆ ซึ่งก็จะมีข้อความ กราฟฟิค ตารางข้อมูลหรือวิดีโอประกอบอยู่ และหากมีเมนูแบบเฉพาะกลุ่มก็จะถูกจัดไว้ในหน้านี้เช่นกัน และที่สำคัญเนื้อหาในส่วนนี้ควรจะมีความกระชับ เข้าใจง่าย มีการใช้รูปแบบตัวอักษรแบบเรียบง่ายและเป็นระเบียบ

3. ส่วนท้ายของหน้า (Footer) : อยู่ล่างสุดของหน้าเว็บ ซึ่งจะมีหรือไม่มีก็ได้ ส่วนนี้จะแสดงถึงข้อมูลต่าง ๆ เพิ่มเติมเข้าไป เช่น ข้อความที่แสดงถึงการเป็นลิขสิทธิ์ ข้อมูลเจ้าของเว็บไซต์ วิธีการติดต่อและคำแนะนำต่าง ๆ เกี่ยวกับการใช้งานเว็บไซต์อย่างถูกต้อง เป็นต้น

2.2.3 ทฤษฎีเกี่ยวกับการทำนอร์มัลไลเซชัน (Normalization)

2.2.3.1 ความหมายของการทำนอร์มัลไลเซชัน (Normalization)

การทำนอร์มัลไลเซชัน เป็นวิธีการในการกำหนดแอตทริบิวต์ให้กับแต่ละเอนทิตี เพื่อให้ได้โครงสร้างของตารางที่ดี สามารถควบคุมความซ้ำซ้อนของข้อมูลหลีกเลี่ยงความผิดปกติของข้อมูล โดยทั่วไปผลลัพธ์ของการนอร์มัลไลเซชัน จะได้ตารางที่มีโครงสร้างซับซ้อนน้อยลง แต่จำนวนของตารางจะมากขึ้น การทำนอร์มัลไลเซชัน จะประกอบด้วยนอร์มัลฟอร์ม (Normal

Form) แบบต่าง ๆ ที่มีเงื่อนไขของการทำให้อยู่ในรูปของนอร์มัลฟอร์มที่แตกต่างกันไป ขึ้นอยู่กับผู้ออกแบบฐานข้อมูลว่า ต้องการลดความซ้ำซ้อนในฐานข้อมูลให้อยู่ในระดับใด ซึ่งประกอบด้วยนอร์มัลฟอร์มแบบต่าง ๆ ดังต่อไปนี้

–นอร์มัลฟอร์มที่ 1 (First Normal Form : 1NF)

–นอร์มัลฟอร์มที่ 2 (Second Normal Form : 2NF)

–นอร์มัลฟอร์มที่ 3 (Third Normal Form : 3NF)

–บอยซ์คอดด์นอร์มัลฟอร์ม (Boyce–Codd Normal Form : BCNF)

–นอร์มัลฟอร์มที่ 4 (Fourth Normal Form : 4NF)

–นอร์มัลฟอร์มที่ 5 (Fifth Normal Form : 5NF)

1. การแปลงให้อยู่ในรูปนอร์มัลฟอร์มที่ 1 (First Normal Form : 1NF)

คุณสมบัติของรีเลชันของแบบจำลองข้อมูลเชิงสัมพันธ์ ก็คือ ข้อมูลในแต่ละทัปเพิลจะต้องไม่ซ้ำกันและค่าในแต่ละแอตทริบิวต์จะต้องไม่สามารถถูกแบ่งแยกย่อยลงไปได้อีกหรือมีความเป็นอะตอมมิค (Atomic) รวมถึงจะต้องมีค่าเพียงค่าเดียวที่อยู่ในแต่ละแอตทริบิวต์หรือมีความเป็นซิงเกิลแวลู (Single Value) ซึ่งในการทำนอร์มัลไลเซชันให้อยู่ในนอร์มัลฟอร์ที่ 1 ก็อาศัยคุณสมบัติดังที่กล่าวไว้ข้างต้น

1.1 รีพีทติ้งกรุ๊ป (Repeating Group)

การที่ข้อมูลใน 1 ทัปเพิล สามารถมีค่าในแต่ละแอตทริบิวต์ได้มากกว่าหนึ่งค่า (Multivalued) จะทำให้เกิดรีพีทติ้งกรุ๊ป ดังตารางที่แสดงในภาพข้างล่าง ซึ่งเลขที่โครงการหนึ่งหมายเลขประกอบด้วยกลุ่มข้อมูลหลายกลุ่ม ซึ่งทำให้รีเลชันดังกล่าว ขาดคุณสมบัติซิงเกิลแวลู

การทำงานของพนักงานในโครงการ

เลขที่โครงการ	ชื่อโครงการ	รหัสนักงาน	ชื่อพนักงาน	ตำแหน่งงาน	ค่าแรง/ ชม.	จำนวน ชม.
11	RFID	103	สมชาย	Engineer	500	23.8
		101	วิชา	Programmer	500	19.4
		102	สุรชัย	Administrator	200	12.6

ภาพที่ 2.18 การรีพีทติ้งกรุ๊ป

ที่มา : https://msit5.wordpress.com

1.2 นิยามของนอร์มัลฟอร์มที่ 1

รีเลชันจะอยู่ในรูปของนอร์มัลฟอร์มที่ 1 ก็ต่อเมื่อมีคุณสมบัติตามเงื่อนไขดังต่อไปนี้

1. มีการกำหนดแอตทริบิวต์ที่เป็นคีย์
2. ต้องไม่มีรีพีทติ้งกรุ๊ป แต่ละแถวหรือคอลัมน์จะมีค่าได้เพียง 1 ค่าเท่านั้น
3. แอตทริบิวต์ทุกตัวต้องขึ้นอยู่กับคีย์หลัก

จากภาพข้างบน เมื่อการการนอร์มัลไลเซชันให้อยู่ในรูปนอร์มัลฟอร์มที่ 1 จะได้ตารางที่แตกย่อยออกมาเป็น 2 ตาราง ดังภาพข้างล่าง ซึ่งมีคุณสมบัติตามนอร์มัลฟอร์มที่ 1 แล้ว

การทำงานของพนักงาน

รหัสนักงาน	ชื่อพนักงาน	ตำแหน่งงาน	ค่าแรง/ ชม.	จำนวน ชม.
103	สมชาย	Engineer	500	23.8
101	วิชา	Programmer	500	19.4
102	สุรชัย	Administrator	200	12.6

ชื่อโครงการ

เลขที่โครงการ	ชื่อโครงการ
11	RFID

ภาพที่ 2.19 รีเลชันที่อยู่ในรูปนอร์มัลฟอร์มที่ 1

ที่มา : https://msit5.wordpress.com

2. การแปลงให้อยู่ในรูปนอร์มัลฟอร์มที่ 2 (Second Normal Form : 2NF)

ในหนึ่งรีเลชันจะประกอบด้วยแอตทริบิวต์ต่าง ๆ ที่มีความสัมพันธ์ที่ขึ้นต่อกัน ซึ่งความสัมพันธ์ดังกล่าวจะเป็นตัวกำหนดว่าแอตทริบิวต์ใดเป็นตัวกำหนดข้อมูล หรือ คีย์แอตทริบิวต์ (Key Attribute) และและแอตทริบิวต์ใดเป็นข้อมูลที่ถูกกำหนดหรือนอนคีย์แอตทริบิวต์ (Nonkey Attribute)

2.1 ฟังก์ชันนัลดีเพนเดนซี (Functional Dependency: FD)

ในการทำนอร์มัลไลเซชัน จะต้องมีความเข้าใจหลักการของฟังก์ชันดีเพนเดนซี (Function Dependency : FD) เสียก่อน โดยมีคำจำกัดความคือ B ขึ้นอยู่กับ A ถ้าทราบค่าของ A ก็จะทำให้รู้ค่าของ B ได้ ฟังก์ชันนัลดีเพนเดนซี สามารถแสดงด้วยการใช้เครื่องหมายลูกศร (–>) ตัวอย่างเช่น A–>B แสดง B เป็นฟังก์ชันนัลดีเพนเดนต์กับ A กล่าวคือ ถ้ารู้ค่า A ก็จะทำให้ทราบค่าของ B ด้วย ทุกค่าของ A ที่มีค่าเท่ากัน จะได้ค่าเท่ากันเสมอ

2.2 พาเชียลดีเพนเดนซี (Partial Dependency)

พาร์เชียลดีเพนเดนซี หมายถึง การที่มีแอตทริบิวต์บางแอตทริบิวต์ ที่ขึ้นอยู่กับเพียงบางส่วนของคีย์หลักเท่านั้น ตัวอย่างเช่น จากตารางในภาพข้างล่าง แอตทริบิวต์ชื่อพนักงานจะขึ้นอยู่กับคีย์รหัสพนักงาน ในขณะที่แอตทริบิวต์ชื่อแผนก จะขึ้นอยู่กับคีย์รหัสแผนก จะเห็นว่า ข้อมูลที่อยู่ในรีเลชันเดียวกัน แต่ไม่ได้ขึ้นอยู่กับคีย์ใดคีย์หนึ่งทั้งหมด แต่จะขึ้นอยู่กับคีย์ใดคีย์หนึ่งเพียงบางส่วนเท่านั้น

พนักงานในแผนก

รหัสนักงาน	ชื่อพนักงาน	รหัสแผนก	ชื่อแผนก
103	สมชาย	501	บัญชี
101	วิชา	601	การตลาด
102	สุรชัย	301	สารสนเทศ

ภาพที่ 2.20 รีเลชันที่มีพาเชียลดีเพนเดนซี

ที่มา : https://msit5.wordpress.com

2.3 นิยามของนอร์มัลฟอร์มที่ 2

รีเลชันจะอยู่ในรูปของนอร์มัลฟอร์มที่ 2 ก็ต่อเมื่อมีคุณสมบัติตามเงื่อนไขดังต่อไปนี้

1. รีเลชันนั้นเป็นนอร์มัลฟอร์มที่ 1 อยู่แล้ว
2. รีเลชันนั้นไม่มีพาร์เชียลดีเพนเดนซี

ตัวอย่างรีเลชันพนักงานในแผนกในภาพข้างบน เมื่อทำการแตกออกเป็นรีเลชันย่อยที่ไม่มีพาร์เชียลดีเพนเดนซีแล้ว จะได้เป็นรีเลชันสองรีเลชัน คือ รีเลชันพนักงานและรีเลชันแผนก ซึ่งอยู่ในรูปของนอร์มัลฟอร์มที่ 2 แล้ว ดังภาพข้างล่าง

พนักงาน

รหัสนักงาน	ชื่อพนักงาน
03	สมชาย
101	วิชา
102	สุรชัย

แผนก

รหัสแผนก	ชื่อแผนก
501	บัญชี
601	การตลาด
301	สารสนเทศ

ภาพที่ 2.21 รีเลชันที่อยู่ในรูปนอร์มัลฟอร์มที่ 2 แล้ว

ที่มา : https://msit5.wordpress.com

3. การแปลงให้อยู่ในรูปนอร์มัลฟอร์มที่ 3 (Third Normal Form : 3NF)

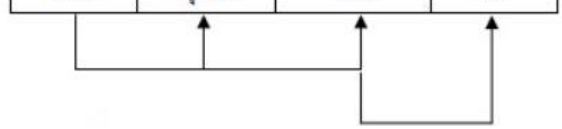
ในหนึ่งรีเลชันจะประกอบคีย์แอตทริบิวต์และนอนคีย์แอตทริบิวต์ คีย์แอตทริบิวต์จะต้องเป็นตัวกำหนดความหมายหรือการมีอยู่ของแอตทริบิวต์อื่น ๆ ที่อยู่ในรีเลชันเสมอ

3.1 ทรานซิทีฟดีเพนเดนซี (Transitive Dependency)

ทรานซิทีฟดีเพนเดนซี หมายถึง การที่มีฟังก์ชันนัลดีเพนเดนซี ระหว่างแอตทริบิวต์ที่ไม่ได้เป็นส่วนของคีย์ใด ๆ แต่มีแอตทริบิวต์อื่น ๆ มาขึ้นกับแอตทริบิวต์นั้นตัวอย่างเช่น จากตารางในภาพข้างล่าง แอตทริบิวต์ชื่อพนักงาน และรหัสตำแหน่งงานจะขึ้นอยู่กับคีย์รหัสพนักงาน ในขณะที่แอตทริบิวต์ค่าแรงต่อชั่วโมงของพนักงาน จะขึ้นอยู่กับแอตทริบิวต์รหัสตำแหน่งงานซึ่งไม่ใช่คีย์อีกต่อหนึ่งทำให้มีทรานซิทีฟดีเพนเดนซีเกิดขึ้นในรีเลชันนี้

การทำงานของพนักงาน

รหัสนักงาน	ชื่อพนักงาน	รหัสตำแหน่ง	ค่าแรง/ ชม.
103	สมชาย	702	500
101	วิชา	704	500
102	สุรชัย	705	200



ภาพที่ 2.22 ตารางที่มีทรายซิทีฟดีเพนเดนซี

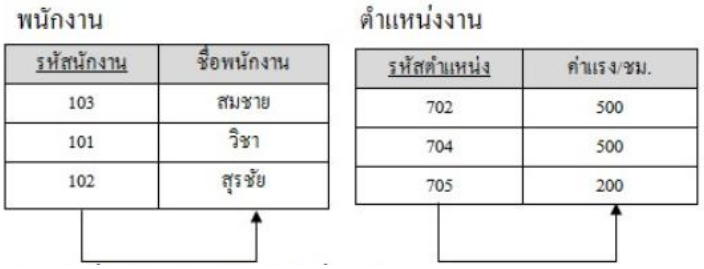
ที่มา : https://msit5.wordpress.com

3.2 นิยามของนอร์มัลฟอร์มที่ 3

รีเลชันจะอยู่ในรูปของนอร์มัลฟอร์มที่ 3 ก็ต่อเมื่อมีคุณสมบัติตามเงื่อนไขดังต่อไปนี้

1. รีเลชันนั้นเป็นนอร์มัลฟอร์มที่ 2 อยู่แล้ว
2. รีเลชันนั้นไม่มีทรานซิทีฟดีเพนเดนซี

ตัวอย่างรีเลชัน การทำงานของพนักงาน ในภาพข้างบน เมื่อทำการแตกออกเป็นรีเลชันย่อยที่ไม่มีทรานซิทีฟดีเพนเดนซีแล้ว จะได้เป็นรีเลชันสองรีเลชัน คือรีเลชันพนักงาน และรีเลชันตำแหน่งงาน ซึ่งอยู่ในรูปของนอร์มัลฟอร์มที่ 3 แล้ว ดังภาพข้างล่าง



ภาพที่ 2.23 รีเลชันที่อยู่ในรูปนอร์มัลฟอร์มที่ 3 แล้ว

ที่มา : https://msit5.wordpress.com

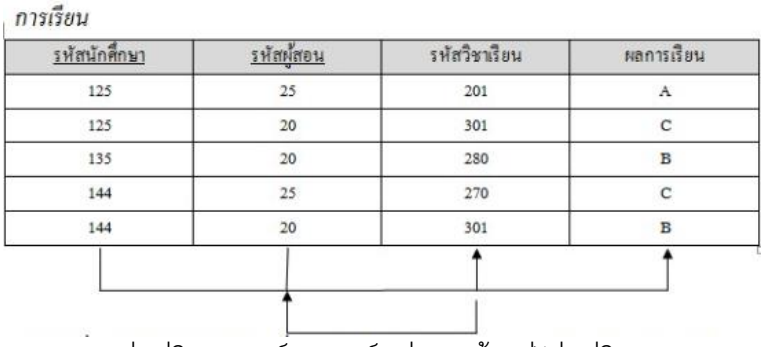
4. การแปลงให้อยู่ในรูปบอยซ์คอดด์นอร์มัลฟอร์ม (Boyce–Codd Normal Form : BCNF)

ในหนึ่งรีเลชันอาจจะประกอบด้วยหลายแคนดิเดตคีย์ (Candidate Key) ทุกแอตทริบิวต์ในรีเลชันจะต้องขึ้นอยู่กับแคนดิเดตคีย์เสมอ เราสามารถกำหนดนิยามของรีเลชันที่อยู่ในรูปของบอยซ์คอดด์นอร์มัลฟอร์ม ก็ต่อเมื่อรีเลชันมีคุณสมบัติตามเงื่อนไขดังต่อไปนี้

1. รีเลชันนั้นเป็นนอร์มัลฟอร์มที่ 3 อยู่แล้ว
2. ทุกแอตทริบิวต์ในรีเลชันจะต้องขึ้นกับแคนดิเดตคีย์

รีเลชันจะอยู่ในรูปบอยซ์คอดด์นอร์มัลฟอร์ม ถ้าทุกแอตทริบิวต์ขึ้นอยู่กับแคนดิเดตคีย์ (Candidate Key) ดังนั้นถ้าใน 1 รีเลชันมีแคนดิเดตคีย์เพียงตัวเดียวแล้ว นอร์มัลฟอร์มที่ 3 และบอยซ์คอดด์นอร์มัลฟอร์ม จะเหมือนกัน โอกาสที่คุณสมบัติของบอยซ์คอดด์นอร์มัลฟอร์มจะถูกละเมิดนั้น เกิดขึ้นได้น้อย และจะเกิดได้กับรีเลชันที่มีแคนดิเดตคีย์มากกว่าหนึ่งเท่านั้น ดังตัวอย่างในภาพข้างล่าง รีเลชันการลงทะเบียนเรียน รีเลชันดังกล่าวอยู่ในรูปนอร์มัลฟอร์มที่ 3 แล้ว แต่ก็ยังมีบางส่วนมีปัญหาอยู่ ตรงจุดที่แอตทริบิวต์รหัสวิชาเรียน และผลการเรียนขึ้นอยู่กับคีย์นักศึกษา และคีย์ผู้สอน แต่ในขณะเดียวกันรหัสผู้สอนก็ขึ้นอยู่กับรหัสวิชาเรียน ทำให้ถ้าต้องการเปลี่ยนแปลงผู้สอนในวิชา 301 จะต้องมีการเปลี่ยนแปลงถึง 2 ทัปเพิล ซึ่งผลลัพธ์ที่ได้อาจจะทำให้เกิดความ

ผิดพลาดหากทำการแก้ไขไม่ครบถ้วน และถ้านักศึกษารหัส 135 ถอนการลงทะเบียนวิชา 280 ข้อมูลของผู้ที่สอนวิชานี้จะหายไปจากระบบเลย ถ้าเราลบข้อมูลนี้



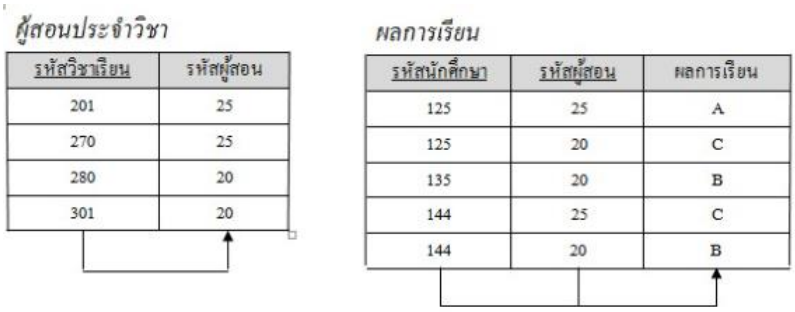
การเรียน

รหัสนักศึกษา	รหัสผู้สอน	รหัสวิชาเรียน	ผลการเรียน
125	25	201	A
125	20	301	C
135	20	280	B
144	25	270	C
144	20	301	B

ภาพที่ 2.24 รีเลชันที่อยู่ในรูปนอร์มัลฟอร์มที่ 3 แล้วแต่ไม่อยู่ในรูปของบอยซ์คอดด์นอร์มัลฟอร์ม

ที่มา : https://msit5.wordpress.com

เราสามารถทำการแตกตารางออกมาให้อยู่ในรูปของบอยซ์คอดด์นอร์มัลฟอร์มได้ โดยการแยกแอตทริบิวต์รหัสวิชาเรียนและรหัสผู้สอนซึ่งขึ้นอยู่กับแอตทริบิวต์รหัสวิชาเรียน ออกมาเป็นอีกหนึ่งรีเลชัน และแยกแอตทริบิวต์ รหัสนักศึกษา รหัสผู้สอน และผลการเรียนออกมาเป็นอีกหนึ่งรีเลชัน ดังแสดงในภาพข้างล่าง



ผู้สอนประจำวิชา

รหัสวิชาเรียน	รหัสผู้สอน
201	25
270	25
280	20
301	20

ผลการเรียน

รหัสนักศึกษา	รหัสผู้สอน	ผลการเรียน
125	25	A
125	20	C
135	20	B
144	25	C
144	20	B

ภาพที่ 2.25 รีเลชันที่ได้รับการนอร์มัลไลเซชันให้อยู่ในรูปบอยซ์คอดด์นอร์มัลฟอร์มแล้ว

ที่มา : https://msit5.wordpress.com

5. การแปลงให้อยู่ในรูปนอร์มัลฟอร์มที่ 4 (Fourth Normal Form : 4NF)

ในขณะที่การทำให้อยู่ในรูปของนอร์มัลฟอร์มต่าง ๆ ที่ผ่านมา จะเกี่ยวข้องกับการขึ้นตรงต่อกันของข้อมูลในแต่ละแอตทริบิวต์หรือฟังก์ชันนัลดีเพนเดนซี แต่การทำให้อยู่ในรูปของนอร์มัลฟอร์มที่ 4 จะเกี่ยวข้องกับรูปแบบของการขึ้นตรงต่อกันของข้อมูลในระดับที่ซับซ้อนกว่า

5.1 มัลติแวลูดีเพนเดนซี (Multivalued Dependency)

ถ้าแต่ละแอตทริบิวต์ในหนึ่งรีเลชัน แบ่งออกเป็นกลุ่มของข้อมูลอิสระ เช่นแอตทริบิวต์ X, Y และ Z แบ่งออกเป็นกลุ่มข้อมูลของ X, Y และ Z ที่เป็นอิสระต่อกัน มัลติแวลลูดีเพนเดนซี X –>> Y หมายถึงว่าค่า X หนึ่งค่าสามารถที่จะบอกค่า Y ได้หลาย ๆ (X Multi–Determinse Y) ไม่ว่า Z จะมีค่าเป็นอะไรก็ตาม โดยปกติ ถ้า R ประกอบด้วย Attribute X, Y และ Z (Z = R – {XY}) ดังนั้น ถ้า X –>> Y แล้ว X –>> Z เสมอ สามารถเขียนใหม่เป็น X –>> Y I Z ถ้า Y เป็นสับเซทของ X หรือ X ยูเนี่ยน Y = R แล้ว เราเรียก X –>> Y ว่า ทริเวียลมัลติแวลูดีเพนเดนซี (Trivial Multivalued Dependency) ซึ่งจะต่างจากฟังก์ชันนัลดีเพนเดนซี X –> Y ที่ X จะสามารถบอกค่า Y ได้แค่เพียงค่าเดียว ดังตัวอย่างภาพข้างล่าง เนื่องจากแอตทริบิวต์ รหัสโครงการ รหัสบริษัท และที่ตั้งโครงการล้วนเป็นคีย์แอตทริบิวต์ ดังนั้นรีเลชันในภาพ จึงถือว่าอยู่ในรูป BCNF แล้ว แต่ยังไม่อยู่ในรูปของ 4NF เนื่องจากรีเลชันดังกล่าวยังมีทริเวียลมัลติแวลูดีเพนเดนซีอยู่ในรีเลชัน ตัวอย่างเช่น รหัสโครงการA001 สามารถบอกค่าของรหัสบริษัทที่เป็นผู้รับผิดชอบได้มากกว่าหนึ่งบริษัท คือ รหัสบริษัท B001 และ B002 ในขณะเดียวกันรหัสโครงการ A001 ก็บอกถึงที่ตั้งของโครงการสองแห่งคือ จันทบุรี และระยอง ซึ่งถ้ามีการเพิ่มบริษัทที่รับผิดชอบโครงการเข้าไปในโครงการ A001 อีกหนึ่งบริษัทก็จะต้องมีการเพิ่มข้อมูลถึงสองทัปเพิลเนื่องจากโครงการดังกล่าวมีที่ตั้งอยู่ถึงสองแห่งคือระยอง และจันทบุรี ส่งผลให้เกิดความซ้ำซ้อนของข้อมูลขึ้นในรีเลชันดังกล่าว และอาจจะเกิดความผิดพลาดในการเพิ่มข้อมูลได้ เนื่องจากที่ตั้งโครงการไม่ได้ขึ้นอยู่กับรหัสบริษัทที่เป็นผู้รับผิดชอบแต่ขึ้นอยู่กับรหัสโครงการ ดังนั้น ถ้าหากมีการเพิ่มบริษัทผู้รับผิดชอบเพิ่มขึ้นอีกหนึ่งบริษัท เราจำเป็นที่จะต้องทำการเพิ่มข้อมูลที่ตั้งโครงการเข้าไปอีกสองแห่งด้วยเสมอ ซึ่งเป็นผลจากความสัมพันธ์ในรูปแบบของ ทริเวียลมัลติแวลูดีเพนเดนซี นั่นเอง

โครงการก่อสร้าง

รหัสโครงการ	รหัสบริษัท	ที่ตั้งโครงการ
A001	B001	จันทบุรี
A001	B001	ระยอง
A001	B002	จันทบุรี
A001	B002	ระยอง

ภาพที่ 2.26 รีเลชันที่อยู่ในรูปบอยซ์คอดด์นอร์มัลฟอร์มแล้ว แต่ยังไม่ได้อยู่ในรูปของนอร์มัลฟอร์มที่ 4

ที่มา : https://msit5.wordpress.com

5.2 นิยามของนอร์มัลฟอร์มที่ 4

รีเลชันจะอยู่ในรูปของนอร์มัลฟอร์มที่ 4 ก็ต่อเมื่อมีคุณสมบัติตามเงื่อนไขดังต่อไปนี้

1. รีเลชันนั้นเป็นบอยซ์คอดด์นอร์มัลฟอร์มอยู่แล้ว
2. รีเลชันนั้นไม่มีทริเวียลมัลติแวลูดีเพนเดนซี

จากรีเลชันในภาพข้างบน เราสามารถขจัดทริเวียลมัลติแวลูดีเพนเดนซี โดยการแตกรีเลชันดังกล่าวออกเป็นรีเลชันย่อย 2 รีเลชัน ซึ่งจะทำให้ทั้งสองรีเลชันอยู่ในรูปของนอร์มัลฟอร์มที่ 4 ดังภาพข้างล่าง

บริษัทในโครงการ

รหัสโครงการ	รหัสบริษัท
A001	B001
A001	B002

ที่ตั้งโครงการ

รหัสโครงการ	ที่ตั้งโครงการ
A001	จันทบุรี
A001	ระยอง

ภาพที่ 2.27 รีเลชันที่รับการนอร์มัลไลเซชันให้อยู่ในรูปของนอร์มัลฟอร์มที่ 4 แล้ว

ที่มา : https://msit5.wordpress.com

6. การแปลงให้อยู่ในรูปนอร์มัลฟอร์มที่ 5 (Fifth Normal Form : 5NF)

การแปลงให้อยู่ในรูปของนอร์มัลฟอร์มที่ 5 จะพิจารณาถึงการขึ้นต่อกันของข้อมูลในการแยกข้อมูลในรีเลชันออกเป็นรีเลชันย่อย และประกอบรีเลชันย่อยกลับเป็นรีเลชันใหญ่เช่นเดิม ซึ่งเป็นการตรวจสอบว่าเมื่อรวมกันใหม่ด้วยวิธีการจอยน์แล้ว จะได้รีเลชันกลับมาเหมือนเดิมทุกประการหรือไม่

6.1 จอยน์โอเปอรชัน (Join Operation)

ถ้ามี R1(X,Y) และ R2(Y,Z) R1 JOIN R2 = R3(X, Y, Z) โดยที่ t(x, y, z) อยู่ใน R3 ก็ต่อเมื่อมี t1(x,y) อยู่ใน R1 และ t2(y,z) อยู่ใน R2

6.2 จอยน์ดีเพนเดนซี (Join Dependency)

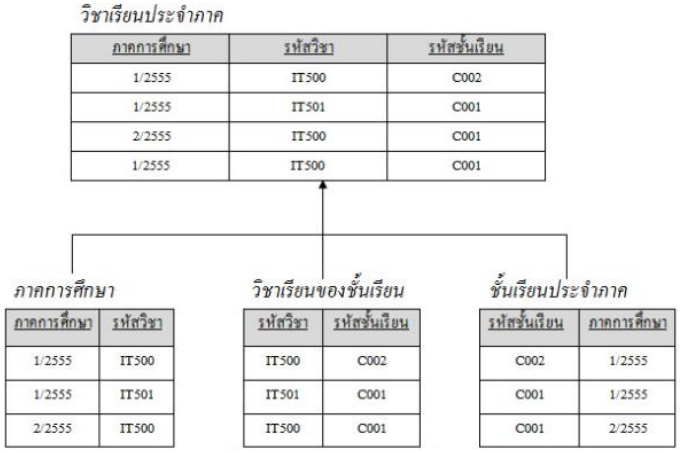
ในการแยกรีเลชันออกเป็นส่วนย่อย (Decomposition) R1, R2, R3, Rn มีคุณสมบัติจอยน์ดีเพนเดนซี ก็ต่อเมื่อ R1 JOIN R2 JOIN R3 ... JOIN Rn = R นั่นคือเมื่อเอารีเลชันย่อยมารวมกันก็ต้องได้รีเลชันเดิม ที่ไม่มีข้อมูลสูญหาย และไม่มีทัปเพิลที่เกินมา ที่เรียกว่า สพิวเรียสทัปเพิล (Spurious Tuple)

6.3 นิยามของ 5NF รีเลชันจะเป็น 5NF ถ้า

1. รีเลชันนั้นเป็นนอร์มัลฟอร์มที่ 4 อยู่แล้ว

2. การแบ่งแยกรีเลชันมีคุณสมบัติจอยน์ดีเพนเดนซี

จากตัวอย่างในภาพข้างล่าง รีเลชัน วิชาเรียนประจำภาคอยู่ในรูปของนอร์มัลฟอร์มที่ 4 แล้ว เนื่องจากแอตทริบิวต์ภาคการศึกษาเป็นตัวกำหนดแอตทริบิวต์รหัสวิชาหลาย ค่าในขณะที่แอตทริบิวต์รหัสวิชา ก็เป็นตัวกำหนดแอตทริบิวต์รหัสชั้นเรียนหลายค่า รีเลชันนี้จึงไม่มีทริเวียลมัลติแวลูดีเพนเดนซี ต่อไปเราจึงทำการทดสอบคุณสมบัตินอร์มัลฟอร์มที่ 5 ของรีเลชันวิชาเรียนประจำภาค โดยเมื่อนำรีเลชันดังกล่าวมาทำการแตกย่อยออกเป็นสามรีเลชันคือ รีเลชันภาคการศึกษา รีเลชันวิชาเรียนของชั้นเรียนและรีเลชันชั้นเรียนประจำภาค และทำการจอยน์ทั้งสามรีเลชันรวมกลับเป็นหนึ่งรีเลชันอีกครั้ง จะได้จำนวนข้อมูลเท่ากันกับรีเลชันก่อนที่จะมีแตกเป็นรีเลชันย่อยทุกประการ ซึ่งก็คือรีเลชันดังกล่าวมีคุณสมบัติจอยน์ดีเพนเดนซีและอยู่ในรูปของนอร์มัลฟอร์มที่ 5 แล้ว



ภาพที่ 2.28 แสดงถึงคุณสมบัติของรีเลชันที่อยู่ในรูปของนอร์มัลฟอร์มที่ 5 แล้ว

ที่มา : https://msit5.wordpress.com

2.3 เครื่องมือในการออกแบบ และวิเคราะห์ข้อมูล

2.3.1 แบบจำลองการพยากรณ์ (Forecasting)

2.3.1.1 การพยากรณ์แบบอนุกรมเวลา (Time Series Forecasting)

1. การพยากรณ์สำหรับค่าเฉลี่ย

เป็นการพยากรณ์โดยคำนวณจากการนำข้อมูลยอดขายจริงมาหาค่าเฉลี่ย จะทำให้เมื่อนำผลการพยากรณ์มาพลอตกราฟ จะได้เส้นกราฟที่มีลักษณะเป็นเส้นเรียบขึ้น เทคนิคการพยากรณ์สำหรับค่าเฉลี่ยมีหลายวิธี ดังต่อไปนี้

1.1 การพยากรณ์อย่างง่าย

วิธีการพยากรณ์นี้เป็นวิธีที่ง่ายที่สุดเมื่อต้องการพยากรณ์ความต้องการสินค้าในช่วงเวลาใด ๆ ก็ตาม ให้นำข้อมูลการขายจริงของช่วงเวลาก่อนหน้ามาใช้เป็นค่าพยากรณ์ ดังแสดงได้โดยสมการ ดังนี้

$$F_t = A_{t-1}$$

เมื่อกาหนด F_t = ค่าพยากรณ์ความต้องการสินค้า ณ ช่วงเวลา t

A_{t-1} = ข้อมูลยอดขายสินค้าจริง ณ ช่วงเวลา t–1

1.2 วิธีค่าเฉลี่ยเคลื่อนที่

การพยากรณ์โดยวิธีค่าเฉลี่ยเคลื่อนที่ เขียนย่อ ๆ ว่า MA ทำได้โดยนำข้อมูลยอดขายจริงในอดีตจำนวน n ข้อมูลมาคำนวณหาค่าเฉลี่ยเพื่อพยากรณ์ความต้องการสินค้าในช่วงเวลาที่ต้องการ โดยมีการกาหนดสมการที่ใช้คำนวณและตัวแปรดังนี้

$$MA_n = \frac{\Sigma A_t}{n}$$

เมื่อกำหนด t = ช่วงเวลา

n = จำนวนข้อมูลที่จะคำนวณค่าเฉลี่ยเคลื่อนที่

A_t = ข้อมูลจริง ณ ช่วงเวลา t

MA_n = ค่าพยากรณ์วิธีค่าเฉลี่ยนเคลื่อนที่ n ช่วงเวลา

1.3 วิธีค่าเฉลี่ยเคลื่อนที่ถ่วงน้ำหนัก (Weighted Moving Average)

เป็นวิธีที่เหมาะสำหรับข้อมูลยอดขายในอดีตที่มีแนวโน้ม โดยจะให้ความสำคัญของข้อมูลไม่เท่ากันคิดเป็นค่าถ่วงน้ำหนัก ซึ่งมีค่าอยู่ระหว่าง 0 ถึง 1 ผู้พยากรณ์เป็นผู้กำหนดค่าถ่วงน้ำหนักในการพยากรณ์โดยจะให้ค่าน้ำหนักมากที่สุดแก่ข้อมูลยอดขายที่ใหม่ที่สุดและลดลงเรื่อย ๆ ตามลำดับ ผลรวมของค่าถ่วงน้ำหนักนิยมให้เป็น 1.0 แต่จะเป็นอย่างอื่นก็ได้ วิธีการคำนวณให้ใช้สมการ ต่อไปนี้

$$WMA_n = \frac{\Sigma W_t A_t}{\Sigma W_t}$$

เมื่อกำหนด t = ช่วงเวลา

n = จำนวนข้อมูลที่จะคำนวณค่าเฉลี่ยเคลื่อนที่

W_t = ค่าถ่วงน้ำหนัก ณ ช่วงเวลา t

A_t = ข้อมูลจริง ณ ช่วงเวลา t

WMA_n = ค่าพยากรณ์ค่าเฉลี่ยเคลื่อนที่ถ่วงน้ำหนัก n ช่วงเวลา

1.4 วิธีเอกซ์โปแนนเชียลปรับเรียบ (Exponential Smoothing)

เป็นวิธีที่ใช้หลักการเดียวกันกับการพยากรณ์แบบค่าเฉลี่ยเคลื่อนที่ถ่วงน้ำหนัก นั่นคือให้ความสำคัญกับข้อมูลชุดใหม่ที่สุดมากที่สุด (ค่าถ่วงน้ำหนักสูงสุด) และค่อย ๆ ลดค่าถ่วงน้ำหนักลง วิธีเอกซ์โปแนนเชียลปรับเรียบจะทำการพยากรณ์โดยนำค่าพยากรณ์ของช่วงเวลาที่ผ่านมาบวกเข้ากับอัตราส่วนความแตกต่างระหว่างข้อมูลจริงกับค่าพยากรณ์ ณ ช่วงเวลานั้น ๆ ดังแสดงโดยสมการต่อไปนี้

$$F_t = F_{t-1} + \alpha(A_{t-1} - F_{t-1})$$

หรือ

$$F_t = F_{t-1} + \alpha(e_{t-1})$$

เมื่อกำหนด F_t = ค่าพยากรณ์ความต้องการสินค้า ณ ช่วงเวลา t

F_{t-1} = ค่าพยากรณ์ความต้องการสินค้า ณ ช่วงเวลา t–1

α = ค่าคงที่ปรับเรียบระหว่าง 0ถึง 1

A_{t-1} = ข้อมูลจริง ณ ช่วงเวลา t–1

e_{t-1} = ค่าผิดพลาด ณ ช่วงเวลา t–1

แปลความหมายจากสมการได้ว่า ค่าพยากรณ์ ณ ช่วงเวลา t ใด ๆ คำนวณได้จากค่าพยากรณ์ของช่วงเวลาก่อนหน้าหนึ่งช่วงเวลา (t–1) บวกกับอัตราส่วนหรือเปอร์เซ็นต์ของค่าผิดพลาด (Errore) ที่เกิดขึ้น ณ ช่วงเวลาก่อนหน้าหนึ่งช่วงเวลา

2. การพยากรณ์สำหรับแนวโน้ม

อิทธิพลของแนวโน้มที่มีผลต่อความต้องการสินค้าของลูกค้านั้นมักมีผลมาจากปัจจัยระยะยาวใด ๆ เช่น การเพิ่มจำนวนประชากรของประเทศ การเพิ่มรายได้ต่อครอบครัว เป็นต้น ดังนั้นในการพยากรณ์ความต้องการสินค้าเมื่อมองเห็นแนวโน้มในยอดขาย จึงจำเป็นต้องหาสมการทางคณิตศาสตร์ที่เหมาะสมที่สุด ลักษณะของแนวโน้มสามารถเป็นได้ทั้งเชิงเส้นตรงหรือแบบอื่น ๆ แนวโน้มที่ไม่ใช่เชิงเส้นนั้นอาจพบในลักษณะเส้นโค้งพาราโบลิกคล้ายถ้วยหงายหรือถ้วยคว่ำหรือมีลักษณะเป็นเส้นโค้งเอกซ์โปแนนเชียลในลักษณะเพิ่มขึ้นหรือลดลง

2.1 วิธีการสมการเชิงเส้น

วิธีการสมการเชิงเส้น (Linear Trend Equation) รูปแบบสมการเชิงเส้นโดยทั่ว ๆ ไปจะอยู่ในรูปสมการ y = a + bx หรือ y = mx + c โดยที่ x เป็นตัวแปรอิสระ (Independent Variable) และ y เป็นตัวแปรตาม (Dependent Variable) ส่วนค่าคงที่ a คือจุดตัดแกน y (Y – Intercept) และ b คือค่าความชันของเส้นตรง (Slope) เมื่อนำมาประยุกต์ใช้ในการพยากรณ์แบบอนุกรมเวลา จึงกำหนดให้ t เป็นตัวแปรอิสระ (แทนค่า x) และ Ft เป็นตัวแปรตามหรือค่าพยากรณ์ความต้องการสินค้า (แทนค่า y) ณ ช่วงเวลา t ใด ๆ ดังนั้นสมการที่ใช้ในการพยากรณ์ความต้องการสินค้า เมื่อมีอิทธิพลของแนวโน้มที่เป็นลักษณะเส้นตรง ได้แก่

$$F_t = a + b_t$$

เมื่อกำหนด F_t = ค่าพยากรณ์ความต้องการสินค้า ณ ช่วงเวลา t ใด ๆ

t = ช่วงเวลา มีค่าตั้งแต่1, 2...

a = จุดตัดแกน y (Y – Intercept) หรือ ค่า F, เมื่อ t = 0

b = ค่าความชันของเส้นตรง

ค่า a และ b สามารถคำนวณได้จากสูตรต่อไปนี้

$$b = \frac{n\Sigma tA_t - \sum t\Sigma A_t}{n\Sigma t^2 - (\Sigma t)^2}$$

$$a = \frac{\Sigma A_t - b\Sigma_t}{n}$$

เมื่อกำหนด F_t = ค่าพยากรณ์ความต้องการสินค้า ณ ช่วงเวลา t ใด ๆ

t = ช่วงเวลา มีค่าตั้งแต่ 1, 2...

a = จุดตัดแกน y *(Y – Intercept)* หรือ ค่า Ft เมื่อ t = 0

b = ค่าความชันของเส้นตรง

n = จำนวณข้อมูล

A_t = ข้อมูลยอดขายจริง ณ ช่วงเวลา t ใด ๆ

2.2 วิธีเอกซ์โปเนนเชียลปรับเรียบแบบมีแนวโน้ม (Exponential Smoothing With Trend*)*

เรียกได้อีกชื่อหนึ่งว่าเอกซ์โปแนนเชียลปรับเรียบสองชั้น (Double Smoothing*)* จะใช้ในกรณีที่ข้อมูลยอดขายมีแนวโน้มเพิ่มขึ้นหรือลดลง ซึ่งจะต่างจากวิธีเอกซ์โปแนนเชียลปรับเรียบที่เหมาะสมกับข้อมูลยอดขายที่ค่อนข้างจะคงที่ สมการที่ใช้ในการพยากรณ์มีรูปแบบดังนี้

$$FIT_t = F_t + T_t$$
$$F_t = \alpha(A_{t-1}) + (1-\alpha)F_{t-1} + T_{t-1}$$
$$T_t = \beta(F_t - F_{t-1}) + (1-\beta)T_{t-1}$$

เมื่อกำหนด FIT_t = ค่าพยากรณ์ความต้องการสินค้าเมื่อมีแนวโน้ม ณ ช่วงเวลา t

F_t = ค่าความต้องการสินค้าเอกซ์โปแนนเชียลปรับเรียบ ณ ช่วงเวลา t

T_t = ค่าแนวโน้มเอกซ์โปเนนเชียลปรับเรียบ ณ ช่วงเวลา t

A_t = ข้อมูลจริง ณ ช่วงเวลา t

α = ค่าคงที่ปรับเรียบสำหรับค่าเฉลี่ย

β = ค่าคงที่ปรับเรียบสำหรับแนวโน้ม

2.3 วิธีสมการพาราโบลิก

วิธีสมการพาราโบลิก ในกรณีที่เมื่อนำค่าข้อมูลจริงของยอดขายมาพลอตกราฟเทียบกับเวลาที่ผ่านไปแล้วพบว่ามีลักษณะเป็นเส้นโค้งพาราโบลิก คล้ายถ้วยค่าหรือหงาย ผู้พยากรณ์สามารถใช้สมการพาราโบลิกมาเป็นเครื่องมือช่วยในการพยากรณ์ได้ สมการมีรูปทั่วไป คือ y = a + bx + cx2 เมื่อจะใช้พยากรณ์จึงปรับรูปสมการเป็นดังนี้

$$F_t = a + bt + ct^2$$

คำนวณค่า a, b และ c ได้จากสมการต่อไปนี้

$$b = \frac{gh - ij}{gk - j^2}$$

$$c = \frac{i - bj}{g}$$

$$a = \frac{\Sigma A_t - b\sum t - c\Sigma t^2}{n}$$

2.4 วิธีสมการเอกซ์โปแนนเชียล

วิธีสมการเอกซ์โปแนนเชียล รูปแบบสมการที่ใช้ในการพยากรณ์ความต้องการสินค้าเมื่อมีอิทธิพลของแนวโน้มที่เป็นลักษณะเส้นโค้งที่เพิ่มขึ้นเรื่อย ๆ หรือลดลงเรื่อย ๆ โดยมีรูปแบบสมการ คือ

$$F_t = ae^{bt}$$

เมื่อกำหนด F_t = ค่าพยากรณ์ความต้องการสินค้า ณ ช่วงเวลา t ใด ๆ

T = ช่วงเวลา มีค่าตั้งแต่1, 2...

a, b = ค่าคงที่

ค่า a และ b สามารถคำนวณได้จากสูตรต่อไปนี้

$$F_t = ae^{br}$$

$$In\ F_t = In\ a + bt$$

$$b = \frac{n\sum t.InA_t - \sum t \sum In\ A_t}{n\Sigma t^2 - (\Sigma t)^2}$$

$$In\ a = \frac{\sum In\ A_t - b\sum t}{n}$$

เมื่อกำหนด F_t = ค่าพยากรณ์ความต้องการสินค้า ณ ช่วงเวลา t ใด ๆ

t = ช่วงเวลา มีค่าตั้งแต่1, 2...

a, b = ค่าคงที่

n = จำนวนข้อมูล

A_t = ข้อมูลยอดขายจริง ณ ช่วงเวลา t ใด ๆ

2.5 การวัดความผิดพลาดของการพยากรณ์

ความแม่นยำและการควบคุมค่าพยากรณ์ เป็นสิ่งสำคัญในการพยากรณ์ ถ้าบริษัทสามารถพยากรณ์ได้แม่นยำมากจะทำให้สามารถวางแผนการผลิตและส่งมอบสินค้าได้ทันตามกำหนด ในสภาพการณ์ปัจจุบันมีปัจจัยหลายอย่างที่อาจทำให้การพยากรณ์ทำได้ยากขึ้น ดังนั้นผู้พยากรณ์จะต้องมีความสามารถที่จะเลือกวิธีการพยากรณ์ที่เหมาะสมที่ให้ผลลัพธ์ที่มีความแม่นยำสูงที่สุด

ค่าผิดพลาด (Error) ของการพยากรณ์ คือ ค่าความแตกต่างระหว่างข้อมูลยอดขายจริงกับค่าพยากรณ์ความต้องการสินค้า ณ ช่วงเวลานั้น ๆ โดยสามารถคำนวณค่าผิดพลาดได้จากสมการต่อไปนี้

$$e_t = |A_t - F_t|$$

เมื่อกำหนด F_t = ค่าพยากรณ์ความต้องการสินค้า ณ ช่วงเวลา t

A_t = ข้อมูลยอดขายสินค้าจริง ณ ช่วงเวลา t

e_t = ค่าความผิดพลาด ณ ช่วงเวลา t

ความแม่นยำของการพยากรณ์ (Forecasting Accuracy) คือ การบอกค่าความแม่นยำในการพยากรณ์ โดยสามารถทำได้สองลักษณะ คือ

–ค่าเฉลี่ยความเบี่ยงเบนสัมบูรณ์ (Mean Absolute Deviation : MAD) สามารถคำนวณได้จาก

$$\text{MAD} = \frac{\sum |actual - forecast|}{n} = \frac{\sum |A_t - F_t|}{n} = \frac{\sum |e_t|}{n}$$

–ค่าเฉลี่ยความผิดพลาดกำลังสอง (Mean Square Error : MSE) สามารถคำนวณได้จาก

$$\text{MSE} = \frac{\sum (actual - forecast)^2}{n} = \frac{\sum (A_t - F_t)^2}{n} = \frac{\sum (e_t)^2}{n}$$

โดยที่ ค่าเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่างจะเท่ากับ$\sqrt{MSE}$

–เปอร์เซ็นต์ความเบี่ยงเบนเฉลี่ยสมบูรณ์ (Mean Absolute Percent Deviation MAPD)

$$\text{MAPE} = \frac{100}{n} \sum_{i=1}^{N} \left[\left| \frac{A_t - F_t}{A_t} \right| \right]$$

วิธีนี้กำจัดปัญหาด้านการแปลความหมายผิดพลาด เพราะค่าที่ได้จากสมการเป็นค่าเฉลี่ยของความแตกต่างระหว่างความต้องการจริง และค่าพยากรณ์ที่สัมพันธ์กับขนาดของความต้องการจริงในรูปของเปอร์เซ็นต์ โดยที่ค่าต่ำกว่าบ่งชี้ความเหมาะสม (ถูกต้อง) ของการพยากรณ์ที่มากกว่า

2.3.2 การแสดงผลแบบ (Visualization)

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

2.3.2.1 ประเภทของ Visualization

1. แผนภูมิแท่ง (Bar Charts) เป็นแผนภูมิที่ประกอบด้วยแกนนอน แกนตั้ง ที่นิยมแสดงออกมาในรูปแท่งสี่เหลี่ยมที่สามารถบอกความสูงได้ เหมาะสำหรับใช้การเปรียบเทียบจำนวนของข้อมูลในแต่ละชุด เช่น รายรับในแต่ละเดือน, ยอดขายที่ขายได้จริงเปรียบเทียบกับเป้าหมายยอดขายที่ตั้งไว้ เป็นต้น ซึ่งแผนภูมิแท่งยังสามารถแบ่งออกเป็นหลายประเภท

1.1 แผนภูมิแท่งแบบจัดกลุ่ม : แผนภูมิแท่งแบบจัดกลุ่ม เหมาะสมกับการนำเสนอข้อมูลที่มีข้อมูลย่อย ๆ อยู่ภายใต้ข้อมูลใหญ่ เป็นการเน้นให้เห็นข้อมูลย่อยนั้น ๆ

1.2 แผนภูมิแท่งแบบวางซ้อนกัน : แผนภูมิแท่งแบบวางซ้อนกัน เหมาะสมกับการนำเสนอข้อมูลที่ต้องการให้เห็นข้อมูลย่อยในแต่ละข้อมูลใหญ่และยังแสดงให้เห็นสัดส่วนของข้อมูลย่อยต่าง ๆ เหล่านั้นได้ด้วย ใช้แผนภูมินี้เมื่อมีชุดข้อมูล หลายชุดและต้องการเน้นผลรวมทั้งหมด

2. แผนภูมิเส้น (Line Charts) มีลักษณะคล้ายแผนภูมิแท่ง ซึ่งประกอบด้วยแกนตั้งและนอน เพียงแต่เปลี่ยนจากแท่งข้อมูลเป็นจุดบนแผนภูมินั่นเอง แผนภูมิประเภทนี้เหมาะกับการนำเสนอข้อมูลตัวเลขที่มีความสัมพันธ์กับข้อมูลที่มีลักษณะเป็นช่วง ใช้แสดงให้เห็นการเปลี่ยนแปลงที่เกิดขึ้นตามเวลาเพื่อดูแนวโน้ม รวมถึงสามารถใช้พยากรณ์แนวโน้มในอนาคตได้ เช่น ข้อมูลของยอดขายในแต่ละปี หรือไตรมาส และนำมาวิเคราะห์เพื่อดูแนวโน้ม เป็นต้น

3. แผนภูมิวงกลม (Pie Charts) เหมาะกับการนำเสนอข้อมูลที่มีส่วนประกอบย่อยที่รวมกันเป็นส่วนใหญ่ มีการแบ่งส่วนให้ดูง่าย และสวยงาม แต่ในทางกลับกันอาจจะดูยากในเรื่องของการประมาณขนาดของแต่ละชิ้น ยิ่งถ้ามีจำนวนชิ้นมาก จะยิ่งแยกยาก เพราะต้องใช้หลายสี ในการนำเสนอข้อมูล เช่น ส่วนแบ่งทางการตลาด (Market Share), ข้อมูลแสดงส่วนผสมต่าง ๆ เป็นต้น

4. แผนภูมิโดนัท (Doughnut Charts) มีหลักการออกแบบเช่นเดียวกับแผนภูมิวงกลม แต่สามารถแสดงชุดข้อมูลได้มากกว่า 1 ชุด โดยนำเสนอข้อมูลเป็นวงกลมซ้อนกันหลายๆ ชั้น นั่นเอง

5. แผนภูมิพื้นที่ (Area Charts) มีหน้าตาคล้ายแผนภูมิเส้น แต่มีการแรเงาพื้นที่ใต้เส้นข้อมูล หรือระหว่าง 2 เส้น เพื่อแสดงให้เห็นปริมาณความแตกต่างระหว่างเส้น เหมาะสำหรับเน้นความสำคัญของการเปลี่ยนแปลงในช่วงเวลา แสดงให้เห็นผลรวมของความแตกต่างระหว่างข้อมูล เช่น ข้อมูลของการซื้อสินค้าในห้างสรรพสินค้าในแต่เดือน ตามหมวดหมู่ต่าง ๆ ไล่ไปเครื่องสำอาง เสื้อผ้าแฟชั่น อาหาร ตามลำดับ

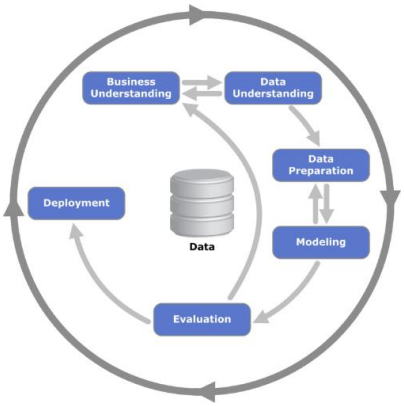
6. แผนภูมิเรดาร์ (Radar Charts) มีลักษณะคล้ายแผนภูมิเส้นที่มีการแสดงผลแบบวงกลม จำนวนเหลี่ยมของเรดาร์เท่ากับจำนวนหัวข้อของข้อมูล แผนภูมินี้ไม่ได้บอกถึงความต่อเนื่องของข้อมูล แต่เหมาะสำหรับการนำเสนอข้อมูลเป็นหัวข้อ แล้วนำมาวิเคราะห์หาจุดอ่อนจุดแข็งของข้อมูล เช่น นำเสนอข้อมูลเกี่ยวกับคุณสมบัติของการรับพนักงานใหม่ เพื่อดูจุดอ่อนจุดแข็งของแต่ละคน เป็นต้น

7. แผนภูมิต้นไม้ (Tree Maps) คือการนำเสนอข้อมูลแบบแสดงให้เห็นพื้นที่แสดงผลได้ในแบบลำดับชั้น เหมือนแบบโครงสร้างต้นไม้ อาจจะนำเสนอข้อมูลที่ต้องการให้เห็นถึงเขตพื้นที่ แสดงพื้นที่สีที่แตกต่างกันได้

8. แผนภูมิรูปภาพ (Picture Graph) ประกอบไปด้วยแกนนอน และแกนตั้ง แต่เลือกใช้รูปภาพ หรือไอคอนแทนจำนวนของสิ่งของนั้น ๆ เช่น การแสดงผลจำนวนของนมที่ขายได้ในแต่ละเดือน โดยนำเสนอทั้งนมรสจืด รสช็อกโกแลต เปรียบเทียบในแต่ละเดือน ซึ่งมีการนำเสนอลักษณะคล้ายกับกราฟแท่ง แต่เปลี่ยนจากแท่งเป็นรูปกราฟิกของนม 2 รสชาติแทน ก็ทำให้การนำเสนอข้อมูลน่าสนใจมากยิ่งขึ้น ซึ่งแนวทางการนำเสนอข้อมูลลักษณะนี้ต้องอาศัยความคุ้นชินของคนดู เพื่อแทนสัญลักษณ์ภาพลงไป เช่น เมื่อพูดถึงจำนวนคน อาจจะแทนด้วยภาพไอคอนคน หรือเมื่อพูดถึงจำนวนเงิน ควรแทนภาพเป็นเหรียญเงิน หรือแบงค์แทน ก็จะทำให้คนดูเข้าใจง่ายจากสัญลักษณ์ภาพที่คุ้นเคยอยู่แล้วและยังดึงดูดความสนใจได้มากกว่าการใช้กราฟแท่งสี่เหลี่ยมอีกด้วย

2.3.3 กระบวนการวิเคราะห์ข้อมูลด้วย (CRISP-DM)

กระบวนการ CRISP-DM นี้จะประกอบด้วย 6 ขั้นตอน แต่ละขั้นตอนจะเป็นขั้นตอนที่ต่อเนื่องกันนั่นคือขั้นตอนถัดไปจะรอผลลัพธ์จากขั้นตอนก่อนหน้าซึ่งแสดงด้วยลูกศรที่เชื่อมระหว่างกล่องสี่เหลี่ยมแต่ละกล่อง ตัวอย่างเช่นเมื่อได้ผลลัพธ์จากขั้นตอนการเตรียมข้อมูล (Data Preparation) แล้วจะนำไปสร้างโมเดลจำแนกประเภทข้อมูลในขั้น Modeling และหลังจากนั้นอาจจะย้อนกลับมาเปลี่ยนแปลงข้อมูลให้ถูกต้องมากขึ้นเพื่อหวังว่าจะโมเดลที่ให้ความถูกต้องมากขึ้น เป็นต้น



ภาพที่ 2.29 กระบวนการ CRISP-DM

ที่มา : https://datacubeth.ai/crisp-dm/

2.3.3.1 ขั้นตอนของกระบวนการ CRISP-DM

1. การทำความเข้าใจธุรกิจ (Business Understanding)

ขั้นตอนแรกมุ่งไปที่การทำความเข้าใจธุรกิจ ปัญหาและวัตถุประสงค์ของโครงการจากมุมมองทางธุรกิจ จากนั้นแปลงปัญหาให้อยู่ในรูปของโจทย์สำหรับการวิเคราะห์ข้อมูล และวางแผนการดำเนินงานเบื้องต้น

2. การทำความเข้าใจข้อมูล (Data Understanding)

ขั้นตอนนี้เริ่มต้นด้วยการรวบรวมข้อมูล จากนั้นทำความเข้าใจ ตรวจสอบคุณภาพ และเลือกข้อมูลที่เก็บรวบรวมมาว่าจะใช้ข้อมูลใดบ้างในการวิเคราะห์และพิจารณาว่าจะใช้ข้อมูลทั้งหมดหรือจำเป็นต้องเลือกข้อมูลบางส่วนมาใช้ในการวิเคราะห์

3. การเตรียมข้อมูล (Data Preparation)

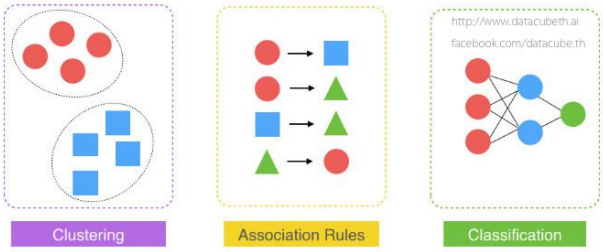
ขั้นตอนนี้เป็นขั้นตอนที่ทำการแปลงข้อมูลที่ได้ทำการเก็บรวบรวมมา (raw data) ให้กลายเป็นข้อมูลที่สามารถนำไปวิเคราะห์ในขั้นถัดไปได้ โดยการแปลงข้อมูลนี้อาจจะต้องมีการทำข้อมูลให้ถูกต้อง (data cleaning) เช่น การแปลงข้อมูลให้อยู่ในช่วง (scale) เดียวกัน หรือการเติมข้อมูลที่ขาดหายไป เป็นต้น โดยขั้นตอนนี้จะเป็นขั้นตอนที่ใช้เวลามากที่สุดของกระบวนการ CRISP-DM

4. การสร้างโมเดล (Modeling)

ขั้นตอนนี้จะเป็นขั้นตอนการวิเคราะห์ข้อมูลด้วยเทคนิคทางดาต้า ไมน์นิง ที่ได้แนะนำไปแล้ว เช่น การจำแนกประเภทข้อมูล หรือ การแบ่งกลุ่มข้อมูล ซึ่งในขั้นตอนนี้หลายเทคนิคจะถูกนำมาใช้เพื่อให้ได้คำตอบที่ดีที่สุด ดังนั้นในบางครั้งอาจจะต้องมีการย้อนกลับไปที่ขั้นตอนที่ 3 Data Preparation เพื่อแปลงข้อมูลบางส่วนให้เหมาะสมกับแต่ละเทคนิคด้วย ตัวอย่างเทคนิคในการวิเคราะห์ข้อมูลต่าง ๆ เช่น

1. การแบ่งกลุ่มข้อมูล (Clustering)
2. การหากฎความสัมพันธ์ (Association Rules)
3. การจำแนกประเภทข้อมูล (Classification) ตัวอย่างเช่น
   - 3.1 เทคนิค Decision Tree
   - 3.2 เทคนิค Naive Bayes
   - 3.3 เทคนิค Neural Network

3.4 เทคนิค Support Vector Machines (SVM)



ภาพที่ 2.30 การสร้างโมเดล

ที่มา : https://datacubeth.ai/crisp-dm/

5. การวัดประสิทธิภาพของโมเดล (Evaluation)

การวัดประสิทธิภาพของโมเดลที่ได้จากขั้นตอนที่ 4 เพื่อวัดว่าโมเดลมีประสิทธิภาพเพียงพอต่อการนำไปใช้งานแล้วหรือไม่ ซึ่งโมเดลแต่ละประเภทมีการวัดประสิทธิภาพของผลลัพธ์ที่ได้ว่าตรงกับวัตถุประสงค์ที่ได้ตั้งไว้ในขั้นตอนแรก หรือ มีความน่าเชื่อถือมากน้อยเพียงใด ซึ่งอาจจะย้อนกลับไปยังขั้นตอนก่อนหน้าเพื่อเปลี่ยนแปลงแก้ไขเพื่อให้ได้ผลลัพธ์ตามที่ต้องการได้ สำหรับการสร้างโมเดลด้วยเทคนิค Classification มีการทดสอบประสิทธิภาพของโมเดลอยู่ 3 แบบใหญ่ คือ

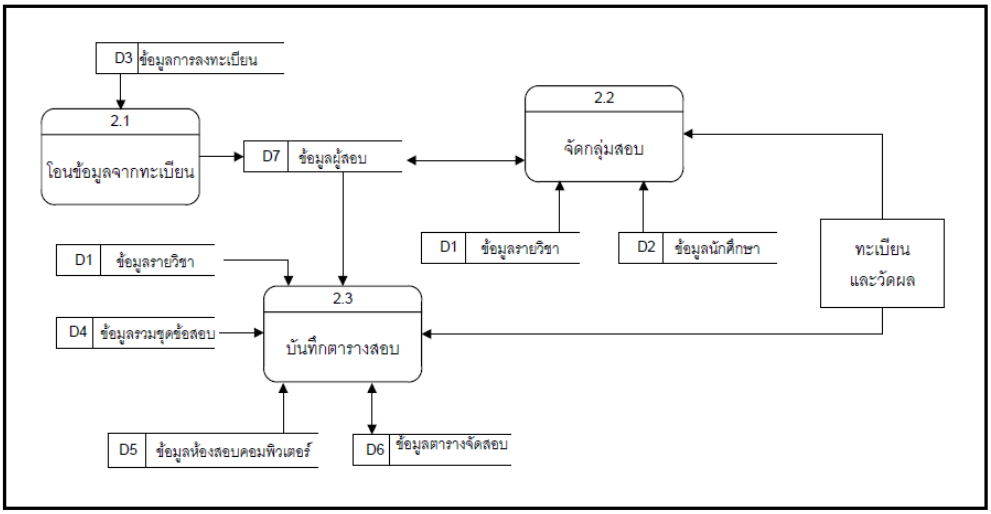
1. Self-consitency test
2. Split test
3. Cross-validation test

6. การนำโมเดลไปใช้งานจริง (Deployment)

เป็นการนำโมเดลที่เหมาะสมที่สุดไปใช้งานจริง เพื่อวิเคราะห์และแก้ปัญหาที่ต้องการ ในกระบวนการทำงานของ CRISP-DM นั้นไม่ได้หยุดเพียงแค่ผลลัพธ์ที่ได้จากการวิเคราะห์ข้อมูลด้วยเทคนิคทางดาต้า ไมน์นิงเท่านั้น แม้ว่าผลลัพธ์ที่ได้จะแสดงถึงองค์ความรู้ที่มีประโยชน์ แต่จะต้องนำองค์ความรู้ที่ได้เหล่านี้ไปใช้ได้จริงในองค์กรหรือบริษัท

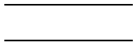
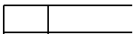
2.3.4 แผนภาพกระแสข้อมูล (Data Flow Diagram: DFD)

แผนภาพกระแสข้อมูล (Data Flow Diagram : DFD) หมายถึง แผนภาพที่แสดงให้เห็นถึงทิศทางการไหลของข้อมูลที่มีอยู่ในระบบ และการดำเนินงานที่เกิดขึ้นในระบบ โดยข้อมูลในแผนภาพทำให้ทราบถึง ข้อมูลมาจากไหน, ข้อมูลไปที่ไหน, ข้อมูลเก็บที่ใด, เกิดเหตุการณ์ใดกับข้อมูลในระหว่างทาง แผนภาพกระแสข้อมูลจะแสดงภาพรวมของระบบ (Overall picture of a system) และรายละเอียดบางอย่าง แต่ในบางครั้งหากต้องการกำหนดรายละเอียดที่สำคัญในระบบ นักวิเคราะห์ระบบอาจจำเป็นต้องใช้เครื่องมืออื่น ๆ ช่วย เช่น ข้อความสั้นๆที่เข้าใจ หรืออัลกอริทึม, ตารางการตัดสินใจ (Decision Table), Data Model, Process Description ทั้งนี้ก็ขึ้นอยู่กับความต้องการในรายละเอียด



ภาพที่ 2.31 ตัวอย่างแผนภาพกระแสข้อมูล

ที่มา : https://alaska.reru.ac.th

DeMarco & Yourdon	Gane & Sarson	ความหมาย
		Process : ขั้นตอนการทำงานภายในระบบ
		Data Store : แหล่งข้อมูลสามารถเป็นได้ทั้งไฟล์ข้อมูลและฐานข้อมูล (File or Database)
		External Agent : ปัจจัยหรือสภาพแวดล้อมที่มีผลกระทบต่อระบบ
		Data Flow : เส้นทางการไหลของข้อมูล แสดงทิศทางของข้อมูลจากขั้นตอนการทำงานหนึ่งไปยังอีกขั้นตอนหนึ่ง

ภาพที่ 2.32 สัญลักษณ์ที่ใช้ในแผนภาพกระแสข้อมูล

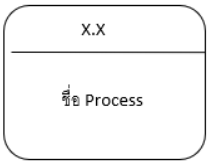
ที่มา : https://alaska.reru.ac.th

2.3.4.1 แนวคิดของแบบจำลองขั้นตอนการทำงานของระบบ

การสร้างแบบจำลองขั้นตอนการทำงานของระบบโดยใช้แผนภาพกระแสข้อมูล (Data Flow Diagram) มีแนวคิดต่าง ๆ ดังนี้ 1. ขั้นตอนการทำงานของระบบ (Process), 2. เส้นทางการไหลของข้อมูล (Data Flow), 3. ตัวแทนข้อมูล (External Agent), 4. แหล่งจัดเก็บข้อมูล (Data Store)

2.3.4.2 ขั้นตอนการทำงานของระบบ

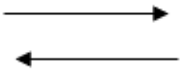
1. Process หรือ ขั้นตอนการดำเนินงาน คือ งานที่ดำเนินการ/ตอบสนองข้อมูลที่รับเข้า หรือดำเนินการ/ตอบสนองต่อเงื่อนไข/สภาวะใด ๆ ที่เกิดขึ้น ไม่ว่าขั้นตอนการดำเนินงานนั้นจะกระทำโดยบุคคล หน่วยงาน หุ่นยนต์ เครื่องจักร หรือ เครื่องคอมพิวเตอร์ก็ตาม โดยจะเป็นกริยา (Verb) เช่น ลงทะเบียน เพิกถอนวิชา เพิ่มวิชา พิมพ์รายงาน เป็นต้น



ภาพที่ 2.33 สัญลักษณ์แสดง Process

ที่มา : https://alaska.reru.ac.th

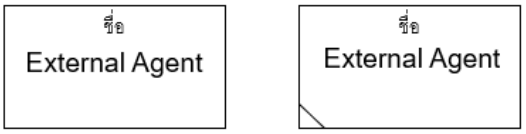
2. เส้นทางการไหลของข้อมูล (Data Flows) เป็นการสื่อสารระหว่างขั้นตอนการทำงาน (Process) ต่าง ๆ และสภาพแวดล้อมภายนอกหรือภายในระบบ โดยแสดงถึงข้อมูลที่นำเข้าไปในแต่ละ Process และข้อมูลที่ส่งออกจาก Process ใช้ในการแสดงถึงการบันทึกข้อมูล การลบข้อมูล การแก้ไขข้อมูลต่าง ๆ ในไฟล์หรือในฐานข้อมูล ซึ่งใน Data Flow Diagram เรียกว่า “Data Store” สัญลักษณ์ที่ใช้อธิบายเส้นทางการไหลของข้อมูลคือ เส้นตรงที่ประกอบด้วยหัวลูกศรตรงปลายเพื่อบอกทิศทางการเดินทางหรือการไหลของข้อมูล ดังรูป



ภาพที่ 2.34 สัญลักษณ์ของ Data Flow

ที่มา : https://alaska.reru.ac.th

3. ตัวแทนข้อมูล (External Agents) หมายถึง บุคคล หน่วยงานในองค์กร องค์กรอื่น ๆ หรือระบบงานอื่น ๆ ที่อยู่ภายนอกขอบเขตของระบบ แต่มีความสัมพันธ์กับระบบ โดยมีการส่งข้อมูลเข้าสู่ระบบเพื่อดำเนินงาน และรับข้อมูลที่ผ่านการดำเนินงานเรียบร้อยแล้วจากระบบ ในบางครั้งเรียกว่า “External Entity” สัญลักษณ์ที่ใช้อธิบาย คือ สี่เหลี่ยมจัตุรัส หรือสี่เหลี่ยมผืนผ้า ภายในจะต้องแสดงชื่อของ External Agent โดยสามารถทำการซ้ำ (Duplicate) ได้ด้วยการใช้เครื่องหมาย \ (back slash) ตรงมุมล่างซ้าย

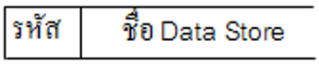


ภาพที่ 2.35 สัญลักษณ์ของ External Agents

ที่มา : https://alaska.reru.ac.th

4. แหล่งจัดเก็บข้อมูล (Data Store) เป็นแหล่งเก็บ/บันทึกข้อมูล เปรียบเสมือนคลังข้อมูล (เทียบเท่ากับไฟล์ข้อมูล และฐานข้อมูล) โดยอธิบายรายละเอียดและคุณสมบัติเฉพาะตัวของสิ่งที่ต้องการเก็บ/บันทึก สัญลักษณ์ที่ใช้อธิบายคือสี่เหลี่ยมเปิดหนึ่งข้าง แบ่งออกเป็นสองส่วน ได้แก่ ส่วนที่ 1 ทางด้านซ้ายใช้แสดงรหัสของ Data Store อาจจะเป็นหมายเลขลำดับหรือ

ตัวอักษรได้เช่น D1, D2 เป็นต้น สำหรับส่วนที่ 2 ทางด้านขวา ใช้แสดงชื่อ Data Store หรือชื่อไฟล์ เช่น Employee, Application, Member เป็นต้น ดังรูป



ภาพที่ 2.36 สัญลักษณ์ของ Data Store

ที่มา : https://alaska.reru.ac.th

2.4 วรรณกรรมที่เกี่ยวข้อง

รวิพิมพ์ ฉวีสุข และ ธนีกานต์ จุฑาเจริญวงศ์ (2553) ได้เปรียบเทียบเทคนิคการพยากรณ์แบบอนุกรมเวลาทั้งหมด 8 วิธี สำหรับพยากรณ์ปริมาณความต้องการและวางแผนการผลิตชิ้นส่วนไก่สำ หรับผลิตภัณฑ์ไก่แปรรูปแช่แข็งเพื่อวางแผนในการหาจำนวนของไข่และกำหนดการไข่เข้าฟัก ซึ่งพบว่าวิธีการพยากรณ์ที่เหมาะสมสำหรับสินค้าที่ใช้ชิ้นส่วนเนื้อไก่เป็นวัตถุดิบได้แก่ วิธีปรับให้เรียบแบบโฮลท์และวินเทอร์สที่มีฤดูกาล โดยพิจารณาจากค่าความคลาดเคลื่อนของการพยากรณ์ Mean Squared Error (MSE), Mean Absolute Deviation (MAD) และ Mean Absolute Percent Error (MAPE) นอกจากนั้น การพยากรณ์ที่เหมาะสมสามารถช่วยเพิ่มประสิทธิภาพในกระบวนการผลิตได้

ณรงค์เดช เดชทวิสุทธิ์ (2555) ได้ประยุกต์ใช้เทคนิคของการพยากรณ์ร่วมกับการวางแผนการจัดตารางการผลิตหลักในอุตสาหกรรมเฟอร์นิเจอร์ไม้ เมื่อนำค่าเฉลี่ยเปอร์เซ็นต์ความคลาดเคลื่อน (MAPE) มาพิจารณาพบว่าเฟอร์นิเจอร์ไม้แต่ละชนิดมีความเหมาะสมสำหรับเทคนิคของการพยากรณ์ที่แตกต่างกันและการนำโปรแกรมทางคอมพิวเตอร์มาประยุกต์ใช้ในการวางแผนการจัดตารางการผลิตหลักพบว่า ประสิทธิภาพในการผลิตเพิ่มขึ้นจากเดิมมีประสิทธิภาพการผลิตเพียง 51.57 เปอร์เซ็นต์ เพิ่มขึ้นเป็น 90.73 เปอร์เซ็นต์ หรือประสิทธิภาพการผลิตเพิ่มขึ้น 39.16 เปอร์เซ็นต์ ทั้งนี้ การพยากรณ์สำหรับการวางแผนการผลิตยังสามารถช่วยลดขนาดของความคลาดเคลื่อนในการวางแผนการผลิตของผลิตภัณฑ์ได้

พูนสิริ ยิ้มหนองโพ และ ปรารถนา ปรารถนาดี (2555) วิเคราะห์การพยากรณ์ราคาขายผักเมืองหนาว ณ ตลาดสี่มุมเมือง จำนวน 10 ชนิดที่มีราคาขายเฉลี่ยสูงสุด พบว่าการพยากรณ์โดยใช้ชุดข้อมูลราคาขายรายวันในการพยากรณ์ผักเมืองหนาวทั้ง 10 ชนิดมีความเหมาะสมมากที่สุด เนื่องจากให้ค่าความคลาดเคลื่อนต่ำที่สุด เมื่อเปรียบเทียบกับการพยากรณ์ด้วยชุดข้อมูลราย

สัปดาห์และรายเดือน เทคนิคที่เหมาะสมสำหรับการพยากรณ์ราคาขายมะนาวเหลือง พริกหวานสีเหลือง พริกหวานสีแดง เบบี้แครอท พริกหวานสีเขียว ถั่วหวาน และคะน้าฮ่องกง คือ วิธี Single Moving Average โดยให้ค่าความคลาดเคลื่อนในการพยากรณ์ MAPE เท่ากับ 1.52%, 1.84%, 1.93%, 1.47%, 1.67%, 4.22%, 3.08% ตามลำดับ ขณะที่เทคนิคการพยากรณ์ที่เหมาะสมสำหรับราคาขายกะหล่ำดอกม่วงและเรดโอ้คลีฟ คือ SARIMA โดยแบบจำลองพยากรณ์ราคาขายกะหล่ำดอกม่วงที่เหมาะสม คือ SARIMA (2,1,2) (1,0,1) ให้ค่า MAPE เท่ากับ 2.09% และแบบจำลองพยากรณ์ราคาขายเรดโอ้คลีฟที่เหมาะสม คือ SARIMA (1,1,0) (0,0,1) ให้ค่า MAPE เท่ากับ 1.63% สำหรับราคาขายพาร์สเลย์เทคนิคการพยากรณ์ที่เหมาะสม คือ Double Exponential Smoothing โดยมีค่า MAPE เท่ากับ 4.12% แบบจำลองมีความน่าเชื่อถือเพียงพอที่จะนำไปใช้เพื่อการวางแผนการผลิตและเก็บเกี่ยวผักเมืองหนาวแต่ละชนิดตามความเหมาะสมต่อไป โดยเกษตรกรสามารถพิจารณาช่วงเวลาที่ราคาผลผลิตสูงและนำไปวางแผนเริ่มต้นการเพาะปลูกให้เหมาะสมเพื่อให้สามารถเก็บเกี่ยวผลผลิตได้ทันในช่วงเวลาดังกล่าว หรือเกษตรกรอาจพิจารณาการวางแผนการปลูกผักที่เหมาะสมภายใต้ข้อจำกัดด้านทรัพยากร

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

ในบทนี้เป็นการนำเสนอเกี่ยวกับ แนวคิด ทฤษฎี เครื่องมือและวรรณกรรมที่เกี่ยวข้องของกับการวิเคราะห์ข้องมูลการใช้งานเน็ตประชารัฐ เพื่อเผยแพร่ข้อมูลลงบนเว็บไซต์ ซึ่งได้รวบรวมการศึกษาเอกสารงานวิจัยที่เกี่ยวข้องกับระบบ ผู้วิเคราะห์ได้เลือกใช้ขั้นตอนกระบวนการปรับปรุงโครงสร้างข้อมูล (Normalization) ในการจัดการข้อมูลบนฐานข้อมูล เพื่อลดความซ้ำซ้อนของข้อมูล และกระบวนการวิเคราะห์ CRISP-DM จากเทคนิคการพยากรณ์ข้อมูลโดยใช้โมเดล Time-Series Forecasting เพื่อจัดกลุ่มข้อมูลที่มีความสัมพันธ์กัน และยังมีเอกสารทฤษฎีที่เกี่ยวข้องกับการนำเสนอแผนภาพข้อมูล การออกแบบและพัฒนาเว็บไซต์ ที่ใช้สำหรับการจัดทำโครงงานในครั้งนี้