การวิเคราะห์ระยะเวลาที่ใช้ในการการประมวลผล SQL SERVER AGENT JOB โดยอาศัยแผงหน้าปัดธุรกิจ

ANALYZING PROCESSING TIME OF SQL SERVER AGENT JOB USING BUSINESS DASHBOARD PANEL

กฤตบุญ นิยม

สหกิจศึกษานี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญา
วิทยาศาสตร์บัณฑิต(สถิติประยุกต์)
ภาควิชาสถิติ คณะวิทยาศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ปีการศึกษา 2566

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ANALYZING PROCESSING TIME OF SQL SERVER AGENT JOB USING BUSINESS DASHBOARD PANEL

KITTABOON NIYOM

A COOPERATIVE EDUCATION SUBMITTED IN PARTIALFUFILLMENT

OF THE REQUIREMENT FOR

THE DEGREE OF BACHELOR OF SCIENCE (APPLIED STATISTICS)

DEPARTMENT OF STATISTICS, SCHOOL OF SCIENCE

KING MONGKUT'S INSTITUTE TECHNOLOGY LADKRABANG

ACADEMIC YEAR 2023

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อสหกิจศึกษา การวิเคราะห์ระยะเวลาที่ใช้ในการการประมวลผล SQL SERVER AGENT JOB โดยอาศัยแผงหน้าปัดธุรกิจ
ANALYZING PROCESSING TIME OF SQL SERVER AGENT JOB USING BUSINESS DASHBOARD PANEL

ชื่อนักศึกษา นาย กฤตบุญ นิยม รหัสนักศึกษา 63050601

ปริญญา วิทยาศาสตร์บัณฑิต (สถิติประยุกต์)

ภาควิชา สถิติ

ปีการศึกษา 2566

อาจารย์ที่ปรึกษา ผศ.ดร.พรพิมล ชัยวุฒิศักดิ์

คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง (สจล.) อนุมัติให้สหกิจศึกษานี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรบัณฑิต (สถิติประยุกต์) ประจำปีการศึกษา 2566

คณะกรรมการสอบ	ลายมือชื่อ
ดร.สกุณา ศรีอโนมัย ประธานกรรมการ	
คุณ นิรุจ จันทร์แสง กรรมการ	
ผศ.ดร.พรพิมล ชัยวุฒิศักดิ์ กรรมการและอาจารย์ที่ปรึกษา	Pornpimol Chaiwuttisak

ลิขสิทธิ์ของคณะวิทยาศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อสหกิจศึกษา การวิเคราะห์ระยะเวลาที่ใช้ในการการประมวลผล SQL SERVE AGENT JOB โดยอาศัยแผงหน้าปัดธุรกิจ
ANALYZING PROCESSING TIME OF SQL SERVER AGENT JOB USING BUSINESS DASHBOARD PANEL
ชื่อนักศึกษา นาย กฤตบุญ นิยม รหัสนักศึกษา 63050601
ปริญญา วิทยาศาสตร์บัณฑิต (สถิติประยุกต์)
ภาควิชา สถิติ
ปีการศึกษา 2566
อาจารย์ที่ปรึกษา ผศ.ดร.พรพิมล ชัยวุฒิศักดิ์

บทคัดย่อ

แผนกพัฒนาบริหารงานขายสาขาของบริษัทแห่งหนึ่งที่ดำเนินธุรกิจด้านสินเชื่อและโบรกเกอร์ประกันวินาศภัย ประสบปัญหาเกี่ยวกับจำนวนฐานข้อมูลด้านการขายที่เพิ่มขึ้นอย่างรวดเร็ว อันเป็นผลมาจากการจัดเก็บข้อมูลไว้ในฐานข้อมูลที่จำแนกตามประเภทของการขาย โดย SQL Job เป็นลำดับของคำสั่งสำหรับประมวลผลข้อมูลในฐานข้อมูลด้านการขายแต่ละประเภท การวิเคราะห์ข้อมูลเหล่านี้จะช่วยให้ผู้ดูแลระบบสามารถมองเห็นภาพรวมการทำงานของ SQL Job ได้อย่างรวดเร็วและชัดเจน และช่วยในการปรับปรุงเพิ่มประสิทธิภาพในการจัดการระบบฐานข้อมูลในอนาคต ในการศึกษาวิจัยครั้งนี้มีวัตถุประสงค์เพื่อออกแบบแผงหน้าปัดธุรกิจ หรือ Dashboard ที่ช่วยให้ผู้ปฏิบัติงานสามารถเห็นภาพรวมเกี่ยวกับการประมวลผล SQL Job ได้แบบเรียลไทม์ และวิเคราะห์จำนวน SQL Job และระยะเวลาการประมวลผลของของ SQL Job ผลการศึกษาพบว่าแผงหน้าปัดธุรกิจ ถูกออกแบบให้แสดงข้อมูลที่สำคัญและชัดเจน เพื่อช่วยในการวิเคราะห์และการตัดสินใจในการจัดการงานการประมวลผล SQL Job และแก้ไขปัญหาที่เกิดขึ้นได้อย่างรวดเร็ว นอกจากนี้ยังพบว่าค่าเฉลี่ยของจำนวน SQL Job ที่เกิดขึ้นในแต่ละวันไม่มีความแตกต่าง ในขณะที่ค่าเฉลี่ยของจำนวน SQL Job ในช่วงเวลาเช้าสูงกว่าค่าเฉลี่ยของจำนวน SQL Job ในช่วงอื่น ๆ โดยอาศัยการทดสอบของแมนน์-วิทนีย์ นอกจากนี้ค่ามัธยฐานของระยะเวลาในการประมวลผล SQL Job ของทั้ง 7 วันไม่มีความแตกต่างกัน ส่วนค่ามัธยฐานของระยะเวลาในการประมวลผล SQL Job ของช่วงเวลาเช้าสูงกว่าช่วงเวลาอื่น รวมทั้งระยะเวลาในการประมวลผล SQL Job ของ 4 ช่วงเวลามีความแตกต่างแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ได้แก่ ช่วงเวลาเช้าแตกต่างกับช่วงเวลาบ่าย ช่วงเวลากลางคืนแตกต่างกับช่วงเวลาบ่าย ช่วงเวลาเช้าแตกต่างกับช่วงเวลาเย็น และช่วงเวลากลางคืนแตกต่างกับช่วงเวลาเย็น โดยอาศัยการทดสอบของครัสคาล-วอลลิส

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คำสำคัญ : SQL Job, แผงหน้าปัดธุรกิจ, การทดสอบของแมนน์-วิทนีย์, การทดสอบครัสคาล-วอลิส

Title	ANALYZING PROCESSING TIME OF SQL SEVER AGENT JOB USING BUSINESS DASHBOARD PANEL
Student	Mr. Kittaboon Niyom Student ID 63050601
Degree	Bachelor of Science (Applied Statistics)
Department	Statistics
School	Science
University	King Mongkut's Institute of Technology Ladkrabang (KMITL)
Academic Year	2023
Advisor	Asst Prof. Dr. Pornpimol Chaiwuttisak

Abstract

The Sales Branch and Development Department of a company operating in the lending and non-life insurance brokerage business faced a problem with the rapidly increasing number of sales databases. This was due to storing data in databases classified by sales type. An SQL Job is a sequence of commands for processing data in each sales database. Analyzing this data would allow system administrators to quickly see an overview of SQL Job operations, aiding in improving the efficiency of database management in the future. This research aimed to design a business dashboard to allow operators to see a real-time overview of SQL Job processing and analyze the number of SQL Jobs and their processing times. The study found that the business dashboard was designed to display important and clear information to help analyze and make decisions in managing SQL Job processing and quickly solve problems. Additionally, it was found that the average number of SQL Jobs occurring per day did not differ, while the average number of SQL Jobs in the morning was higher than the average at other times, using the Mann-Whitney test. Moreover, the median processing time for SQL Jobs across all Seven days did not differ, but the median processing time in the morning was higher than in other periods. The processing times for SQL Jobs in the four periods were also significantly different statistically, namely morning differed from the afternoon, night differed from the afternoon, morning differed from the evening, and night differed from the evening, using the Kruskal-Wallis One-Way Analysis of Variance by Rank Test.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Keywords: SQL Job, Dashboard, The Mann-Whitney test, The Kruskal-Wallis One-Way Analysis of Variance by Rank Test

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

สหกิจศึกษาฉบับนี้สำเสร็จลุล่วงได้ด้วยความอนุเคราะห์จาก ผศ.ดร.พรพิมล ชัยวุฒิศักดิ์ อาจารย์ที่ปรึกษาสหกิจศึกษาที่ให้คำปรึกษา แนวคิด และข้อคิดเห็นต่าง ๆ เกี่ยวกับกระบวนการทำวิจัย การวิเคราะห์ผล และปรับปรุงแก้ไขข้อบกพร่องต่าง ๆ จนกระทั่งวิจัยครั้งนี้สำเร็จเรียบร้อยด้วยดี ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ ที่นี้

ขอขอบพระคุณ ดร.สกุณา ศรีอโนมัย ที่ให้เกียรติเป็นประธานกรรมการสอบสหกิจให้คำปรึกษา คำแนะนำ ข้อคิดเห็น และชี้แนะแนวทางในการดำเนินการวิจัย จนทำให้วิจัยในครั้งนี้สำเร็จลุล่วงไปได้ด้วยดี

ขอขอบพระคุณ คุณ นิรุจ จันทร์แสง เป็นอย่างสูงที่ให้การสนับสนุนการทำสหกิจศึกษาและช่วยตรวจสอบข้อมูลที่ใช้การทำวิจัย รวมทั้งให้คำปรึกษาเพื่อแก้ไขในการทำสหกิจให้มีคุณภาพ คอยมอบความรู้ในการทำงานต่างๆ ให้แก่ผู้วิจัย

ขอขอบพระคุณบิดา มารดา และครอบครัวที่สนับสนุนในด้านการศึกษาเล่าเรียน และให้กำลังใจผู้วิจัยเสมอมา รวมถึงขอบคุณเพื่อนนักศึกษา และพี่ร่วมงานที่ให้คำปรึกษาให้กำลังใจในการทำสหกิจศึกษาครั้งนี้ สุดท้ายขอบคุณบริษัทกรณีศึกษาที่อนุญาตให้นำข้อมูลมาใช้งานในการทำสหกิจศึกษาครั้งนี้

กฤตบุญ นิยม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	ก
บทคัดย่อภาษาอังกฤษ	ข
กิตติกรรมประกาศ	ง
สารบัญ	ฉ
สารบัญตาราง	จ
สารบัญรูปภาพ	ฌ
บทที่ 1 บทนำ	1
1.1 ความสำคัญและที่มาของการทำวิจัย	1
1.2 วัตถุประสงค์ของงานวิจัย	2
1.3 ขอบเขตของงานวิจัย	2
1.4 ประโยชน์ที่คาดว่าจะได้รับ	3
1.5 นิยามศัพท์เฉพาะ	3
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง	4
2.1 แนวคิดและทฤษฎีที่เกี่ยวข้องกับการวิเคราะห์ทางสถิติ (Statistics)	4
2.1.1 สถิติเชิงพรรณนา (Descriptive Statistics)	4
2.1.2 สถิติเชิงอนุมาน (Inferential Statistics)	5
2.1.3 Data Visualization	14
2.1.4 แดชบอร์ด (Dashboard)	30
2.1.5 ความแตกต่างของ Visualization, Dashboard และ Data Storytelling	34
2.2 ภาษาโปรแกรมที่เกี่ยวข้อง	35
2.2.1 ภาษาเอสคิวแอล (Structured Query Language: SQL)	35
2.2.2 ภาษาแซส (Statistical Analysis Software: SAS)	35
2.3 งานวิจัยที่เกี่ยวข้อง	37
บทที่ 3 วิธีดำเนินงานวิจัย	40
3.1 ขั้นตอนการดำเนินงานวิจัย	40
3.2 เครื่องมือที่ใช้ในการวิจัย	42
3.2.1 ซอฟต์แวร์ที่ใช้ในการวิจัย (Software)	42

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

3.2.2 ฮาร์ดแวร์ที่ใช้ในการวิจัย (Hardware)..... 42
3.3 การสร้าง Log File..... 42
3.4 การเก็บรวบรวมข้อมูล..... 44
3.4.1 การนำข้อมูลเข้า..... 44
3.4.2 การทำความเข้าใจข้อมูล..... 44
3.5 การออกแบบรูปแบบแผงหน้าปัดธุรกิจ หรือ แดชบอร์ด (Dashboard)..... 45
3.5.1 การ์ด (Card) แสดงจำนวนข้อมูล SQL Job..... 45
3.5.2 แผนภูมิเส้นและแผนภูมิแท่ง (Line and Clustered Column Chart) 46
3.5.3 แผนภูมิแท่งแบบคลัสเตอร์ (Clustered Bar Chart)..... 48
3.5.4 แผนภูมิคอลัมน์แบบคลัสเตอร์ (Clustered Column Chart)..... 49
3.5.5 แผนภูมิกรวย (Funnel)..... 50
3.5.6 ตัวกรองข้อมูล (Slicer)..... 51
3.6 การจัดเตรียมข้อมูล..... 51
3.6.1 การสกัดข้อมูล (Data Extraction)..... 51
3.6.2 การแปลงข้อมูล (Data Transformations)..... 52
3.7 การวิเคราะห์ข้อมูล..... 56
3.7.1 สถิติเชิงพรรณนา..... 56
3.7.2 สถิติเชิงอนุมาน..... 56
บทที่ 4 ผลการวิจัยและอภิปรายผล..... 63
4.1 ผลการออกแบบแผงหน้าปัดธุระกิจหรือ Dashboard..... 63
4.2 ผลการวิเคราะห์ข้อมูลด้วยสถิติพรรณนา..... 65
4.2.1 ข้อมูลของ SQL Job..... 65
4.2.2 จำนวน SQL Job ที่มีการประมวลผลในแต่ละวัน..... 65
4.2.3 จำนวน SQL Job ที่มีการประมวลผลในแต่ละช่วงเวลา..... 66
4.2.4 ระยะเวลาการประมวลผล SQL Job ในแต่ละวัน..... 67
4.2.5 ข้อมูลระยะวลาที่ใช้ในการประมวลผลและระยะห่างการมาถึงของแต่ละ SQL Job..... 69

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

4.3 ผลการวิเคราะห์ความแตกต่างระหว่างกลุ่มโดนสถิติอนุมาณ (Inferential Statistic 70
4.3.1 การวิเคราะห์แจกแจงข้อมูล..... 70
4.3.2 การทดสอบความเป็นอิสระ..... 70
4.3.3 การทดสอบสมมติฐานจำนวนการประมวลผลของ SQL Job..... 71
4.3.4 การทดสอบสมมติฐานระยะเวลาที่ใช้ในการประมวลผลของ SQL Job. 72
4.3.5 การเปรียบเทียบพหุคูณระยะเวลาที่ใช้ในการประมวลผลของ SQL Job ในแต่ละช่วงเวลา..... 73
4.4 อภิปรายผลการวิจัย..... 76
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ..... 77
5.1 สรุปผลการวิจัย..... 77
5.2 ข้อจำกัดและข้อเสนอแนะ..... 78
5.2.1 ข้อจำกัด..... 78
5.2.2 ข้อเสนอแนะ..... 78
เอกสารอ้างอิง..... 79
ภาคผนวก..... 81
ภาคผนวก ก 82

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตาราง		หน้า
3.1	รายละเอียดชุดคำสั่งออกแบบการเก็บ Log File	42
3.2	รายละเอียดข้อมูล Log File ของ SQL Job จากฐานข้อมูลบริษัท	45
3.3	ข้อมูลสถานะที่เริ่มการประมวลผลและสถานะที่สิ้นสุดการประมวลผลของ SQL Job	51
3.4	ชุดคำสั่งการแปลงข้อมูลวันที่และเวลาที่ใช้ในการประมวลผลของ SQL Job	53
3.5	ชุดคำสั่งเก็บข้อมูลที่แสดงเป็นช่วงเวลาและวันของสัปดาห์ที่ใช้ในการประมวลผลของ SQL Job	54
4.1	สถิติพรรณนาจำนวน SQL Job ทุกช่วงเวลาในแต่ละวัน	66
4.2	สถิติพรรณนาระยะเวลาการประมวลผลของ SQL Job ในแต่ละวัน	67
4.3	สถิติพรรณนาระยะเวลาการประมวลผลของ SQL Job ในแต่ละช่วงเวลา	68
4.4	สถิติพรรณนาแสดงค่าเฉลี่ยระยะเวลาในการประมวลผล SQL Job ทุกช่วงเวลาในแต่ละวัน	68
4.5	สถิติพรรณนาระยะวลาที่ใช้ในการประมวลผลและระยะห่างการมาถึงของแต่ละ SQL Job	69
4.6	ผลการทดสอบความเป็นอิสระของตัวแปรวัน และตัวแปรช่วงเวลา	71
4.7	ผลการทดสอบค่าฉลี่ย SQL Job ของกลุ่มวันพุธ และกลุ่มวันอื่นๆ ที่ไม่ใช่วันพุธ	71
4.8	ผลการทดสอบค่าเฉลี่ย SQL Job ของกลุ่มช่วงเวลาเช้า และ กลุ่มที่ไม่ใช่ช่วงเวลาเช้า	72
4.9	ผลการทดสอบความแตกต่างของระยะเวลาที่ใช้ในการประมวลผลของ SQL Job แต่ละวันของสัปดาห์	72
4.10	ผลการทดสอบความแตกต่างของระยะเวลาที่ใช้ในการประมวลผลของ SQL Job แต่ละช่วงเวลา	73
4.11	การจัดกลุ่มข้อมูลโดยใช้วิธี Fisher LSD และช่วงความเชื่อมั้น 95%	74
4.12	ผลการทดสอบพหุคูณแบบ LSD ของระยะเวลาที่ใช้ในการประมวลผลของ SQL Job แต่ละช่วงเวลา	74

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูปภาพ

รูปที่		หน้า
3.1	ขั้นตอนดำเนินงาน	41
3.2	ข้อความที่แสดงจำนวนตัวเลข	45
3.3	กราฟแท่งและกราฟเส้นแสดงการเปรียบเทียบจำนวนของ SQL Job	46
3.4	คำสั่งในการคำนวนจำนวนของ SQL Job ที่ “Succeeded”	46
3.5	คำสั่งในการคำนวนจำนวนของ SQL Job ที่ “Failed”	46
3.6	คำสั่งในการคำนวนจำนวนเฉลี่ยของ SQL Job ใน 7 วันก่อน	47
3.7	คำสั่งในการคำนวนจำนวนของ SQL Job ที่ “Delayed”	47
3.8	คำสั่งในการคำนวนจำนวนของ SQL Job ที่ “On Time”	47
3.9	กราฟแท่งแบบคลัสเตอร์ (Clustered Bar Chart) แสดงการเปรียบเทียบจำนวนของ SQL Job	48
3.10	คำสั่งในการคำนวนจำนวนของ SQL Job ที่ล้มเหลว (Failed) และกำหนด top 10 จำนวนที่ SQL Jobมีสถานะล้มเหลวของข้อมูลมากที่สุด	48
3.11	กราฟแท่งแบบคลัสเตอร์ (Clustered Column Chart) แสดงการเปรียบเทียบจำนวนข้อมูล SQL Job ที่ Delayed และ On Time และกำหนด 10 Job ที่มีจำนวนข้อมูลล่าช้าและตามเวลา	49
3.12	ชุดคำสั่งในการคำนวณที่แสดงผลล่าช้า (Delayed) และ ทันเวลา (On Time)	49
3.13	แผนภูมิกรวย (Funnel) แสดงค่าเฉลี่ยของ SQL Job	50
3.14	คำสั่งในการคำนวนค่าเฉลี่ยของ SQL Job และกำหนด Top 10 จำนวนค่าเฉลี่ยที่ SQL Job การประมวลผลมากที่สุด	50
3.15	ตัวกรองข้อมูล (Slicer) วันที่ (เดือน/วัน/ปี)	51
3.16	ตัวกรองข้อมูล (Slicer) SQL Job	51
3.17	ตัวอย่างข้อมูลสถานะที่เริ่มการประมวลผลและสถานะที่สิ้นสุดการประมวลผลของ SQL Job	52
3.18	ตัวอย่างชุดข้อมูลหลังจากแปลงวันที่และเวลาที่ใช้ในการประมวลผลของ SQL Job	53
3.19	ตัวอย่างชุดข้อมูลหลังจากสร้างคอลัมน์ 2 คอลัมน์เก็บข้อมูลที่แสดงเป็นช่วงเวลาและวันของสัปดาห์ที่ใช้ในการประมวลผลของ SQL Job	55
3.20	หน้าต่างการวิเคราะห์การแจกแจงข้อมูลด้วยโปรแกรมสำเร็จรูป Minitab Statistical Software	56

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูปภาพ (ต่อ)

รูปที่		หน้า
3.21	หน้าต่างการวิเคราะห์ทดสอบค่าเฉลี่ยของจำนวน SQL Job ด้วยโปรแกรมสำเร็จรูป Minitab Statistical Software	58
3.22	หน้าต่างการเลือกข้อมูลเพื่อใช้ทดสอบความแตกต่าง Job ด้วยโปรแกรมสำเร็จรูป Minitab Statistical Software	59
3.23	หน้าต่างการเลือกข้อมูลเพื่อใช้ทดสอบความเป็นอีสระของตัวแปรทั้ง 2 ด้วยโปรแกรมสำเร็จรูป Minitab Statistical Software	60
4.1	แดชบอร์ดการวิเคราะห์ประวัติการทำงานของ SQL Job	63
4.2	แดชบอร์ดการวิเคราะห์ประวัติการทำงานของ SQL Job หลังทำการปรับ Ink Ratio	64
4.3	จำนวนข้อมูลของ SQL Job	65
4.4	จำนวน SQL Job ที่มีการประมวลผลในแต่ละวันของสัปดาห์	65
4.5	จำนวน SQL Job ที่มีการประมวลผลแต่ละช่วงเวลา	66
4.6	ผลลัพธ์การวิเคราะห์แจกแจงข้อมูล	70
4.7	กราฟแสดงผลวิธีการเปรียบเทียบพหุคูณแบบ LSD หรือ Fisher's Least – Significant Different	75

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่1
บทนำ

1.1 ความเป็นและความสำคัญของปัญหา

การนำข้อมูลมาใช้ในการวางแผนเพื่อการบริหารงานให้เกิดประสิทธิภาพสูงสุด ถือเป็นกุญแจสำคัญในการผลักดันให้ธุรกิจประสบความสำเร็จ การขับเคลื่อนธุรกิจด้วยข้อมูลคือการนำข้อมูลที่มีรูปแบบที่หลากหลายจำนวนมากที่ธุรกิจได้จากการดำเนินงานมาทำให้เกิดมิติที่ตอบสนองต่อความต้องการของธุรกิจ ดังนั้น หากองค์กรสามารถใช้เครื่องมือที่จะนำเสนอข้อมูลในเชิงเปรียบเทียบเพื่อให้เห็นมุมมองที่นำไปสู่การตัดสินใจได้อย่างรวดเร็วย่อมส่งผลเชิงบวกต่อการดำเนินการขององค์กร รูปแบบของการนำเสนอข้อมูลในเชิงเปรียบเทียบได้อย่างชัดเจนคือ การแสดงผลข้อมูลบนแผงหน้าปัจจุบันธุรกิจหรือแดชบอร์ด (Dash Board) ที่สามารถปรับเปลี่ยนเงื่อนไขและปัจจัยในการเปรียบเทียบได้สะดวกและรวดเร็ว (นภสร, 2564)

โดยในปัจจุบันภาคธุรกิจได้เปลี่ยนรูปแบบวิธีการจัดเก็บข้อมูลเดิมที่เป็นการจดลงในสมุดหรือกระดาษมาอยู่ในรูปฐานข้อมูลแทน ทำให้สามารถเก็บข้อมูลได้เป็นจำนวนตั้งแต่หลักร้อยจนถึงหลักล้านรายการต่อวัน ไม่ว่าจะเป็นข้อมูลประเภทด้านการให้สินเชื่อ ข้อมูลของลูกค้า เป็นต้น นอกจากนี้การใช้ SQL ซึ่งเป็นภาษามาตรฐานสำหรับการจัดการและบริหารฐานข้อมูลเชิงสัมพันธ์ เข้ามาช่วยภาคธุรกิจที่ทำงานกับฐานข้อมูลได้อย่างมีประสิทธิภาพ ไม่ว่าจะเป็นการดึงข้อมูล การอัพเดทข้อมูล การเพิ่มข้อมูลใหม่ หรือการลบข้อมูลที่ไม่ต้องการออกจากฐานข้อมูล นอกจากนี้ SQL ยังมีความสามารถในการประมวลผลข้อมูลที่ซับซ้อนและการทำงานร่วมกับข้อมูลจากหลายตาราง (tables) ได้อย่างรวดเร็ว ทำให้สามารถใช้ฐานข้อมูลสำหรับแหล่งจัดเก็บข้อมูลเป็นจำนวนมาก โดยนำฐานข้อมูลเหล่านี้มาทำการวิเคราะห์และประมวลผลจัดทำแผงหน้าปัดธุรกิจหรือ Dashboard ที่มีการดึงฐานข้อมูลการเรียกใช้ฐานข้อมูลทางด้านการขาย และการอัพเดทข้อมูลทางด้านการขาย โดยนำเสนอภาพรวม ที่เป็นปัจจุบัน โดยสามารถดูข้อมูลภาพรวมได้ทั้งหมด และสามารถสื่อสารให้เข้าใจได้ง่ายไม่สับสน การอัพเดทสถานะระยะเวลาของข้อมูลด้วย SQL ที่ถูกต้องและทันเวลาเป็นองค์ประกอบสำคัญที่ช่วยให้ Dashboard ทำงานได้เป็นประสบการณ์ที่มีประสิทธิภาพและมีประสิทธิผลสูง ทำให้ทราบถึงสถานะของการเรียกใช้ฐานข้อมูลเพื่อเพิ่มศักยภาพในการในการอัพเดทของข้อมูล อีกในแง่มุมหนึ่งที่สำคัญที่ต้องพิจารณาในการออกแบบและพัฒนา Dashboard BI คือคุณภาพของข้อมูล(Buchana and Idrissi 2015; Wieder and Ossimitz 2015) ความถูกต้องของการส่งออกข้อมูล (Buchana & Idrissi 2015; Yigitbasioglu and Velcu 2012) Visualization (Jooste et al . 2014; Orlovskyi and Kopp 2020; Scholtz et al . 2018; Schulze and Kromker 2010; Yigitbasioglu and Velcu 2012; Young and Kitchin 2020) ความต้องการ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(วัตถุประสงค์แดชบอร์ด (Scholtz et al. 2018) ผู้ใช้แดชบอร์ดความต้องการของผู้ใช้ (Schulze and Krömker 2010; Yigitbasioglu and Velcu 2012)

ปัจจุบันแผนกพัฒนาบริหารงานขายสาขาของบริษัทแห่งหนึ่งที่ทำด้านสินเชื่อและโบรกเกอร์ประกันวินาศภัย กำลังเผชิญกับปัญหาการเพิ่มขึ้นอย่างรวดเร็วของจำนวนฐานข้อมูลด้านการขาย การจัดเก็บและจัดประเภทฐานข้อมูลเหล่านี้มีความซับซ้อน เนื่องจากมีการประมวลผลด้วย SQL Job ที่แตกต่างกันไป ทำให้การติดตามการประมวลผลแบบเรียลไทม์เป็นเรื่องยาก อีกทั้งการจัดการฐานข้อมูลด้านการขายโดยไม่มีเครื่องมือสนับสนุนก็ทำได้ยากเช่นกัน ดังนั้นผู้วิจัยจึงได้รับมอบหมายให้ออกแบบแดชบอร์ดติดตามการประมวลผล SQL Job ของฐานข้อมูลแบบเรียลไทม์ เพื่อให้สามารถตัดสินใจในการวางแผนกำหนดตารางการประมวลผล SQL Job ของแต่ละช่วงเวลาได้อย่างเหมาะสม ช่วยลดความล่าช้าในการประมวลผลของฐานข้อมูลและทำการวิเคราะห์การประมวลผลของ SQL Job ของฐานข้อมูลในแต่ละวัน เพื่อให้องค์กรสามารถได้รับข้อมูลที่ถูกต้องและรวดเร็วในการทำธุรกิจได้อย่างมีประสิทธิภาพ

1.2 วัตถุประสงค์ของงานวิจัย

1. เพื่อออกแบบแผงหน้าปัดธุรกิจ หรือ Dashboard ที่สามารถติดตามจำนวนและระยะเวลาการประมวลผล SQL Job รวมถึงสถานะการทำงานแบบเรียลไทม์ได้

2. เพื่อวิเคราะห์ระยะเวลาการประมวลผลของ SQL Job โดยอาศัยการทดสอบแมนน์-วิทนีย์ (The Mann-Whitney Test) และการทดสอบครัสคาล-วอลลิส (The Kruskal-Wallis One-Way Analysis of Variance by Rank Test)

1.3 ขอบเขตของงานวิจัย

1. ขอบเขตด้านข้อมูล

การศึกษาครั้งนี้ได้นำเนินการศึกษากับบริษัทด้านสินเชื่อ และโบรกเกอร์ประกันวินาศภัย โดยทำการสร้าง Log File ของ SQL Job โดยข้อมูลประกอบด้วยชื่อตาราง ชื่อฐานข้อมูล ประเภทของคีย์ วันที่เวลา SQL Job ทำการประมวลผล จำนวนรวมทั้งสิ้น 5,167 รายการ

2. ขอบเขตด้านระยะเวลา

ข้อมูลรายละเอียดด้านการขายในระบบฐานข้อมูลของบริษัทที่ใช้ศึกษาวิจัยและเก็บรวบรวมตั้งแต่วันที่ 1 ธันวาคม 2566 ถึงวันที่ 29 มีนาคม 2567

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.4 ประโยชน์ที่คาดว่าจะได้รับ

1. การสร้างแผงหน้าปัดธุรกิจหรือ (Dashboard) ที่แสดงผลแบบเรียลไทม์และย้อนหลัง สามารถติดตามและบริหารจัดการ SQL Job ข้อมูลเหล่านี้ช่วยให้ผู้ดูแลระบบสามารถมองเห็นภาพรวมการทำงานและประสิทธิภาพของ SQL Job ได้อย่างรวดเร็วและชัดเจน

2. การวิเคราะห์ทางสถิติเพื่อให้เกิดความเข้าใจในระยะเวลาการประมวลผล SQL Job และปัจจัยของวันและช่วงเวลาที่มีผลต่อความแตกต่างระยะเวลาการประมวลนี้ เพื่อช่วยในการปรับปรุงและเพิ่มประสิทธิภาพในการจัดการระบบฐานข้อมูลในอนาคต

3. การมีข้อมูลและการวิเคราะห์อย่างละเอียดช่วยเพิ่มความมั่นใจให้กับผู้ดูแลระบบในการดำเนินการ SQL Job โดยสามารถคาดการณ์และป้องกันปัญหาที่อาจเกิดขึ้นได้ล่วงหน้า

1.5 นิยามศัพท์เฉพาะ

1. SQL Job หมายถึงภาษาโปรแกรมจัดเก็บและประมวลผลข้อมูลให้ได้ผลลัพธ์ที่ต้องการและถูกกำหนดให้ทำงานตามเวลาที่กำหนด

2. แผงหน้าปัดธุรกิจ (Dashboard) หมายถึงเป็นคำศัพท์ที่ใช้ในทางเทคโนโลยีและธุรกิจเพื่ออธิบายหน้าจอหรือพื้นที่ที่รวบรวมและแสดงข้อมูลหรือข้อมูลสถานะต่าง ๆ ได้รับความนิยมมากในระบบสารสนเทศและการบริหารจัดการถูกออกแบบเพื่อให้ผู้ใช้งานสามารถรับข้อมูลในรูปแบบกราฟิกหรือตารางอย่างที่ต้องการเพื่อให้การตัดสินใจของผู้ใช้งานง่ายขึ้นด้วยข้อมูลที่สรุปไว้ในรูปแบบที่กระชับและเข้าใจง่าย

3. Log file เป็นไฟล์ที่ใช้ในการบันทึกเหตุการณ์หรือกิจกรรมที่เกิดขึ้นในระบบคอมพิวเตอร์หรือแอปพลิเคชันต่างๆ โดยมักจะเป็นข้อมูลที่รวบรวมจากการทำงานของระบบ เช่น ข้อผิดพลาด (errors) ที่เกิดขึ้น การเข้าถึงข้อมูล (access) ข้อมูลการเชื่อมต่อ (connections) หรือกิจกรรมอื่นๆ ที่สำคัญ โดยทั่วไปแล้ว Log file จะถูกเก็บเป็นข้อความ (text) หรือรูปแบบอื่น ๆ ที่สามารถอ่านได้ง่ายๆ เพื่อให้สามารถตรวจสอบและวิเคราะห์ข้อมูลได้ในภายหลัง เพื่อใช้ในการดำเนินการต่างๆ เช่น การวิเคราะห์ปัญหา การตรวจสอบการปฏิบัติตามนโยบาย หรือการเก็บข้อมูลเพื่อการตรวจสอบความปลอดภัยและการควบคุม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในการศึกษาวิจัยครั้งนี้ ผู้วิจัยได้มีการศึกษาค้นคว้าแนวความคิดและทฤษฎีที่เกี่ยวข้อง ก่อนทำการวิจัยเพื่อระกอบการศึกษาวิจัยเรื่อง “การวิเคราะห์ระยะเวลาที่ใช้ในการการประมวลผล SQL Server Agent JOB” โดยอาศัยแผงหน้าปัดธุรกิจ” โดยมีรายละเอียดดังต่อไปนี้

2.1 แนวคิดและทฤษฎีที่เกี่ยวข้องกับการวิเคราะห์ทางสถิติ (Statistics)

2.1.1 สถิติเชิงพรรณนา (Descriptive Statistics)

สถิติที่ใช้วิเคราะห์ข้อมูลพื้นฐานในงานวิจัยเชิงปริมาณ เริ่มต้นที่การวิเคราะห์เพื่อบรรยายลักษณะข้อมูลทั่วไป และวิเคราะห์เพื่อตอบคำถามหรือวัตถุประสงค์ของการวิจัยในกรณีที่เป็นคำถามเชิงพรรณนา (สุทิน, 2560) โดยการวิเคราะห์เชิงพรรณนาที่ใช้ประกอบด้วย

1. ร้อยละหรือเปอร์เซ็นต์ (Percentage)

ร้อยละหรือเปอร์เซ็นต์ คือ การคำนวณเพื่อหาสัดส่วนของข้อมูลในแต่ละตัวเทียบกับข้อมูลทั้งหมดโดยให้ข้อมูลรวบรวมมีค่าทั้งหมดเท่ากับ 100 ใช้สัญลักษณ์ % แทนคำว่าร้อยละ ซึ่งใช้วิเคราะห์ตัวแปรประเภทมาตรานมบัญญัติ (Nominal Scale) หรือมาตราอันดับ (Ordinal Scale) ซึ่งมีสูตรคำนวณ ดังสมการที่ 2.1

$$\text{ร้อยละ} = \frac{x}{n} \times 100 \qquad (2.1)$$

เมื่อ x คือ จำนวนข้อมูลความถี่

n คือ ขนาดตัวอย่าง

2. ค่าเฉลี่ย (Mean)

ค่าเฉลี่ย คือ ค่ากลางของข้อมูลรูปแบบหนึ่งจะใช้ค่าเฉลี่ยเป็นตัวแทนของข้อมูลที่มาคำนวน ใช้วิเคราะห์ตัวแปรประเภทมาตราช่วง (Interval Scale) หรือมาตราส่วน อัตราส่วน (Ratio Scale) ซึ่งมีสูตรคำนวณดังสมการที่ 2.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \qquad (2.2)$$

เมื่อ $\bar{x}$ คือ ค่าเฉลี่ยของข้อมูล

x_i คือ คะแนนของตัวอย่างชุดที่ i

n คือ ขนาดตัวอย่าง

3. ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation)

ส่วนเบี่ยงเบนมาตรฐาน คือ ผลรวมของค่าทุกค่าที่ห่างจากค่ากลางของข้อมูล $\left(X - \overline{X}\right)$ ที่ยกกำลังสองหารด้วยจำนวนข้อมูลลบด้วย 1 แล้วนำค่าที่ได้มาหารากที่สองใช้วิเคราะห์ตัวแปรประเภทมาตราวัดส่วนช่วง (Interval Scale) หรือมาตราอัตราส่วน (Ratio Scale) ซึ่งมีสูตรคำนวณดังสมการที่ 2.3

$$S.D. = \sqrt{\frac{\left(\sum_{i=1}^{n} X_i - \overline{X}\right)^2}{n-1}} \qquad (2.3)$$

เมื่อ $S.D.$ คือ ค่าส่วนเบี่ยงเบนมาตรฐาน

X_i คือ คะแนนของตัวอย่างชุดที่ i

n คือ ขนาดตัวอย่าง

2.1.2 สถิติเชิงอนุมาน (Inferential Statistics)

สถิติเชิงอนุมาณ เป็นการนำข้อมูลที่เก็บมาได้จากกลุ่มตัวอย่างไปใช้อ้างอิงหรืออธิบายกลุ่มประชากร ได้แก่ การประเมินค่าพารามิเตอร์ในประชากร (Estimation) และการทดสอบสมมุติฐานHypothesis Testing) แบ่งออกเป็น Parametric และ Nonparametric Statistics (วีระศักดิ์, 2557) แม้ว่าสถิติเชิงอนุมานจะมีอยู่หลายตัว แต่ในงานวิจัยกล่าวถึงเพียงสถิติที่ใช้ในงานวิจัยเท่านั้น ได้แก่ การทดสอบการแจกแจงปรกติ การวิเคราะห์ความแตกต่างของกลุ่มสองกลุ่ม ด้วยสถิติทดสอบของแมนน์-วิตนีย์ สถิติทดสอบครัสคาล-วอลลิส และการวิเคราะห์ความเป็นอิสระด้วยสถิติทดสอบแบบไคสแควร์ (The Chi-Square Test) ซึ่งมีรายละเอียดดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. การทดสอบการแจกแจงปรกติ (Normality Test)

1) การทดสอบชาพิโร-วิลก์ (Shapiro-Wilk Test)

การทดสอบชาพิโร-วิลก์เป็นสถิติทดสอบการแจกแจงของตัวแปรเชิงปริมาณ การทดสอบชาพิโร-วิลก์ใช้ในกรณีไม่ทรายค่าเฉลี่ย หรือค่าความแปรปรวนของประชากรก็ได้ และตัวอย่างมีขนาดไม่เกิน 50 หน่วย (Pearson and Hartley, 1972)

ข้อจำกัดเบื้องต้น (Assumption)

ข้อมูลประกอบไปด้วยตัวอย่างสุ่ม $X_1,...,X_n$ ที่มาจากประชากรที่มีฟังก์ชันการแจกแจงเหมาะสม $F(X)$ ที่ไม่ทราบค่า

สมมติฐานการทดสอบ

H_0 : ตัวอย่างจากประชากรที่มีการแจกแจงปรกติ

H_1 : ตัวอย่างจากประชากรที่ไม่ได้มีการแจกแจงปรกติ

สถิติทดสอบ

$$T=\frac{1}{D}\left[\sum_{i=1}^{k}a_i\left(X^{(n-i+1)}-X^{(i)}\right)\right]^2 \text{ โดยที่ } \sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2 \qquad (2.4)$$

เมื่อ $\overline{X}$ คือ ค่าเฉลี่ยของตัวอย่าง

a_i คือ ค่าสัมประสิทธิ์ที่ได้จากการเปิดตารางของชาพิโร-วิลก์

$X^{(i)}$ คือ ค่าสังเกตตัวอย่างแบบเรียงลำดับ

การสรุปผล

ถ้าค่า T มีค่าเข้าใกล้ 1 แสดงว่าตัวอย่างมีการแจกแจงปรกติ แต้ถ้า T มีค่าน้อยกว่า 1 มาก แสดงว่าไม่ได้มีการแจกแจงปกรติ

2) การทดสอบลิลลีโฟรส์ (Lilliefors Test)

การทดสอบลิลลีโฟรส์เป็นการทดสอบการแจกแจงของประชากรว่าเป็นการแจกแจงปรกติหรือไม่ โดยเป็นวิธีทดสอบที่มีการกำลังทดสอบสูงกว่าการทดสอบอื่นๆ ซึ่งจะเหมือนการทดสอบของคอลโมโกรอฟ-สมีร์นอฟ (Kolmogorov-Smirnov (K-S) Test) แต่การทดสอบลิลลีโฟรส์จะไม่กำหนดค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐานของประชากร จึงต้องประมาณค่าด้วย μ ด้วย $\overline{X}$ และประมาณค่า σ ด้วย s การทดสอบนี้จะใช้เมื่อประชากรมีขนาดมากกว่า 50 หน่วย (สุจิตรา, 2564)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สมมติฐานการทดสอบ

H_0 : ตัวอย่างจากประชากรที่มีการแจกแจงปรกติ

H_1 : ตัวอย่างจากประชากรที่ไม่ได้มีการแจกแจงปรกติ

สถิติทดสอบ

$$D = \max\left|F(X) - S(X)\right| \tag{2.5}$$

เมื่อ $\max$ คือ การเลือกค่าสูงสุด

$F(X)$ = ความน่าจะเป็นสะสมของตัวอย่าง

$S(X)$ = ความน่าจะเป็นสะสมภายใต้สมมติฐานว่าง

การสรุปผล

ปฏิเสธ H_0 ถ้าค่า p-value ของการทดสอบมีค่าน้อยกว่า α ที่กำหนด

2. การทดสอบของแมนน์-วิตนีย์ (The Mann – Whitney test)

การทดสอบของแมนน์-วิตนีย์ (The Mann – Whitney test) มีชื่อเรียกอีกอย่างหนึ่งว่าการทดสอบวิลคอกซัน (Wilcoxon test) วิธีของวิลคอกซัน (Wilcoxon test) ใช้ในกรณีที่ตัวอย่างมีขนาดเท่ากัน ในขณะที่วิธีแมนน์-วิตนีย์(The Mann – Whitney test) ใช้ในกรณีที่ตัวอย่างมีขนาดไม่เท่ากัน ในกรณีตัวอย่างขนาดใหญ่สามารถประมาณด้วยการแจกแจงปรกติมาตรฐาน (Z) (อุมาพร, 2542)

ข้อกำหนดเบื้องต้น

1) ข้อมูลประกอบด้วยตัวอย่างสุ่ม ด้วยค่า $X_1, X_2, ..., X_{n1}$ จากประชากรที่ 1 และตัวอย่างสุ่มอีก 1 ชุด ด้วยค่าสังเกต $Y_1, Y_2, ..., Y_{n2}$ จากประชากรที่ 2 ซึ่งเป็นอิสระกัน

2) ตัวอย่าง 2 ชุดนี้เป็นอิสระกัน

3) ค่าตัวแปรสุ่มมีค่าต่อเนื่อง

4) มาตราวัดอย่างน้อยเป็นแบบเรียงลำดับ (ordinal scale)

5) ฟังก์ชันการแจกแจงของ 2 ประชากร ต่างกันเฉพาะค่ากลาง (ซึ่งนิยมวัดด้วยมัธยฐาน) นั่นคือประชากรทั้ง 2 ต้องมีการแจกแจงที่เหมือนกัน ต่างกันเฉพาะค่ากลางเท่านั้น

หมายเหตุ ในทางปฏิบัติไม่จำเป็นต้องทราบว่ามีการแจกแจงแบบใด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สมมติฐานการทดสอบ

ถ้าให้ M_x และ M_y แทนค่ามัธยฐานของประชากรที่ 1 และ 2 ตามลำดับ อาจทำการทดสอบสองหางหรือหางเดียว ได้ดังนี้

$H_0 : M_x = M_y$

$H_1 : M_x \neq M_y$ หรือ

$H_0 : M_x \geq M_y$

$H_1 : M_x < M_y$ หรือ

$H_0 : M_x \leq M_y$

$H_1 : M_x > M_y$

สถิติทดสอบ

วิธีของ Wilcoxon และ Mann-Whitney ได้แสดงความสัมพันธ์ระหว่างสถิติที่ใช้ทดสอบของ Mann-Whitney กับของ Wilcoxon พบว่า ถ้าให้ $T = S - \frac{n_1(n_1+1)}{2}$ แล้วค่า T ที่ได้จะมีค่าเท่ากับ U นั้นเอง เมื่อ U = Mann-Whitney U Statistics หลักในการหาอาณาเขตวิกฤต ค่า T ที่มากเกินไปหรือน้อยเกินไปจะทำให้ ปฏิเสธ H_0 เพื่อยอมรับ H_1

ดังนั้นสถิติที่ใช้ทดสอบคือ $T = S - \frac{n_1(n_1+1)}{2}$ (2.6)

เมื่อ S = ผลรวมลำดับที่ของตัวอย่างขนาด n_1 ในข้อมูลรวมทั้งหมดที่เรียงลำดับแล้ว

การตัดสินใจ

ในกรณีการทดสอบสองหางจะปฏิเสธ H_0 ถ้าพบว่า T น้อยเกินไปหรือใหญ่เกินไป

อาณาเขตวิกฤตคือ $T < W_{\frac{\alpha}{2}}$ หรือ $T > W_{\frac{\alpha}{2}}$ เมื่อ $W_{1-\frac{\alpha}{2}} = n_1 n_2 - W_{\frac{\alpha}{2}}$ เมื่อเป็นการทดสอบหางเดียงด้านน้อยกว่า คือ $H_1 : M_x < M_y$ จะปฏิเสธ H_0 เมื่อพบว่าค่า T น้อยเกินไป

อาณาเขตวิกฤตคือ $T < W_\alpha$ เมื่อเป็นการทดสอบหางเดียงด้านมากกว่า $H_1 : M_x > M_y$ เมื่อพบว่าค่า T ใหญ่เกินไปอาณาเขตวิกฤตคือ $T > W_\alpha$ เมื่อ $W_{1-\alpha} = n_1 n_2 - W_\alpha$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. การทดสอบครัสคาล-วอลลิส (The Kruskal-Wallis One-Way Analysis of Variance by Rank Test)

การทดสอบครัสคาล-วอลลิส เป็นการทดสอบประชากร k กลุ่มที่เป็นอิสระกัน มีวิธีการสำคัญคือ ค่าคาดหมายของลำดับที่ของข้อมูลตัวอย่างแต่ละกลุ่ม ถ้าประชากรทั้ง k กลุ่มมีค่ากลางไม่ต่างกันควรจะมีค่าพอๆกัน ข้อมูลที่นำมาทดสอบประกอบด้วยข้อมูลจากตัวอย่างสุ่ม k ชุด แต่ละชุดอาจมีขนาดตัวอย่างแตกต่างกัน (อุมาพร, 2542)

สมมติฐานการทดสอบ

H_0 : ค่ามัธยฐานของประชากร K กลุ่มไม่แตกต่างกัน

H_1 : ค่ามัธยฐานของประชากรอย่างน้อย 1 คู่แตกต่างกัน

วิธีการทดสอบ

1. จัดลำดับของข้อมูลทั้งหมดร่วมกัน จากน้อยไปหามาก โดยให้คะแนนต่ำสุดมีลำดับที่ 1 และคะแนนสูงสุดเป็นลำดับที่ n เมื่อ n เป็นจำนวนข้อมูลทั้งหมด
2. หาผลรวมขอลำดับที่ในข้อมูลแต่ละชุด คือ $R_i, i=1,2,...,k$

สถิติทดสอบ

$$H=\left[\frac{12}{n(n+1)}\sum_{i=1}^{k}\frac{R_i}{n_i}\right]-3(n-1) \qquad (2.7)$$

เมื่อ k คือ จำนวนประชากรที่เป็นอิสระต่อกัน

R_i คือ ผลรวมของลำดับที่ในตัวอย่างที่ $i, j=1,2,...,k$

n_i คือ ขนาดตัวอย่างชุดที่ $i, j=1,2,...,k$

$$n=\sum_{i=1}^{k}n_i$$

สรุปผล

ถ้า H_0 เป็นจริง H จะมีการแจกแจงประมาณด้วย χ^2 ที่ $d.f=k-1$ ถ้า $n_i n_i$ มีค่าใหญ่พอสมควร การหสอณาเขตวิกฤตและการสรุปผล สามารถแบกได้ตามขนาดตัวอย่างคือ

เมื่อ $n_i>5$ การแตกต่างของค่าสถิติ H ประมาณด้วย χ^2 ที่ $d.f=k-1$ เมื่อกำหนดระดับนัยสำคัญ = α หาอาณาเขตวิกฤตจากตาราง χ^2 ที่ $d.f=k-1$ จะปฏิเสธ H_0 เมื่อค่าของ H มากกว่าหรือเท่ากับ χ^2 จากตาราง

เมื่อ $k=3$ และ $n_i\le 5$ ในแต่ละ k ใช้ตารางที่ Kruskal สร้างไว้โดยแสดงค่าวิกฤตของ H พร้อมทั้งความน่าจะเป็นที่จะเกิดค่า H นั้นๆตารางของ Kruskal จะสามารเปรียบเทียบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ค่า H หรือ p-value ได้ คือ จะปฏิเสธ H_0 เมื่อค่า H จากตัวอย่างมากกว่าหรือเท่ากัน H จากตารางที่ระดับนัยสำคัญ α

4. การเปรียบเทียบพหุคูณ (Multiple comparison)

เทคนิคการวิเคราะห์ความแปรปรวนเป็นการทดสอบว่าจะมีค่าเฉลี่ยของประชากร k กลุ่มแตกต่างกันหรือไม่ ถ้าแตกต่างกันอย่างมีนัยสำคัญทางสถิติ(significant) ก็จะบอกเพียงว่ามีค่าเฉลี่ยอย่างน้อย 1 คู่ที่มีค่าแตกต่างกันแต่จะไม่บอกว่าเป็นคู่ใด ซึ่งเราจะต้องทำการทดสอบหลังการวิเคราะห์ (Post hoc test) โดยวิธีการเปรียบเทียบพหุคูณ (Multiple Comparison)

1) วิธี Least -Significant Different (LSD)

วิธีการเปรียบเทียบพหุคูณแบบ LSD หรือ Fisher's Least – Significant Different เป็นเทคนิคที่ R.A. Fisher ได้พัฒนาขึ้นหรือเปรียบเทียบค่าเฉลี่ยประชากรครั้งละหลายคู่ (Kirk, 1995)

สถิติทดสอบ

$$LSD = t_{\frac{\alpha}{2}} \sqrt{MSE} \sqrt{\frac{1}{n_i} + \frac{1}{n_j}} \qquad (2.8)$$

เมื่อ MSE คือ ความแปรปรวนจาก one way ANOVA

n_i คือ แทนจำนวนข้อมูลกลุ่มที่ i

n_j คือ แทนจำนวนข้อมูลกลุ่มที่ j

วิธี LSD มีขั้นตอนดังนี้

1) คำนวณค่า LSD

2) คำนวณความแตกต่างระหว่างค่าเฉลี่ย $\overline{X}_i - \overline{X}_j$

3) นำค่า $\left|\overline{X}_i - \overline{X}_j\right|$ เปรียบเทียบกับค่า LSD

3.1) ถ้าค่า $\left|\overline{X}_i - \overline{X}_j\right|$ > ค่า LSD แสดงว่า $\mu_i \neq \mu_j$

3.2) ถ้าค่า $\left|\overline{X}_i - \overline{X}_j\right|$ < ค่า LSD แสดงว่า $\mu_i = \mu_j$

2) วิธีของทูกีย์ (Tukey's Test)

การทดสอบจะหาค่าค่าหนึ่งเป็นตัวเปรียบเทียบกับความแตกต่างของทรีทเมนตเป็นรายคูทุกคู คานี้เรียกวา HSD กรณีที่ทุกทรีทเมนตมีจำนวนคาสังเกตเทากันทุกทรีทเมนต HSD หาได้ดังนี้ (Tukey, 1956)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สถิติทดสอบ

$$HSD = q_{\alpha,a,N-a}\sqrt{\frac{MSE}{n}} \tag{2.9}$$

เมื่อ α คือ ระดับนัยสำคัญ

a คือ จำนวนทรีทเมนตในการทดลอง

N คือ จำนวนคาสังเกตทั้งหมดในการทดลอง

n คือ จำนวนคาสังเกตในแต่ละทรีทเมนต

MSE คือ mean square error หรือ within mean square ที่ได้จากตาราง ANOVA

q คือ เปิดจากตาราง H ที่ $\alpha - a$ และ $N - a$

คำนวณหาความแตกต่างของคาเฉลี่ยของทรีทเมนตทุกคูที่เป็นไปได้เป็นค่าสัมบูรณ์ นำคาที่คำนวนได้เปรียบเทียบกับ HSD ถ้าค่าที่คำนวณมากกวา HSD จะปฏิเสธสมมติฐานสูญ

สำหรับกรณีที่ตัวอย่างแต่ละกลุมมีจำนวนไม่เทากัน การคำนวณหา HSD หาได้ดังนี้

$$HSD^* = q_{\alpha,a,N-a}\sqrt{\frac{MSE}{n_i^*}} \tag{2.10}$$

ค่าสัมบูรณ์ของความแตกต่างระหว่างค่าเฉลี่ยของตัวอย่าง 2 กลุม โดยที่ n_i^* คือ จำนวนของกลุมตัวอย่างที่เล็กกว่าอีกกลุ่มหนึ่ง เปรียบเทียบคาที่คำนวณได้กับ HSD^* ถ้าค่าที่คำนวณมากกวา HSD จะปฏิเสธ H_0

3) วิธีของงดันแคน (Duncan's Multiple Range Test)

ดันแคนได้พัฒนาการทดสอบการเปรียบเทียบพหุเพื่อเปรียบเทียบคาเฉลี่ยของแต่ละทรีทเมนตกับทรีทเมนตอื่น ๆ ทั้งหมดที่เหลือ วิธีการทดสอบประกอบด้วย 3 ขั้นตอน และเมื่อจำนวนทรีทเมนตเทากับ 10 หรือมากกวา ในป 1955 ดันแคนได้พัฒนาวิธีใหม่คือ new multiple range test เพื่อใหง่ายขึ้น (Duncan, 1955)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สถิติทดสอบ

$$S_{\bar{y}_{i.}} = \sqrt{\frac{MSE}{n}} \qquad (2.11)$$

ถ้าขนาดตัวอย่างไม่ทากันทุกทรีทเมนต แทน n ด้วย n_h

$$n_h = \frac{a}{\sum_{i}^{a}\left(\frac{1}{n_i}\right)}$$

การคำนวณค่าวิกฤต

โดยการเปิดตารางคา Significant Ranges for Duncan's Multiple Range Test จะได้คา $r_{\alpha(p,f)}$ สำหรับ $p = 2,3,...,a$ และทำการเปรียบเทียบทรีทเมนต์แต่ละคู่กับค่าวิกฤต

เมื่อ α คือ ระดับนัยสำคัญ

f คือ จำนวนชั้นอิสระของ error

p คือ จำนวนทรีทเมนต์ที่อยู่ในกลุ่มหนึ่ง

การคำนวณคาวิกฤติของทรีทเมนตที่มีขนาดแตกต่างกัน

หาคา least significant ranges $(R_p); p,2,3,...,a$

$$R_p = r_{\alpha(p,f)} S_{\bar{y}_{i.}} \qquad ; p = 2,3,...,a$$

5. การทดสอบแบบไคสแควร์ (The Chi-Square Test)

การทดสอบเกี่ยวกับข้อมูลที่มีลักษณะสองทางจากข้อมูลที่สนใจศึกษาซึ่งมีลักษณะ 2 ทางนำมาสร้างตารางความถี่ของตัวแปรสุ่มทั้ง 2 และเรียกว่า ตารางการณ์จรณ์ ทำการทดสอบถึงความเป็นอิสระของตัวแปรทั้ง 2 และเรียกชื่อการทดสอบสำหรับข้อมูลประเภทนี้ว่า "การทดสอบความเป็นอิสระ" (Test for Independence) (อุมาพร, 2542)

ตารางการจร จะจัดในลักษณะแกนนอน (row) r แถว และแกนตั้ง (column) c แถว หรือเรียกว่าตารางชนิด r x c ใช้เพื่ออธิบายตัวแปรสุ่ม 2 ตัวซึ่งเป็นตัวแปรคุณภาพในลักษณะเหตุการณ์ย่อยหลายๆเหตุการณ์ แล้วสังเกตจำนวนครั้งของการเกิดขึ้นของเหตุการณ์ย่อยเหล่านั้น และคำนวณจำนวนครั้งที่คาดหวังตามนิยามของเหตุการณ์อิสระในการทดสอบนี้จะทดสอบถึงความเป็นอิสระของ 2 เหตุการณ์เท่านั้น คือ เหตุการณ์ A และเหตุการณ์ B และจำแนกเหตุการณ์ A ออกเป็นเหตุการณ์ย่อย $A_1, A_2, ..., A_r$ และจำแนกเหตุการณ์ B ออกเป็นเหตุการณ์ย่อย $B_1, B_2, ..., B_c$ ทำการทดลองหรือสังเกตจำนวนครั้งของการเกิดเหตุการณ์ $A_i, B_j; i = 1,2,...,r \; ; j = 1,2,...,c$ ดังตารางการณ์จรณ์ ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่2.1 ตารางการจร

ลักษณะตัวแปร A (i=1,2,3,...,r)	ลักษณะตัวแปร B (j=1,2,3,...,r)				รวม
	B_1	B_2	B_j	B_c	
A_1	O_{11}	O_{12}	.	O_{1c}	R_1
A_2	O_{21}	O_{22}	.	O_{2c}	R_2
.	.	.	.	.	.
A_i	.	.	.	.	R_i
.	.	.	.	.	.
A_r	O_{r1}	O_{r2}	.	O_{rc}	R_r
รวม	c_1	c_2	c_j	c_c	N

เมื่อ O_{ij} = ความถี่ที่ได้จากการทดสองหรือสังเกตของเหตุการณ์ A_i และเหตุการณ์ $B_i; = 1,2,...,r; = 1,2,...,c$

R_i = ผลรวมของความถี่ที่ได้จากการทดลองหรือสังเกตุของเหตุการณ์

C_j = ผลรวมของความถี่ที่ได้จากการทดลองหรือสังเกตของเหตุการณ์

N = ผลรวมของความถี่ที่ได้จากการทดลองทั้งหมด

ซึ่ง $N = \sum_{i=1}^{r} R_i = \sum_{i=1}^{c} R_j$

สมมติฐานการทดสอบ

H_0 : ตัวแปร A และ B เป็นอิสระต่อกัน

H_1 : ตัวแปร A และ B ไม่เป็นอิสระต่อกัน

หรือ H_0 : ตัวแปร A ไม่ขึ้นกับตัวแปร B

H_1 : ตัวแปร A ขึ้นกับตัวแปร B

หรือ H_0 : ตัวแปร A และ B ไม่มีความสัมพันธ์กัน

H_1 : ตัวแปร A และ B มีความสัมพันธ์กัน

สถิติทดสอบ

$$\chi^2_{cal} = \sum_{i=1}^{r}\sum_{i=1}^{c} \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}} \quad (2.12)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อ O_{ij} คือ ความถี่ที่ได้จากการสังเกตใน cell (i , j)

E_{ij} คือ แทนความถี่คาดหวังใน cell (i , j)

และ χ^2 จะมีการแจกแจงเข้าใกล้การแจกแจงไคสแควร์ด้วยชั้นแห่งความเป็นอิสระ = $(c-1)(r-1)$

การคำนวณหาความถี่คาดหวัง คือ การหาค่า E_{ij} สามารถคำนวณได้ตามนิยามของความเป็นอิสระ ดังนี้ ถ้า A และ B เป็นเหตุการณ์ที่เป็นอิสระต่อกันแล้ว

$$P(AB) = P(A)P(B)$$

นั่นคือ $P(A_iB_j) = P(A_i)P(B_j)$

ดังนั้น ความถี่คาดหวังของเหตุการณ์ A_iB_j คือ E_{ij} มีค่าดังนี้

$$E_{ij} = NP(A_iB_j) = NP(A_i)P(B_j) = N \times \frac{R_i}{N} \times \frac{C_j}{N}$$

$$E_{ij} = \frac{R_iC_j}{N}$$

2.1.3 Data Visualization

Visualization คือการสร้างภาพข้อมูลเป็นการแสดงข้อเท็จจริงในรูปแบบกราฟิกที่ทำให้เข้าใจได้ง่ายและรวดเร็ว ข้อมูลจะถูกนำเสนอในลักษณะที่เข้าใจได้ง่ายโดยใช้การสร้างภาพข้อมูล ซึ่งอาจรวมถึงแผนภูมิ กราฟ และแผนภาพ การสร้างภาพข้อมูลช่วยให้ค้นพบรูปแบบ แนวโน้ม และความเชื่อมโยงของข้อมูลที่อาจมองไม่เห็นได้ง่ายจากฐานข้อมูลที่มีอยู่

การแปลงข้อมูลเป็นภาพเช่นการสร้างกราฟ เช่น กราฟแท่ง กราฟเส้น กราฟวงกลม มาก่อนแล้วเหล่านี้ถือเป็น Visualization อย่างหนึ่ง แต่ก็ยังมีกราฟรูปแบบอื่นๆ อีกมากมาย ที่มีประสิทธิภาพ ในการสื่อสารมากกว่า ทำอย่างไรให้ผู้ใช้งานทั่วไปสามารถสร้าง Visualization เหล่านี้ได้ด้วยตนเอง โดยไม่ต้องใช้ Skill ทางด้านภาษาคอมพิวเตอร์ที่ซับซ้อน เพื่อให้เข้าถึงคนหมู่มากได้ดังนี้

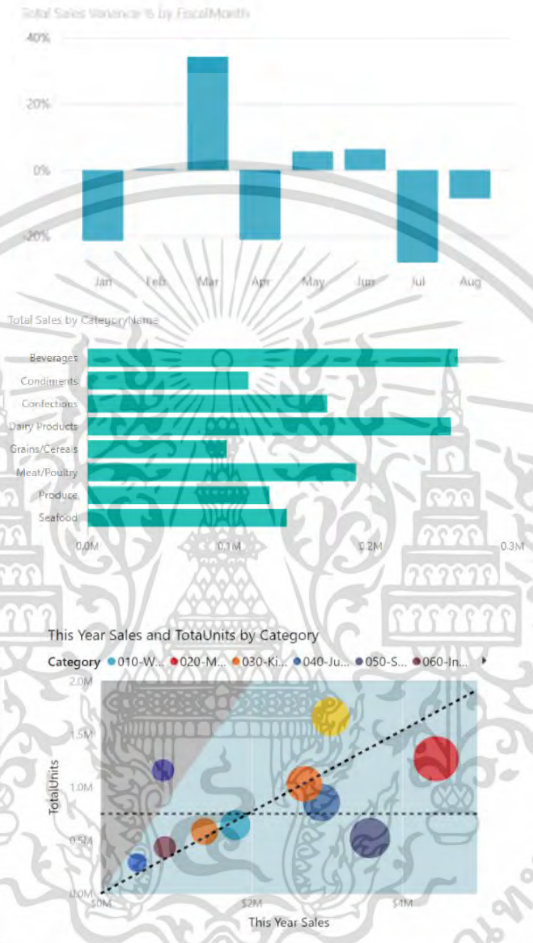
2.1.3.1 Visualization ในการนำเสนอข้อมูล

การนำเสนอข้อมูลในรูปแบบต่าง ๆ มีมากมาย หลากหลายรูปแบบนี้ผู้วิเคราะห์ควรพิจารณาจากข้อมูลที่ใช้ และรายละเอียดอื่น ๆ ประกอบการออกแบบการนำเสนอข้อมูล ในหัวข้อนี้เป็นการเลือกใช้แผนภูมิ แผนภาพ กราฟต่าง ๆ ในการนำเสนอข้อมูล ให้เหมาะสมกับชนิดและประเภทของข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1) แผนภูมิแสดงปริมาณ (Amount)

แผนภูมิชนิดนี้มักใช้ในการนำเสนอข้อมูลที่เป็นจำนวนที่มีชนิดของข้อมูลเพียงชุดเดียว เช่น การแสดงจำนวนต่าง ๆ แผนภูมิชนิดนี้ประกอบไปด้วยแผนภูมิรูปแท่ง (Bar) ทั้งแบบแนวนอนและแนวตั้ง และแผนภูมิแบบจุด (Dot)

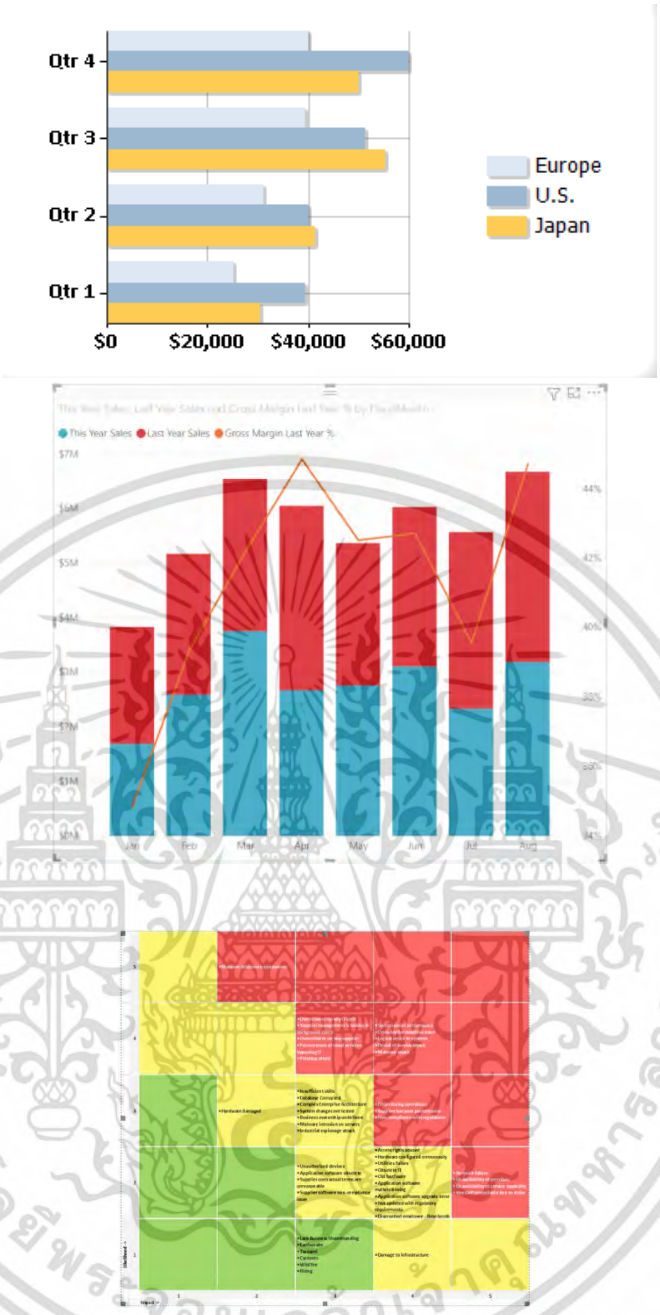


รูปที่ 2.1 แผนภูมิแสดงปริมาณ

จากรูปที่ 2.1 เป็นตัวอย่างการแสดงจำนวนของแผนภูมิปริมาณชนิดต่าง ๆ ประกอบไปด้วยแผนภูมิรูปแท่ง (Bar) ทั้งแบบแนวนอนและแนวตั้ง และแผนภูมิแบบจุด (Dot)

ในกรณีที่ต้องการแสดงจำนวนชุดของข้อมูลตั้งแต่ 2 ชุดหรือมากกว่า มักนิยมใช้แผนภูมิในรูปแบบ Group แผนภูมิแบบ Stack ที่ใช้แกน x และ แกน y ในการกำหนดการแสดงข้อมูล หรือ การใช้แผนภูมิแบบ Heat Map เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.2 แผนภูมิแสดงปริมาณที่มีข้อมูลมากกว่าหนึ่งชุด

จากรูปที่ 2.2 เป็นตัวอย่างการแสดงจำนวนของแผนภูมิปริมาณชนิดต่างๆที่มีข้อมูลมากกว่า 1ชุด ประกอบไปด้วยแผนภูมิในรูปแบบ Group แผนภูมิแบบ Stack และแผนภูมิแบบ Heat map

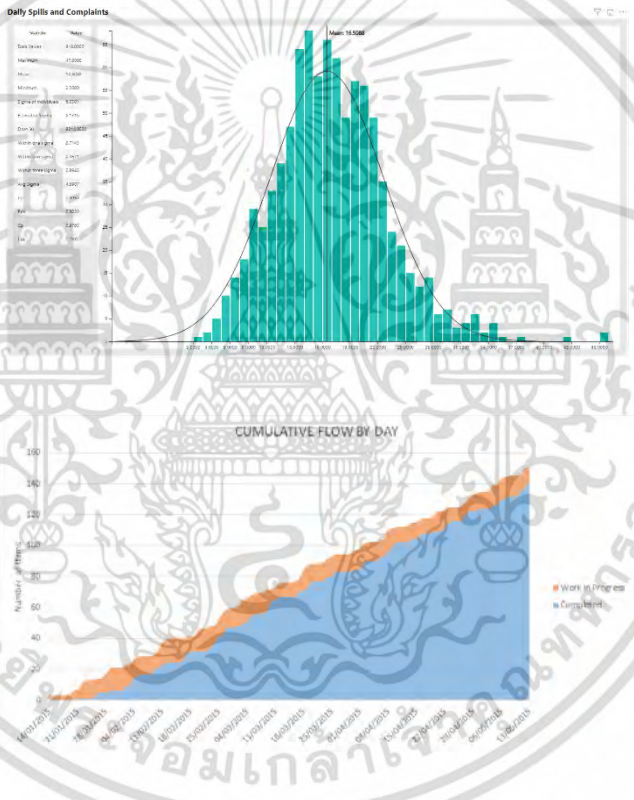
แผนภูมิแท่ง เป็นการนำเสนอเพื่อเน้นแนวเปรียบเทียบของแต่ละอย่าง โดยพยายามสร้างกราฟที่เรียบง่าย และเข้าใจได้ง่าย โดยสามารถเปรียบเทียบได้มากกว่า 1 ด้าน (ใน Excel จะเรียกว่า Series) เป็นกราฟที่ใช้สำหรับเปรียบเทียบข้อมูล หากเปรียบเทียบข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คนละประเภทกัน ควรใช้คนละแกน เช่น การเปรียบเทียบยอดขายกับกำไรกราฟแท่งแบบต่อกัน เป็นการนำเสนอเพื่อให้เห็นการรวมกันของข้อมูลเพื่อนำมาต่อๆ กัน ว่าได้เท่าใด ไม่ได้ต้องการนำข้อมูลแต่ละด้าน (หรือSeries) มาเปรียบเทียบกัน แต่ต้องการนำทุก ๆ Series มารวมกัน กราฟชนิดนี้เหมาะสำหรับการนำยอดรวมมาเปรียบเทียบกัน ข้อมูลที่นำเสนอควรเป็นข้อมูลในหน่วยเดียวกัน

2) แผนภูมิแสดงการกระจายตัวของข้อมูล (Distribution)

แผนภูมิฮิสโตแกรมและแผนภูมิที่แสดงความหนาแน่นของข้อมูลเป็นการนำเสนอข้อมูลที่ใช้งานง่าย ไม่ซับซ้อนสำหรับการนำเสนอข้อมูลในรูปแบบการกระจายตัวของข้อมูล แต่ในรูปแบบการนำเสนอข้อมูลแบบรูปนี้จำเป็นต้องมีตัวแปรอย่างน้อยสองตัว



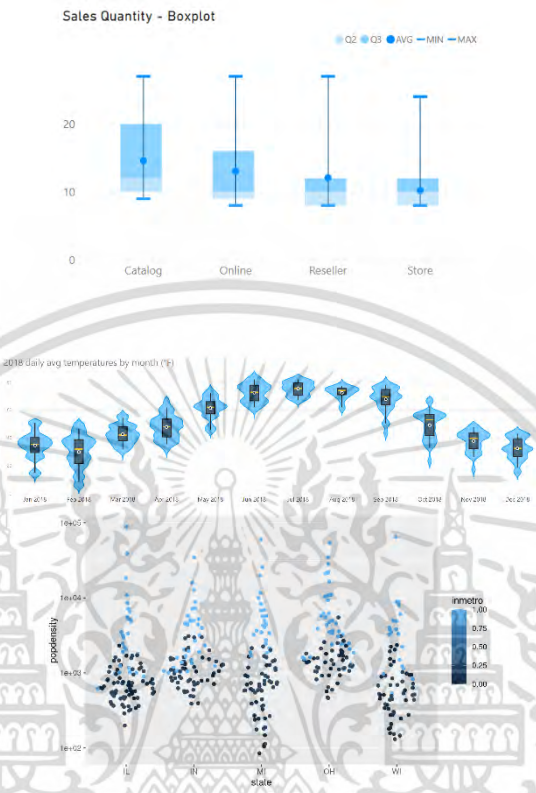
รูปที่ 2.3 แผนภูมิแสดงการกระจายตัวของข้อมูล

จากรูปที่ 2.3 เป็นตัวอย่างแผนภูมิแสดงการกระจายตัวของข้อมูล ประกอบไปด้วยแผนภูมิฮิสโตแกรม และแผนภูมิที่

แผนภูมิแบบบ็อกซ์, ไวโอลิน, แผนภูมิแถบ และแผนภูมิซินาเหมาะสำหรับการนำเสนอข้อมูลที่มีการกระจายตัวของข้อมูลมาก แผนภูมิแบบเนินเป็นอีกหนึ่งแผนภูมิที่สามารถใช้แทนกันกับแผนภูมิไวโอลินได้ และยังเป็นแผนภูมิที่ได้รับความนิยมมากอีกด้วยเช่นกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เพราะสามารถนำเสนอข้อมูลที่จำนวนการกระจายของข้อมูลได้มากหรือข้อมูลที่มีการกระจายตัวอยู่ตลอดเวลา



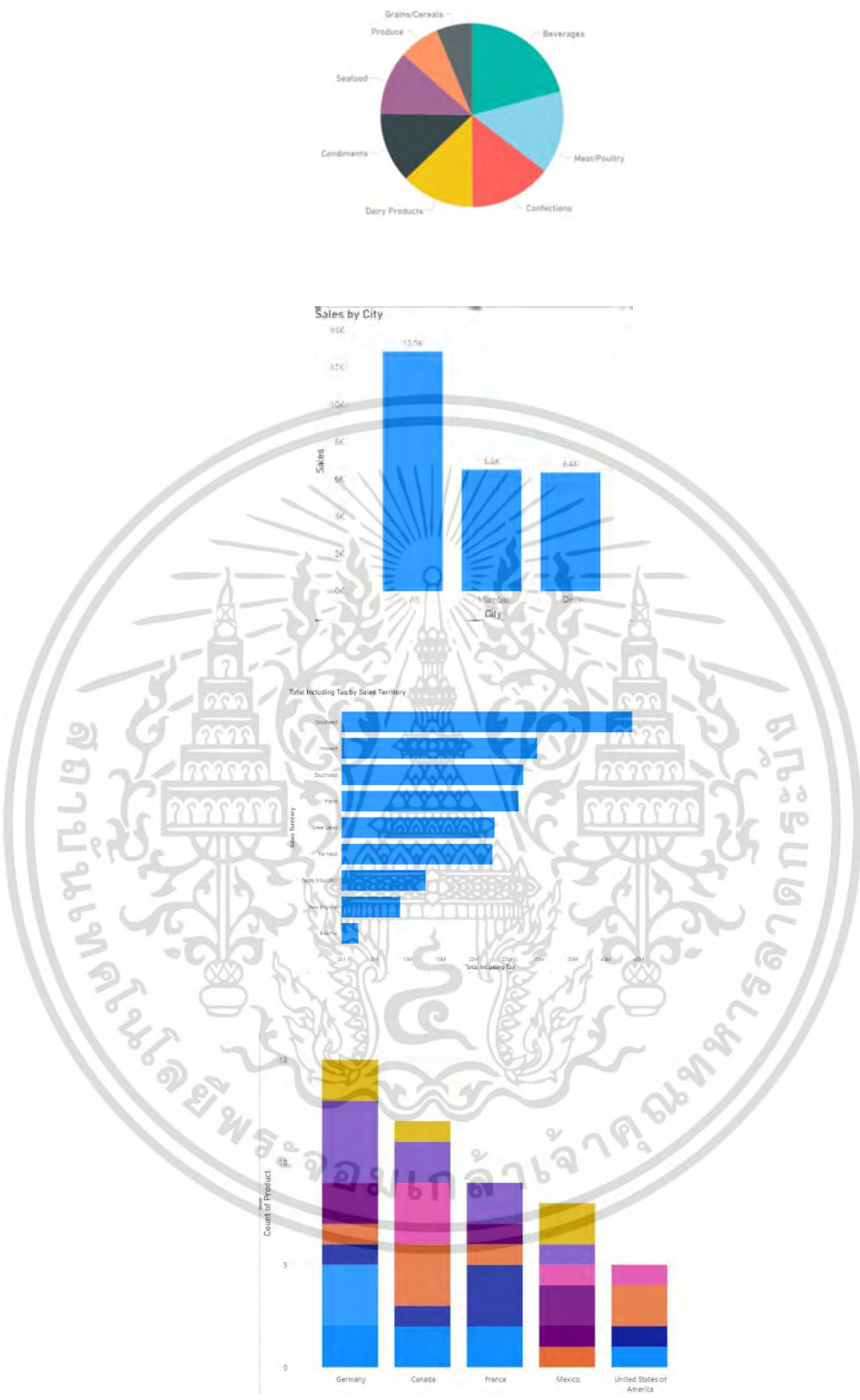
รูปที่ 2.4 แผนภูมิแสดงการกระจายตัวของข้อมูล

จากรูปที่ 2.4 เป็นตัวอย่างแผนภูมิแสดงการกระจายตัวของข้อมูล ประกอบไปด้วยแผนภูมิแบบบ็อกซ์ แผนภูมิแบบไวโอลิน และแผนภูมิซินา

3) แผนภูมิแสดงสัดส่วน

แผนภูมิแสดงสัดส่วนสามรถนำเสนอข้อมูลได้ในรูปแบบของ แผนภูมิวงกลม (Pie Chart), แผนภูมิแท่งเปรียบเทียบ (Side-by-Side), หรือแผนภูมิแบบซ้อนกัน (Stackedbar) แผนภูมิแบบแท่งสามารถจัดรูปแบบในอยู่ในรูปแบบแนวตั้งหรือแนวนอนก็ได้ แผนภูมิวงกลมจะเป็นการนำเสนอข้อมูลทั้งหมดและเน้นส่วนที่จะต้องการนำเสนอ แต่อย่างไรก็ตามแผนภูมิวงกลมสามารถเปรียบเทียบข้อมูลได้ง่ายกว่าแผนภูมิแบบแท่งเปรียบเทียบ แผนภูมิแบบซ้อนกันมีลักษณะที่หนาแน่นมากเกินไป ไม่เหมาะสมที่จะนำมานำเสนอข้อมูลแบบเดี่ยว มีจำนวนของชุดข้อมูลไม่มากแต่แผนภูมิชนิดนี้เหมาะสำหรับนำมาใช้ได้ในการนำเสนอข้อมูลที่มีจำนวนมาก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



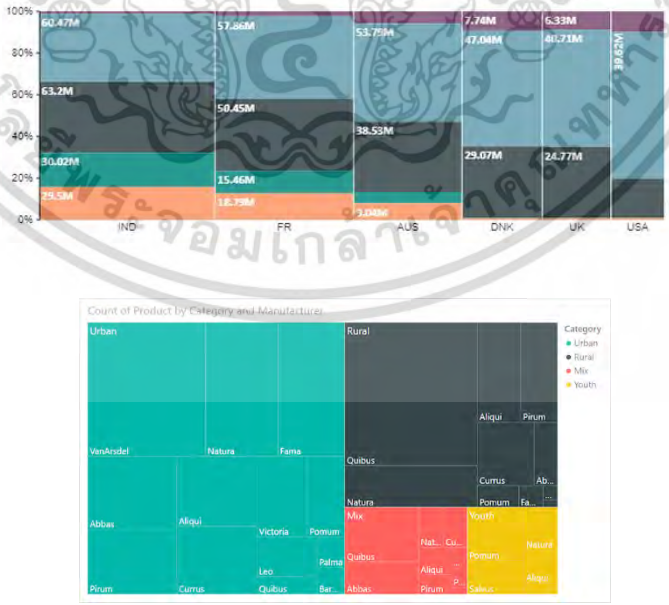
รูปที่ 2.5 แผนภูมิแสดงการกระจายตัวของข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 2.5 เป็นตัวอย่างแผนภูมิแสดงการกระจายตัวของข้อมูล ประกอบไปด้วยแผนภูมิวงกลม (Pie Chart), แผนภูมิแท่งเปรียบเทียบ (Side-by-Side), หรือแผนภูมิแบบซ้อนกัน (Stackedbar) แผนภูมิแบบแท่ง

แผนภูมิแท่ง เป็นการนำเสนอเพื่อเน้นแนวเปรียบเทียบของแต่ละอย่าง โดยพยายามสร้างกราฟที่เรียบง่าย และเข้าใจได้ง่าย โดยสามารถเปรียบเทียบได้มากกว่า 1 ด้าน (ใน Excel จะเรียกว่า Series) เป็นกราฟที ใช้สำหรับเปรียบเทียบข้อมูล หากเปรียบเทียบข้อมูลคนละประเภทกัน ควรใช้คนละแกน เช่นการเปรียบเทียบยอดขายกับกำไรกราฟแท่งแบบต่อกัน เป็นการนำเสนอเพื่อให้เห็นการรวมกันของข้อมูลเพื่อนำมาต่อ ๆ กัน ว่าได้เท่าใดไม่ได้ต้องการนำข้อมูลแต่ละด้าน (หรือ Series) มาเปรียบเทียบกัน แต่ต้องการนำทุก ๆ Series มารวมกัน กราฟชนิดนี้เหมาะสำหรับการนำยอดรวมมาเปรียบเทียบกัน ข้อมูลที่น ามาต่อกันควรเป็นข้อมูลในหน่วยเดียวกันแผนภูมิวงกลม ใช้สำหรับการเปรียบเทียบสัดส ่วนของข้อมูลโดยเทียบกับผลรวมของทุกๆ ข้อมูลรวมกัน เพื่อเปรียบเทียบเชิงสัดส่วน เช่น ส่วนแบ่งทางการตลาด (Market Share) เป็นต้น โดยกราฟที่อยู่ในกลุ่มนี้ที่เลือกใช้ได้ เช่น Donut, Tree Map เป็นต้น กราฟวงกลมเน้นดูสัดส่วนเทียบกับยอดรวม การใช้กราฟชนิดนี้ไม่ควรใช้ในการเปรียบเทียบข้อมูลเกิน 10รายการ เพราะจะทำให้ดูยากหรือรายละเอียดเยอะเกินไป

แผนภูมิโมเสค, แผนภูมิต้นไม้ หรือ แผนภูมิแบบขนาน เหมาะสำหรับการนำเสนอข้อมูลในรูปแบบตัวแปรหลายกลุ่ม แผนภูมิโมเสคสามารถนำเสนอข้อมูลรวมกันได้หลายกลุ่ม แต่ในขณะที่แผนที่ต้นไม้ไม่สามารถนำเสนอข้อมูลลักษณะนั้นได้



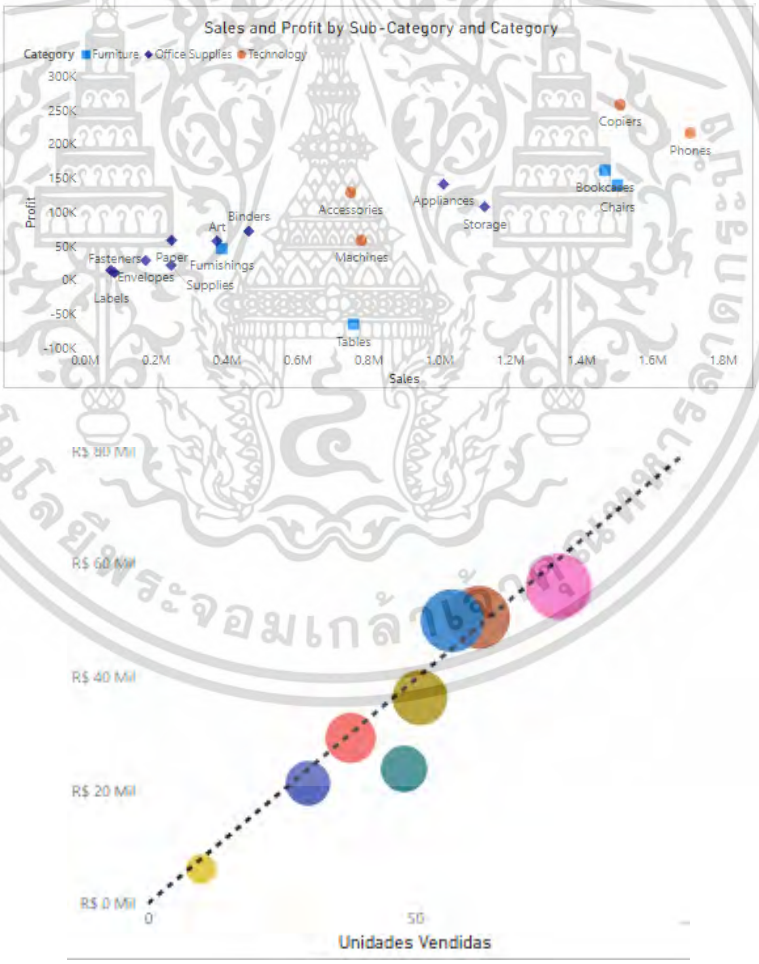
รูปที่ 2.6 แผนภูมิแสดงสัดส่วน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 2.6 เป็นตัวอย่างแผนภูมิแสดงสัดส่วน ประกอบไปด้วย แผนภูมิโมเสค และแผนภูมิต้นไม้

4) แผนภาพแสดงความสัมพันธ์แกน X และ แกน Y

แผนภูมิแบบกระจาย (Scatter Plot) เป็นต้นแบบในการนำเสนอข้อมูลที่ตัวแปรแบบเชิงปริมาณที่มีความสัมพันธ์กับตัวแปรอื่น ๆ แต่ถ้าเรามีตัวแปรของข้อมูล 3 ตัวแปร เราสามารถนำตัวแปรมาสร้างเป็นจุดและเพิ่มขนาดของจุดตามขนาดของตัวแปรนั้นได้ โดยที่เรียกแผนภาพดังกล่าวนั้นว่า แผนภูมิแบบบับเบิล (Bubble Chart) สำหรับการจับคู่ข้อมูล จะใช้การจับคู่ด้วยการใช้แกน X และ แกน Y เพื่อวัดค่าที่เหมือนกันในแต่ละหน่วย ข้อมูลที่จับคู่สามารถแสดงเป็นกราฟความชันของจุดที่จับคู่ที่เชื่อมต่อด้วยเส้นตรง เป็นกราฟที่สามารถให้ข้อมูลได้ถึง 3-4 มิติด้วยกัน ทั้งแกน X แกน Y และ ขนาด ของวงกลมด้วย และสีของวงที่อาจจะไม่เหมือนกันก็ได้

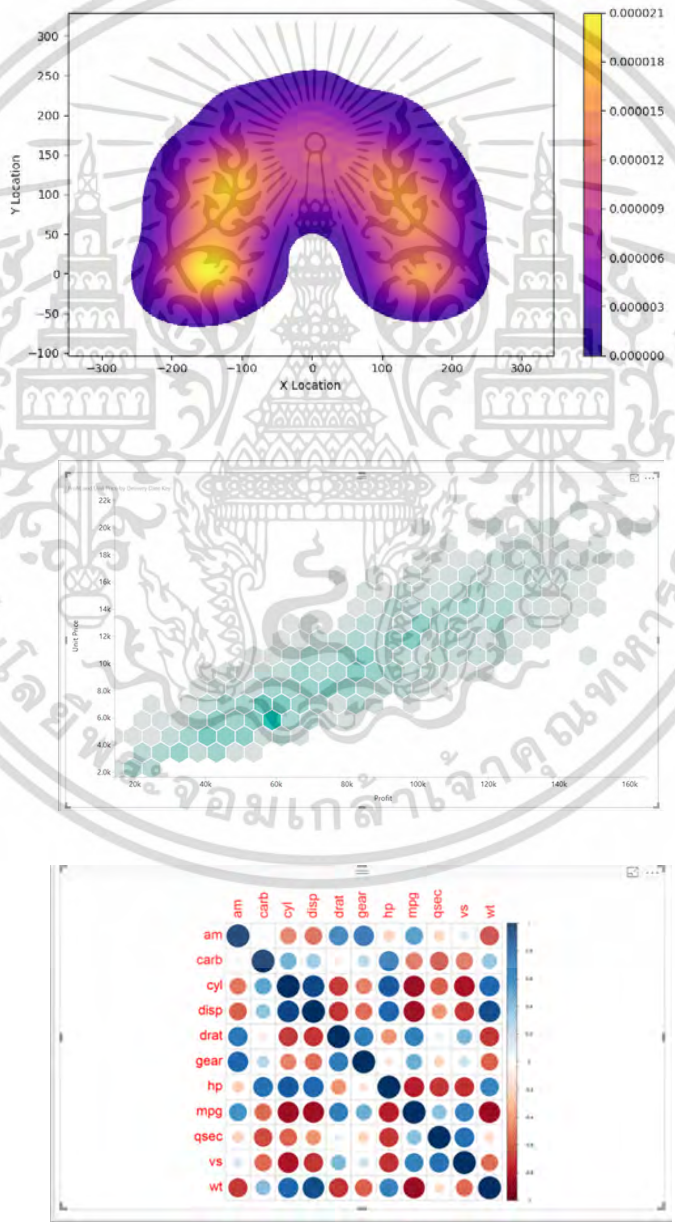


รูปที่ 2.7 แผนภูมิแสดงความสัมพันธ์ของข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 2.7 เป็นตัวอย่างแผนภูมิแสดงความสัมพันธ์ของข้อมูล ประกอบไปด้วยแผนภูมิแบบกระจาย (Scatter Plot) และแผนภูมิแบบบับเบิล (Bubble Chart)

สำหรับการนำเสนอข้อมูลที่มีจำนวนมาก แผนภาพกระจายไม่สามารถสื่อความหมายของข้อมูลได้ ในลักษณะนี้ควรใช้แผนภาพเส้นชั้นความสูง (Density Contour), แผนภาพแบบเส้นสองมิติ หรือแบบแผนภาพรังผึ้งเป็นทางเลือกในการนำเสนอข้อมูลที่ดีกว่า เพราะแผนภาพชนิดนี้สามารถแสดงข้อมูลได้ตั้งแต่สองชุดขึ้นไป แต่ในทางกลับกันเราสามารถเลือกใช้แผนภาพสหสัมพันธ์ ที่เป็นแผนภาพที่ใช้ความสัมพันธ์ของข้อมูลแบบสหสัมพันธ์ (Correlation)

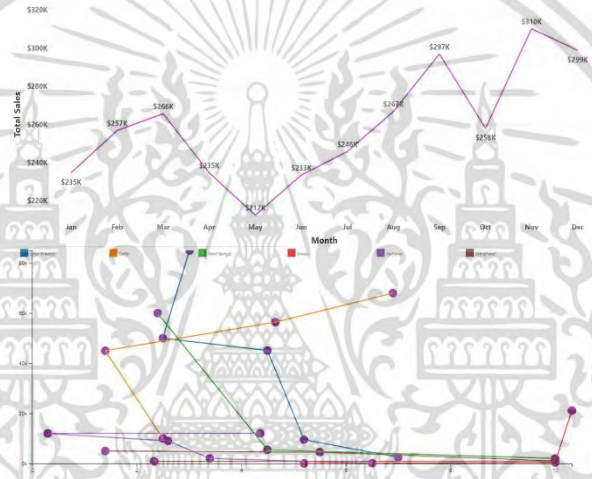


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 2.8 แผนภูมิแสดงความสัมพันธ์ของข้อมูล

จากรูปที่ 2.8 เป็นตัวอย่างแผนภูมิแสดงความสัมพันธ์ของข้อมูล ประกอบไปด้วยแผนภาพเส้นชั้นความสูง (Density Contour) แผนภาพรังผึ้ง และแผนภาพสหสัมพันธ์

การเปรียบเทียบข้อมูลเพื่อดูแนวโน้ม (Trends) โดยอาจจะเทียบกับมิติของเวลา (Time Series) เช่น กราฟยอดขายรายไตรมาส กราฟจำนวนลูกค้าใหม่ในแต่ละเดือน เป็นต้น กราฟในกลุ่มนี้ จะมี แผนภูมิเส้น (Line Chart) , แผนภูมิแสดงพื้นที่ (Area Chart) ซึ่งจะมีทั้งแบบ Cluster และ แบบ Stack กราฟชนิดนี้ใช้ดูแนวโน้มของข้อมูล แผนภูมิเส้นสามารถกำหนดแนวโน้ม (Trend line) ได้แต่ต้องเลือกความสัมพันธ์ให้เหมาะสม เช่น ความสัมพันธ์เชิงเส้น (Linear), ความสัมพันธ์แบบเลขชี้กำลัง (Exponential) หรือ Power เป็นต้น



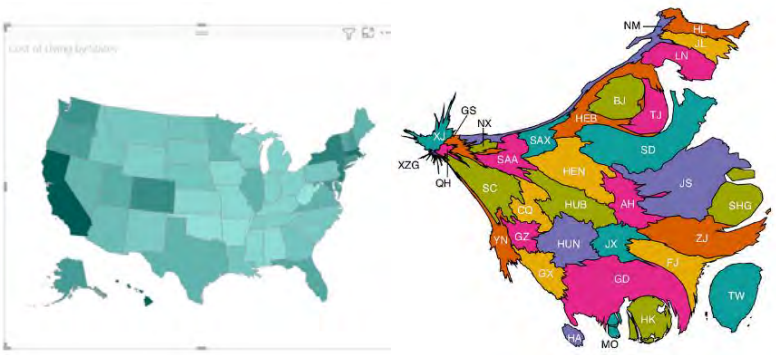
รูปที่ 2.9 แผนภูมิแสดงความสัมพันธ์ของข้อมูล

จากรูปที่ 2.9 เป็นตัวอย่างแผนภูมิแสดงความสัมพันธ์ของข้อมูล ประกอบไปด้วยแผนภูมิเส้น (Line Chart) และแผนภูมิแสดงพื้นที่ (Area Chart)

5) แผนภูมิภูมิสารสนเทศเชิงพื้นที่

แผนภาพลักษณะนี้จะแสดงเป็นข้อมูลออกมาในรูปแบบของแผนที่ แผนภาพลักษณะนี้จะไปสัมพันธ์กับพื้นผิวโลก ตัวอย่างเช่น พื้นผิวและระยะทางโดยเฉลี่ย นำเสนอในรูปแบบของภาพสองมิติ นอกจากนี้เรายังสามารถแสดงความแตกต่างของข้อมูลด้วยการใส่สีลงไปในแผนที่ ตัวอย่างเช่น แผนที่โคโรเพลทที่แสดงความแตกต่างของข้อมูล หรือแผนที่ที่แสดงข้อมูลในรูปแบบของรูปทรง เช่น แผนภูมิคาร์โตแกรม (Cartogram) เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.10 แผนภูมิสารสนเทศ

จากรูปที่ 2.10 เป็นตัวอย่างแผนภูมิสารสนเทศ ประกอบไปด้วยแผนภูมิแผนภูมิคาร์โตแกรม (Cartogram)

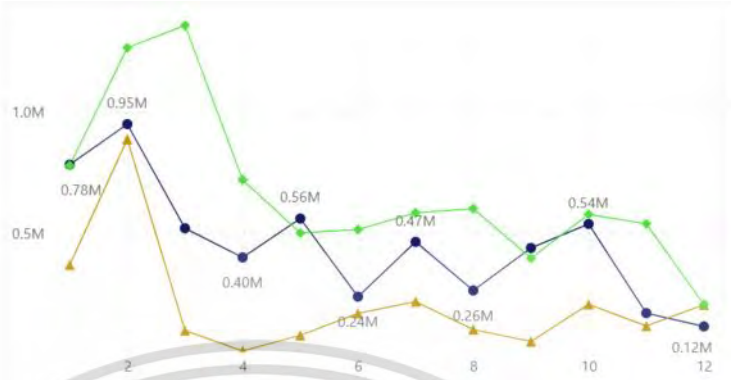
ในหัวข้อลักษณะการนำเสนอข้อมูลในรูปแบบต่าง ๆ นั้น เป็นสิ่งที่จำเป็นสำหรับในงานธุรกิจอัจฉริยะมาก เพราะจะทำให้ทุกคนที่เกี่ยวข้องกับข้อมูลที่นำมาวิเคราะห์นั้นสามารถเห็นข้อมูลที่ซ่อนอยู่ เพื่อนำข้อมูลที่ได้มาใช้ประโยชน์ให้ได้มากที่สุด การเลือกรูปแบบในการนำเสนอ การใช้สีและรายละเอียดต่างๆก็สำคัญมากในงานการนำเสนอข้อมูล ผู้ใช้ควรพิจารณาจากข้อมูลที่มีอยู่และเลือกใช้แผนภูมิให้เหมาะสมกับการนำเสนอข้อมูล

2.3.1.2 การเลือกใช้สีและตัวอักษรในการนำเสนอข้อมูล

นอกจากวิธีการนำเสนอข้อมูลให้เหมาะสมกับข้อมูลที่ได้เพื่อให้ผู้เกี่ยวข้องสามารถนำข้อมูลที่ผ่านการวิเคราะห์จากข้อมูลดิบไปใช้ประโยชน์ตามวัตถุประสงค์หรือตามเป้าหมายที่ต้องการได้แล้ว การเลือกใช้สีที่เหมาะสมก็เป็นอีกหนึ่งปัจจัยที่จะทำให้สามารถมองเห็นข้อมูลได้อย่างชัดเจน เป็นที่สังเกตได้ง่าย สื่อความหมายได้ตรงประเด็น การใช้สีในการนเสนอข้อมูลควรมีลักษณะดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1) การใช้สีแบบโค้ดดิง รีดันแดนท์ (Coding Redundant)



รูปที่ 2.11 ตัวอย่างการใช้สีแบบโค้ดดิงรีดันแดนท์

ในรูปที่ 2.11 รูปทางซ้ายมือใช้สีเพียงอย่างเดียว แต่รูปทางขวามือใช้ทั้งสีและรูปทรงต่างๆทำให้ข้อมูลแต่ละชุดมีความแตกต่างกัน ทำให้สามารถสังเกตได้ง่ายขึ้น

- รูปทรง สัญลักษณ์ที่มีขนาดใหญ่กว่าเส้นทำให้สามารถจำแนกสีได้ง่าย
- กำหนดรูปทรง สัญลักษณ์ให้แตกต่างกัน ตัวอย่างในรูปทางขวาที่มีการใช้วงกลม สามเหลี่ยม สี่เหลี่ยม ในการจำแนกข้อมูลไม่ให้สับสน

หลังจากที่ออกแบบรูปแบบการนำเสนอได้รูปแบบอย่างที่กล่าวมาก่อนหน้านี้แล้ว มาถึงขั้นตอนการเลือกใช้สีในการนำเสนอให้อ่านง่ายมากขึ้น สีไม่ใช่ปัจจัยเดียวที่นักออกแบบต้องใส่ใจเลือกให้เหมาะสมกับผู้ที่ใช้ การเลือกใช้สี รูปทรงและสัญลักษณ์ที่แตกต่างกันช่วยทำให้อ่านง่ายขึ้น ใช้เวลาในการทำความเข้าใจน้อยลงและนักออกแบบยังสามารถนำวิธีการต่างๆนำไปประยุกต์ใช้ได้กับทุกคนอีกด้วย หลักการมีทั้งหมด 3 ข้อดังนี้

1) หลีกเลี่ยงที่จะใช้สีในการเน้นข้อมูลเพียงอย่างเดียว
2) หลีกเลี่ยงการใช้ข้อความหรือวัตถุที่มีสีกลมกลืนกับพื้นหลัง
3) ใช้ตัวอักษรและสัญลักษณ์ให้หนาหรือใหญ่ สำหรับการเลือกใช้รูปแบบตัวอักษรควรใช้รูปแบบ แอเรียล (Arial) หรือ เฮลเวทิกา (Helvetica) หรือควรเลือกใช้รูปแบบที่ไม่เล็กกว่ารูปตัวอักษรแบบ ไทมส์ (Times) หรือนิวยอร์ก (New York)

ตัวอักษร หรือ ฟอนต์สามารถแบบออกได้เป็น 2 ประเภทใหญ่ๆ คือ แบบมีเชิง (Serif) และแบบไม่มีเชิง (San Serif) มีรายละเอียดดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1) ตัวอักษรแบบมีเชิง (Serif)

เป็นตัวอักษรที่มีเส้นยื่นของฐานและปลายตัวอักษรในทางราบที่เรียกว่า Serif ลักษณะของตัวอักษรจะมีเส้นตัวอักษรเป็นแบบหนา บางไม่เท่ากัน เหมือนการเขียนประดิษฐ์ด้วยขนนก หรือปากกาปากแบน มีหลายรูปแบบแตกต่างกันไป ตัวอักษรแบบนี้ ให้ความรู้สึกถึงความเก่า ความขลัง จึงมักใช้กับงานที่เป็นทางการ กึ่งไปทางพิธีรีตองหรือเรื่องราวในเชิงอนุรักษ์นิยม ถ้าเป็นงานพิมพ์ ก็จะใช้เป็นตัวพาดหัวเรื่องมากกว่าที่เป็นเนื้อหาให้อ่าน

2) ตัวอักษรแบบไม่มีเชิง (San Serif)

เป็นลักษณะของตัวอักษรที่มีรูปแบบเรียบง่าย ต่างจากแบบแรก คือ ไม่มีเชิง หมายถึงไม่มีเส้นยื่นออกมาจากปลายฐานและปลายของตัวอักษรในทางราบ เป็นตัวอักษรที่นิยมเป็นอย่างมาก เพราะเรียบง่าย ทันสมัย เหมาะกับการนำไปใช้ออกแบบหลายชนิด ในหนังสือมักนำไปใช้วางเป็นเนื้อหาให้อ่าน เพราะมีรูปร่างที่อ่านง่าย โดยทั่วไปนิยมใช้ในงานพิมพ์ทั่วไปและงานประชาสัมพันธ์ ฟอนต์แบบไม่มีเชิงได้แก่ คาลิบรี (Calibri), แอเรียล (Arial), เวอร์ดาน่า(Verdana), เซนจูรี่ กอธิค (Century Gothic)

รูปที่ 2.12 แผนภูมิภูมิสารสนเทศ

จากรูปที่ 2.12 ตัวอักษรภาษาอังกฤษที่นิยมใช้และมีความเหมาะสมกับสถานการณ์ต่างๆ มีดังต่อไปนี้

- คาลิบรี (Calibri) เป็นอักษรที่มีขนาดเล็ก ดูภูมิฐาน และดูเป็นงานเป็นการ
- แอเรียล (Arial) เป็นสีที่ใช้ได้ในเกือบทุกสถานการณ์
- เวอร์ดาน่า (Verdana) เป็นตัวอักษรที่ดูสะอาดตา และ พบเห็นได้บ่อยแต่น้อยกว่า แอเรียล (Arial) เพียงเล็กน้อย แต่มีความคลาสสิกอยู่ในตัว
- เซนจูรี่ กอธิค (Century Gothic) แบบอักษรดั้งเดิม ดูทันสมัย ภูมิฐานและดูมีความมั่นใจ

ข้อความภาษาไทยควรเลือกใช้ฟอนต์สำหรับภาษาไทยโดยเฉพาะ เนื่องจากให้ความสวยงามมากกว่าฟอนต์ที่ถูกออกแบบมาสำหรับการพิมพ์ภาษาอังกฤษ ตัวอย่างเช่น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การใช้ฟอนต์คาลิบรี (Calibri) กับข้อความที่เป็นภาษาอังกฤษและเป็นมาใช้ฟอนต์คณิต (Kanit) เมื่อเป็นข้อความที่เป็นภาษาไทย

หลักการพื้นฐานการออกแบบสัดส่วนของการนำเสนอข้อมูลมีทั้งหมด 3 ประการ คือ

1) สีคือเครื่องมือในการจำแนก

ผู้ใช้สามารถใช้สีในการแยกแยะข้อมูลที่แตกต่างกัน เช่น การใช้สีเพื่อแยกประเทศแต่ละประเทศออกจากกัน หรือการใช้สีเพื่อแสดงถึงผลิตภัณฑ์ต่างๆที่ผลิตในโรงงาน เป็นต้น

2) สีคือการนำเสนอข้อมูล

สีเป็นสิ่งที่เกี่ยงข้องกับการนำเสนอข้อมูล เช่น การนำเสนอข้อมูลที่เกี่ยวกับรายได้, อุณหภูมิ หรือความเร็ว ในกรณีนี้ จะเลือกใช้สีที่มีเฉดสีที่มีความต่อเนื่องกัน ชุดสีที่มีความต่อเนื่องกันมีที่มาจากสี (ยกตัวอย่างเช่น การไล่เฉดสีน้ำเงินไปถึงเฉดสีฟ้า หรือมากกว่านั้น ตัวอย่างเช่น การไล่เฉดจากสีแดงเข้มไปจนถึงสีเหลืองอ่อน ดังตัวอย่างในรูปที่ 2.13

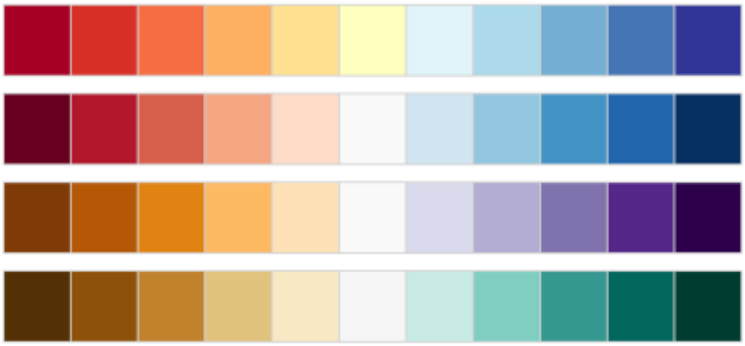
รูปที่ 2.13 ตัวอย่างของชุดสีที่มีความต่อเนื่องของเฉดสี

จากรูป 2.13 แสดงตัวอย่างของชุดสีที่มีความต่อเนื่องของเฉดสีเพื่อบอกลักษณะของข้อมูลแต่ละประเภท

3) สีคือเครื่องมือในการเน้นข้อความหรือข้อมูลที่สำคัญ

สีเป็นอีกหนึ่งเครื่องมือที่ใช้ในการเน้นข้อมูลที่สำคัญ สีสามารถจำแนกประเภทชุดของข้อมูลที่สำคัญๆ โดยการเน้นข้อมูลที่ต้องการนำเสนอมักจะใช้สีสดใสในการนำเสนอเพื่อให้ข้อมูลมีความโดดเด่น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.14 ตัวอย่างชุดสีในแบบต่าง ๆ

จากรูป 2.14 แสดงตัวอย่างของชุดตัวอย่างชุดสีในแบบต่าง ๆ ที่ใช้ในการเน้นข้อมูลที่สำคัญ สีสามารถจำแนกประเภทชุดของข้อมูล

นอกจากการใช้สีให้มีความเหมาะสมแล้ว การใช้สีก็ยังมีการใช้ในลักษณะการสื่อความหมาย โดยการใช้สีเพื่อสื่อความหมายมีทั้งหมด 5 ประเภท ดังนี้

1) การใช้สีแสดงความถี่ (Sequential Color)

เป็นการใช้เพียง 1 สีโดยไล่ลำดับความเข้มของสีจากเข้มที่สุดไปยังความเข็มน้อยที่สุด ตัวอย่างการใช้ เช่นการแสดงจำนวนยอดขายสินค้าต่อพื้นที่จังหวัด โดยสีเข้มที่สุดแทนด้วยพื้นที่ที่มียอดขายสูงสุดและสีเข้มน้อยที่สุดแทนด้วยยอดขายที่ต่ำที่สุด

2) การใช้สีแสดงความแตกต่าง (Diverging Color)

ในส่วนนี้จะคล้ายกับ การใช้สีแสดงความถี่ (Sequential Color) แต่ จะแตกต่างโดยใช้สำหรับเปรียบเทียบข้อมูลที่แบ่งออกเป็น 2 กลุ่มโดยเริ่มจากจุดศูนย์กลาง (Midpoint)ตัวอย่างเช่นการใช้วัดผลของข้อมูลระหว่างด้านบวกและด้านลบ หรือจำนวนข้อมูลจาก 2 กลุ่มเปรียบเทียบกัน เช่นผลนับคะแนนเลือกตั้งระหว่าง 2 พรรคการเมือง หรือการแสดงผลของอุณหภูมิจากร้อนที่สุดไปหาเย็นที่สุดโดยจุดศูนย์กลางคือค่าของอุณหภูมิโดยเฉลี่ย

3) **การใช้สีเพื่อการจัดกลุ่ม แบ่งประเภท (Categorical Color)**

ใช้แทนการใช้สีเพื่อแยกความแตกต่างของข้อมูลในแต่ละประเภทออกจากกันให้ชัดเจน ดังนั้นจึงสามารถใช้สีอะไรก็ได้แทนประเภทของข้อมูลแต่ละชนิด เช่น ตัวอย่างการใช้สีแทนด้วยชื่อจังหวัดต่างๆ

4) การใช้สีเพื่อเน้นข้อมูลสำคัญ (Highlight Color)

ใช้ในกรณีที่ต้องการเน้นข้อมูลที่ต้องการนำเสนอเป็นพิเศษ โดยการแทนข้อมูลด้วยสีเดียวกันทั้งหมดยกเว้นข้อมูลที่ต้องการให้เน้นเป็นพิเศษ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5) การใช้สีเพื่อแสดงความผิดปกติ (Alerting Color)

คล้ายกับการใช้สีเพื่อเน้นข้อมูลสำคัญ (Highlight Color) แต่ใช้ในกรณีที่ต้องการแสดงข้อมูลที่เน้นให้เห็นถึงความผิดปกติ เพื่อวัตถุประสงค์สำหรับการแจ้งเตือน โดยเน้นการใช้สีแดงแทนเหตุที่อันตรายหรือใช้สีส้มเป็นสัญลักษณ์ในการแจ้งเตือน

เมื่อเริ่มสร้าง Data Visualization ในการนำข้อมูลไปประมวลผลนั้นมีหลากหลายวิธีการ ดังนั้นควรมีความเข้าใจเกี่ยวกับตัวชี้วัดและลักษณะของข้อมูลที่รวบรวมมา ยกตัวอย่างเช่น ข้อมูลเชิงตัวเลข (Numeric) ข้อมูลหมวดหมู่ (Categorical) และข้อมูลที่จัดเรียงลำดับเวลา (Time Series) เพื่อป้องกันไม่ให้การวิเคราะห์เกิดข้อผิดพลาด หรือคลาดเคลื่อนจากชุดข้อมูลที่เลือกมาใช้งานและสามารถนำเสนอได้อย่างแม่นยำ ประโยชน์ที่ได้จากการทำ Data Visualization นั้นไม่ได้มีเพียงการอธิบายข้อมูลให้ง่ายขึ้นเพียงอย่างเดียว แต่ยังช่วยให้สามารถคาดการณ์แนวโน้ม เปรียบเทียบความแตกต่าง และหาความสัมพันธ์ที่เชื่อมโยงกันของชุดข้อมูลแม้กระทั่งการคาดเดาพฤติกรรมต่างๆ ของลูกค้า คู่แข่งของธุรกิจ หรือแม้แต่สภาวะตลาดในอนาคต ทำให้ข้อมูลที่เก็บรวบรวมมานั้นมีคุณค่ามากกว่าเป็นแค่เพียงชุดตัวเลข ซึ่งเครื่องมือที่คนส่วนใหญ่นึกถึงอาจจะเป็นการคำนวณข้อมูลในโปรแกรม MS-Excel ออกมาในรูปของแผนภูมิหรือกราฟต่าง ๆ

การสรุปข้อมูลและแสดงออกมาเป็นภาพมาจากการทำ Information Visualization และ Visual Analytics โดยแสดงผลลัพธ์ให้อยู่ในรูปของแผนภูมิและกราฟแบบต่าง ๆ แม้กระทั่งแผนที่ โดยมีจุดประสงค์เพื่อให้สามารถเข้าใจง่ายอธิบายได้ชัดเจน และเห็นภาพรวมของผลลัพธ์ที่ได้ แม้การทำ Data Visualization มักจะพบในรายงานการวิจัยต่าง ๆ แต่ยังไม่ค่อยได้นำมาประยุกต์ใช้เพื่อนำเสนอในด้านของธุรกิจ ซึ่งในปัจจุบันธุรกิจส่วนใหญ่ได้ปรับตัวเข้าสู่การเปลี่ยนแปลงทางดิจิทัล (Digital Transformation) มากขึ้น ทำให้ธุรกิจครอบครองข้อมูลในจำนวนมหาศาล ซึ่งการนำข้อมูลไปใช้ประโยชน์ด้วยการเล่าข้อมูลออกมาเป็นรูปภาพนั้น จะช่วยข้อมูลที่วิเคราะห์ออกมานั้นชัดเจนไม่ซับซ้อน และทำให้มีประสิทธิภาพมากยิ่งขึ้น ที่สำคัญยังสามารถนำไปต่อยอดในการทำ Data Storytelling

2.3.1.2 Data-Ink Ratio

Data-Ink Ratio หมายถึงอัตราส่วนของ "หมึก" ที่ใช้ในการแสดงข้อมูลที่สำคัญ (data ink) เมื่อเทียบกับ "หมึก" ทั้งหมดที่ใช้ในกราฟหรือแผนภูมิ (total ink) ซึ่งรวมถึงทั้งหมึกที่ใช้ในการแสดงข้อมูลและหมึกที่ใช้ในการตกแต่งหรือส่วนประกอบอื่น ๆ ที่ไม่เกี่ยวข้องกับข้อมูลโดยตรง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หลักการของ Data-Ink Ratio

- เพิ่ม Data-Ink: ใช้หมึกในการแสดงข้อมูลที่มีความสำคัญและมีความหมายให้มากที่สุด
- ลด Non-Data-Ink: ลดหรือกำจัดหมึกที่ใช้ในการตกแต่ง, กรอบ, เส้นกริด หรือองค์ประกอบอื่น ๆ ที่ไม่เกี่ยวข้องกับข้อมูลที่แสดง
- เน้นความชัดเจน: การแสดงผลควรเน้นความชัดเจนและการตีความข้อมูลได้ง่าย ไม่ทำให้ผู้ใช้สับสนหรือเสียเวลาในการเข้าใจข้อมูล
- ลดสิ่งรบกวน: องค์ประกอบที่ไม่จำเป็นหรือซับซ้อนเกินไปควรถูกกำจัดออกไปเพื่อลดการรบกวนการแสดงผลข้อมูล

การสร้าง Visualizations หรือ Dashboards ที่ทำการประเมิน ink ratio สามารถช่วยให้การนำเสนอข้อมูลมีประสิทธิภาพและสื่อความหมายได้ดีขึ้น ตัวอย่างการปรับปรุงได้แก่ การเลือก ใช้กราฟหรือแผนภูมิที่แสดงข้อมูลได้ชัดเจนและตรงประเด็นที่สุด การออกแบบที่เรียบง่าย ลดการใช้สีที่ไม่จำเป็น, การใช้กริดที่ไม่จำเป็น, และการตกแต่งที่ซับซ้อน และจัดเรียงข้อมูลให้เป็นระเบียบและสื่อความหมายได้ชัดเจน

2.1.4 แดชบอร์ด (Dashboard)

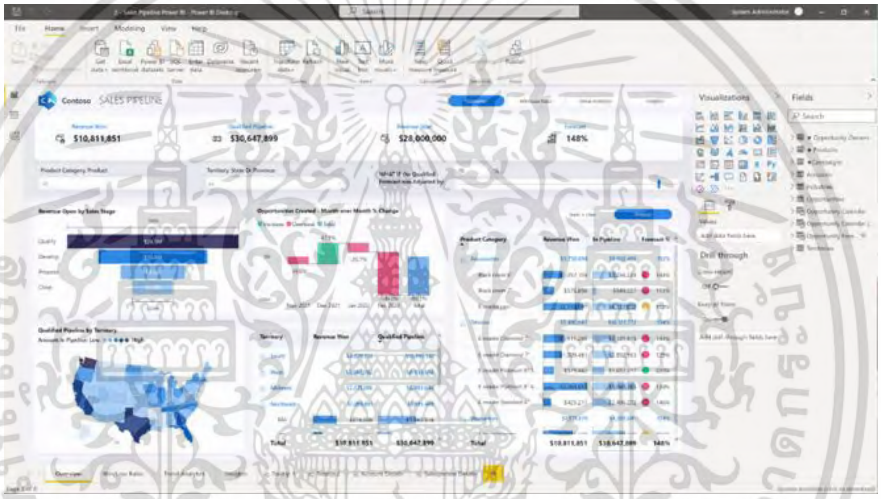
แดชบอร์ด (Dashboard) เป็นการนำเสนอข้อมูลตัวชี้วัดสำคัญต่างๆ โดยการแปลงข้อมูลเหล่านี้ให้เป็นภาพด้วย Data Visualization ที่ช่วยให้ผู้ใช้สามารถรวบรวมผลการวิเคราะห์มาเสนอบนแดชบอร์ด ได้เข้าใจง่ายยิ่งขึ้น ซึ่งช่วยให้องค์กรสามารถติดตามความคืบหน้าในการบรรลุเป้าหมายต่างๆ ได้ แดชบอร์ดมักประกอบด้วยการนำเสนอข้อมูลในรูปแบบต่างๆ ที่มีจุดประสงค์ให้ผู้ใช้สามารถมองเห็นภาพรวมของตัวชี้วัดสำคัญได้อย่างรวดเร็วและง่าย โดยแดชบอร์ดสามารถสร้างได้จากโปรแกรมสำเร็จรูปสำหรับงานสเปรดชีต โปรแกรมเฉพาะทาง หรือภาษาคอมพิวเตอร์ เช่น ไพธอนและอาร์ แดชบอร์ดมีหลายประเภท เช่น แดชบอร์ดสำหรับผู้บริหาร แดชบอร์ดสำหรับการปฏิบัติงาน และแดชบอร์ดสำหรับการวิเคราะห์ข้อมูล ในปัจจุบันมีซอฟต์แวร์ต่างๆ ที่สามารถช่วยผู้ใช้แสดงผลลัพธ์ในรูปแบบแผนภาพ หรือแผนภูมิต่าง ๆ ได้สะดวกและรวดเร็วมากขึ้นได้แก่โปรแกรม Microsoft Power BI โปรแกรม Tableau และ โปรแกรม Google Data Studio (Looker Studio) เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.1.4.1 โปรแกรมในการทำแดชบอร์ด (Dashboard)

2.4.4.1.1 Microsoft Power BI

โปรแกรม Microsoft Power BI เป็นเครื่องมือ BI ของ Microsoft ซึ่งมีอินเทอร์เฟซที่เรียบง่ายพร้อม Visualization แบบ Interactive ช่วยให้ผู้ใช้งานสามารถสร้างรายงานและแดชบอร์ดของตนเองได้ Power BI แตกต่างจากเครื่องมืออื่นตรงที่ว่ามันสามารถจัดการพร้อมทั้งประมวลผลข้อมูลจำนวนมากได้แบบเรียลไทม์ และยังสามารถเชื่อมต่อที่ทำงานเพื่อเปลี่ยนแหล่งข้อมูลที่ไม่เกี่ยวข้อง ให้เป็นข้อมูลเชิงลึกที่สอดคล้องกัน แสดงข้อมูลได้ และโต้ตอบได้โดยข้อมูลอาจเป็นสเปรดชีต Excel หรือจากของระบบ Cloud ข้อมูลภายในองค์กร เพื่อสร้างรายงานและแดชบอร์ดเฉพาะสำหรับองค์กรและเก็บไว้ในเซิร์ฟเวอร์ได้



รูปที่ 2.15 ภาพโปรแกรม Microsoft Power BI

จากรูปที่ 2.15 แสดงภาพการทำงานของโปรแกรม Microsoft Power BI ในการ Visualization รูปแบบต่าง และสามารถบอกข้อดีและข้อเสียของโปรแกรม Microsoft Power BI ได้ดังนี้

ข้อดี

1) รองรับข้อมูลได้หลายรูปแบบเช่น ฐานข้อมูล (Database), ไฟล์ (File) และระบบต่าง ๆ อย่าง Google Analytics, Facebook และ Twitter เป็นต้น

2) สามารถแสดงผลของรายงานได้หลากหลายรูปแบบ อันประกอบด้วย Web, Apps โดยสามารถแสดงผลได้ทั้งบน Desktop, Tablet, Mobile ทำให้เกิดความสะดวกสบายเรียกได้ว่าอยู่ที่ไหนก็ดูรายงานได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3) สามารถทำการ Update รายงาน Dashboard ได้อย่างอัตโนมัติ และสามารถใช้งานได้ง่าย ไม่มีการใช้คำสั่งหรือการเขียนโค้ดที่ยากมาก ผู้ที่ไม่มีความรู้ด้านการเขียนโปรแกรมมาก่อนก็สามารถใช้งานได้

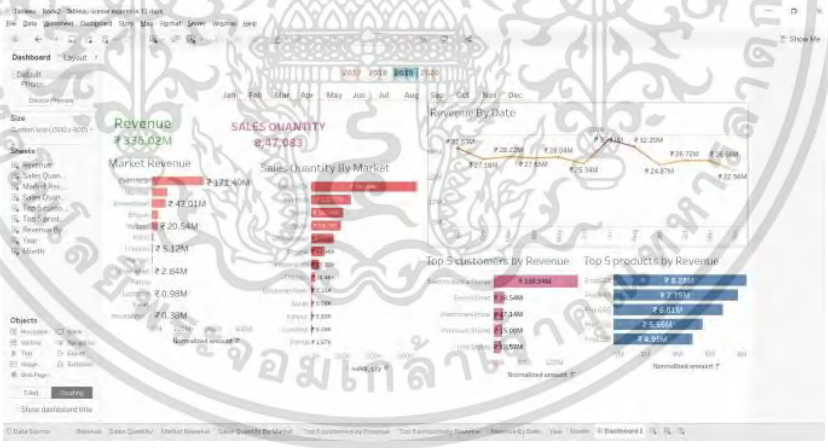
ข้อเสีย

1) ไม่สามารถปรับจูนค่า Parameter ได้

2) หากใช้แบบฟรีจะไม่สามารถแชร์ข้อมูลให้ผู้อื่นได้ (ยกเว้นส่งไฟล์ Power BI Desktop ให้ทั้งไฟล์ ให้ แต่วิธีนี้จะไม่มีความปลอดภัยของข้อมูล อาจเกิดปัญหาข้อมูลรั่วไหลได้)

2.4.4.1.2 Tableau

โปรแกรม Tableau เป็นซอฟต์แวร์ที่ใช้ในวงการ BI, Analytics ใช้ในการสร้างและแสดงผลข้อมูลในรูปแบบของกราฟซึ่งดูแล้วเข้าใจง่าย สามารถใช้ในการวิเคราะห์ได้ เป็นซอฟต์แวร์ที่ใช้งานง่าย และสามารถเชื่อมต่อกับแหล่งข้อมูลจากแหล่งต่างๆ เช่น ฐานข้อมูล (Oracle, SQL server, Azure, Redshift เป็นต้น) , ไฟล์ (csv, excel, PDF, Spatial file เป็นต้น), และเว็บเซอร์วิสต่างๆ เพื่อให้ผู้ใช้สามารถดึงข้อมูลมาใช้และจัดการได้ง่ายขึ้น นอกจากนี้ยังสามารถแชร์รายงานและการวิเคราะห์กับผู้ใช้อื่นในทีมหรือองค์กรได้



รูปที่ 2.16 ภาพโปรแกรม Microsoft Power BI

จากรูปที่ 2.16 แสดงภาพการทำงานของโปรแกรม Tableau ในการ Visualization รูปแบบต่าง และสามารถบอกข้อดีและข้อเสียของโปรแกรม Tableau ได้ดังนี้

ข้อดี

1) การเตรียมข้อมูล (Data Preparation) มีส่วนช่วยในการ Blending Data เพื่อให้สามารถทำการ Cleansing Data และบันทึกขั้นตอนทุกอย่างเป็น Flow

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

และเมื่อ Data ต้นทางเปลี่ยนแปลง ไม่จำเป็นต้องเตรียมข้อมูลใหม่ ทำให้เกิดความง่ายในการเตรียมข้อมูล

2) ใช้งานที่ง่าย และรองรับการเข้าถึงข้อมูลจากหลายฐานข้อมูล

3) รูปแบบการนำเสนอรายงานที่สวยงาม เข้าใจง่าย และช่วยให้เห็นภาพของข้อมูลขนาดใดก็ได้เพียงการลากและวาง (Drag and Drop)

ข้อเสีย

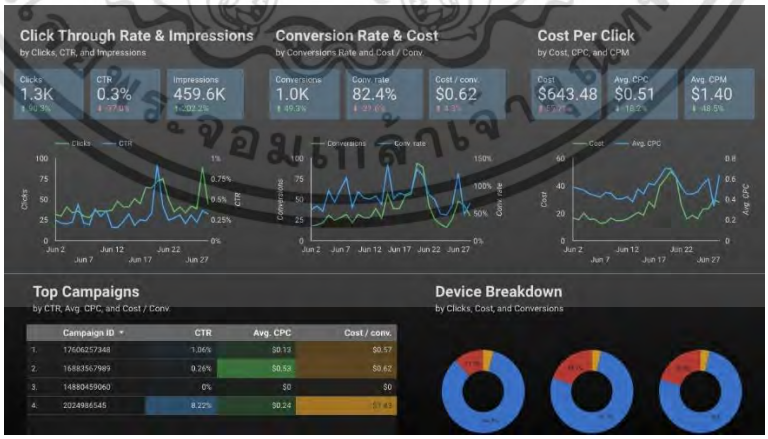
1) มีค่าใช้จ่ายค่อนข้างสูงในการซื้อมาใช้งาน

2) มีการนำเข้าการแสดงภาพแบบกำหนดเองนั้นค่อนข้างยาก ต่างจาก Power BI ที่สามารถ Customization เพิ่มเติมได้

3) ฝังรายงานไปยังแอปพลิเคชันอื่นนั้นค่อนข้างยาก

2.4.4.1.3 Google Data Studio (Looker Studio)

โปรแกรม Google Data Studio (Looker Studio) คือ Business Intelligence Tool จาก Google สำหรับทำ Data Visualization ที่ช่วยเปลี่ยนข้อมูลอ่านยากๆ ในตารางมาเป็นภาพให้เข้าใจง่ายขึ้น โดยเป็นเครื่องมือที่ใช้งานง่าย ไม่ซับซ้อน มีกราฟให้เลือกมากกว่า 30 แบบ สามารถเชื่อมตอกับข้อมูลจากเครื่องมือของทาง Google ได้ทันทีเชน Google Analytics, Google Ads, Google Sheets, YouTube Analytics, Google big query ฯลฯ และเครื่องมือของ Partners Connectors อีกกวา 600 รายการที่สามารถใชงานร่วมกันสามารถปรับแต่ง Dashboard ได้หลากหลาย รวมไปถึงทำ Interactive Dashboard ได้



รูปที่ 2.17 ภาพโปรแกรม Google Data Studio (Looker Studio)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 2.1 แสดงภาพการทำงานของโปรแกรม Google Data Studio (Looker Studio) ในการ Visualization รูปแบบต่าง และสามารถบอกข้อดีและข้อเสียของโปรแกรม Google Data Studio (Looker Studio) ได้ดังนี้

ข้อดี

1) รองรับฐานขอมูลได้หลายรูปแบบ และดึงขอมมูลได้แบบ Real Time ไม่ตองนำขอมูลมากรอกเอง

2) สามารถแชรใหผู้อื่นได้อย่างรวดเร็ว และ Responsive ในทุกอุปกรณ์

3) สร้างการนำเสนอที่หลากหลายสวยงามง่ายต่อการใชงาน

ข้อเสีย

1) ถ้าไม่ใช่ข้อมูล (Data) จาก Google จะไม่ฟรี ต้องมีค่าใช้จ่าย Super Metric เพิ่ม

2) การ Combine หรือ Merge ของข้อมูลเข้าด้วยกันทำได้ยากเมื่อเทียบกับ Power BI ที่เชื่อม Data หลาย Source และหลาย Sheet เข้าด้วยกันได้ง่ายกว่า

2.1.5 ความแตกต่างของ Visualization, Dashboard และ Data Storytelling

การทำ Visualization, Dashboard และ Data Storytelling อาจคล้ายกัน แต่จุดประสงค์หลักของการ Visualization คือการนำเสนอข้อมูลในรูปแบบที่เข้าใจง่าย มักถูกนำมาใช้เพื่อค้นพบแนวโน้ม ความสัมพันธ์ และรูปแบบในข้อมูลที่อาจไม่สามารถมองเห็นได้อย่างชัดเจนจากสเปรดชีทหรือฐานข้อมูล Visualization นั้นสามารถช่วยทำให้ชุดข้อมูลที่ซับซ้อนเข้าใจง่ายขึ้น เพิ่มการเข้าถึงได้สำหรับกลุ่มผู้ใช้ ในส่วนจุดประสงค์หลักของ Dashboard คือการให้ภาพรวมที่รวดเร็วและเข้าใจง่ายของตัวชี้วัดที่สำคัญ Dashboard มักถูกนำมาใช้เพื่อติดตามผลการดำเนินงานของบริษัทแบบเรียลไทม์ ช่วยให้บริษัทสามารถมองเห็นปัญหาหรือโอกาสต่างๆ ได้ทันที Dashboard สามารถนำมาใช้เพื่อติดตามตัวแปรต่างๆ เช่น ยอดขาย รายได้ ความพึงพอใจของลูกค้า และผลการปฏิบัติงานของพนักงาน อย่างสุดท้าย จุดประสงค์ของ Data Storytelling นั้นเป็นการ “เล่าเรื่อง” โดยมีข้อมูลเป็นตัวละครหลัก ข้อมูลของเราผ่านการวิเคราะห์มาแล้ว ได้ Insights ที่จะเล่าเรื่อง นำมาร้อยเรียงด้วยเทคนิค Storytelling เพื่อดึงความสนใจผู้ฟัง ซึ่งผู้รับสารกลุ่มนี้มักอาจไม่คุ้นเคยกับข้อมูลในรายละเอียดแต่มีส่วนเกี่ยวข้องในการตัดสินใจหรือ Stakeholders เช่นเป็นผู้บริหาร เพื่อนพนักงานแผนกอื่นที่ทำงานเกี่ยวข้องกับเรา หรือลูกค้าขององค์กร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2 ภาษาโปรแกรมที่เกี่ยวข้อง

2.2.1 ภาษาเอสคิวแอล (Structured Query Language: SQL)

ภาษาเอสคิวแอล (SQL) เป็นภาษาโปรแกรมสำหรับจัดเก็บและประมวลผลข้อมูลในฐานข้อมูลแบบเชิงสัมพันธ์ ฐานข้อมูลแบบเชิงสัมพันธ์เก็บข้อมูลในรูปแบบตารางที่มีแถวและคอลัมน์ที่เป็นตัวแทนของหมวดหมู่ข้อมูลที่แตกต่างกันและความสัมพันธ์ต่างๆ ระหว่างค่าข้อมูล สามารถใช้คำสั่ง SQL ในการจัดเก็บ ปรับปรุง ลบ ค้นหา และดึงข้อมูลจากฐานข้อมูล นอกจากนี้สามารถใช้ SQL ในการรักษาเพิ่มประสิทธิภาพการทำงานของฐานข้อมูล ดังนี้

1) การจัดการเกี่ยวกับการกำหนดโครงสร้างข้อมูล มีหน้าที่กำหนดโครงสร้างของข้อมูลที่ผู้ใช้แต่ละคนมองเห็น โครงสร้างข้อมูลที่นักออกแบบฐานข้อมูลเห็น และโครงสร้างพื้นฐานที่จัดเก็บในอุปกรณ์เก็บในอุปกรณ์เก็บข้อมูลซึ่งผลของการแปล DDL จะเก็บในไฟล์พิเศษเรียกว่าพจนานุกรมของข้อมูล (Data Dictionary) ตัวอย่างได้แก่ คำสั่ง CREATE, ALTER, DROP, TRUNCATE เป็นต้น

2) การจัดการเกี่ยวกับข้อมูล มีหน้าที่ในด้านจัดการเข้าถึงข้อมูล ได้แก่ การสอบถามหรือค้นหาข้อมูล (Select) ที่อยู่ในฐานข้อมูล การเพิ่มเติมข้อมูลใหม่ (Insert) เข้าไปในฐานข้อมูล การลบIข้อมูล (Delete) ออกจากฐานข้อมูล การเปลี่ยนแปลงแก้ไขข้อมูล (Update) ที่อยู่ใยฐานข้อมูล ตัวอย่างคำสั่งได้แก่ SELECT, INDERT, DELETE, UPDATE เป็นต้น

3) การจัดการรวมข้อมูล มีหน้าที่รวมข้อมูลจากหลายๆ ตารางเข้าด้วยกันโดยอ้างอิงกับค่าในคอลัมน์ที่กำหนด โดยมีหลายรูปแบบตามต้องการ ซึ่งจะมีการใช้งานประเภทต่าง ๆ ขึ้นอยู่กับการต้องการผลลัพธ์ข้อมูลที่ต้องการ ตัวอย่างคำสั่งได้แก่ คำสั่ง INNER JOIN, LEFT JOIN, RIGHT JION, FULL JOIN เป็นต้น

4) การเลือกข้อมูล มีหน้าที่กรองข้อมูลจากตารางตามเงื่อนไขที่ระบุ โดยเฉพาะข้อมูลที่ตรงตามเงื่อนไขนั้นๆ ซึ่งมักจะประกอบด้วยการเปรียบเทียบค่าข้อมูลในคอลัมน์กับค่าที่กำหนดไว้ ตัวอย่างคำสั่งได้แก่ WHERE

2.2.2 ภาษาแซส (Statistical Analysis Software: SAS)

ภาษาโปรแกรมที่ใช้ในการวิเคราะห์ข้อมูลและการประมวลผลทางสถิติที่เป็นเครื่องมือที่มีประสิทธิภาพสำหรับการทำงานกับข้อมูลในงานที่ต้องการความถูกต้องและประสิทธิภาพสูง ใช้ในการจัดการข้อมูล การวิเคราะห์ทางสถิติ และการสร้างรายงานต่าง ๆ ประเภทและชนิดของ SAS สามารถแบ่งได้ตามลักษณะการทำงานได้ ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. Base SAS

เป็นแกนหลักของ SAS ที่ประกอบด้วยเครื่องมือพื้นฐานในการจัดการข้อมูล การวิเคราะห์ข้อมูล การสร้างรายงานใช้ในการเขียนโปรแกรมพื้นฐาน การจัดการข้อมูล การจัดการตาราง และการทำรายงาน

2. SAS/STAT

ใช้สำหรับการวิเคราะห์ทางสถิติขั้นสูงมีฟังก์ชันการทำงานหลากหลาย เช่น การวิเคราะห์ความแปรปรวน การวิเคราะห์การถดถอย การวิเคราะห์การอยู่รอด และอื่น ๆ

3. SAS/GRAPH

ใช้สำหรับการสร้างกราฟ การแสดงผลข้อมูลในรูปแบบกราฟฟิก และรองรับการสร้างกราฟหลายรูปแบบ เช่น กราฟเส้น กราฟแท่ง กราฟพาย และอื่น ๆ

4. SAS/ETS (Econometrics and Time Series)

ใช้สำหรับการวิเคราะห์ข้อมูลเชิงเศรษฐศาสตร์และอนุกรมเวลามีฟังก์ชันในการวิเคราะห์การพยากรณ์ การวิเคราะห์วงจรธุรกิจ และการจำลองข้อมูลทางเศรษฐศาสตร์

5. SAS/IML (Interactive Matrix Language)

ใช้สำหรับการคำนวณทางคณิตศาสตร์ที่ซับซ้อนรองรับการทำงานกับเมทริกซ์และเวกเตอร์

6. SAS/OR (Operations Research)

ใช้สำหรับการวิจัยเชิงปฏิบัติการและการแก้ปัญหาเกี่ยวกับการจัดการทรัพยากร รองรับการแก้ปัญหาการปรับแบบเชิงเส้น การแก้ปัญหาการขนส่ง และการวางแผนทรัพยากร

7. SAS/AF (Application Facility)

ใช้สำหรับการพัฒนาแอปพลิเคชันแบบกำหนดเองรองรับการสร้างอินเตอร์เฟสกราฟิกและการสร้างแอปพลิเคชันแบบอินเตอร์แอคทีฟ

8. SAS/QC (Quality Control)

ใช้สำหรับการควบคุมคุณภาพและการวิเคราะห์การควบคุมคุณภาพรองรับการสร้างแผนภูมิการควบคุม การวิเคราะห์ความสามารถของกระบวนการ และการตรวจสอบการสุ่มตัวอย่าง

9. SAS Enterprise Guide

เป็นเครื่องมือที่ใช้ในการสร้างโปรแกรมและการวิเคราะห์ข้อมูลแบบมีอินเตอร์เฟสกราฟิกเหมาะสำหรับผู้ที่ไม่เชี่ยวชาญในการเขียนโค้ด

10. SAS Visual Analytics

ใช้สำหรับการวิเคราะห์ข้อมูลแบบกราฟิกและการสร้างแดชบอร์ดรองรับการวิเคราะห์ข้อมูลแบบเรียลไทม์และการสร้างภาพกราฟิกที่ซับซ้อน

เอกสารนี้เป็นลิขสิทธิ์ของมหาวิทยาลัยเกษตรศาสตร์ ที่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

นอกจากนี้มีชุดคำสั่งที่แบ่งลักษณะการทำงานได้ 5 ส่วน คือ

1) มุ่งเน้นการจัดการและประมวลผลข้อมูล โดยมีเครื่องมือและคำสั่งที่ช่วยในการโหลดข้อมูลเข้าสู่ระบบ การจัดการข้อมูล การทำความสะอาดข้อมูล และการวิเคราะห์ข้อมูลต่าง ๆ อย่างมีประสิทธิภาพ

2) โครงสร้างของโปรแกรม มีโครงสร้างที่สามารถแบ่งเป็นส่วนย่อย ๆ ซึ่งช่วยให้ผู้ใช้สามารถจัดการโปรแกรมได้อย่างมีระเบียบและง่ายต่อการจัดการ

3) การประมวลผลข้อมูลและสร้างรายงาน ซึ่งเป็นคำสั่งที่มีโครงสร้างและสัญลักษณ์ที่เฉพาะเจาะจง ได้แก่คำสั่ง DATA, PROC, RUN, BY เป็นต้น

4) การควบคุมการทำงานของโปรแกรม ซึ่งช่วยให้ผู้ใช้สามารถสร้างโปรแกรมที่มีการตรวจสอบเงื่อนไขและการทำซ้ำได้อย่างยืดหยุ่น ได้แก่คำสั่ง IF-THEN-ELSE, DO-WHILE, DO-END เป็นต้น

5) การสร้างรายงานและการแสดงผลมีคำสั่งและเครื่องมือที่ช่วยในการสร้างรายงานและการแสดงผลข้อมูลออกมาในรูปแบบที่ต้องการ ได้แก่คำสั่ง PRINT, REPORT, PLOT เป็นต้น

2.3 งานวิจัยที่เกี่ยวข้อง

ภูวิดล (2550) กล่าวว่า Log File เป็นไฟล์ที่ใช้สำหรับบันทึกการกระทำต่างๆ บนระบบคอมพิวเตอร์ ซึ่งมีการเก็บข้อมูลจำนวนมาก Log File ส่วนใหญ่จะเก็บข้อมูลในรูปแบบของข้อความ (Text File) เป็นการบันทึกข้อมูลทีละบรรทัด โดยโปรแกรรมระบบหือโปแกมประยุกต์ต่างๆ นะบบคอมพิวเตอร์ เช่น แต่ละครั้งที่ผู้ใช้ทำการล็อกอินเข้าใช้งานอินเทอร์เน็ต ก็จะมี Log File ที่ชื่อว่า Squid Log โดนจะบันทึกข้อมูลทั้งหมดไม่ว่าการกระทำนั้นสำเร็จหรือไม่สำเร็จ Log File ที่เกิดขึ้นจากระบบปฏิบัติการรวมทั้งโปรแกรมประยุกต์ จะแสดงข้อความต่าง ๆ ที่เกิดขึ้นในระหว่างการทำงานเพื่ออำนวยความสะดวกในการตรวจสอบ

อภิยศ (2563) ได้ทำการรวบรวมและวิเคราะห์ข้อมูล โดยนำข้อมูลงานบริการมาจัดทำเป็นคลังข้อมูล และจัดทำแผงหน้าปัดรายงาน เพื่อใช้สนับสนุนการตัดสินใจในด้านต่าง ๆ เป็นผลให้การให้บริการเป็นไปอย่างมีระบบมากขึ้น ผู้บริหารสามารถเข้าถึงข้อมูลได้อย่างสะดวก และสามารถนำไปวิเคราะห์และบริหารจัดการการให้บริการได้อย่างมีประสิทธิภาพมากยิ่งขึ้น ผู้ใช้บริการสามารถทราบถึงข้อมูลการดำเนินงาน แนวโน้มการใช้บริการในแต่ละวันที่มีการอัพเดตข้อมูลตลอดเวลาผ่านทางการแสดงผลด้วยภาพหรือแผนภูมิ ทำให้ดูทันสมัย มีความน่าสนใจยิ่งขึ้น

วรกานต์ (2563) กล่าวว่า แดชบอร์ด (Dashboard) คือเครื่องมือหนึ่งในการจัดการสารสนเทศเพื่อให้สามารถติดตามผลวิเคราะห์แสดงดัชนีชี้วัดความสำเร็จ หรือที่เรียกว่า KPI (Key Performance Indicator) รวมไปถึงข้อมูลที่สำคัญ ในการตรวจสอบสถานการณ์ปัจจุบันขององค์กรหรือหน่วยงานเฉพาะได้ซึ่งแดชบอร์ดนั้นสามารถปรับแต่งรูปแบบเพื่อให้เข้ากับความต้องการของแผนกหรือบริษัทได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ซึ่งในเบื้องหลังนั้น แดชบอร์ดมักจะเชื่อมต่ออยู่กับข้อมูลที่ถูกเก็บไว้ แต่ในเบื้องหน้าแล้วข้อมูลมักจะถูกแสดงในรูปแบบของตารางแผนภูมเส้น แผนภูมิแท่ง หรือแผนภูมิวงกลม เป็นต้น ทั้งนี้แดชบอร์ดเป็นวิธีที่มีประสิทธิภาพมากที่สุดในการติดตามขอ้ มูลที่มาจาหลายแหล่ง เพราะสามารถ นำข้อมูลหลายข้อมูลมารวมไว้ในจุดเดียวกันเพื่อติดตามและวิเคราะห์ประสิทธิภาพ อีกทั้งการติดตามแบบเรียลไทม์จะสามารถช่วยลดระยะเวลาในการวิเคราะห์ข้อมูลได้

ทศพล และจรัญ (2564) ได้ทำการรวบรวมและวิเคราะห์ข้อมูลการแพร่ระบาดของเชื้อไวรัสโควิด-19 (COVID-19) ที่ระบาดไปทั่วโลกได้รับการพิสูจน์แล้วว่าส่งผลกระทบต่อทุกคน การใช้ประโยชน์จาก Tableau จึงเป็นสิ่งสำคัญที่ช่วยให้การวิเคราะห์ข้อมูลที่มีจำนวนมากและมีความซับซ้อนของสถานการณ์โควิด-19 ทำให้แสดงออกมาเป็นภาพแดชบอร์ดเชิงโต้ตอบ (Interactive Dashboard) ได้ง่ายและรวดเร็ว โดยที่ผู้ใช้ไม่จำเป็นต้องมีความรู้ทางด้านการเขียนโปรแกรมก็สามารถทำได้ด้วยตนเอง การสร้างภาพข้อมูลจึงเป็นวิธีการที่ใช้งานง่ายสำหรับผู้ใช้ในการอ่านและทำความเข้าใจข้อมูล โดยเฉพาะในการวิเคราะห์ข้อมูลขนาดใหญ่ ซึ่งจะช่วยปรับปรุงคุณภาพการบริหารงาน นโยบาย หรือการจัดการบริการต่าง ๆ ได้ด้วยการนำเสนอมุมมองแบบรอบด้าน ที่สามารถเชื่อมต่อกับแหล่งข้อมูลได้หลากหลายและช่วยให้สามารถสร้างภาพข้อมูลทั้ง แผนภูมิ แผนที่ แดชบอร์ด และ สตอรีเพื่อเล่าเรื่องราวการนำเสนอผ่านอินเทอร์เฟซด้วยการลาก และวางแบบง่าย ๆ ภายใน Tableau

วนิดา (2562) กล่าวว่า “ระบบข่าวกรองธุรกิจ สำหรับธุรกิจให้บริการขนส่งกรณีศึกษาห้างหุ้นส่วนจำกัด เก้าเลี้ยว ทรานสปอร์ต” ได้นำระบบธุรกิจอัจฉริยะเข้ามาช่วยวิเคราะห์การบริการขนส่ง นำเสนอในรูปแบบรายงานอัจฉริยะด้วยโปรแกรม Tableau สามารถเรียกดูข้อมูลเป็นลำดับชั้นได้ ช่วยให้ผู้บริหารวางแผนและตัดสินใจการจัดส่งสินค้า ช่วยประหยัดเวลาในการจัดสินค้าและช่วยบริหารคนในการจัดสินค้า

รัตนา (2560) กล่าวว่า “การพัฒนาระบบรายงานรูปแบบหลายมิติเพื่อสนับสนุนการตัดสินใจเชิงนโยบายของสำนักงานคณะกรรมการวิจัยแห่งชาติ” มีจุดประสงค์เพื่อพัฒนาระบบรายงานหลายมิติเพื่อนำเสนอรายงานต่อผู้บริหารได้มองเห็นภาพรวมข้อมูลวิจัยของประเทศได้ และช่วยสนับสนุนการตัดสินใจ โดยใช้โปรแกรม Tableau ผลที่ได้คือ สามารถวิเคราะห์และจัดทำรายงานได้อย่างรวดเร็ว ค้นพบข้อมูลที่น่าสนใจที่ยังไม่เคยนำเสนอมาก่อน เช่น ข้อมูลเฉลี่ยของนักวิจัยจำแนกตามมิติของระดับการศึกษา ตำแหน่งวิชาการ และผู้ใช้มีความพึงพอใจต่อระบบรายงานนี้

Bennyhoff (2023) กล่าวว่า SQL Agent เป็นการทำงานอย่างต่อเนื่องในพื้นหลัง และสามารถกำหนดค่าให้เริ่มทำงานโดยอัตโนมัติ เมื่อเซิร์ฟเวอร์เริ่มทำงาน ทำการควบคุมการจัดการส่วนกลางใน SQL Server Management Studio (SSMS) ที่ช่วยให้ผู้ดูแลระบบสามารถสร้าง กำหนดเวลา และจัดการงานได้ SQL Agent ยังมีคุณสมบัติสำหรับการตรวจสอบสถานะของงาน การตั้งค่าการแจ้งเตือนสำหรับเหตุการณ์เฉพาะ และการบันทึกประวัติงาน เมื่อใช้ SQL Agent ผู้ดูแลระบบสามารถกำหนดเวลางานที่เกิดซ้ำได้ เช่น การสำรองข้อมูล การบำรุงรักษาดัชนี และการซิงโครไนซ์ข้อมูล งานสามารถกำหนดเวลาให้ทำงานตามช่วงเวลาที่ระบุ ในวันที่ระบุของสัปดาห์หรือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ ไม่ว่ากรณีใด ๆ ทั้งสิ้น

เดือน หรือเพื่อตอบสนองต่อเหตุการณ์เฉพาะได้ SQL Agent ยังสนับสนุนการสร้างงานหลายขั้นตอนที่ดำเนินการตามลำดับงานตามลำดับเฉพาะโดยรวมแล้ว SQL Agent เป็นเครื่องมือสำหรับการทำงานด้านการดูแลระบบใน SQL Server โดยอัตโนมัติ และงานสำคัญจะได้รับการปฏิบัติอย่างเชื่อถือได้และตรงเวลา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

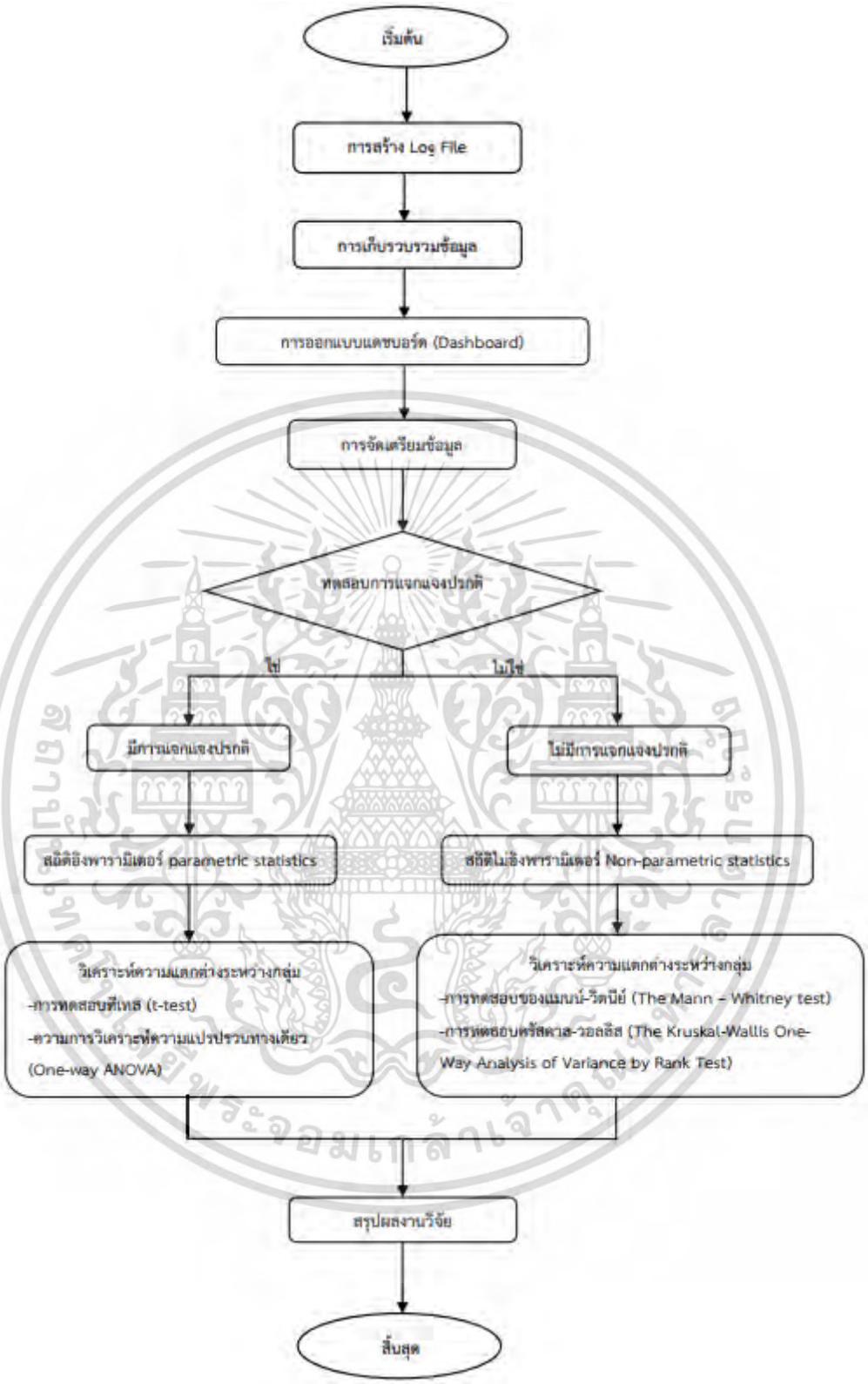
บทที่ 3
วิธีดำเนินงานวิจัย

การศึกษา “การวิเคราะห์ระยะเวลาที่ใช้ในการการประมวลผล SQL Server Agent JOB โดยอาศัยแผงหน้าปัดธุรกิจ” ผู้วิจัยนำชุดข้อมูลจากฐานข้อมูลของบริษัทมาวิเคราะห์ ระยะเวลาการทำงานของ SQL Job เพื่อให้เกิดความเข้าใจถึงประสิทธิภาพการทำงานในระบบ โดยรายระเอียดการดำเนินการวิจัยสามารถแบ่งออกเป็นหัวข้อ ดังนี้

3.1 ขั้นตอนการดำเนินงานวิจัย

รูปที่ 3.1 แสดงให้เห็นถึงตอนการดำเนินงานโดยเริ่มจากทำการสร้าง Log File ของ SQL Job ในฐานข้อมูลและทำการเก็บรวบรวมข้อมูล Log File ของ SQL Job จากฐานข้อมูล ต่อมานำข้อมูลที่เก็บรวบรวมทำการออกแบบแผงหน้าปัดธุรกิจหรือ Dashboard หลังจากนั้นจัดเตรียมข้อมูล Log File เพื่อเก็บจำนวนการประมวลผลของข้อมูลของ SQL Job ตั้งแต่วันที่ 1 ธันวาคม พ.ศ. 2566 ถึงวันที่ 29 มีนาคม พ.ศ 2567 และทำการทดสอบการแจกแจงปรกติของข้อมูล ถ้าผลลัพธ์ของการทดสอบของข้อมูลมีการแจกแจงปรกติ จะทำการวิเคราะห์ความแตกต่างระหว่างกลุ่มด้วยการทดสอบ T-test และการทดสอบ ANOVA หรือการทดสอบของข้อมูลไม่มีการแจกแจงปรกติจะทำการวิเคราะห์ความแตกต่างระหว่างกลุ่มด้วยการทดสอบของแมนน์-วิตนีย์ (The Mann – Whitney test) และการทดสอบครัสคาล-วอลลิส (The Kruskal-Wallis One-Way Analysis of Variance by Rank Test) จากนั้นทำการสรุปผลการวิจัย เพื่ออธิบายระยะเวลาการประมวลผลของ SQL Job

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.1 ขั้นตอนดำเนินงาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2 เครื่องมือที่ใช้ในการวิจัย

3.2.1 ซอฟต์แวร์ที่ใช้ในการวิจัย (Software)

1. SAS Enterprise Guide 7.1 (64-bit) เป็นโปรแกรมสำหรับการวิเคราะห์สถิติ และใช้สำหรับจัดการฐานข้อมูลที่ใช้สำหรับเก็บรวบรวมที่ข้อมูลต้องการศึกษารวมไปถึงการสร้าง Log File

2. Microsoft Power BI เป็นโปรแกรมนำข้อมูลจากแหล่งต่าง ๆ เช่น ฐานข้อมูล ไฟล์แบบต่าง ๆ หรือแหล่งข้อมูลอื่น ๆ การสร้างรายงาน และ Dashboardได้

3. Minitab Statistical Software เป็นโปรแกรมสำเร็จรูปทางสถิติที่ใช้ในการวิเคราะห์ข้อมูลเชิงสถิติ

3.2.2 ฮาร์ดแวร์ที่ใช้ในการวิจัย (Hardware)

หน่วยประมวลผล Intel Core i5 11TH Gen หน่วยความจำ 8.0 GB

3.3 การสร้าง Log File

การเก็บข้อมูลโดยทำการสร้าง Log File เพื่อเก็บข้อมูลการดำเนินงานหรือการประมวลผลต่างๆ ที่เกิดขึ้นในระบบของ SQL Job ใช้ในการติดตาม ตรวจสอบ และวิเคราะห์กระบวนการทำงานในโปรแกรม SAS Enterprise Guide7.1 โดยการออกแบบการเก็บ Log File มีรายละเอียดแสดงดังตารางที่ 3.1

ตารางที่ 3.1 รายละเอียดชุดคำสั่งออกแบบการเก็บ Log File

LOG START	data work.STEP1_LOAN; KEY = 'xxx_'||put(today(),date9.)||put(hour(time()),$2.)||'H'||put(minute(time()),$2.)||'M'; DATABASE = 'xxx'; TABLE_NAME = 'xxx'; TYPE = 'START'; format RUNNING_DATETIME DATETIME.; RUNNING_DATETIME = datetime() ; FORMAT VALUE_DATETIME DATETIME.; /*VALUE_DATETIME = datetime();*/ FORMAT VALUE_NUMURIC 10.; FORMAT VALUE_CHARACTER $50.;

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.2 รายละเอียดชุดคำสั่งออกแบบการเก็บ Log File (ต่อ)

LOG START	FORMAT ERROR $50.;
	run;
	data _null_ ;
	set WORK.STEP1_LOAN;
	call symput('KEY',KEY);
	call symput('TABLE_NAME',TABLE_NAME);
	run;
	%put &KEY &TABLE_NAME;
	proc sql;
	insert into TOOLS.LOG_TABLE_CHECK
	select * from STEP1_LOAN
	;quit;

LOG END	data work.STEP_LOAN_END;
	KEY = "&KEY";
	DATABASE = 'xxx';
	TABLE_NAME = "xxx";
	TYPE = 'END';
	format RUNNING_DATETIME DATETIME.;
	RUNNING_DATETIME = datetime() ;
	FORMAT VALUE_DATETIME DATETIME.;
	/*VALUE_DATETIME = datetime();*/
	FORMAT VALUE_NUMURIC 10.;
	VALUE_NUMURIC = "&NOB";
	FORMAT VALUE_CHARACTER $50.;
	FORMAT ERROR $50.;
	run;
	data _null_ ;
	set WORK.STEP1_LOAN_END;
	call symput('KEY',KEY);
	call symput('TABLE_NAME',TABLE_NAME);
	run;

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.3 รายละเอียดชุดคำสั่งออกแบบการเก็บ Log File (ต่อ)

LOG	%put &KEY &TABLE_NAME;
END	proc sql;
	insert into TOOLS.LOG_TABLE_CHECK
	select * from STEP1_LOAN_END
	;quit;

จากตารางที่ 3.1 แสดงชุดคำสั่งออกแบบการเก็บ Log File ที่มีสองส่วนคือ LOG START ที่ บอกการเก็บข้อมูล SQL Job ที่เริ่มการประมวลผลและ LOG END ที่บอกการเก็บข้อมูล SQL Job ที่สิ้นสุดการประมวลผล

3.4 การเก็บรวบรวมข้อมูล

การศึกษาและเก็บรวบรวมข้อมูลที่ใช้ในการวิเคราะห์ โดยนำเนินการเก็บข้อมูล SQL Job ที่ได้มากจากการสร้าง Log File ในฐานข้อมูลของบริษัทด้านสินเชื่อ และโบรกเกอร์ประกันวินาศภัยตั้งแต่วันที่ 1 ธันวาคม พ.ศ. 2566 ถึงวันที่ 29 มีนาคม พ.ศ. 2567

3.4.1 การนำข้อมูลเข้า

ขั้นตอนการนำชุดข้อมูลจากฐานข้อมูลเข้าโปรแกรม Microsoft Power BI เพื่อเชื่อมต่อกับฐานข้อมูลและดูข้อมูลจากไฟล์ฐานข้อมูล

3.4.2 การทำความเข้าใจข้อมูล

ผู้วิจัยได้ทำการศึกษาข้อมูลจากฐานข้อมูลของบริษัทมาเก็บไว้เพื่อศึกษารายละเอียดของข้อมูลก่อนนำไปทำการวิเคราะห์ข้อมูล โดยผู้วิจัยได้นำข้อมูลระยะเวลาการประมวลผลของ SQL Job จากฐานข้อมูลบริษัท มีรายละเอียด แสดงดังตารางที่ 3.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.2 รายละเอียดข้อมูล Log File ของ SQL Job จากฐานข้อมูลบริษัท

ข้อมูล	รายละเอียด
KEY	KEY ไว้สำหรับเชื่อม TYPE และระบุครั้งที่รันในวันที่รัน
DATABASE	ชื่อแหล่งที่มาในการเก็บ
TABLE_NAME	ชื่อ TABLE ที่เก็บข้อมูล
TYPE	ประเภทของ KEY
RUNNING_DATETIME	วันเวลาที่ประมวลผล
VALUE_DATETIME	ข้อมูลที่เก็บเป็น DATETIME
VALUE_NUMURIC	ข้อมูลที่เก็บเป็น NUM

3.5 การออกแบบรูปแบบแผงหน้าปัดธุรกิจ หรือ แดชบอร์ด (Dashboard)

จากการรวบรวมข้อมูลของ SQL Job เพื่อทำการออกแบบแผงหน้าปัดธุรกิจ หรือ แดชบอร์ด (Dashboard) ในโปรแกรม Microsoft Power BI เพื่อนำเสนอรายละเอียดการประมวลผล SQL Job จาก Log File

ความต้องการของผู้ใช้ทีทำการมอนิเตอร์ Dashboard นั้น ได้แก่ ต้องการทราบจำนวน SQL Job ที่มีการประมวลผล จำนวน SQL Job ที่มีการประมวลผลสำเร็จและล้มเหลว เวลาการในประมวลผล และค่าเฉลี่ยในการประมวล 7 วันก่อนหน้า สามารถค้นหา table ใน แต่ละ SQL Job และเลือกดูวันที่ SQL Job ที่มีการประมวลผลได้

3.5.1 การ์ด (Card) แสดงจำนวนข้อมูล SQL Job แสดงดังรูปที่ 3.2 และ 3.3

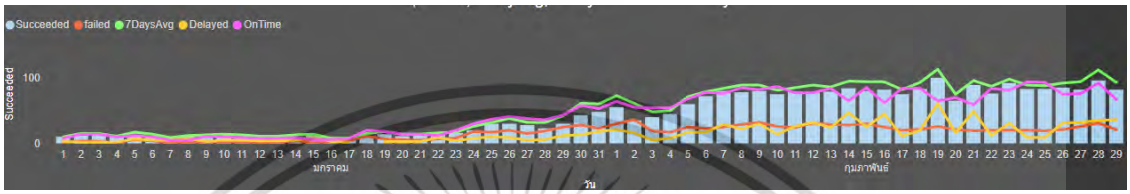
รูปที่ 3.2 ข้อความที่แสดงจำนวนตัวเลข

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 3.2 ผู้วิจัยเลือกใช้เลือกใช้ visualization ที่เป็น การ์ด (Card) เพื่อแสดงจำนวนข้อมูล SQL Job ที่มีอยู่แบบเป็นตัวเลข

3.5.2 แผนภูมิเส้นและแผนภูมิแท่ง (Line and Clustered Column Chart)

แสดงการเปรียบเทียบจำนวนของ SQL Job ในแต่ละวัน ดังรูปที่ 3.3



รูปที่ 3.3 กราฟแท่งและกราฟเส้นแสดงการเปรียบเทียบจำนวนของ SQL Job

จากรูปที่ 3.3 ผู้วิจัยเลือกใช้เลือกใช้ Visualization ที่เป็นแผนภูมิเส้นและแผนภูมิแท่ง (Line and Clustered Column Chart) เพื่อง่ายต่อการแสดงการเปรียบเทียบจำนวนข้อมูล SQL Job ที่มีแนวโน้ม การเพิ่ม ลด และการเปลี่ยนแปลงของจำนวนข้อมูล

โดยกำหนดแผนภูมิแท่งและแผนภูมิเส้นแสดงจำนวนข้อมูล SQL Job ที่ Succeed ซึ่งมีชุดคำสั่งในการคำนวณดังรูปที่ 3.4-3.8

```
CountSucceededJobs =
CALCULATE(
    COUNTROWS('SQL_Job'),
    'SQL_Job'[Status] = "succeeded"
)
```

รูปที่ 3.4 คำสั่งในการคำนวนจำนวนของ SQL Job ที่ “Succeeded”

```
CountfailedJobs =
CALCULATE(
    COUNTROWS('SQL_Job'),
    'SQL_Job'[Status] = "failed"
)
```

รูปที่ 3.5 คำสั่งในการคำนวนจำนวนของ SQL Job ที่ “Failed”

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
AvgDuration7DaysPerJob =
CALCULATE(
    AVERAGE('SQL_Job'[DurationInMinutes]),
    FILTER(
        ALL('SQL_Job'),
        'SQL_Job'[SQL_Name] = EARLIER('SQL_Job'[SQL_Name]) &&
        'SQL_Job'[START] <= EARLIER('SQL_Job'[START]) &&
        'SQL_Job'[START] > EARLIER('SQL_Job'[START]) - 7
    )
)
```

รูปที่ 3.6 คำสั่งในการคำนวนจำนวนเฉลี่ยของ SQL Job ใน 7 วันก่อน

```
CountDelayedJobs =
CALCULATE(
    COUNTROWS('SQL_Job'),
    'SQL_Job'[DelayedStatus] = "Delayed"
)
```

รูปที่ 3.7 คำสั่งในการคำนวนจำนวนของ SQL Job ที่ “Delayed”

```
CountOnTimeJobs =
CALCULATE(
    COUNTROWS('SQL_Job'),
    'SQL_Job'[DelayedStatus] = "On Time"
)
```

รูปที่ 3.8 คำสั่งในการคำนวนจำนวนของ SQL Job ที่ “On Time”

ในการกำหนดสีของกราฟจะบอกประเภทของ SQL Job ดังนี้

1) กราฟแท่งสีฟ้าแสดงจำนวนข้อมูล SQL Job ที่มีสถานะการประมวลผลสำเร็จ (Succeed)

2) กราฟเส้นสีแดงแสดงจำนวนข้อมูล SQL Job ที่มีการประมวลผลล้มเหลว (Failed)

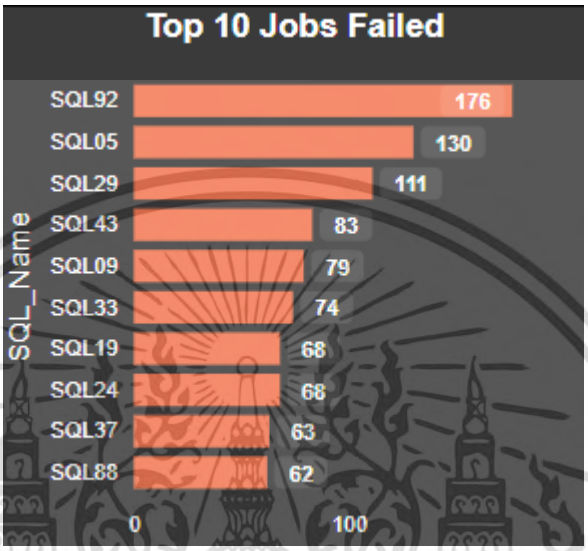
3) กราฟเส้นสีเขียวแสดงจำนวนข้อมูล SQL Job ที่มีการประมวลผลเฉลี่ย 7 วันก่อนหน้า (7DayAvg)

4) กราฟเส้นสีเหลืองแสดงจำนวนข้อมูล SQL Job ที่ทำการประมวลผลที่มากกว่าการประมวลผลเฉลี่ย 7 วันก่อนหน้า และมีสถานะประมวลผล่าช้า (Delayed)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5) กราฟเส้นสีม่วงแสดงจำนวนข้อมูล SQL Job ที่ทำการประมวลผลที่น้อยกว่าหรือเท่ากับการประมวลผลเฉลี่ย 7 วันก่อนหน้า และมีสถานะประมวลผลตามเวลา (On Time)

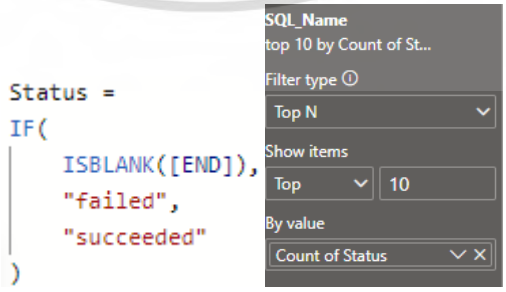
3.5.3 แผนภูมิแท่งแบบคลัสเตอร์ (Clustered Bar Chart) แสดงจำนวน SQL Job ดังรูปที่ 3.9



รูปที่ 3.9 กราฟแท่งแบบคลัสเตอร์ (Clustered Bar Chart) แสดงการเปรียบเทียบจำนวนของ SQL Job

จากรูปที่ 3.9 ผู้วิจัยเลือกใช้เลือกใช้ visualization ที่เป็น แผนภูมิแท่ง (Clustered Bar Chart) เพื่อแสดงการเปรียบเทียบจำนวนข้อมูลแต่ละ SQL Job ที่ล้มเหลว (Failed) โดยกำหนด 10 Job ที่มีสถานะล้มเหลวของข้อมูลมากที่สุด

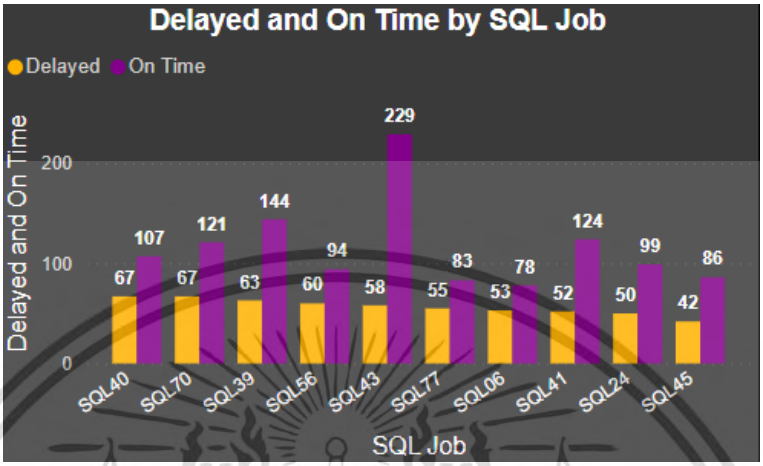
โดยสามารถคำนวณได้การทำชุดคำสั่งและกำหนด top 10 จำนวนที่ SQL Jobมีสถานะล้มเหลวของข้อมูลมากที่สุดดังรูปที่ 3.10



รูปที่ 3.10 คำสั่งในการคำนวนจำนวนของ SQL Job ที่ล้มเหลว (Failed) และกำหนด top 10 จำนวนที่ SQL Jobมีสถานะล้มเหลวของข้อมูลมากที่สุด

เอกสารนี้เป็นลิขสิทธิ์ของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของลิขสิทธิ์ทุกครั้งที่มีการนำไปใช้

3.5.4 แผนภูมิคอลัมน์แบบคลัสเตอร์ (Clustered Column Chart) แสดงจำนวน SQL Job ดังรูปที่ 3.11



รูปที่ 3.11 กราฟคอลัมน์แบบคลัสเตอร์ (Clustered Column Chart) แสดงการเปรียบเทียบจำนวนข้อมูล SQL Job ที่ Delayed และ On Time และกำหนด 10 Job ที่มีจำนวนข้อมูลล่าช้าและตามเวลา

จากรูปที่ 3.11 ผู้วิจัยเลือกใช้เลือกใช้ visualization ที่เป็น แผนภูมิแท่ง (Clustered Column Chart) เพื่อแสดงการเปรียบเทียบจำนวนข้อมูลแต่ละ SQL Job ที่ ล่าช้า(Delayed) และทันเวลา (On Time)

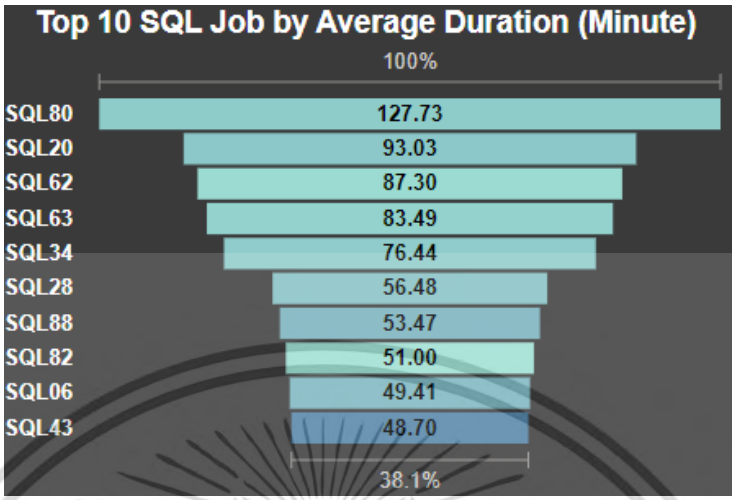
โดยกำหนดว่าถ้า Duration ของ SQL Job มากกว่า 7DayAvg จะแสดงผลที่เป็น Delayed และ Duration ของ SQL Job น้อยกว่าหรือเท่ากับ 7DayAvg จะแสดงผลที่เป็น On Time โดยสามารถใช้ชุดคำสั่งในการคำนวณได้ดังรูปที่ 3.12

```
DelayedStatus =
IF(
    'SQL_Job'[DurationInMinutes] > [AvgDuration7DaysPerJob],
    "Delayed",
    "On Time"
)
```

รูปที่ 3.12 ชุดคำสั่งในการคำนวนที่แสดงผลล่าช้า(Delayed) และ ทันเวลา (On Time)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

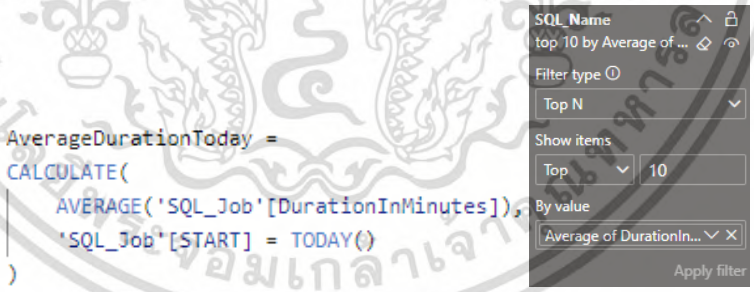
3.5.5 แผนภูมิกรวย (Funnel) แสดงค่าเฉลี่ยของ SQL Job ดังรูปที่ 3.8



รูปที่ 3.13 แผนภูมิกรวย (Funnel) แสดงค่าเฉลี่ยของ SQL Job

จากรูปที่ 3.13 ผู้วิจัยเลือกใช้เลือกใช้ visualization ที่เป็น แผนภูมิกรวย (Funnel) เพื่อแสดงการเปรียบเทียบค่าเฉลี่ยของ SQL Job โดยกำหนด 10 Job ที่มีการประมวลผลมากที่สุด

โดยสามารถคำนวณได้การทำชุดคำสั่งและกำหนด top 10 จำนวนค่าเฉลี่ยที่ SQL Job การประมวลผลมากที่สุดดังรูปที่ 3.14



รูปที่ 3.14 คำสั่งในการคำนวนค่าเฉลี่ยของ SQL Job และกำหนด Top 10 จำนวนค่าเฉลี่ยที่ SQL Job การประมวลผลมากที่สุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.5.6 ตัวกรองข้อมูล (Slicer) ดังรูปที่ 3.15 และรูปที่ 3.16

รูปที่ 3.15 ตัวกรองข้อมูล (Slicer) วันที่ (เดือน/วัน/ปี)

รูปที่ 3.16 ตัวกรองข้อมูล (Slicer) SQL Job

จากรูปที่ 3.15 และ รูปที่ 3.16 ผู้วิจัยเลือกใช้ตัวกรองข้อมูล (Slicer) ที่สามารถแสดงดูเป็นรายวันได้หรือจะกำหนดช่วงวันที่เราต้องการทราบ และตัวกรองกรอง SQL Job ที่เราต้องการทราบได้หรือจะกำหนดช่วงวันที่เราต้องการดู

3.6 การจัดเตรียมข้อมูล

3.6.1 การสกัดข้อมูล (Data Extraction)

ขั้นตอนการสกัดข้อมูลที่จำเป็นจากฐานข้อมูลของบริษัท เพื่อใช้สำหรับหารวิเคราะห์ ซึ่งการสกัดข้อมูลจำเป็นต้องทำความเข้าใจของข้อมูลก่อนเพื่อตัดข้อมูลที่ไม่จำเป็นโดยการตรวจสอบระยะเวลาการประมวลผลของ SQL JOB โดยที่ข้อมูลเป็นข้อมูลที่เก็บสถานะที่เริ่มการประมวลผลและสถานะที่สิ้นสุดการประมวลผลของ SQL Job จำนวน 5,161 รายการ ดังตารางที่ 3.3

ตารางที่ 3.3 ข้อมูลสถานะที่เริ่มการประมวลผลและสถานะที่สิ้นสุดการประมวลผลของ SQL Job

ข้อมูล	รายละเอียด
TABLE_NAME	ชื่อ TABLE ที่เก็บข้อมูล
RUNNING_DATETIME	วันเวลาที่ประมวลผล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

	TABLE_NAME	RUNNING_DATETIME_START	RUNNING_DATETIME_END
1	A01	1/12/23 9:25:00	01/12/2023 10:41:00
2	A02	1/12/23 12:23:00	1/12/23 13:06:00
3	A03	1/12/23 13:06:00	1/12/23 13:07:00
...	...	...	...
5159	A91	29/03/2024 09:00:00	29/03/2024 09:09:00
5160	A92	29/03/2024 09:45:00	29/03/2024 09:47:00
5161	A93	29/03/2024 10:00:00	29/03/2024 10:20:00

รูปที่ 3.17 ตัวอย่างข้อมูลสถานะที่เริ่มการประมวลผลและสถานะที่สิ้นสุดการประมวลผลของ SQL Job

รูปที่ 3.17 ประกอบไปด้วย ชื่อตาราง (TABLE_NAME) วันเวลาที่ประมวลผลเริ่มต้น (RUNNING_DATETIME_START) และ วันเวลาที่ประมวลผลสิ้นสุด (RUNNING_DATETIME_STOP)

3.6.2 การแปลงข้อมูล (Data Transformations)

ขั้นตอนการแปลงข้อมูลให้พร้อมสำหรับนำไปใช้ในการวิเคราะห์ ในขั้นตอนนี้จะทำหลังจากสกัดข้อมูลที่จำเป็น โดยนำข้อมูลไปประมวลผลบนโปรแกรม Tableau เพื่อทำการแปลงข้อมูลให้อยู่ในรูปที่สามารถนำไปวิเคราะห์ต่อได้ โดยมีขั้นตอนดังนี้

ขั้นตอนที่ 1 ทำการแปลงข้อมูล วันที่เวลาให้อยู่ในรูป Time (minute) โดยมีวัตถุประสงค์เพื่อให้สามารถนับจำนวนเวลาที่ใช้ในการประมวลผลของ SQL Job ได้ โดยมีชุดคำสั่งและรายละเอียดดังตารางที่ 3.4 ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.4 ชุดคำสั่งการแปลงข้อมูลวันที่และเวลาที่ใช้ในการประมวลผลของ SQL Job

1	DurationInMinutes = DATEDIFF('SQL_Job'[STRT], 'SQL_Job'[END], MINUTE)

	TABLE_NAME	RUNNING_DATETIME_START	RUNNING_DATETIME_END	TIME
1	A01	1/12/23 9:25:00	01/12/2023 10:41:00	76
2	A02	1/12/23 12:23:00	1/12/23 13:06:00	43
3	A03	1/12/23 13:06:00	1/12/23 13:07:00	1
...	...	...	...	...
5159	A91	29/03/2024 09:00:00	29/03/2024 09:09:00	9
5160	A92	29/03/2024 09:45:00	29/03/2024 09:47:00	2
5161	A93	29/03/2024 10:00:00	29/03/2024 10:20:00	20

รูปที่ 3.18 ตัวอย่างชุดข้อมูลหลังจากแปลงวันที่และเวลาที่ใช้ในการประมวลผลของ SQL Job

จากรูปที่ 3.18 แสดงผลที่ได้จากการแปลงข้อมูลวันที่และเวลา โดยที่จะสามารถนำไปวิเคราะห์และตรวจสอบการประมวลผลของ SQL Job ได้

ขั้นตอนที่ 2 ทำการสร้างคอลัมน์ขึ้นมา 2 คอลัมน์ เพื่อเก็บข้อมูลที่แสดงเป็นช่วงเวลาและวันของสัปดาห์ที่ใช้ในการประมวลผลของ SQL Job โดยมีชุดคำสั่งและรายละเอียดดังตารางที่ 3.5 ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.5 ชุดคำสั่งเก็บข้อมูลที่แสดงเป็นช่วงเวลาและวันของสัปดาห์ที่ใช้ในการประมวลผลของ SQL Job

Time Period	TimePeriod = SWITCH(TRUE(), HOUR('SQL_Job'[STRT]) >= 5 && HOUR('SQL_Job'[STRT]) < 12, "morning", HOUR('SQL_Job'[STRT]) >= 12 && HOUR('SQL_Job'[STRT]) < 17, "Afternoon", HOUR('SQL_Job'[STRT]) >= 17 && HOUR('SQL_Job'[STRT]) < 21, "Evening", HOUR('SQL_Job'[STRT]) >= 21 ‖ HOUR('SQL_Job'[STRT]) < 5, "Night", "ไม่ระบุ")
Day Of Week	DayOfWeek = SWITCH(WEEKDAY('SQL_Job'[STRT], 2), 1, "Monday", 2, "Tuesday", 3, "Wednesday", 4, "Thursday", 5, "Friday", 6, "Saturay", 7, "Sunday")

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

	TABLE_ NAME	RUNNING_DATE TIME_START	RUNNING_DATE TIME_END	TIME	TIME OF DAY	WEEKDAY
1	A01	1/12/23 9:25:00	01/12/2023 10:41:00	76	Morning	Friday
2	A02	1/12/23 12:23:00	1/12/23 13:06:00	43	Afternoon	Friday
3	A03	1/12/23 13:06:00	1/12/23 13:07:00	1	Afternoon	Friday
...	...	...	...	...		
5159	A91	29/03/2024 09:00:00	29/03/2024 09:09:00	9	Morning	Friday
5160	A92	29/03/2024 09:45:00	29/03/2024 09:47:00	2	Morning	Friday
5161	A93	29/03/2024 10:00:00	29/03/2024 10:20:00	20	Morning	Friday

รูปที่ 3.19 ตัวอย่างชุดข้อมูลหลังจากสร้างคอลัมน์ 2 คอลัมน์เก็บข้อมูลที่แสดงเป็นช่วงเวลาและวันของสัปดาห์ที่ใช้ในการประมวลผลของ SQL Job

จากรูปที่ 3.19 ได้ทำการจัดรูปแบบข้อมูลใหม่ของข้อมูลวันที่และเวลาที่เป็นช่วงเวลาและวันของสัปดาห์โดยที่จะสามารถนำไปวิเคราะห์และตรวจสอบการประมวลผลของ SQL Job ได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.7 การวิเคราะห์ข้อมูล

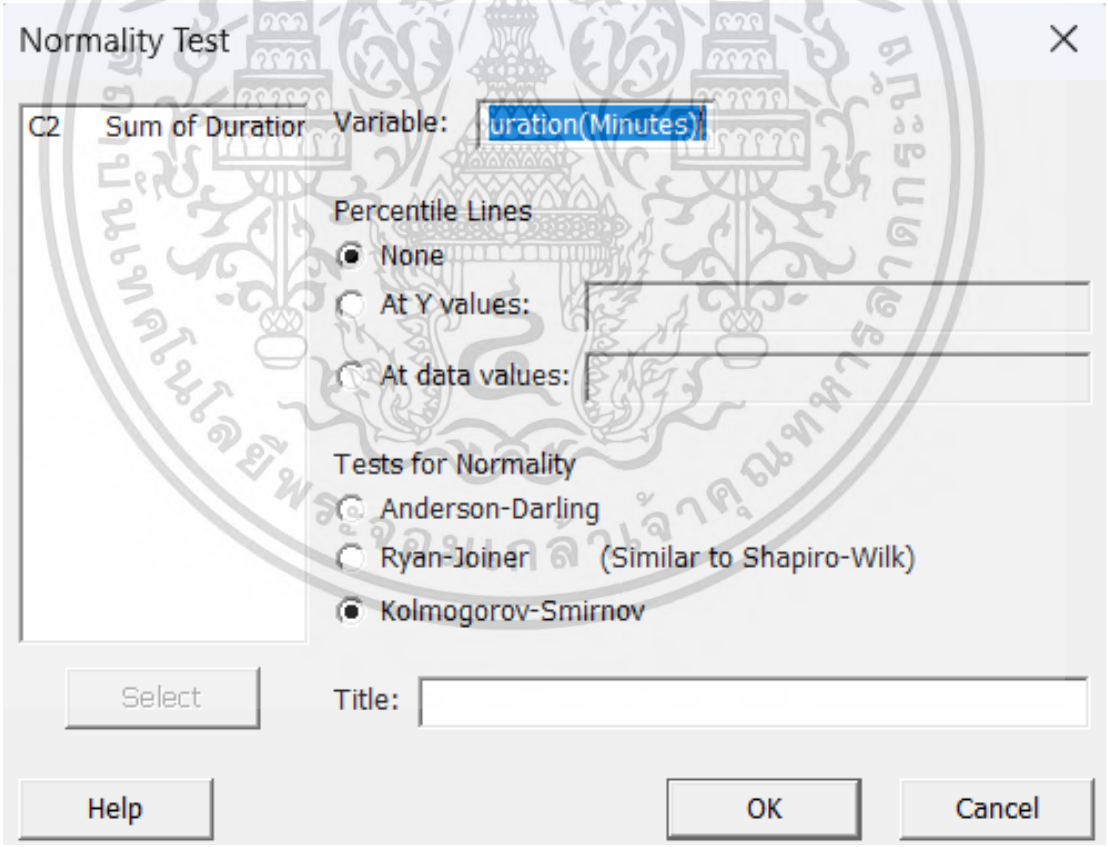
3.7.1 สถิติเชิงพรรณนา ใช้สำหรับวิเคราะห์ระยะเวลาการประมวลผลของ SQL JOB ในแต่ละช่วงเวลาและวันของสัปดาห์ ได้แก่ ค่าเฉลี่ย ค่าร้อยละ ส่วนเบี่ยงเบนมาตรฐาน

3.7.2 สถิติเชิงอนุมาน ได้แก่ การวิเคราะห์ความสัมพันธ์ระหว่างกลุ่มโดยสถิติอนุมาน (Inferential Statistics) ด้วยโปรแกรมสำเร็จรูป Minitab Statistical Software มีขั้นตอนดังต่อไปนี้

3.7.2.1 ข้อตกลงเบื้องต้น

เพื่อตรวจสอบข้อมูลว่าข้อมูลมีการแจกแจงปรกติหรือไม่ ถ้าข้อมูลไม่มีการแจกแจงปรกติควรจะใช้การวิเคราะห์สถิติแบบไม่ใช้พารามิเตอร์ โดยขั้นตอนการวิเคราะห์ด้วย Minitab มีรายละเอียดดังนี้

1) Stat
2) Basic Statistics
3) Normality Test



รูปที่ 3.20 หน้าต่างการวิเคราะห์การแจกแจงข้อมูลด้วยโปรแกรมสำเร็จรูป Minitab Statistical Software

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 3.20 ทำการทดสอบการแจกแจงปรติของข้อมูลไม่เจอการทดสอบลิลลีโฟรส์ จึงเลือก Kolmogorov-Smirnov โดยเป็นวิธีทดสอบที่มีการกำลังทดสอบสูงกว่าการทดสอบอื่นๆ เหมือนกับการทดสอบลิลลีโฟรส์ (Lilliefors Test)

ข้อมูลส่วนใหญ่ไม่มีการแจกแจงปรติ จึงเลือกการใช้สถิติไม่อิงพารามิเตอร์ (Nonparametric Statistics)

3.7.2.2 การทดสอบสมมติฐาน

เนื่องจากต้องการทดสอบค่าเฉลี่ยของจำนวน SQL Job โดยใช้การทดสอบของแมนน์-วิทนีย์ (The Mann – Whitney test) และทดสอบค่ามัธยฐานของระยะเวลาการประมวลผลของ SQL Job โดยใช้การทดสอบครัสคาล-วอลลิส (The Kruskal-Wallis One-Way Analysis of Variance by Rank Test) เพื่อตรวจสอบว่าคุณลักษณะ หรือตัวแปรสามารถแบ่งค่าตามกลุ่มได้อย่างมีนัยสำคัญทางสถิติกรณีข้อมูลไม่มีการแจกแจงปรกติ

1) สมมติฐานโดยอาศัยการทดสอบของแมนน์-วิทนีย์

ถ้าให้ $M_{Morning.}$ และ $M_{otherdays.}$ แทนค่าเฉลี่ยของประชากรที่ 1 และ 2 ตามลำดับ ทำการทดสอบหางเดียวเพื่อทดสอบว่าค่าเฉลี่ยของจำนวน SQL Job ในวันพุธ มากกว่าค่าเฉลี่ยของจำนวน SQL Job ในวันอื่น ๆ ที่ไม่ใช่วันพุธ ได้ดังนี้

$$H_0 : M_{Wed.} \leq M_{otherdays.}$$
$$H_1 : M_{Wed.} > M_{otherdays.}$$

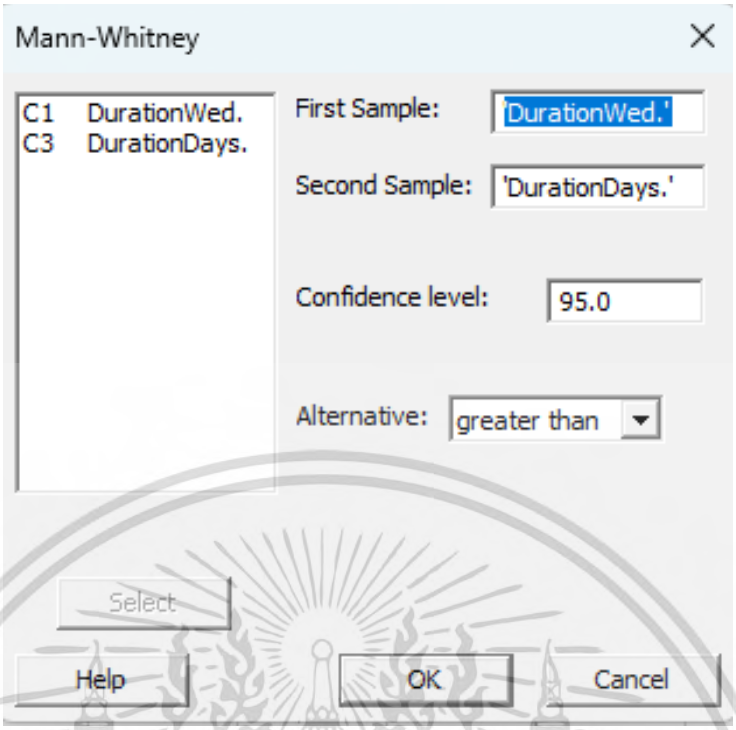
ถ้าให้ $M_{Morning.}$ และ $M_{othertimes.}$ แทนค่าเฉลี่ยของประชากรที่ 1 และ 2 ตามลำดับ ทำการทดสอบหางเดียวเพื่อทดสอบว่าค่าเฉลี่ยของจำนวน SQL Job ในช่วงเวลาเช้า มากกว่าค่าเฉลี่ยของจำนวน SQL Job ในช่วงเวลาอื่น ๆ ที่ไม่ใช่ช่วงเวลาเช้า ได้ดังนี้

$$H_0 : M_{Morning} \leq M_{othertimes}$$
$$H_1 : M_{Morning.} > M_{othertimes.}$$

ขั้นตอนการวิเคราะห์ด้วย Minitab

1) Stat
2) Nonparametrics
3) Mann – Whitney

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.21 หน้าต่างการวิเคราะห์ทดสอบค่าเฉลี่ยของจำนวน SQL Job ด้วยโปรแกรมสำเร็จรูป Minitab Statistical Software

จากรูปที่ 3.21 ทำการทดสอบค่าเฉลี่ยของจำนวน SQL Job ทำการเลือกลุ่มที่เปรียบเทียบและต้องการทดสอบหางเดียวว่าค่าเฉลี่ยของจำนวน SQL Job กลุ่มใดมากกว่า จึงเลือก greater than

2) สมมติฐานโดยอาศัยการทดสอบครัสคาล-วอลลิส

การทดสอบสองหางเพื่อทดสอบว่าค่ามัธยฐานของระยะเวลาการประมวล SQL Job ทั้ง 7 วัน มีความแตกต่างระหว่างกลุ่ม สามารถเขียนเป็นสมติฐานได้ดังนี้

H_0 : ค่ามัธยฐานของระยะเวลาการประมวล SQL Job ทั้ง 7 วัน ไม่แตกต่างกัน

H_1 : ค่ามัธยฐานของระยะเวลาการประมวล SQL Job อย่างน้อย 1 วันแตกต่าง กัน

การทดสอบสองหางเพื่อทดสอบว่าค่ามัธยฐานระยะเวลาการประมวผล SQL Job ทั้ง 4 ช่วงเวลา มีความแตกต่างระหว่างกลุ่ม สามารถเขียนเป็นสมติฐานได้ดังนี้

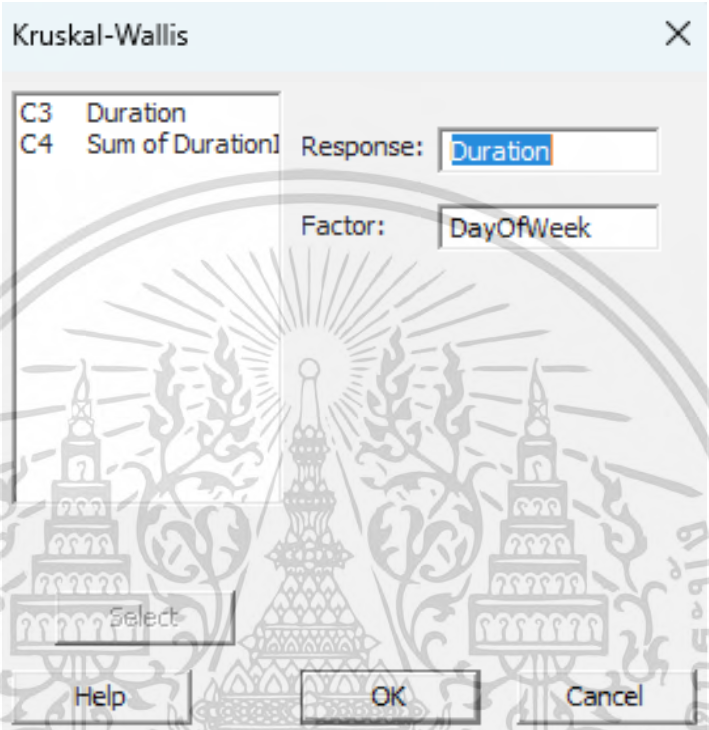
H_0 : ค่ามัธยฐานระยะเวลาการประมวผล SQL Job ทั้ง 4 ช่วงเวลาไม่แตกต่างกัน

H_1 : ค่ามัธยฐานระยะเวลาการประมวผล SQL Job อย่างน้อย 1 ช่วงเวลาแตกต่างกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนการวิเคราะห์ด้วย Minitab

1) Stat
2) Nonparametrics
3) Kruskal-Wallis...



รูปที่ 3.22 หน้าต่างการเลือกข้อมูลเพื่อใช้ทดสอบความแตกต่าง Job ด้วยโปรแกรมสำเร็จรูป Minitab Statistical Software

จากรูปที่ 3.22 การทดสอบว่าค่ามัธยฐานของระยะเวลาในการประมวลผลของ SQL Job ในแต่ละกลุ่มมีความแตกต่างระหว่าง โดยใช้สถิติทดสอบ Kruskal-Wallis

3) สมมติฐานของการทดสอบแบบไคสแควร์(The Chi-Square Test)

เนื่องจากต้องการทดสอบความเป็นอิสระของตัวแปรวันและช่วงเวลา โดยใช้การทดสอบแบบไคสแควร์ (The Chi-Square Test) เพื่อตรวจสอบว่าคุณลักษณะหรือตัวแปรสามารถบอกความเป็นอิสระของตัวแปรทั้ง 2 อย่างมีนัยสำคัญทางสถิติที่ข้อมูลไม่มีการแจกแจงปรกติ

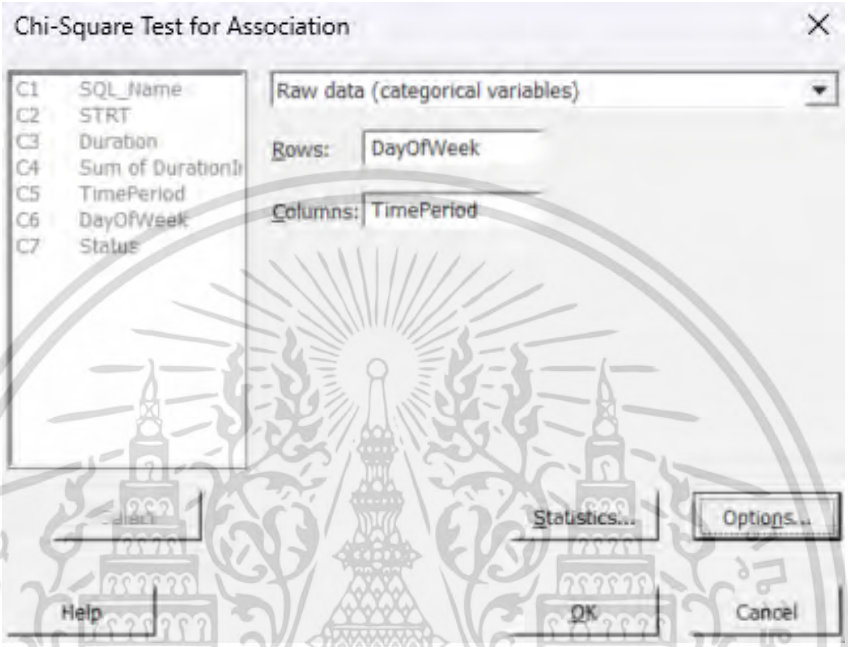
H_0 : จำนวน SQL Job ในแต่ละวันของแต่ละช่วงเวลา เป็นอิสระต่อกัน

H_1 : จำนวน SQL Job ในแต่ละวันของแต่ละช่วงเวลา ไม่เป็นอิสระต่อกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนการวิเคราะห์ด้วย Minitab

1) Stat
2) Tables
3) Chi–Square Test for Association...



รูปที่ 3.23 หน้าต่างการเลือกข้อมูลเพื่อใช้ทดสอบความเป็นอิสระของตัวแปรทั้งสองด้วยโปรแกรมสำเร็จรูป Minitab Statistical Software

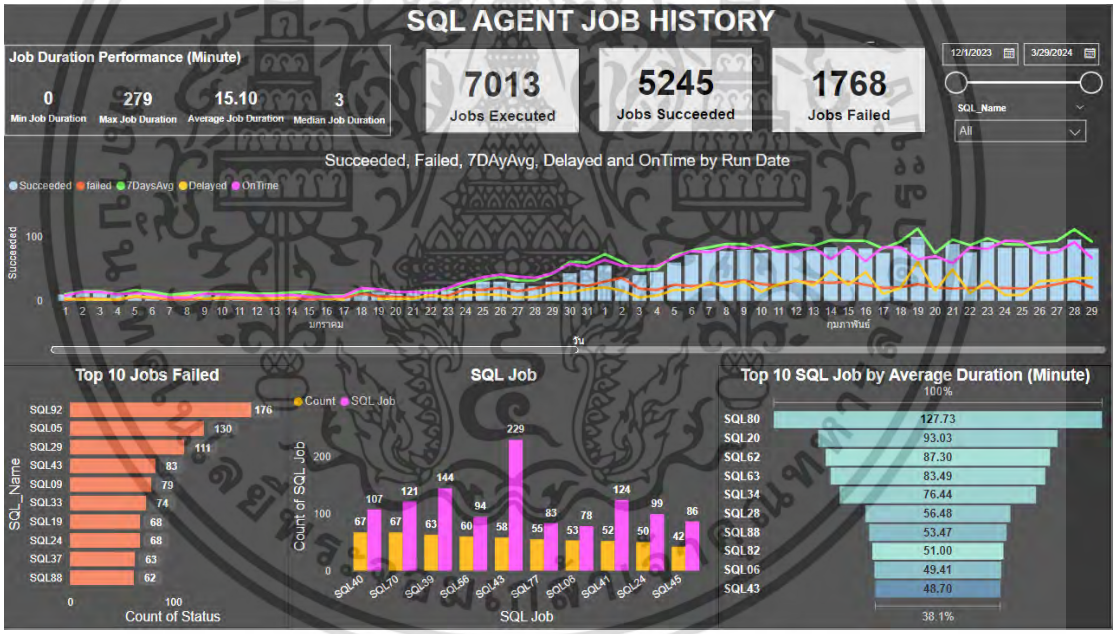
จากรูปที่ 3.22 ทำการทดสอบว่าตัวแปรวันและตัวแปรช่วงเวลาของทั้ง 2 ตัวแปร มีความเป็นเป็นอิสระ โดยใช้สถิติไคสแควร์ (The Chi-Square Test)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4
ผลการวิจัยและอภิปรายผล

ในบทนี้จะกล่าวถึงผลของการออกแบบแผงหน้าปัดธุรกิจ หรือ Dashboard และผลการวิเคราะห์ความสัมพันธ์ระหว่างกลุ่มโดยสถิติอนุมาน (Inferential Statistics) ซึ่งแบ่งออกเป็น 7 วัน ได้แก่ วันจันทร์ วันอังคาร วันพุธ วันพฤหัสบดี วันศุกร์ วันเสาร์ และวันอาทิตย์ ในส่วนช่วงเวลาต่างๆ แบ่งออกเป็น 4 ช่วงเวลา ได้แก่ ช่วงเวลาเช้า [05.00-11.59] ช่วงเวลาบ่าย [12.00-16.59] ช่วงเวลาเย็น [17.00-20.59] และช่วงเวลาดึก [21.00-04.59] พร้อมทั้งผลการวิเคราะห์ เพื่อให้เกิดความเข้าใจถึงระยะเวลาการประมวลผลของ SQL Job โดยเนื้อหาของบทนี้ประกอบด้วยหัวข้อย่อย ๆ ดังนี้

4.1 ผลการออกแบบแผงหน้าปัดธุระกิจหรือ Dashboard

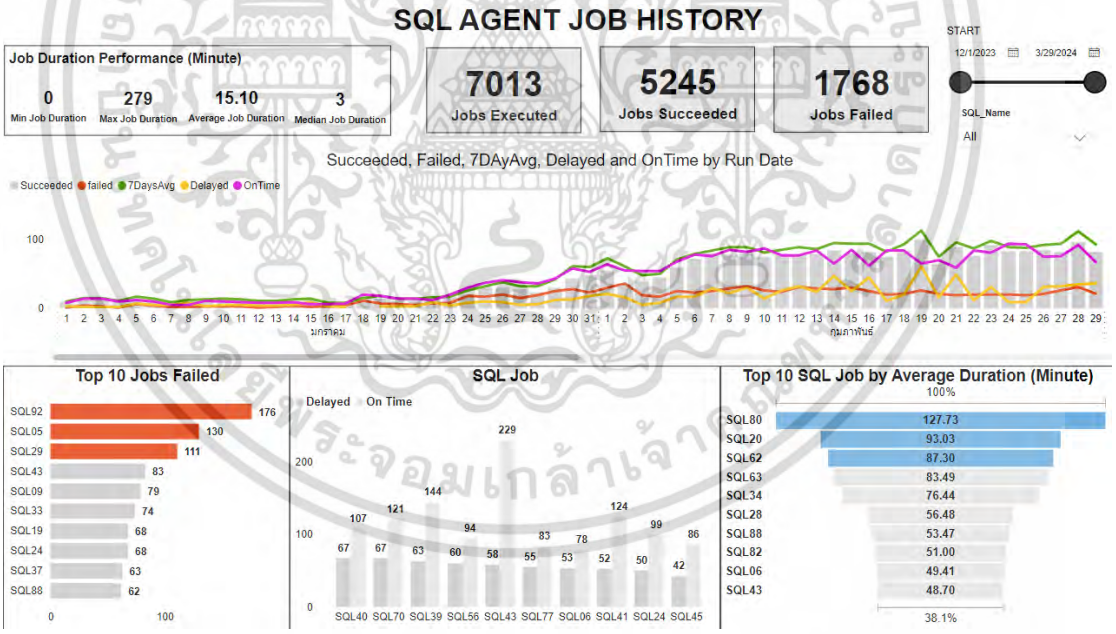


รูปที่ 4.1 แดชบอร์ดการวิเคราะห์ประวัติการทำงานของ SQL Job

จากรูปที่ 4.1 แดชบอร์ดการวิเคราะห์การการทำงานของ SQL Job ทำให้ให้เห็นภาพรวมโดยย่อเกี่ยวกับ SQL Job แบบเรียลไทม์และย้อนหลังได้ในแต่ละวัน แสดงจำนวน Min Job Duration จำนวน Max Job Duration จำนวน Average Job Duration Median Job Duration Job Executed Job Succeeded Job Failed โดยแสดงผลในรูปแบบตัวเลข และแผนภูมิเพื่อให้ทราบจำนวน SQL Job แต่ละกลุ่มที่มีการประมวลผลในแต่ละครั้ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากการทำแดชบอร์ดการวิเคราะห์ประวัติการทำงานของ SQL Job หลังจากนั้นมีการปรับปรุงตามสัดส่วนของ Ink Ratio แล้วสามารถสรุปผล และอัตราส่วนของแต่ละกราฟได้ดังนี้ กราฟหลักที่แสดง (Succeeded, Failed, 7DayAvg, Delayed and OnTime by Run Date) ทำการปรับปรุงเปลี่ยนจากสีทึบเป็นสีโปร่งแสง และใช้แท่งสีเทาแทนสีเข้ม ในอัตราส่วน ink ratio ประมาณ 40% ต่อมากราฟ Top 10 Jobs Failed ทำการปรับปรุงเปลี่ยนแถบสีส้มเป็นสีเทาอ่อน ยกเว้น 3 อันดับแรกที่ยังคงสีส้มเพื่อยังคงแสดงจำนวนนวน SQL Job ที่มากที่สุด ในอัตราส่วน ink ratio ประมาณ 30% ถัดไปเป็นกราฟ SQL Job (Delayed and On Time) ทำการปรับปรุงเปลี่ยนจากแท่งสีสดเป็นแท่งสีเทาทั้งหมดและแสดงตัวเลขจำนวนบนแท่งกราฟแทน ในอัตราส่วน ink ratio ประมาณ 20% และลำดับสุดท้าย กราฟ Top 10 SQL Job by Average Duration (Minute) ทำการปรับปรุงเปลี่ยนแถบสีฟ้าเข้มเป็นสีเทาอ่อนสำหรับอันดับ 4-10 คงสีฟ้าไว้สำหรับ 3 อันดับแรกที่มีค่าเฉลี่ยที่มากที่สุด ในอัตราส่วน ink ratio ประมาณ 40% โดยรวม การปรับปรุง Ink Ratio ทำให้การใช้หมึกลดลงโดยเฉลี่ยประมาณ 60-70% เมื่อเทียบกับก่อนปรับ ซึ่งหมายความว่าแดชบอร์ดใหม่ใช้หมึกเพียง 30-40% ของแดชบอร์ดเดิม ทำให้ประหยัดหมึกได้อย่างมีประสิทธิภาพ แต่ยังคงรักษาความชัดเจนของข้อมูลสำคัญไว้ได้ ดังรูปที่ 4.2

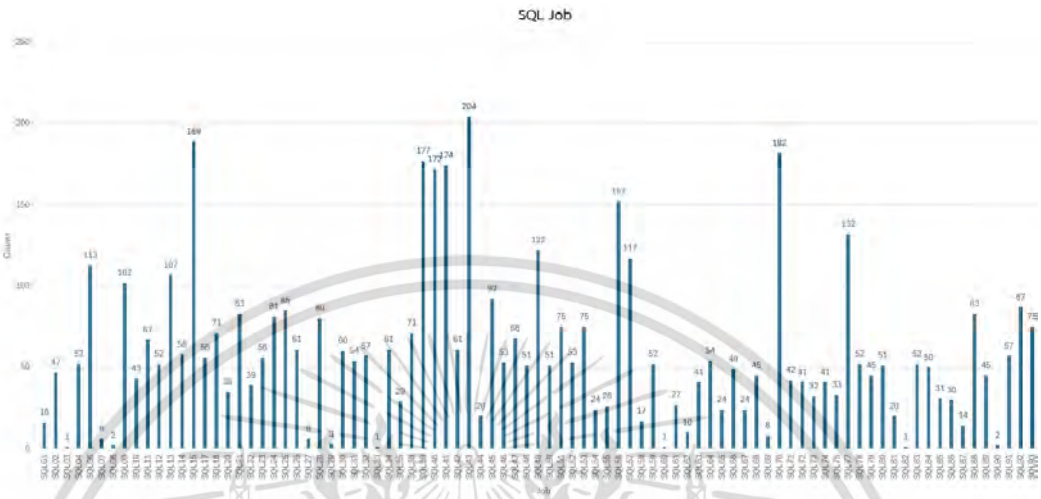


รูปที่ 4.2 แดชบอร์ดการวิเคราะห์ประวัติการทำงานของ SQL Job หลังทำการปรับ Ink Ratio

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2 ผลการวิเคราะห์ข้อมูลด้วยสถิติพรรณนา

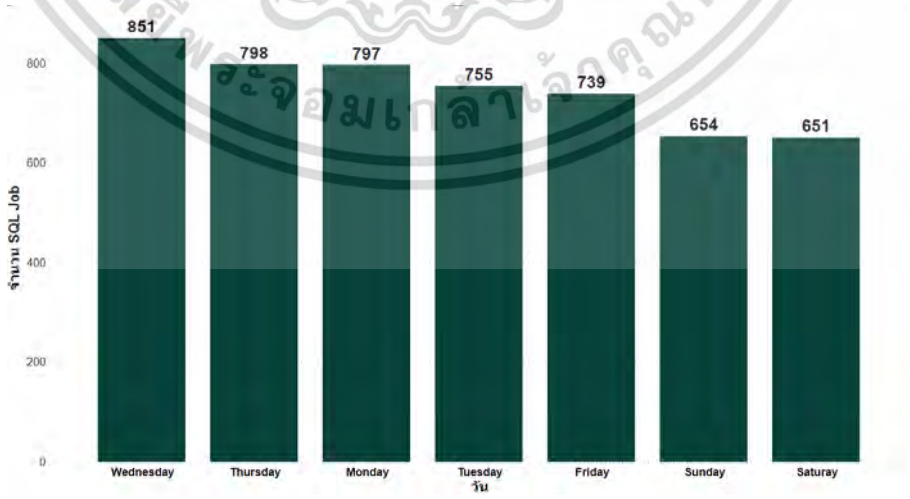
4.2.1 ข้อมูลของ SQL Job



รูปที่ 4.3 จำนวนข้อมูลของ SQL Job

รูปที่ 4.3 ในแกน X แสดงชื่อ และในแกน Y แสดงจำนวนข้อมูลของ SQL Job พบว่ามีจำนวนที่เรียกใช้ในการประมวลผลมากที่สุดอยู่ที่ 204 ครั้ง มาจาก Job ที่ชื่อว่า SQL43 ซึ่งเป็นข้อมูลเกี่ยวกับข้อมูลการขายประกัน

4.2.2 จำนวน SQL Job ที่มีการประมวลผลในแต่ละวัน

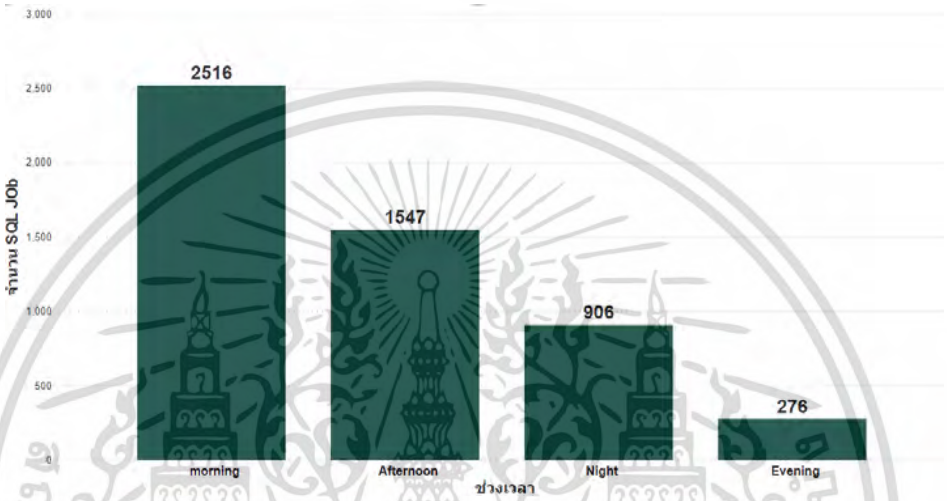


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 4.4 จำนวน SQL Job ที่มีการประมวลผลในแต่ละวันของสัปดาห์

รูปที่ 4.4 แสดงจำนวน SQL Job ที่มีการประมวลผลในแต่ละวันของสัปดาห์ จากรูปจะเห็นได้ว่า SQL Job ที่มีการประมวลผลในแต่ละวันมีจำนวนใกล้เคียงกัน โดยวันพุธมีจำนวน SQL Jobสูงที่สุด คิดเป็นร้อยละ 16.22

4.2.3 จำนวน SQL Job ที่มีการประมวลผลในแต่ละช่วงเวลา



รูปที่ 4.5 จำนวน SQL Job ที่มีการประมวลผลแต่ละช่วงเวลา

รูปที่ 4.5 แสดงจำนวน SQL Job ที่มีการประมวลผลในแต่ละช่วงเวลา ในเวลาช่วงเช้า [05.00-11.59] มีจำนวนรายการค่อนข้างสูงเมื่อเปรียบเทียบกับช่วงเวลาอื่น ๆ โดยคิดเป็นร้อยละ 47.97 สำหรับการประมวลผล SQL Job

ตารางที่ 4.1 สถิติพรรณนาจำนวน SQL Job ของทุกช่วงเวลาในแต่ละวัน

สถิติพรรณนาจำนวน SQL Job ของทุกช่วงเวลาในแต่ละวัน							
วัน / เวลา	อาทิตย์	จันทร์	อังคาร	พุธ	พฤหัสบดี	ศุกร์	เสาร์
ช่วงเวลาเช้า	138	233	245	278	230	217	165
ช่วงเวลาบ่าย	22	57	28	80	44	28	16
ช่วงเวลาเย็น	352	367	352	344	373	357	329
ช่วงเวลาดึก	123	129	123	137	147	122	125

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากตารางที่ 4.1 จะพบว่าจำนวน SQL Job ที่มีการประมวลผลทุกช่วงเวลาในแต่ละวันพบว่า โดยในช่วงเวลาเช้าจะมีจำนวนการประมวลผลใน SQL Job มากที่สุด ตามด้วยช่วงเวลาบ่าย ช่วงเวลาเย็นและช่วงเวลาดึกดึกตามลำดับ

4.2.4 ระยะเวลาการประมวลผล SQL Job ในแต่ละวัน

ตารางที่ 4.2 สถิติพรรณนาระยะเวลาการประมวลผลของ SQL Job ในแต่ละวัน

สถิติพรรณนาของข้อมูลรายการในการประมวลผลSQL Job ในแต่ละวันของสัปดาห์							
ค่าสถิติ / วัน	Min	Max	Mean	SD	Median	IQR	Skewness
วันอาทิตย์	0	275	16.96	30.15	3	18.75	3.3
วันจันทร์	0	234	16.23	29.17	3	17	3.2
วันอังคาร	0	115	11.189	16.72	4	14	2.42
วันพุธ	0	129	12.759	18.577	5	17	2.49
วันพฤหัสบดี	0	250	15.31	29.21	3	16	3.58
วันศุกร์	0	238	15.88	28.55	3	17.75	3.07
วันเสาร์	0	279	17.78	33.48	3	19	3.22

จากตารางที่ 4.2 จะพบว่าค่าเฉลี่ย SQL Job ที่มีการประมวลผลในแต่ละวันใกล้เคียงกัน นอกจากนี้ยังพบว่าส่วนเบี่ยงเบนมาตรฐานในแต่ละช่วงมีค่าสูง แสดงเห็นการกระจายตัวของรายการในการประมวลผลมากสำหรับ SQL Job และข้อมูลที่มีการประมวลผลมีลักษณะเบ้ขวา สามารถดูได้จากค่า Skewness ที่คำนวณได้มีค่าเป็นบวก นั้นคือ รายการในการประมวลผลที่มีจำนวนน้อยกว่าความถี่ มากกว่ารายการในการประมวลผลที่มีจำนวนครั้งมาก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.5 ระยะเวลาการประมวลผล SQL Job ในแต่ละช่วงเวลา

ตารางที่ 4.3 สถิติพรรณนาระยะเวลาการประมวลผลของ SQL Job ในแต่ละช่วงเวลา

สถิติพรรณนาระยะเวลาการประมวลผล SQL Job ในแต่ละช่วงเวลา							
ค่าสถิติ / ช่วงเวลา	Min	Max	Mean	SD	Median	IQR	Skewness
ช่วงเวลาเช้า	0	234	12.535	23.911	2	16	3.08
ช่วงเวลาบ่าย	0	279	9.6	26.29	2	7	6.09
ช่วงเวลาเย็น	0	275	21.522	32.629	8	26	2.88
ช่วงเวลาดึก	0	89	4.902	8.943	1	7	4.13

จากตารางที่ 4.3 จะพบว่าค่าเฉลี่ย SQL Job ที่มีการประมวลผลในแต่ละช่วงเวลา โดยเฉลี่ยในช่วงเวลาเย็นมีระยะการประมวลผลใน SQL Job มากที่สุด ตามด้วยช่วงเวลาเช้า บ่ายและดึกตามลำดับ นอกจากนี้ยังพบว่าส่วนเบี่ยงเบนมาตรฐานในแต่ละช่วงมีค่าสูง แสดงเห็นการกระจายตัวของรายการในการประมวลผลมากสำหรับ SQL Job และข้อมูลที่มีการประมวลผลมีลักษณะเบ้ขวาสามารถดูได้จากค่า Skewness ที่คำนวณได้มีค่าเป็นบวก นั้นคือ รายการในการประมวลผลที่มีจำนวนน้อยกว่าความถี่ มากกว่ารายการในการประมวลผลที่มีจำนวนครั้งมาก

ตารางที่ 4.4 สถิติพรรณนาแสดงค่าเฉลี่ยระยะเวลาในการประมวลผล SQL Job ทุกช่วงเวลาในแต่ละวัน

สถิติพรรณนาแสดงค่าเฉลี่ยของข้อมูลรายการในการประมวลผล SQL Job ทุกช่วงเวลาในแต่ละวัน							
วัน / เวลา	อาทิตย์	จันทร์	อังคาร	พุธ	พฤหัสบดี	ศุกร์	เสาร์
ช่วงเวลาเช้า	13.17	13.42	14.16	12.34	12.21	12.09	9.72
ช่วงเวลาบ่าย	4.41	7.42	10.29	13.79	10.14	5.82	7.5
ช่วงเวลาเย็น	15.83	23.74	21.33	28.17	21.65	25.04	14.43
ช่วงเวลาดึก	5.016	3.814	5	5.102	5.639	4.566	5.056

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากตารางที่ 4.4 จะพบว่าค่าเฉลี่ยในการประมวลผล SQL Job ทุกช่วงเวลาในแต่ละวันพบว่า โดยเฉลี่ยในช่วงเวลาเช้าจะมีจำนวนการประมวลผลใน SQL Job มากที่สุด ตามด้วยช่วงเวลาบ่าย ช่วงเวลาเย็นและช่วงเวลาดึก

4.2.6 ข้อมูลระยะวลาที่ใช้ในการประมวลผลและระยะห่างการมาถึงของแต่ละ SQL Job

ตารางที่ 4.5 สถิติพรรณนาระยะวลาที่ใช้ในการประมวลผลและระยะห่างการมาถึงของแต่ละ SQL Job

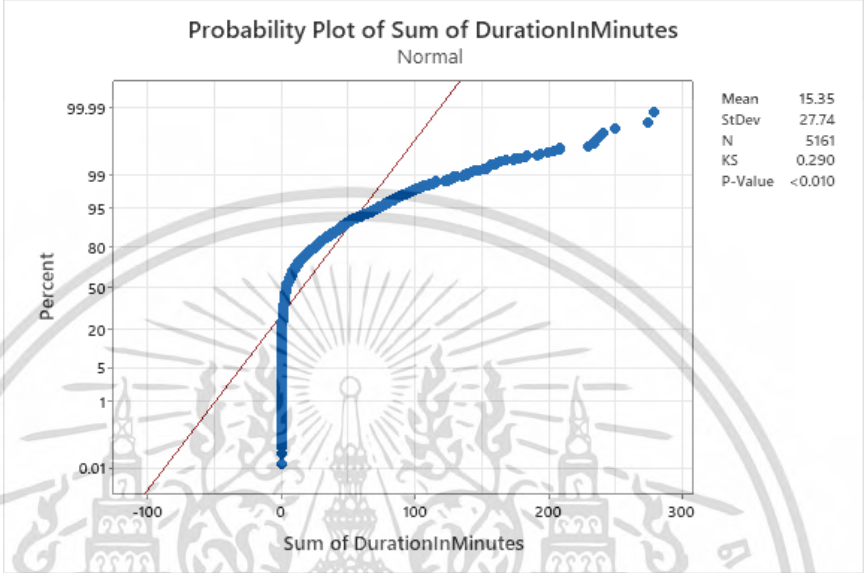
สถิติพรรณนาระยะวลาที่ใช้ในการประมวลผลและระยะห่างการมาถึงของแต่ละ SQL Job							
ค่าสถิติ / ตัวแปร	Min	Max	Mean	SD	Median	IQR	Skewness
TIME_DIFF	0	279	15.347	27.735	3	18	3.44
Arrival(minute)	0	171335	2393	10586	1371	1063	9.8

จากตารางที่ 4.5 จะพบว่าค่าเฉลี่ยระยะห่างการมาถึงของแต่ละ SQL Job สูงกว่าระยะวลาที่ใช้ในการประมวลผล SQL Job นอกจากนี้ยังพบว่าส่วนเบี่ยงเบนมาตรฐานทั้งสองมีค่าสูง แสดงเห็นการกระจายตัวของระยะวลาที่ใช้ในการประมวลผล และระยะห่างการมาถึงของแต่ละ SQL Job ข้อมูลที่มีลักษณะเบ้ขวาสามารถดูได้จากค่า Skewness ที่คำนวณได้มีค่าเป็นบวก นั้นคือระยะเวลาในการประมวลผลและระยะห่างการมาถึงของแต่ละ SQL Job ที่มีจำนวนน้อยกว่าความถี่มากกว่ารายการในการประมวลผลที่มีจำนวนครั้งมาก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3 ผลการวิเคราะห์ความแตกต่างระหว่างกลุ่มโดนสถิติอนุมาณ (Inferential Statistic

4.3.1 การวิเคราะห์แจกแจงข้อมูล



รูปที่ 4.6 ผลลัพธ์การวิเคราะห์แจกแจงข้อมูล

จากรูปที่ 4.6 พบว่าข้อมูลส่วนใหญ่ไม่มีการแจกแจงปรกติ จึงเลือกการใช้สถิติแบบไม่อิงพารามิเตอร์ (Nonparametric Statistics)

4.3.2 การทดสอบความเป็นอิสระ

การทดสอบเกี่ยวกับข้อมูลที่มีลักษณะสองทางจากข้อมูลที่สนใจศึกษาซึ่งมีลักษณะ 2 ทางนำมาสร้างตารางความถี่ของตัวแปรสุ่มทั้ง 2 และเรียกว่า ตารางการจร ทำการทดสอบถึงความเป็นอิสระของตัวแปรทั้ง 2 ในตัวแปรแรกเป็นตัวแปรของวัน และตัวแปรที่สองเป็นตัวแปรช่วงเวลา ด้วยโปรแกรมสำเร็จรูป Minitab Statistical Software ได้ผลลัพธ์ดังต่อไปนี้

สมมติฐานของการทดสอบ

H_0: จำนวน SQL Job ในแต่ละวันของแต่ละช่วงเวลา เป็นอิสระต่อกัน

H_1: จำนวน SQL Job ในแต่ละวันของแต่ละช่วงเวลา ไม่เป็นอิสระต่อกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.6 ผลการทดสอบความเป็นอิสระของตัวแปรวัน และตัวแปรช่วงเวลา

	Chi-Square	DF	P-value
Pearson	99.594	18	<0.001
Likelihood Ratio	98.22	18	<0.001

จากตารางที่ 4.6 พบว่า ค่า P-value มีน้อยกว่า 0.05 จึงปฏิเสธ H_0ที่ระดับนัยสำคัญ 0.05 ดังนั้น จำนวน SQL Job ของวัน และ จำนวน SQL Job ของช่วงเวลา มีความเกี่ยวข้องกัน สามารถดูผลการวิเคราะห์ได้จากโปรแกรม Minitab ที่รูป ก.1 ของภาคผนวก ก

4.3.3 การทดสอบสมมติฐานจำนวนการประมวลผลของ SQL Job

การวิเคราะห์ทางสถิติเพื่อตรวจสอบว่าคุณลักษณะ หรือตัวแปรสามารถแบ่งค่าความแตกต่างระหว่างกลุ่มได้ ของระยะเวลาที่ใช้ในการประมวลผลของ SQL Jobในแต่ละช่วงเวลา โดยใช้สถิติทดสอบแมนน์-วิทนีย์ (The Mann-Whitney Test) ด้วยโปรแกรมสำเร็จรูป Minitab Statistical Software ได้ผลลัพธ์ดังต่อไปนี้

สมมติฐานของการทดสอบ

$$H_0 : M_{Wed.} \leq M_{otherdays.}$$
$$H_1 : M_{Wed.} > M_{otherdays.}$$

ตารางที่ 4.7 ผลการทดสอบค่าเฉลี่ย SQL Job ของกลุ่มวันพุธ และ กลุ่มวันอื่นๆ ที่ไม่ใช่วันพุธ

Method	H-value	P-value
Not adjusted for ties	24438877	0.603

เนื่องจากการทดสอบข้อกำหนดเบื้องต้นแล้ว พบว่า ข้อมูลไม่มีการแจกแจงปรกติ ดังนั้น สถิติที่ใช้ในการทดสอบ คือ The Mann-Whitney Test

จากตารางที่ 4.7 พบว่า ค่า P-value มีค่ามากกว่า 0.05 จึงยอมรับ H_0ที่ระดับนัยสำคัญ 0.05 นั้นคือค่าเฉลี่ยของ SQL Job ในกลุ่มวันพุธน้อยกว่าหรือเท่ากับวันอื่นๆ ที่ไม่ใช่วันพุธ สามารถดูผลการวิเคราะห์ได้จากโปรแกรม Minitab ที่รูป ก.2 ของภาคผนวก ก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สมมติฐานของการทดสอบ

$$H_0 : M_{Morning} \leq M_{othertimes}$$
$$H_1 : M_{Morning.} > M_{othertimes.}$$

ตารางที่ 4.8 ผลการทดสอบค่าเฉลี่ย SQL Job ของกลุ่มช่วงเวลาเช้า และ กลุ่มที่ไม่ใช่ช่วงเวลาเช้า

Method	H-value	P-value
Not adjusted for ties	7530142	<0.001
Adjusted for ties	7530142	<0.001

เนื่องจากการทดสอบข้อกำหนดเบื้องต้นแล้ว พบว่า ข้อมูลไม่มีการแจกแจงปรกติ ดังนั้น สถิติที่ใช้ในการทดสอบ คือ The Kruskal-Wallis One-Way Analysis of Variance by Rank Test

จากตารางที่ 4.8 พบว่า ค่า P-value < 0.001 ซึ่งมีค่าน้อยกว่า 0.05 จึงปฏิเสธ H_0ที่ระดับนัยสำคัญ 0.05 นั้นคือค่าเฉลี่ยของ SQL Job ในกลุ่มช่วงเวลาเช้ามากกว่าช่วงเวลาอื่นๆที่ไม่ใช่ช่วงเวลาเช้า สามารถดูผลการวิเคราะห์ได้จากโปรแกรม Minitab ที่รูป ก.3 ของภาคผนวก ก

4.3.4 การทดสอบสมมติฐานระยะเวลาที่ใช้ในการประมวลผลของ SQL Job

การวิเคราะห์ทางสถิติเพื่อตรวจสอบว่าระยะเวลาการประมวลผล SQL Job ในแต่ละวันของสัปดาห์และในแต่ละช่วงเวลามีความแตกต่างระหว่างกลุ่มได้ โดยใช้สถิติทดสอบทดสอบครัสคาล-วอลลิส (The Kruskal-Wallis One-Way Analysis of Variance by Rank Test) ด้วยโปรแกรมสำเร็จรูป Minitab Statistical Software ได้ผลลัพธ์ดังต่อไปนี้

สมมติฐานของการทดสอบ

H_0 : ค่ามัธยฐานระยะเวลาการประมวล SQL Job ทั้ง 7 วัน ไม่แตกต่างกัน

H_1 : ค่ามัธยฐานระยะเวลาการประมวล SQL Job อย่างน้อย 1 วันแตกต่างกัน

ตารางที่ 4.9 ผลการทดสอบความแตกต่างของระยะเวลาที่ใช้ในการประมวลผลของ SQL Job แต่ละวันของสัปดาห์

Method	DF	H-value	P-value
Not adjusted for ties	6	5.48	0.483
Adjusted for ties	6	5.72	0.456

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เนื่องจากการทดสอบข้อกำหนดเบื้องต้นแล้ว พบว่า ข้อมูลไม่มีการแจกแจงปรกติ ดังนั้น สถิติที่ใช้ในการทดสอบ คือ The Kruskal-Wallis One-Way Analysis of Variance by Rank Test

จากตารางที่ 4.9 พบว่า ค่า P-value มีค่ามากกว่ากว่า 0.05 จึงยอมรับ H_0ที่ระดับนัยสำคัญ 0.05 นั้นคือระยะเวลาที่ใช้ในการประมวลผลของ SQL Job ทั้ง 7 วัน มีค่ามัธยฐานไม่แตกต่างกัน สามารถดูผลการวิเคราะห์ได้จากโปรแกรม Minitab ที่รูป ก.4 ของภาคผนวก ก

สมมติฐานของการทดสอบ

H_0 : ค่ามัธยฐานระยะเวลาการประมวผล SQL Job ทั้ง 4 ช่วงเวลาไม่แตกต่างกัน

H_1 : ค่ามัธยฐานระยะเวลาการประมวผล SQL Job อย่างน้อย 1 ช่วงเวลาแตกต่างกัน

ตารางที่ 4.10 ผลการทดสอบความแตกต่างของระยะเวลาที่ใช้ในการประมวลผลของ SQL Job แต่ละช่วงเวลา

Method	DF	H-value	P-value
Not adjusted for ties	3	431.37	<0.001
Adjusted for ties	3	449.64	<0.001

เนื่องจากการทดสอบข้อกำหนดเบื้องต้นแล้ว พบว่า ข้อมูลไม่มีการแจกแจงปรกติ ดังนั้น สถิติที่ใช้ในการทดสอบ คือ The Kruskal-Wallis One-Way Analysis of Variance by Rank Test

จากตารางที่ 4.10 พบว่า ค่า P-value < 0.001 ซึ่งมีค่าน้อยกว่า 0.05 จึงปฏิเสธ H_0 ที่ระดับนัยสำคัญ 0.05 นั้นคือระยะเวลาที่ใช้ในการประมวลผลของ SQL Job ทั้ง 4 ช่วงเวลา มีค่ามัธยฐานแตกต่างกัน สามารถดูผลการวิเคราะห์ได้จากโปรแกรม Minitab ที่รูป ก.5 ของภาคผนวก ก

4.3.5 การเปรียบเทียบพหุคูณระยะเวลาที่ใช้ในการประมวลผลของ SQL Job ในแต่ละช่วงเวลา

จากการทดสอบ The Kruskal-Wallis One-Way Analysis of Variance by Rank Testระยะเวลาที่ใช้ในการประมวลผลของ SQL Job ทั้ง 4 ช่วงเวลา มีค่ามัธยฐานแตกต่างกัน กรณีอยากทราบว่าระยะเวลาที่ใช้ในการประมวลผลของ SQL Job แต่ละช่วงเวลามีคู่ไหนแตกต่างกัน โดยใช้วิธีการเปรียบเทียบพหุคูณแบบ LSD หรือ Fisher's Least – Significant Different ได้ผลลัพธ์ดังต่อไปนี้

เอกสารนี้สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.11 การจัดกลุ่มข้อมูลโดยใช้วิธี Fisher LSD และช่วงความเชื่อมั้น 95%

Time of Day	N	Mean	Grouping
ช่วงเวลาเช้า	2474	21.522	A
ช่วงเวลาบ่าย	1506	12.535	B
ช่วงเวลาเย็น	275	9.6	B
ช่วงเวลากลางคืน	906	4.902	C

จากตารางที่ 4.11 พบว่าการจัดกลุ่มข้อมูลได้ 3 กลุ่ม ได้แก่ กลุ่ม A ในช่วงเวลาเช้าต่อมากลุ่ม B ในช่วงเวลาบ่ายและช่วงเวลาเย็น และกลุ่ม C ช่วงเวลากลางคืน สามารถดูผลการวิเคราะห์ได้จากโปรแกรม Minitab ที่รูป ก.6 ของภาคผนวก ก

ตารางที่ 4.12 ผลการทดสอบพหุคูณแบบ LSD ของระยะเวลาที่ใช้ในการประมวลผลของ SQL Job แต่ละช่วงเวลา

Difference of Levels	Difference of Mean	SE of Difference	95% CI	T-Value	P-Value
ช่วงเวลาเย็น-ช่วงเวลาบ่าย	-2.94	1.77	(-6.40, 0.53)	-1.66	0.097
ช่วงเวลาเช้า-ช่วงเวลาบ่าย	8.987	0.882	(7.258, 10.716)	10.19	<0.001
ช่วงเวลากลางคืน-ช่วงเวลาบ่าย	-7.63	1.13	(-9.86, -5.41)	-6.73	<0.001
ช่วงเวลาเช้า - ช่วงเวลาเย็น	11.92	1.72	(8.56, 15.28)	6.95	<0.001
ช่วงเวลากลางคืน-ช่วงเวลาเย็น	-4.7	1.86	(-8.34, -1.06)	-2.53	0.011
ช่วงเวลากลางคืน-ช่วงเวลาเช้า	-16.62	1.05	(-18.67, -14.57)	-15.86	<0.001

จากตารางที 4.12 พบว่าค่า P-value <0.001 มีค่าน้อยกว่า 0.05 จึงปฏิเสธ H_0 ที่ระดับนัยสำคัญ 0.05 แสดงว่าค่าเฉลี่ยของ Temp ในแต่ละ Plant แตกต่างกันอย่างน้อย 2 กลุ่ม

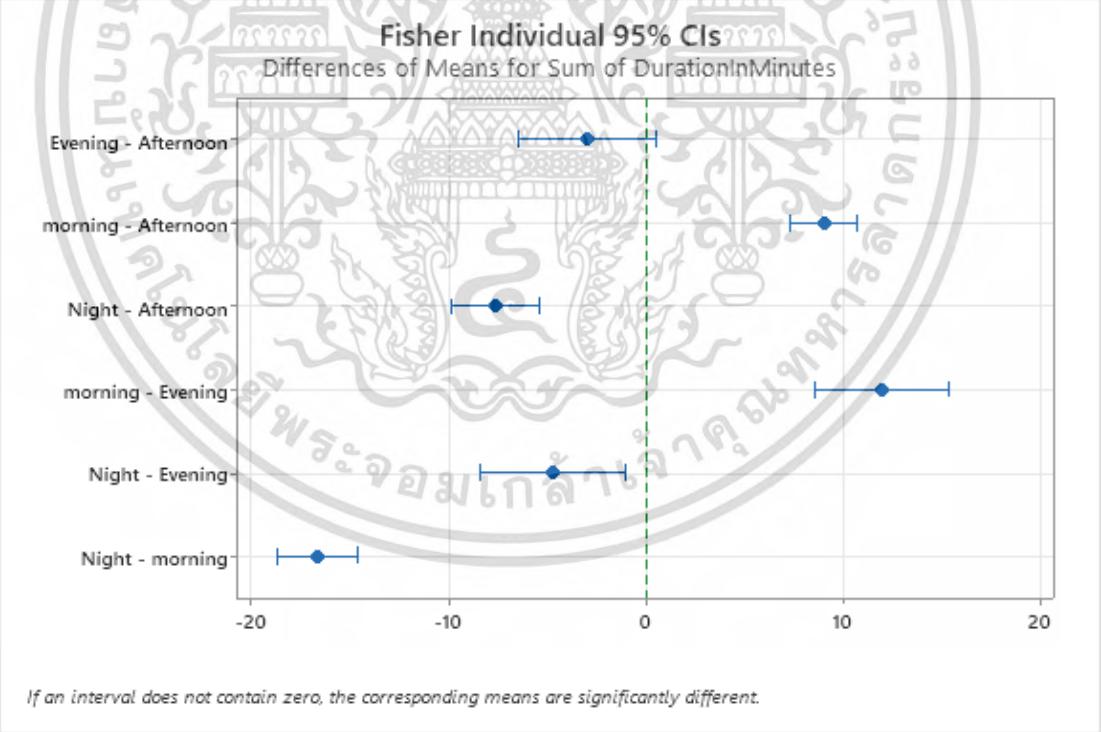
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สามารถดูผลการวิเคราะห์ได้จากโปรแกรม Minitab ที่รูป ก.7 ของภาคผนวก ก และการเปรียบเทียบเปรียบเทียบพหุคูณแบบ LSD จะให้ผลลัพธ์ดังนี้

- ช่วงเวลาเช้า แตกต่างกับ ช่วงเวลาบ่าย ที่ระดับนัยสำคัญ 0.05
- ช่วงเวลากลางคืน แตกต่างกับ ช่วงเวลาบ่าย ที่ระดับนัยสำคัญ 0.05
- ช่วงเวลาเช้า แตกต่างกับ ช่วงเวลาเย็น ที่ระดับนัยสำคัญ 0.05
- ช่วงเวลากลางคืน แตกต่างกับ ช่วงเวลาเย็น ที่ระดับนัยสำคัญ 0.05
- ช่วงเวลากลางคืน แตกต่างกับ ช่วงเวลาเช้า ที่ระดับนัยสำคัญ 0.05

หรือ

Plant	Temp
Morning	21.522^{A}
Afternoon	12.535^{B}
Evening	9.60^{B}
Night	4.902^{C}



รูปที่ 4.7 กราฟแสดงผลวิธีการเปรียบเทียบพหุคูณแบบ LSD หรือ Fisher's Least – Significant Different

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 4.7 แสดงผลลัพธ์วิธีการเปรียบเทียบพหุคูณแบบ LSD หรือ Fisher's Least – Significant Different มีทั้งหมด 5 คู่ที่แตกต่างกันสามารถดูได้จากเส้นของกราฟที่ไม่ครอบคลุมศูนย์ ซึ่งจะมีคู่เดียวที่ไม่แตกต่างกันคือ ช่วงเวลากลางคืนกับช่วงเวลาเย็น

4.4 อภิปรายผลการวิจัย

การสร้างแผงหน้าปัด (Dashboard) ที่แสดงผลแบบเรียลไทม์และย้อนหลังสำหรับการติดตามและบริหารจัดการ SQL Job โดยแสดงข้อมูลสำคัญเช่น Min Job Duration, Max Job Duration, Average Job Duration, Median Job Duration, จำนวนงานที่ประมวลผล, งานที่สำเร็จและงานที่ล้มเหลวเป็นต้น โดยข้อมูลเหล่านี้ช่วยให้ผู้ดูแลระบบสามารถมองเห็นภาพรวมการทำงานและประสิทธิภาพของ SQL Job ได้อย่างรวดเร็วและชัดเจน ซึ่งข้อมูลในวันพุธมีจำนวน SQL Job มากที่สุด อยู่ที่ร้อยละ 16.22 อาจเกิดจากการประมวลของบาง Job ที่จะเกิดขึ้นอาทิตย์ละครั้งในวันพุธ หรืออาจเกิดการประมวลผลของ Job นั้นหลาย ๆ ครั้งเพื่อทำการแก้ไขบางอย่าง หากลดจำนวนจากสาเหตุที่กล่าวมาก็จะให้ทราบว่าในทุก ๆ วันจะมีจำนวนมีการประมวลไม่ต่างกัน และในส่วนช่วงเวลาเช้าจะเป็นช่วงเวลาที่มีการประมวลผลมากที่สุด อยู่ที่ร้อยละ 47.97 อาจเกิดจากการต้องการรายงานของทุก SQL Job ในช่วงเวลาเช้า จึงมีการประมวลของทุก Job ทำให้เกิดความแตกต่างในระยะเวลาการประมวลผลในช่วงเวลานั้นและส่งผลต่อการประมวลผลที่ล่าช้าขึ้นเพราะต้องมีการประมวลผลพร้อมกันทั้งหมดในช่วงเวลาเดียวกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5
สรุปผลการวิจัยและข้อเสนอแนะ

สรุปผลการวิจัยการวิเคราะห์ระยะเวลาที่ใช้ในการการประมวลผล SQL Server Agent JOB โดยอาศัยแผงหน้าปัดธุรกิจ ซึ่งมีวัตถุประสงค์ดังนี้ 1) เพื่อออกแบบแผงหน้าปัดธุรกิจ หรือ Dashboard ที่สามารถระบุระยะเวลาการประมวลผล SQL Job แบบเรียลไทม์ได้ ที่ 2) เพื่อวิเคราะห์ระยะเวลาการประมวลผลของ SQL Job โดยอาศัยการทดสอบแมนน์-วิตนีย์ (The Mann-Whitney Test) และการทดสอบครัสคาล-วอลลิส (The Kruskal-Wallis One-Way Analysis of Variance by Rank Test) ที่เป็นสถิติเชิงอนุมาน บริษัทจะสามารถนำไปใช้และสามารถแก้ปัญหาได้ จึงสามารถสรุปผลและข้อเสนอแนะได้ดังนี้

5.1 สรุปผลการวิจัย

แผงหน้าปัดรายงานการวิเคราะห์การการทำงานของ SQL Job ทำให้ให้เห็นภาพรวมโดยย่อเกี่ยวกับ SQL Job แบบเรียลไทม์และย้อนหลังได้ในแต่ละวัน แสดงจำนวน Min Job Duration จำนวน Max Job Duration จำนวน Average Job Duration Median Job Duration Job Executed Job Succeeded Job Failed โดยแสดงผลในรูปแบบตัวเลข และแผนภูมิเพื่อให้ทราบจำนวน SQL Job แต่ละกลุ่มที่มีการประมวลผลในแต่ละครั้ง ข้อมูลเหล่านี้ช่วยให้ผู้ดูแลระบบสามารถมองเห็นภาพรวมการทำงานและประสิทธิภาพของ SQL Job ได้อย่างรวดเร็วและชัดเจน ใช้ในการสนับสนุนแผนการส่งเสริมการจัดการ SQL Job และใช้ในการปรับปรุงการในอนาคต

การวิเคราะห์ระยะเวลาการประมวลผลของ SQL Job พบว่าในวันพุธมีจำนวน SQL Job มากที่คิดคิดเป็นร้อยละ 16.22% และช่วงเวลาเช้ามากที่สุดเช่นเดียวกันคิดเป็นร้อยละ 47.97% ซึ่งสามารถบอกค่าความแตกต่างระหว่างของกลุ่มจำนวนและระยะเวลาการการประมวลผล โดยใช้การทดสอบของแมนน์-วิตนีย์ (The Mann – Whitney test) ในส่วนผลที่ได้จากการวิเคราะห์คือ ค่าเฉลี่ยของ SQL Job ในกลุ่มวันพุธน้อยกว่าหรือมากกว่าวันอื่นๆที่ไม่ใช่วันพุธ และค่าเฉลี่ยของ SQL Job ในกลุ่มช่วงเวลาเช้ามากกว่ากลุ่มที่ไม่ใช่ช่วงเวลาเช้า ต่อมาใช้สถิติทดสอบ ครัสคาล-วอลลิส (The Kruskal-Wallis One-Way Analysis of Variance by Rank Test) และผลที่ที่ได้จากการวิเคราะห์คือ ค่ามัธยฐานระยะเวลการประมวผล SQL Job ทั้ง 7 วันไม่มีความแตกต่างกัน และค่ามัธยฐานระยะเวลการประมวผล SQL Job ทั้ง 4 ช่วงเวลามีความแตกต่างแตกต่างกัน อย่างมีนัยสำคัญทางสถิติ สามารถบอกเวลาในแต่ละช่วงที่แตกต่างมีดังนี้ ได้แก่ ช่วงเวลาเช้าแตกต่างกับช่วงเวลาบ่าย, ช่วงเวลากลางคืนแตกต่างกับช่วงเวลาบ่าย, ช่วงเวลาเช้าแตกต่างกับช่วงเวลาเย็น, ช่วงเวลากลางคืนแตกต่างกับช่วงเวลาเย็น และช่วงเวลากลางคืนแตกต่างกับช่วงเวลาเช้า ซึ่งสามารถใช้เป็นข้อมูลสำคัญในการจัดการและปรับปรุงการกระมวลผล SQL Job ในแต่ละช่วงเวลาให้มีประสิทธิภาพยิ่งขึ้น

เอกสารนี้สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ ไม่ว่ากรณีใดๆ ทั้งสิ้น ไม่อนุญาตให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารนี้ทุกครั้งที่มีการนำไปใช้

5.2 ข้อจำกัดและข้อเสนอแนะ

5.2.1 ข้อจำกัด

เนื่องจากเวลาในการเก็บข้อมูล เป็นการเริ่มต้นในการเก็บข้อมูลที่ใหม่ทั้งหมดซึ่งไม่มีข้อมูลที่เก็บไว้ก่อนหน้านี้จึงทำให้ข้อมูลที่ทำวิจัยมีจำกัด จึงทำให้การวิเคราะห์ระยะเวลาการประมวลผล SQL Server Agent JOB อาจจำเป็นต้องใช้ข้อมูลจากฐานข้อมูล SQL Server Agent ซึ่งอาจมีข้อมูลที่ไม่ครอบคลุมหรือไม่เพียงพอ

5.2.2 ข้อเสนอแนะ

1) สร้างแผนภาพหรือกราฟเพื่อแสดงความสัมพันธ์ระหว่างระยะเวลาในการประมวลผลและปัญหาที่เกิดขึ้น เพื่อช่วยในการทำนายและการวางแผนจัดกลุ่ม SQL Job ในอนาคต

2) SQL Job ที่มีสถานการณ์ประมวลผลที่ล้มเหลวบ่อย ๆ หรือมีข้อผิดพลาดเกิดขึ้นมากขึ้น อาจส่งผลกระทบต่อประสิทธิภาพของระบบฐานข้อมูลอย่างมาก เนื่องจาก SQL Job ที่ล้มเหลวอาจทำให้ข้อมูลไม่ถูกประมวลผลหรือบันทึกได้ตามที่ควร และอาจทำให้เกิดความเสียหายกับข้อมูลที่มีอยู่ในระบบ ซึ่งอาจส่งผลกระทบต่อการดำเนินงานของธุรกิจที่ใช้งานระบบฐานข้อมูลนั้น

3) จากการวิเคราะห์ข้อมูลเหล่านี้สามารถใช้ในการวางแผนการกระจายงานให้เหมาะสมกับช่วงเวลา เช่น การเลื่อนและลดจำนวน SQL Job ที่ไม่เร่งด่วนไปยังช่วงเวลาที่มีภาระ SQL Job น้อย หรือการเพิ่มทรัพยากรในช่วงเวลาที่มีการใช้งานสูง เพื่อเพิ่มประสิทธิภาพการประมวลผล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง

ทศพล บ้านคลองสี่ และจรัญ แสนราช. 2564. ***การสร้างภาพข้อมูลด้วยแท็บโบลว์เพื่อความเข้าใจเกี่ยวกับสถานการณ์โควิด-19.*** [ออนไลน์]. เข้าถึงได้จาก https://www.ojs.kmutnb.ac.th/index.php/jote/article/view/4831/pdf_17

นภสร ใจทับทิม. 2564. ***การวิเคราะห์ข้อมูลและสร้างแดชบอร์ดด้วยโปรแกรม Tableau.*** [ออนไลน์]. เข้าถึงได้จาก http://IT-Digital-business-2021-coop-Analyzing-data-and-creating- dashboards-with-Tableau_.pdf

ภูวิดล ด่านระหาญ. 2550. ***Syslog-ng (Syslog new generation).*** [ออนไลน์]. เข้าถึงได้จาก http://.www.msit.mut.ac.th/newweb/phpfile/show.php?Qid=776.

เมธาพร เจ็งธรรม. 2563. ***การออกแบบวิธีการนำเสนอโปรแกรม Business Intelligence เพื่อผู้มีภาวะตาบอดสี.*** [ออนไลน์]. เข้าถึงได้จาก http://library.tni.ac.th/thesis/upload/files/IS%20MIT%202020/Methaporn%20Chengtham%20IS%20MIT%202020.pdf

รัตนา สุวรรณวิชนีย์. 2560. ***การพัฒนาระบบรายงานรูปแบบหลายมิติเพื่อสนับสนุนการตัดสินใจเชิงนโยบายของสำนักงานคณะกรรมการวิจัยแห่งชาติ***. วิทยานิพนธ์ปริญญามหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิชาเทคโนโลยีสารสนเทศ มหาวิทยาลัยศรีปทุม

วนิดา ธรรมคุณ. 2562. ***ระบบข่าวกรองธุรกิจทางการตลาด สำหรับธุรกิจให้บริการขนส่ง กรณีศึกษาห้างหุ้นส่วนจำกัด เก้าเลี้ยว ทรานสปอร์ต***. สารนิพนธ์ปริญญามหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิชาเทคโนโลยีสารสนเทศ มหาวิทยาลัยศรีปทุม.

วรกานต์ เชื้อสิงห์. 2563. ***การพัฒนาระบบดาต้าวิชวลไลเซชันเพื่อสนับสนุนการวางแผนเชิงกลยุทธการตลาดขององค์กรประเภทพาณิชย์อิเล็กทรอนิกส์.*** [ออนไลน์]. เข้าถึงได้จาก http://dspace.spu.ac.th/bitstream/123456789/8735/1/วรกานต์ เชื้อสิงห์.pdf

สุจิตรา สุคนธมัต. 2565. ***เอกสารประกอบการสอนบิชาโปรแกรมสำเร็จรูปทางสถิติ.*** ภาควิชาสถิติ คณะวิทยาศาสตร์ ถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.

สมชาย อารยพิทยา. 2559. ***ข้อมูลจราจรคอมพิวเตอร์ (logfile).*** [ออนไลน์]. เข้าถึงได้จาก https://erp.mju.ac.th/acticleDetail.aspx?qid=606.

สุธาพร ล้าเลิศกุล, 2561, เลือกฟอนต์ให้โดนใจและเหมาะสมกับงานนำเสนอ, ใช้ฟอนต์อะไรดี, กรุงเทพฯ: ม.ป.พ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง (ต่อ)

อภิยศ เหรียญวิพัฒน์. (2563). ***การวิเคราะห์ข้อมูลเพื่อสร้าง Dashboard แสดงสถิติการใช้บริการกึ่งเรียลไทม์ของสำนักหอสมุด มหาวิทยาลัยเกษตรศาสตร์***. [ออนไลน์]. เข้าถึงได้จาก http://pulinet2020.tsu.ac.th/Documentation/Proceeding/Oral/IT/09.pdf

อุมาพร จันทศรม. 2542. ***สถิติที่ไม่ใช้พารามิเตอร์.*** บริษัท ฟิสิกส์เซ็นเตอร์, กรุงเทพฯ

Bennyhoff, M., 2023. ***SQL Server Agent Job Schedule Reporting.*** [ออนไลน์]. เข้าถึงได้จาก https://www.bps-corp.com/post/sql-server-agent-job-schedule-reporting

Boards, N., 2020. ***Tableau vs. Power BI vs. Looker – Which tool is better for your Business?.*** [ออนไลน์]. เข้าถึงได้จาก https://nineboards.com/tableau-vs-power-bi-vs-looker-which-tool-is-better-for-your-business/

Devine, M., 2024. ***Dashboards vs Reports: A Comparison.*** [ออนไลน์]. เข้าถึงได้จาก https://www.calxa.com/dashboards-vs-reports/

Kirk. Roger ,E. 1995. ***Experimental design : procedures for the behavioral sciences. 3ND ed***. California : Brooks/Cole Publishing Company.

Pearson, A.V. and Hartley, H.O. 1972. ***Biometrical Tables for Statisticians***, Vol 2, Cambridge, England, Cambridge University Press.

Raj, A., 2022, ***Data-ink Ratio Explained With Example.*** [ออนไลน์]. เข้าถึงได้จาก https://www.codeconquest.com/blog/data-ink-ratio-explained-with-example/

Softnix. 2019. ***แนวทางการใช้สีใน Data Visualization,*** [ออนไลน์]. เข้าถึงได้จาก Available: https://www.softnix.co.th/2019/03/26/แนวทางการใช้สีใน-data-visualization/[Accessed: January 17, 2020].

Wayan, W., ***Improving data visualization using the principle of Data-Ink Ratio.*** [ออนไลน์]. เข้าถึงได้จาก https://www.linkedin.com/pulse/improving-data-visualization-using-principle-data-ink-wijesinghe?utm_source=share&utm_medium=member_ios&utm_campaign=share_via

Wilke, C. O., ***Fundamentals of Data Visualization***. [ออนไลน์]. เข้าถึงได้จาก Available: https://serialmentor.com/dataviz/index.html [Accessed: September 24, 2019].

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ก

ก.1 ผลการทดสอบความเป็นอิสระของตัวแปรวัน และตัวแปรช่วงเวลา

Chi-Square Test

	Chi-Square	DF	P-Value
Pearson	99.594	18	0.000
Likelihood Ratio	98.220	18	0.000

รูปที่ ก.1 ผลการทดสอบความเป็นอิสระของตัวแปรวัน และตัวแปรช่วงเวลา

จากรูปที่ ก.2 พบว่า ค่า P-value มีน้อยกว่า 0.05 จึงปฏิเสธที่ระดับนัยสำคัญ 0.05 ตัวแปรวัน และ ตัวแปรช่วงเวลา ไม่เป็นอิสระต่อกัน

ก.2 ผลการทดสอบค่าเฉลี่ย SQL Job ของกลุ่มวันพุธ และ กลุ่มที่ไม่ใช่วันพุธ

Test

Null hypothesis $H_0: \eta_1 - \eta_2 = 0$
Alternative hypothesis $H_1: \eta_1 - \eta_2 > 0$

Method	W-Value	P-Value
Not adjusted for ties	24438877.00	0.603

รูปที่ ก.2 ผลการทดสอบค่าเฉลี่ย SQL Job ของกลุ่มวันพุธ และ กลุ่มที่ไม่ใช่วันพุธ

จากรูปที่ ก.2 พบว่า ค่า P-value มีค่ามากกว่า 0.05 จึงยอมรับที่ระดับนัยสำคัญ 0.05 นั้นคือ ค่าเฉลี่ยของ SQL Job ในกลุ่มวันพุธน้อยกว่าหรือเท่ากับวันอื่นๆ ที่ไม่ใช่วันพุธ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ก.3 ผลการทดสอบค่าเฉลี่ย SQL Job ของกลุ่มช่วงเวลาเช้า และ กลุ่มที่ไม่ใช่ช่วงเวลาเช้า

Test

Null hypothesis $H_0: \eta_1 - \eta_2 = 0$
Alternative hypothesis $H_1: \eta_1 - \eta_2 > 0$

Method	W-Value	P-Value
Not adjusted for ties	7530142.00	0.000
Adjusted for ties	7530142.00	0.000

รูปที่ ก.3 ผลการทดสอบความแตกต่างของระยะเวลาที่ใช้ในการประมวลผลของ SQL Job แต่ละวันของสัปดาห์

จากรูปที่ ก.3 พบว่า ค่า P-value < 0.001 ซึ่งมีค่าน้อยกว่า 0.05 จึงปฏิเสธที่ระดับนัยสำคัญ 0.05 นั้นคือค่าเฉลี่ยของ SQL Job ในกลุ่มช่วงเวลาเช้ามากกว่าช่วงเวลาอื่นๆที่ไม่ใช่ช่วงเวลาเช้า

ก.4 ผลการทดสอบความแตกต่างของระยะเวลาที่ใช้ในการประมวลผลของ SQL Job แต่ละวันของสัปดาห์

Test

Null hypothesis H_0: All medians are equal
Alternative hypothesis H_1: At least one median is different

Method	DF	H-Value	P-Value
Not adjusted for ties	6	5.48	0.483
Adjusted for ties	6	5.72	0.456

รูปที่ ก.4 ผลการทดสอบความแตกต่างของระยะเวลาที่ใช้ในการประมวลผลของ SQL Job แต่ละวันของสัปดาห์

จากรูปที่ ก.4 พบว่า ค่า P-value มีค่ามากกว่ากว่า 0.05 จึงยอมรับที่ระดับนัยสำคัญ 0.05 นั้นคือระยะเวลาที่ใช้ในการประมวลผลของ SQL Job ทั้ง 7 วัน มีค่ามัธยฐานไม่แตกต่างกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ก.5 ผลการทดสอบความแตกต่างของระยะเวลาที่ใช้ในการประมวลผลของ SQL Job แต่ละช่วงเวลา

Test

Null hypothesis H_0: All medians are equal
Alternative hypothesis H_1: At least one median is different

Method	DF	H-Value	P-Value
Not adjusted for ties	3	431.37	0.000
Adjusted for ties	3	449.64	0.000

รูปที่ ก.5 ผลการทดสอบความแตกต่างของระยะเวลาที่ใช้ในการประมวลผลของ SQL Job แต่ละช่วงเวลา

จากรูปที่ ก.5 พบว่า ค่า P-value < 0.001 ซึ่งมีค่าน้อยกว่า 0.05 จึงปฏิเสธที่ระดับนัยสำคัญ 0.05 นั้นคือระยะเวลาที่ใช้ในการประมวลผลของ SQL Job ทั้ง 4 ช่วงเวลา มีค่ามัธยฐานแตกต่างกัน

ก.6 การจัดกลุ่มข้อมูลโดยใช้วิธี Fisher LSD และช่วงความเชื่อมั้น 95%

Grouping Information Using the Fisher LSD Method and 95% Confidence

TimePeriod	N	Mean	Grouping
morning	2474	21.522	A
Afternoon	1506	12.535	B
Evening	275	9.60	B
Night	906	4.902	C

Means that do not share a letter are significantly different.

รูปที่ ก.6 ผลการจัดกลุ่มข้อมูลโดยใช้วิธี Fisher LSD และช่วงความเชื่อมั้น 95%

จากรูปที่ ก.6 พบว่าการจัดกลุ่มข้อมูลได้ 3 กลุ่ม ได้แก่ กลุ่ม A ในช่วงเวลาเช้า ต่อมากลุ่ม B ในช่วงเวลาบ่ายและช่วงเวลาเย็น และกลุ่ม C ช่วงเวลากลางคืน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ก.7 ผลการทดสอบพหุคูณแบบ LSD ของระยะเวลาที่ใช้ในการประมวลผลของ SQL Job แต่ละช่วงเวลา

Fisher Individual Tests for Differences of Means

Difference of Levels	Difference of Means	SE of Difference	95% CI	T-Value	Adjusted P-Value
Evening - Afternoon	-2.94	1.77	(-6.40, 0.53)	-1.66	0.097
morning - Afternoon	8.987	0.882	(7.258, 10.716)	10.19	0.000
Night - Afternoon	-7.63	1.13	(-9.86, -5.41)	-6.73	0.000
morning - Evening	11.92	1.72	(8.56, 15.28)	6.95	0.000
Night - Evening	-4.70	1.86	(-8.34, -1.06)	-2.53	0.011
Night - morning	-16.62	1.05	(-18.67, -14.57)	-15.86	0.000

Simultaneous confidence level = 79.69%

รูปที่ ก.7 ผลการทดสอบพหุคูณแบบ LSD ของระยะเวลาที่ใช้ในการประมวลผลของ SQL Job แต่ละช่วงเวลา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประวัติผู้วิจัย

กฤตบุญ นิยม (โอม)

ที่อยู่ 28 หมู่ 1 ต.สันทราย อ.แม่จัน จ.เชียงราย 57110

เบอร์โทรศัพท์ : 082-381-4926 E-mail : kittaboon.o@gmail.com

ประวัติการศึกษา

- ระดับมัธยมศึกษา
  โรงเรียนดำรงราษฎร์สงเคราะห์ จังหวัดเชียงราย
- ระดับปริญญาตรี
  วิทยาศาสตรบัณฑิต สาขาสถิติประยุกต์
  ภาควิชาสถิติ คณะวิทยาศาสตร์
  สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

งานทะเบียนคณะวิทยาศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

คำรับรองเล่มสหกิจศึกษา

วันที่ 24 เดือน มิถุนายน พ.ศ. 2567

ข้าพเจ้า นายกฤตบุญ นิยม รหัสประจำตัว 63050601 นักศึกษาหลักสูตร วิทยาศาสตรบัณฑิต สาขาวิชา สถิติประยุกต์ ภาควิชา สถิติ .

ขอรับรอง สหกิจศึกษา เรื่อง

ชื่อภาษาไทย การวิเคราะห์ระยะเวลาที่ใช้ในการการประมวลผล SQL SERVER AGENT JOB โดยอาศัยแผงหน้าปัดธุรกิจ

ชื่อภาษาอังกฤษ ANALYZING PROCESSING TIME OF SQL SEVER AGENT JOB USING BUSINESS DASHBOARD PANEL

ปีการศึกษา 2566

เป็นผลงานวิจัยของข้าพเจ้าแต่เพียงผู้เดียวโดยมิได้คัดลอกหรือละเมิดลิขสิทธิ์ของผู้อื่นและได้ผ่านการตรวจสอบความซ้ำซ้อนโดยใช้โปรแกรมเรียบร้อยแล้ว และได้แนบเอกสารการตรวจสอบการลอกเลียนงานวรรณกรรมที่ตรวจสอบจากวิทยานิพนธ์ฉบับสมบูรณ์แล้ว

โปรแกรมอักขราวิสุทธิ์ 1.09%

ลงชื่อ.....กฤตบุญ นิยม.....

(นายกฤตบุญ นิยม)

นักศึกษา

ข้าพเจ้า ผ.ศ.ดร.พรพิมล ชัยวุฒิศักดิ์ อาจารย์ที่ปรึกษาโครงงานสหกิจศึกษา ได้ตรวจสอบโครงงานสหกิจศึกษาของนักศึกษาข้างต้นแล้ว ขอรับรองว่าเป็นผลงานวิจัยของนักศึกษาจริงและมีเนื้อหาสมบูรณ์จึงขอลงชื่อไว้เป็นหลักฐาน

ลงชื่อ.....Pornpimol Chaiwuttisak.....

(ผศ.ดร.พรพิมล ชัยวุฒิศักดิ์)

อาจารย์ที่ปรึกษา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้