คลังข้อมูล ธุรกิจอัจฉริยะ และการวิเคราะห์ข้อมูลขั้นสูงของธุรกิจร้านอาหารและเครื่องดื่ม

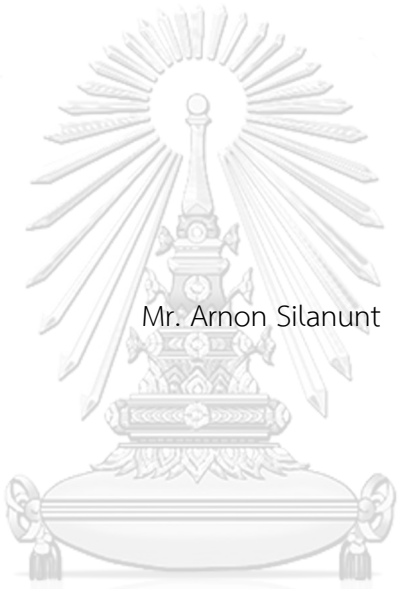
นายอานนท์ ศิลานันท์

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

สารนิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาเทคโนโลยีสารสนเทศทางธุรกิจ
คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2565
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

DATA WAREHOUSE, BUSINESS INTELLIGENCE AND ADVANCED DATA ANALYTICS OF FOOD & BEVERAGE BUSINESS

Mr. Arnon Silanunt



จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

An Independent Study Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Information Technology in Business
FACULTY OF COMMERCE AND ACCOUNTANCY
Chulalongkorn University
Academic Year 2022
Copyright of Chulalongkorn University

หัวข้อสารนิพนธ์	คลังข้อมูล ธุรกิจอัจฉริยะ และการวิเคราะห์ข้อมูลขั้นสูงของธุรกิจร้านอาหารและเครื่องดื่ม
โดย	นายอานนท์ ศิลานันท์
สาขาวิชา	เทคโนโลยีสารสนเทศทางธุรกิจ
อาจารย์ที่ปรึกษาหลัก	รองศาสตราจารย์ ดร.จันทร์เจ้า มงคลนาวิน

คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับสารนิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

คณะกรรมการสอบสารนิพนธ์

..... ประธานกรรมการ
(รองศาสตราจารย์ ดร.วชิระ บุณยเนตร)

..... อาจารย์ที่ปรึกษาหลัก
(รองศาสตราจารย์ ดร.จันทร์เจ้า มงคลนาวิน)

..... กรรมการ
(ผู้ช่วยศาสตราจารย์สุภารัตน์ ตันทนงศักดิ์กุล)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

อานนท์ ศิลานันท์ : คลังข้อมูล ธุรกิจอัจฉริยะ และการวิเคราะห์ข้อมูลขั้นสูงของธุรกิจร้านอาหารและเครื่องดื่ม. (DATA WAREHOUSE, BUSINESS INTELLIGENCE AND ADVANCED DATA ANALYTICS OF FOOD & BEVERAGE BUSINESS) อ.ที่ปรึกษาหลัก : รศ. ดร.จันทร์เจ้า มงคลนาวิน

ในช่วงเวลาที่ผ่านมาธุรกิจร้านอาหารและเครื่องดื่มมีจำนวนเพิ่มมากขึ้น ทั้งในเขตกรุงเทพ ฯ และปริมณฑล อย่างไรก็ตาม สถานการณ์ระบาดโควิด-19 ส่งผลให้การเติบโตของธุรกิจร้านอาหารและเครื่องดื่มชะลอตัวลง การพัฒนาเครื่องมือช่วยในการวิเคราะห์และตัดสินใจในการดำเนินธุรกิจเป็นวิธีหนึ่งที่จะช่วยให้ธุรกิจร้านอาหารและเครื่องดื่มรับมือกับการเปลี่ยนแปลงและสามารถวางแผนดำเนินธุรกิจให้ประสบความสำเร็จมากยิ่งขึ้น

โครงการ “คลังข้อมูล ธุรกิจอัจฉริยะ และการวิเคราะห์ข้อมูลขั้นสูงของธุรกิจร้านอาหารและเครื่องดื่ม” ประกอบด้วย 5 ระบบหลัก ได้แก่ ระบบวิเคราะห์ยอดขาย ระบบการวิเคราะห์การสั่งอาหารและเครื่องดื่ม ระบบแบ่งกลุ่มใบเสร็จค่าอาหารและเครื่องดื่ม ระบบวิเคราะห์กฎความสัมพันธ์ในใบเสร็จค่าอาหารและเครื่องดื่ม และระบบพยากรณ์ยอดขาย ระบบได้พัฒนาขึ้นบนระบบจัดการฐานข้อมูล Google Bigquery พัฒนารายงานวิเคราะห์โดยใช้เครื่องมือ Tableau Desktop และวิเคราะห์ข้อมูลขั้นสูงด้วยภาษาไพธอน

ระบบที่พัฒนาขึ้นจะช่วยให้ผู้บริหารวิเคราะห์ข้อมูลในมุมมองต่าง ๆ ได้อย่างถูกต้องและรวดเร็ว เพื่อนำไปใช้ประกอบการตัดสินใจดำเนินธุรกิจและทำความเข้าใจพฤติกรรมการใช้บริการของลูกค้าเพื่อสร้างความได้เปรียบทางการแข่งขันในธุรกิจ

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

สาขาวิชา	เทคโนโลยีสารสนเทศทางธุรกิจ	ลายมือชื่อนิสิต .....
ปีการศึกษา	2565	ลายมือชื่อ อ.ที่ปรึกษาหลัก .....

6382207026 : MAJOR INFORMATION TECHNOLOGY IN BUSINESS
KEYWORD:

Arnon Silanunt : DATA WAREHOUSE, BUSINESS INTELLIGENCE AND ADVANCED DATA ANALYTICS OF FOOD & BEVERAGE BUSINESS. Advisor: Assoc. Prof. JANJAO MONGKOLNAVIN

In the past years, the number of food and beverage businesses has increased in Bangkok and its surrounding. However, the Covid-19 epidemic has slowed down its growth. The development of an analyzing and decision-supporting tool is one approach that enables the business to handle the change and make a good business plan.

The project 'Data Warehouse, Business Intelligence, and Advanced Data Analytics of Food & Beverage Business' consists of 5 systems: Sales analysis system, Food & beverage order analysis system, Food & beverage receipt segmentation system, Food & beverage receipt association analysis system, and sales forecasting system. The system is developed using Google BigQuery as the database management system, Tableau Desktop as the analytical reporting tool, and Python as the Advanced data analytic programming tool.

This developed system will help executives analyze data from different perspectives accurately and quickly and use the analysis results to support business decisions, understand customer behavior and create a competitive advantage.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Field of Study:	Information Technology in Business	Student's Signature .....
Academic Year:	2022	Advisor's Signature .....

กิตติกรรมประกาศ

โครงการพิเศษเรื่อง “คลังข้อมูล ธุรกิจอัจฉริยะ และการวิเคราะห์ข้อมูลขั้นสูงของธุรกิจร้านอาหารและเครื่องดื่ม” นี้สำเร็จลุเรื่องไปได้ด้วยดี เนื่องมาจากความช่วยเหลือและสนับสนุนจากหลายฝ่าย โดยเฉพาะอย่างยิ่ง รองศาสตราจารย์ ดร. จันทร์เจ้า มงคลนาวิน อาจารย์ที่ปรึกษาโครงการพิเศษที่ได้สละเวลาอันมีค่าในการให้คำแนะนำ ปรึกษา ตรวจสอบ และแก้ไขจุดบกพร่องต่าง ๆ เป็นอย่างดีตลอดมา ผู้จัดทำขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ ที่นี้

ขอขอบพระคุณคณาจารย์ในหลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาเทคโนโลยีสารสนเทศทางธุรกิจทุกท่าน ที่ได้ประสิทธิ์ประสาทวิชาความรู้ต่าง ๆ ตลอดการศึกษาให้แก่ผู้จัดทำ ซึ่งทางผู้จัดทำได้นำความรู้ทั้งหมดที่ได้รับมาประกอบใช้ในการจัดทำโครงการพิเศษนี้ ทำให้โครงการพิเศษนี้สำเร็จ

ขอขอบพระคุณบิดามารดา พี่น้อง เพื่อน ๆ ที่ทำงานที่ให้การสนับสนุนและเป็นกำลังใจให้ตลอดการศึกษาและการจัดทำโครงการพิเศษนี้

ขอขอบคุณเพื่อน ๆ หลักสูตรเทคโนโลยีสารสนเทศทางธุรกิจ รุ่น 28 ภาคนอกเวลา ตลอดจนเจ้าหน้าที่ในหลักสูตรเทคโนโลยีสารสนเทศทางธุรกิจทุกท่าน สำหรับความช่วยเหลือในด้านต่าง ๆ เป็นอย่างดีตลอดระยะเวลาที่ศึกษาจนกระทั่งโครงการพิเศษนี้สำเร็จลุล่วงไปได้ด้วยดี

สุดท้ายนี้ผู้จัดทำหวังว่าโครงการพิเศษนี้จะเป็นประโยชน์ต่อผู้ที่สนใจ และเป็นแนวทางในการดำเนินการพัฒนาระบบอื่น ๆ ต่อไปในภายภาคหน้า ประโยชน์อันใดที่พึงเกิดขึ้นในอนาคต ผู้จัดทำขอมอบแด่ผู้มีพระคุณที่ได้กล่าวถึงทุกท่าน หากโครงการนี้มีจุดบกพร่องประการใด ผู้จัดทำขออภัยมา ณ ที่นี้

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

อานนท์ ศิลานันท์

สารบัญ

หน้า

..... ค

บทคัดย่อภาษาไทย ..... ค

..... ง

บทคัดย่อภาษาอังกฤษ ..... ง

กิตติกรรมประกาศ ..... จ

สารบัญ ..... ฉ

สารบัญตาราง ..... ซ

สารบัญรูปภาพ ..... ฌ

บทที่ 1 1

1.1 ความสำคัญและที่มาของโครงการ 1

1.2 วัตถุประสงค์ของโครงการ 6

1.3 ขอบเขตของโครงการ 7

1.4 วิธีการดำเนินงานโครงการ 8

1.5 เทคโนโลยีที่ใช้ในโครงการ 9

1.6 ประโยชน์ที่คาดว่าจะได้รับ 10

บทที่ 2 11

2.1 แนวคิดด้านคลังข้อมูล (Data Warehouse) 11

2.2 แนวคิดด้านระบบธุรกิจอัจฉริยะ (Business Intelligence: BI) 17

2.3 แนวคิดด้านการวิเคราะห์ข้อมูลเชิงทำนาย 21

บทที่ 3 23

3.1 ข้อมูลเกี่ยวกับองค์กร 23

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

3.2 โครงสร้างองค์กร 24
3.3 การดำเนินงานขององค์กร 25
3.4 ปัญหาที่เกิดขึ้นในปัจจุบัน 25
บทที่ 4 26
4.1 การวิเคราะห์ระบบ 26
4.2 การออกแบบระบบ 51
4.3 การติดตั้งและพัฒนาระบบ 59
4.4 การเชื่อมต่อข้อมูลจากฐานข้อมูลเข้าสู่โปรแกรมที่ใช้สำหรับพัฒนาระบบคลังข้อมูล 62
4.5 การสร้างรายงาน 64
บทที่ 5 76
5.1 บทสรุป 76
5.2 ปัญหา 77
5.3 ข้อเสนอแนะ 78
บรรณานุกรม 79
ประวัติผู้เขียน 81

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

สารบัญตาราง

หน้า

ตารางที่ 1: เทคโนโลยีที่ใช้ในการพัฒนาระบบ 9

ตารางที่ 2: ค่าวัดของระบบวิเคราะห์ยอดขาย (Sales Analysis System) 28

ตารางที่ 3: ตัวชี้วัดผลการดำเนินการหลักของระบบวิเคราะห์ยอดขาย (Sales Analysis System). 29

ตารางที่ 4: คำถามของผู้บริหาร ผู้ใช้ และแดชบอร์ดการวิเคราะห์ของระบบวิเคราะห์ยอดขาย (Sales Analysis System) 31

ตารางที่ 5: แดชบอร์ดการวิเคราะห์ ค่าวัด ตัวชี้วัดผลการดำเนินงานหลัก และมิติของระบบวิเคราะห์ยอดขาย (Sales Analysis System) 32

ตารางที่ 6: ตัวชี้วัดผลการดำเนินการหลักของระบบวิเคราะห์การสั่งอาหารและเครื่องดื่ม (Food & Beverage Order Analysis System) 35

ตารางที่ 7: คำถามของผู้บริหาร ผู้ใช้ และแดชบอร์ดการวิเคราะห์ของระบบการวิเคราะห์การสั่งอาหารและเครื่องดื่ม (Food & Beverage Order Analysis System) 37

ตารางที่ 8: แดชบอร์ดการวิเคราะห์ ค่าวัด ตัวชี้วัดผลการดำเนินงานหลัก และมิติของระบบการวิเคราะห์การสั่งอาหารและเครื่องดื่ม (Food & Beverage Order Analysis System) 38

ตารางที่ 9: ข้อมูลที่จะนำมาใช้แบ่งกลุ่มใบเสร็จค่าอาหารและเครื่องดื่ม 40

ตารางที่ 10: ข้อมูลที่จะนำมาใช้วิเคราะห์กฎความสัมพันธ์ในใบเสร็จค่าอาหารและเครื่องดื่ม 43

ตารางที่ 11: ข้อมูลที่จะนำมาใช้สร้างสมการถดถอยเชิงพหุคูณ 47

ตารางที่ 12: สิทธิ์ในการเข้าถึงระบบ 58

สารบัญรูปภาพ

หน้า

รูปที่ 1: ข้อมูลประเภทร้านอาหารแยกตามจำนวน และการจ้างงานในปี 2563 1

รูปที่ 2: กราฟแสดงข้อมูล 10 จังหวัดที่มีร้านอาหาร SMEs จำนวนมากที่สุดในปี 2563..... 2

รูปที่ 3: กราฟแสดงรายได้ กำไร ขาดทุน ของธุรกิจภัตตาคาร/ร้านอาหารในปี 2562-2563..... 3

รูปที่ 4: กราฟแสดงรายได้ กำไร ขาดทุนของร้านขายเครื่องดื่มที่มีแอลกอฮอล์ ปี 2562-2563..... 3

รูปที่ 5: ตัวอย่างการออกแบบโมเดลข้อมูลหลายมิติแบบ Star Schema..... 16

รูปที่ 6: ตัวอย่างการออกแบบโมเดลข้อมูลหลายมิติแบบ Snowflake Schema 17

รูปที่ 7: Business Intelligence Model 21

รูปที่ 8: โครงสร้างองค์กรของร้าน Bluebird Jazz Bar 24

รูปที่ 9: โมเดลข้อมูลหลายมิติของระบบวิเคราะห์ยอดขาย (Sales Analysis System)..... 30

รูปที่ 10: โมเดลข้อมูลหลายมิติของระบบการวิเคราะห์การสั่งอาหารและเครื่องดื่ม(Food & Beverage Order Analysis System) 36

รูปที่ 11: กราฟแสดงตัวอย่างค่าผลรวมระยะห่างกำลังสอง[1-20] 41

รูปที่ 12: รายงานจำนวนนับใบเสร็จช่วงเวลาที่ขายดีที่สุดในแต่ละวันในสัปดาห์ 52

รูปที่ 13: รายงานการแนวโน้มการเติบโตยอดของสินค้า 52

รูปที่ 14: รายงานจำนวนใบเสร็จแยกตามรายการอาหารและเครื่องดื่มของแต่ละไตรมาส 53

รูปที่ 15: รายงานจำนวนนับใบเสร็จแยกตามประเภทของลูกค้า 53

รูปที่ 16: รายงานยอดขายเฉลี่ยต่อใบเสร็จแยกตามรายปี 54

รูปที่ 17: รายงานยอดขายเฉลี่ยต่อใบเสร็จแยกตามรายวัน 54

รูปที่ 18: รายงานสัดส่วนเปอร์เซ็นต์ยอดขายของสินค้าต่าง ๆ 55

รูปที่ 19: รายงานยอดขายเฉลี่ยต่อใบเสร็จแยกตามรายการอาหารและเครื่องดื่ม 55

รูปที่ 20: รายงานจำนวนนับใบเสร็จแยกตามชื่อเมนูในรายการอาหารและเครื่องดื่ม 56

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

รูปที่ 21: หน้าจอการสร้างรายงาน 57
รูปที่ 22: หน้าจอ Dashboard 57
รูปที่ 23: การสร้างฐานข้อมูลใหม่ 59
รูปที่ 24: หน้าจอการสร้างฐานข้อมูลใหม่ 60
รูปที่ 25: การนำเข้าข้อมูลลงในฐานข้อมูล 60
รูปที่ 26: หน้าจอการ Upload Data เพื่อสร้าง Table ในฐานข้อมูล 61
รูปที่ 27: หน้าจอการนำเข้าข้อมูลในฐานข้อมูลที่สำเร็จ 61
รูปที่ 28: หน้าจอการนำเข้าข้อมูลจาก Google Bigquery 62
รูปที่ 29: Sign In เพื่อเชื่อมต่อข้อมูลจาก Google Biquery 63
รูปที่ 30: หน้าจอการสร้างรายงาน 63
รูปที่ 31: หน้าจอการสร้างรายงาน Tableau Desktop 64
รูปที่ 32: หน้าจอการสร้างรายงานผ่าน Jupyter Lab 65
รูปที่ 33: แดชบอร์ดการวิเคราะห์ทิศทางและแนวโน้มยอดขายของกิจการ 66
รูปที่ 34: แดชบอร์ดการวิเคราะห์ยอดขายแยกตามประเภทรายการอาหารและเครื่องดื่ม 67
รูปที่ 35: แดชบอร์ดการวิเคราะห์ใบเสร็จการสั่งอาหารและเครื่องดื่มที่พบ 68
รูปที่ 36: แดชบอร์ดการวิเคราะห์ใบเสร็จการสั่งอาหารและเครื่องดื่มที่พบ 69
รูปที่ 37: กราฟเส้นแสดงผลลัพธ์การแบ่งกลุ่มที่เหมาะสมโดยใช้เทคนิค Elbow testing 70
รูปที่ 38: กราฟเส้นแสดงผลลัพธ์การแบ่งกลุ่มที่เหมาะสมโดยใช้เทคนิค Average Silhouette 70
รูปที่ 39: กราฟวงกลมแสดงสัดส่วนของข้อมูลแยกตามจำนวนกลุ่ม Cluster 71
รูปที่ 40: รายงานแสดงค่าเฉลี่ยของแต่ละตัวแปรที่สนใจ โดยแยกตาม Cluster 71
รูปที่ 41: กราฟวงกลมแสดงสัดส่วนของตัวแปรที่สนใจ โดยแยกตาม Cluster 72
รูปที่ 42: ตัวอย่างรายงานประเภทกลุ่มสินค้าคู่ที่พบความสัมพันธ์อย่างมีนัยสำคัญ 73
รูปที่ 43: กราฟแสดงยอดค่าใช้จ่ายใบเสร็จรายวันแยกตามรายเดือนและรายปี 74

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

รูปที่ 44: รายงานการประมาณค่าพารามิเตอร์จากการเลือกช่วงเวลามาสร้างสมการพหูคูณด้วยวิธีกำลังสองน้อยที่สุด (Ordinary Least Squares: OLS) 74

รูปที่ 45: รายงานการทำนายผลของชุดทดสอบ test set..... 75

รูปที่ 46: รายงานการวิเคราะห์การทำนายผลผิดพลาดของชุดทดสอบ 75

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

บทที่ 1
บทนำ

ในบทนี้จะกล่าวถึงความสำคัญและที่มาของโครงการ วัตถุประสงค์ของโครงการ ขอบเขตของโครงการ วิธีการดำเนินงานโครงการ เทคโนโลยีที่ใช้ในการพัฒนาระบบและเทคโนโลยีที่ใช้ในดำเนินงาน ตลอดจนประโยชน์ที่จะได้รับจากโครงการนี้

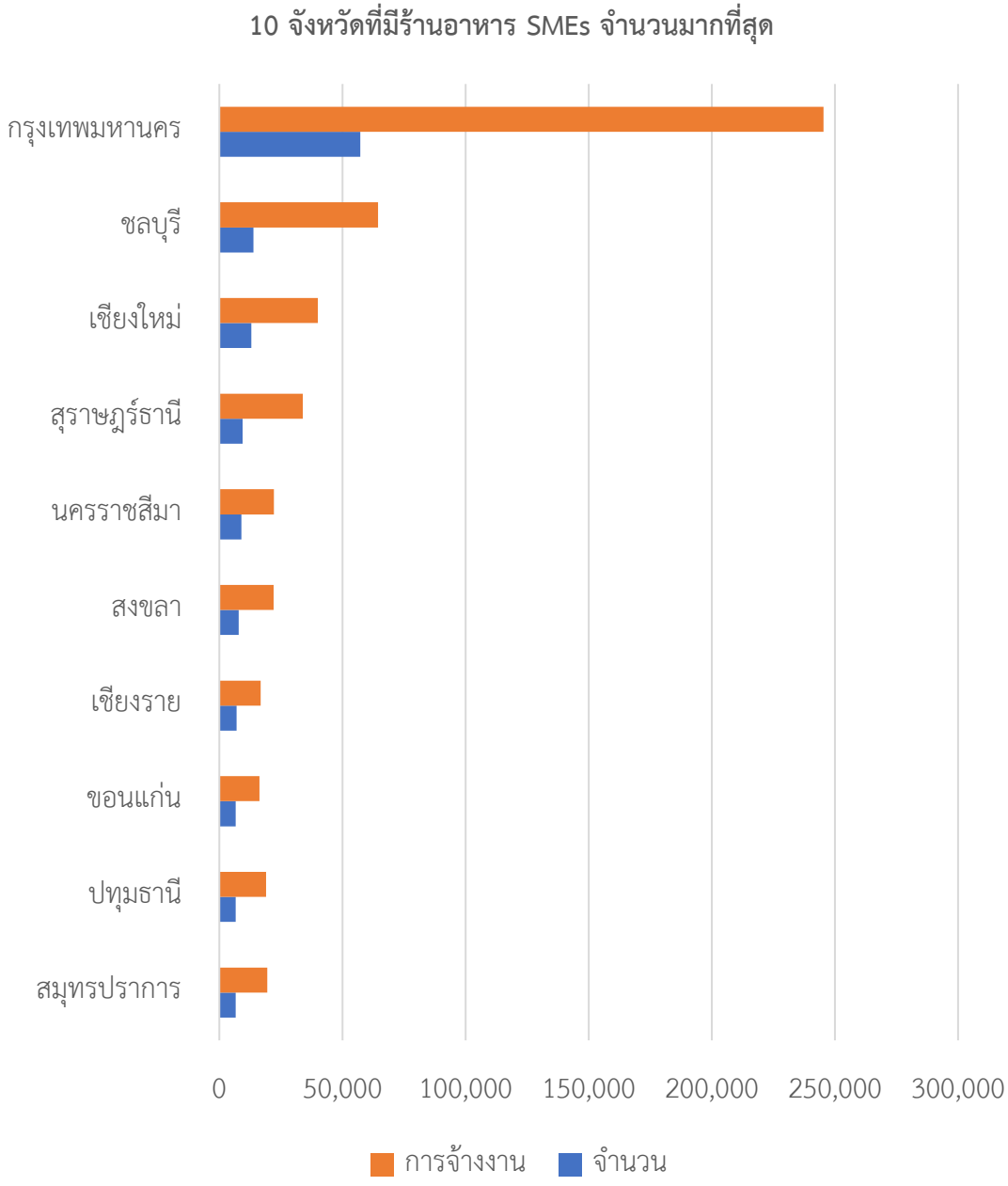
1.1 ความสำคัญและที่มาของโครงการ

ในปี 2563 ธุรกิจร้านอาหารในประเทศมีจำนวนทั้งสิ้น 334,736 ราย เป็นธุรกิจ SMEs 99.9% (334,606 ราย) โดยเป็นธุรกิจประเภทร้านอาหาร/ภัตตาคาร มากที่สุด ตามมาด้วยร้านขายเครื่องดื่ม สำหรับจังหวัดที่มีผู้ประกอบการร้านอาหารมากที่สุด คือ กรุงเทพฯ ชลบุรี และเชียงใหม่ ตามลำดับ (ศุภริน เจริญพานิช, 2563)

ประเภทร้านอาหาร	SMEs	
	จำนวน (ร้าน)	การจ้างงาน (คน)
ภัตตาคาร/ร้านอาหาร	286,417	853,391
ร้านเครื่องดื่มที่ไม่มีแอลกอฮอล์	39,972	78,460
ร้านเครื่องดื่มที่มีแอลกอฮอล์เป็นหลักในร้าน	5,010	34,945
การจัดเลี้ยง	2,653	19,088
การบริการด้านอาหารอื่นๆ	295	1,834
โรงอาหาร	131	1,399
แผงลอยและตลาด	49	321
อาหารสำหรับกิจการขนส่ง	35	286
ร้านอาหารแบบเคลื่อนที่	31	316
ร้านเครื่องดื่มโดยร้านเคลื่อนที่	8	63
ร้านเครื่องดื่มบนแผงลอยและตลาด	5	29
รวม	334,606	990,132

ที่มา : สำนักงานส่งเสริมวิสาหกิจขนาดกลางและขนาดย่อม

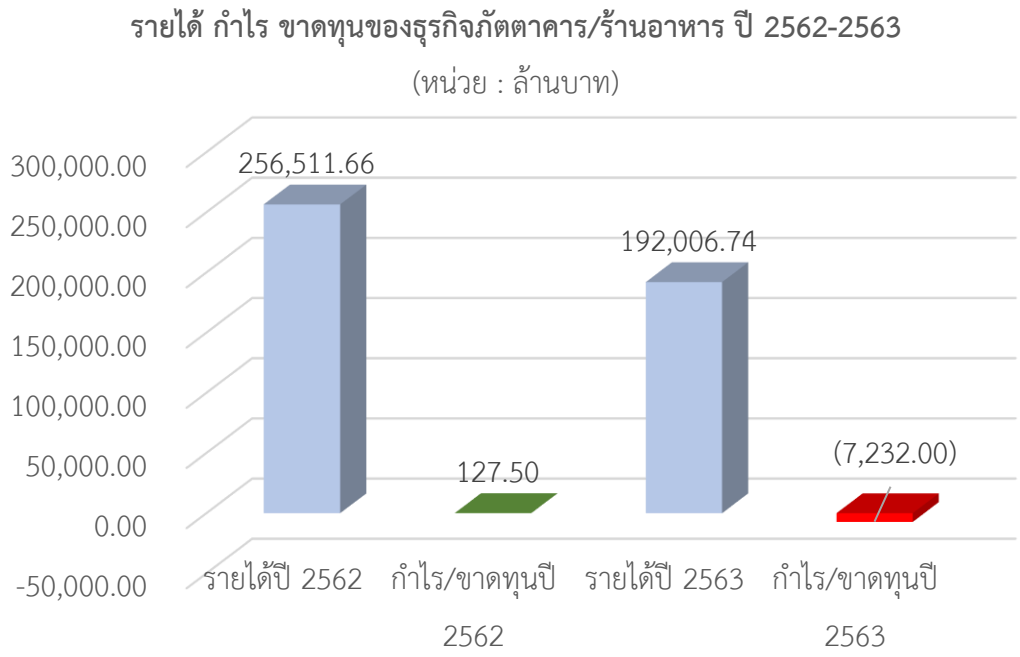
รูปที่ 1: ข้อมูลประเภทร้านอาหารแยกตามจำนวน และการจ้างงานในปี 2563



ที่มา : สำนักงานส่งเสริมวิสาหกิจขนาดกลางและขนาดย่อม

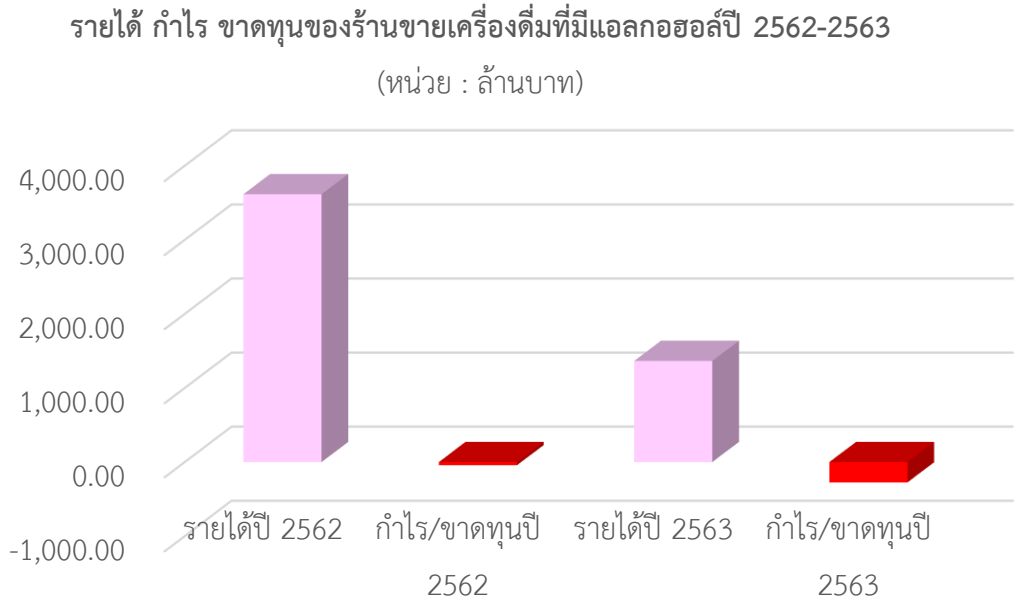
รูปที่ 2: กราฟแสดงข้อมูล 10 จังหวัดที่มีร้านอาหาร SMEs จำนวนมากที่สุดในปี 2563

ในขณะเดียวกันเมื่อเข้าสู่ช่วงวิกฤตจากสถานการณ์ระบาดของโควิด-19 ส่งผลให้ในปี 2563 ธุรกิจร้านอาหารได้รับผลกระทบสูง ซึ่ง GDP SMEs ของร้านอาหารในปี 2563 ลดลง 38% จากปีก่อนหน้า ในส่วนของธุรกิจร้านอาหารในภาพรวม ปี 2563 มีรายได้ลดลง 28% (ศุภริน เจริญพานิช, 2563)



ที่มา : สำนักงานส่งเสริมวิสาหกิจขนาดกลางและขนาดย่อม

รูปที่ 3: กราฟแสดงรายได้ กำไร ขาดทุน ของธุรกิจภัตตาคาร/ร้านอาหารในปี 2562-2563



ที่มา : สำนักงานส่งเสริมวิสาหกิจขนาดกลางและขนาดย่อม

รูปที่ 4: กราฟแสดงรายได้ กำไร ขาดทุนของร้านขายเครื่องดื่มที่มีแอลกอฮอล์ ปี 2562-2563

แต่อย่างไรก็ตามเมื่อเวลาผ่านไป จนถึงช่วงโค้งสุดท้ายของปี 2564 ธุรกิจร้านอาหารให้บริการเต็มรูปแบบ หรือ Full-Service Restaurants ได้รับปัจจัยหนุนเพิ่มขึ้น จากการได้รับการผ่อนคลายมาตรการต่างๆ ได้แก่ การยกเลิกเคอร์ฟิวในพื้นที่นำร่องด้านการท่องเที่ยว และการอนุญาตให้ธุรกิจร้านอาหารสามารถกลับมาขายเครื่องดื่มแอลกอฮอล์ให้กับลูกค้าที่มาใช้บริการนั่งทานในร้านได้ตั้งแต่วันที่ 1 พ.ย. 2564 แบบมีเงื่อนไข ซึ่งช่วยให้ธุรกิจสามารถสร้างรายได้เพิ่มขึ้น และหากสถานการณ์การระบาดของโควิดในประเทศคลี่คลายลง จนทำให้ความเชื่อมั่นของผู้บริโภคในการใช้บริการในร้านอาหารทยอยกลับมา จะส่งผลดีต่อธุรกิจในช่วงที่เหลือของปีนี้และในปี 2565 (อรนุช ภัทรกุล, 2564)

หลังจากสถานการณ์ระบาดของโควิด-19 เบาบางลง ผู้ประกอบการต่าง ๆ จะต้องเตรียมตัวให้พร้อมสำหรับการกลับมาสร้างยอดขายให้กับธุรกิจของตนได้อีกครั้ง โดยการค้นหากลยุทธ์ที่จะสามารถทำให้ธุรกิจดำเนินการต่อไปได้อย่างราบรื่น

6W1H คือทฤษฎีการศึกษาและวิเคราะห์พฤติกรรมผู้บริโภค เพื่อนำมาวางกลยุทธ์ทางการตลาด ตามหากลุ่มเป้าหมายที่แท้จริง สามารถยกระดับธุรกิจเพื่อสร้างยอดขายได้อย่างยั่งยืน กล่าวคือ “การตระหนักถึงผู้บริโภคคือหัวใจสำคัญของการทำธุรกิจ” ยิ่งมีฐานผู้บริโภคติดตามมากเท่าไหร่ก็ยิ่งทำให้ธุรกิจเติบโตได้อย่างมีประสิทธิภาพมากเท่านั้น โดยสามารถอธิบายแต่ละองค์ประกอบของทฤษฎี 6W1H ได้ดังนี้

Who ลูกค้าของเราคือใคร คือ การกำหนดกลุ่มเป้าหมายในเชิงปริมาณประชากร สามารถกำหนดได้จากปัจจัยคร่าว ๆ เช่น เพศ ถิ่นที่อยู่ ช่วงอายุ ฐานะทางการเงิน การศึกษา ฯลฯ เพื่อทำการวิเคราะห์กลุ่มเป้าหมายแบบเบื้องต้น ก่อนวิเคราะห์ลงรายละเอียดในขั้นตอนต่อไป

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

What ลูกค้าต้องการอะไร เป็นการวิเคราะห์ความต้องการของลูกค้า ว่าพวกเขาต้องการอะไร ซึ่งอาจจะดูจากพฤติกรรมการเลือกซื้อสินค้า ไลฟ์สไตล์การใช้ชีวิต คุณภาพสินค้าที่กลุ่มเป้าหมายชอบ รวมไปถึงรูปร่างหน้าตา Packaging อีกด้วย

Where ลูกค้าอยู่ที่ไหน เป็นการวิเคราะห์ช่องทางการเลือกซื้อสินค้าของกลุ่มลูกค้า ในธุรกิจออนไลน์ ลูกค้าก็จะอยู่ในช่องทางออนไลน์เป็นหลัก บนแพลตฟอร์มยอดนิยม อย่างเช่น Facebook Instagram Twitter หรือเว็บไซต์ Market place โดยอาจจะเข้ามาเจอร้านค้าหรือสินค้าผ่านการ Search บน Google Ads หรือโฆษณาบนช่องทางที่ใช้งานเป็นประจำ

Why ทำไมลูกค้าต้องเลือกซื้อสินค้า เป็นการวิเคราะห์ว่าสินค้าหรือบริการธุรกิจ สามารถตอบโจทย์ความต้องการและแก้ปัญหาใดให้กับลูกค้า หรือมอบประโยชน์ใดให้ลูกค้า เช่น ความคุ้มค่าของวัสดุ ราคาที่ถูกกว่าท้องตลาด หรือสินค้าดังกล่าวเป็นสินที่ขาดไม่ได้ เป็นต้น

When กลุ่มเป้าหมายจะซื้อสินค้าเมื่อไหร่ วิเคราะห์ช่วงเวลาที่สินค้าจะเป็นที่ต้องการมากที่สุด ตัวอย่างเช่น หากการจำหน่ายสินค้าประเภทของขวัญ ช่วงเวลาที่สินค้าจะเป็นที่ต้องการก็คือช่วงเทศกาลคริสต์มาสหรือช่วงเทศกาลปีใหม่ที่ลูกค้าต่างต้องเลือกซื้อสินค้าเพื่อนำไปเป็นของขวัญ

Whom ใครคือคนที่ส่งผลต่อการตัดสินใจซื้อสินค้า ลูกค้าบางคนอาจจะไม่ได้ซื้อสินค้าเพราะเชื่อในสรรพคุณที่ร้านค้าโฆษณา แต่ซื้อเพราะคำรีวิว คำบอกเล่าสรรพคุณจากคนที่เคยใช้งานจริง ๆ หรือแม้กระทั่งซื้อตามดาราหรือบิวตี้บล็อกเกอร์ที่มีชื่อเสียง การวิเคราะห์ในส่วนนี้จะทำให้องค์กรวางกลยุทธ์โฆษณาสินค้าเพื่อกระตุ้นยอดขายให้เติบโตได้อย่างมีประสิทธิภาพ

How ลูกค้ามีเงื่อนไขอะไรที่ทำให้ตัดสินใจซื้อ เป็นการวิเคราะห์สิ่งที่ทำให้ลูกค้ายอมตกลงปลงใจซื้อสินค้า ลูกค้าบางคนอาจจะซื้อเฉพาะช่วงที่มีโปรโมชั่นลดราคาหรือมีของแถม การมีบริการที่สะดวกและตอบโจทย์อย่างขั้นตอนเลือกซื้อสินค้าและระบบจ่ายเงินที่ลูกค้าสามารถจบขั้นตอนได้อย่างรวดเร็ว รวมถึงการมีบริการขนส่งที่รองรับความต้องการของลูกค้าอย่างครบวงจร (fillgoods, 2563)

ในการจัดทำโครงการนี้ ผู้จัดทำได้รับความอนุเคราะห์ให้ข้อมูลของกิจการร้านตัวอย่างที่มีอยู่จริง ซึ่งเป็นใบเสร็จค่าอาหารและเครื่องดื่มทั้งหมดมากกว่า 5,000 รายการ เพื่อนำมาศึกษาพฤติกรรมการสั่งซื้อของผู้ใช้บริการ และนำผลการวิเคราะห์มาสนับสนุนการตัดสินใจการส่งเสริมการขายธุรกิจในช่วงที่เศรษฐกิจฟื้นตัวหลังจากสถานการณ์ระบาดโควิด-19

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

1.2 วัตถุประสงค์ของโครงการ

โครงการพัฒนา “คลังข้อมูล ธุรกิจอัจฉริยะ และการวิเคราะห์ข้อมูลขั้นสูงของธุรกิจร้านอาหารและเครื่องดื่ม” ได้จัดทำขึ้นโดยมีวัตถุประสงค์ดังนี้

1) เพื่อพัฒนาคลังข้อมูล โดยรวบรวมข้อมูลที่เก็บในรูปแบบ Excel และข้อมูลที่เก็บแยกกันมาเชื่อมโยงและรวมไว้ที่เดียวกัน เพื่อลดความซ้ำซ้อนของข้อมูล สะดวกต่อการนำข้อมูลมาใช้ในการวิเคราะห์
2) เพื่อสร้างรายงานเชิงวิเคราะห์ที่มีลักษณะ Interactive เพื่อให้เจ้าของกิจการสามารถเรียกดูรายงานในมุมมองที่หลากหลายและใช้สนับสนุนการตัดสินใจและวางแผนกลยุทธ์ต่าง ๆ ตามความต้องการ
3) เพื่อแก้ปัญหาการนำเสนอข้อมูลที่ล่าช้า เนื่องจากในปัจจุบันเทคโนโลยีพัฒนาขึ้นมาก ทำให้ธุรกิจและพฤติกรรมของผู้บริโภคเปลี่ยนแปลงอย่างรวดเร็ว ดังนั้น การจะก้าวนำคู่แข่งจึงต้องมีการตัดสินใจอย่างรวดเร็วเพื่อให้ทันต่อความต้องการของผู้บริโภค

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

1.3 ขอบเขตของโครงการ

1) ระบบวิเคราะห์ยอดขาย (Sales Analysis System)

ระบบนี้เป็นการแสดงผลประกอบการของธุรกิจร้านอาหารและเครื่องดื่มในรูปแบบ Dashboard เพื่อให้ผู้จัดการร้าน และหัวหน้าฝ่ายจัดซื้อใช้ตรวจสอบสถานะยอดขายเฉลี่ยต่อใบเสร็จ (Average sales per receipt) ในมิติต่าง ๆ ได้ทันที เช่น มิติเวลา มิติรูปแบบการให้บริการ และมิติสินค้า เป็นต้น

2) ระบบการวิเคราะห์การสั่งอาหารและเครื่องดื่ม (Food & Beverage Order Analysis System) ระบบนี้เป็นการวิเคราะห์การสั่งอาหารและเครื่องดื่มในรูปแบบของ Dashboard เพื่อให้ผู้จัดการร้าน ใช้ตรวจสอบสัดส่วนของวันหรือเวลาที่ลูกค้าเข้ามาใช้บริการ (Percentage of datetime in receipts) จำนวนลูกค้าเฉลี่ยต่อใบเสร็จ (Average number of customers per receipt) และสัดส่วนวิธีการชำระเงินของลูกค้าในใบเสร็จ (Percentage of payment method in receipts) ตามมิติต่าง ๆ ได้ทันที เช่น มิติใบเสร็จ มิติลูกค้า มิติเวลา และมิติการชำระเงิน เป็นต้น

3) ระบบแบ่งกลุ่มใบเสร็จค่าอาหารและเครื่องดื่ม (Food & Beverage Receipt Segmentation System)

ระบบนี้จะใช้เทคนิค Unsupervised Learning เพื่อจัดกลุ่มที่ได้จากใบเสร็จค่าอาหารและเครื่องดื่มโดยนำเสนอเป็นรายงานให้กับผู้จัดการร้าน เพื่อทำความเข้าใจพฤติกรรมการสั่งอาหารและเครื่องดื่มของลูกค้า

4) ระบบวิเคราะห์กฎความสัมพันธ์ในใบเสร็จค่าอาหารและเครื่องดื่ม (Food & Beverage Receipt Association Analysis System)

ระบบนี้เป็นการค้นหาความสัมพันธ์ที่เกิดขึ้นร่วมกันบ่อยครั้ง จากพฤติกรรมการสั่งอาหารและเครื่องดื่มจากใบเสร็จที่ผ่านมา โดยใช้เทคนิคการค้นหากฎความสัมพันธ์ และประเมินผลกฎความสัมพันธ์ด้วย Support Confidence และ Lift เพื่อนำกฎความสัมพันธ์ที่ได้ไปประกอบการตัดสินใจในการจัดทำกลยุทธ์ส่งเสริมการขายในอนาคต

5) ระบบพยากรณ์ยอดขาย (Sales Forecasting System)

ระบบนี้เป็นการพัฒนาตัวแบบพยากรณ์ยอดขายธุรกิจ โดยใช้สมการการถดถอยเชิงพหุคูณ (Multiple Linear Regression Analysis) ในการช่วยพัฒนาและเลือกตัวแบบที่ดีที่สุดจากตัวแปรต่าง ๆ ที่ได้จากใบเสร็จค่าอาหารและเครื่องดื่ม พร้อมนำเสนอเป็นรายงานให้กับผู้จัดการร้าน ในการคาดการณ์ยอดขายธุรกิจในอนาคต

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

1.4 วิธีการดำเนินงานโครงการ

1) การวิเคราะห์ระบบ (System Analysis)

1.1 ศึกษาลักษณะการดำเนินงานของธุรกิจร้านอาหารและเรื่องดื่มในปัจจุบัน จากการทำงานของฝ่ายต่าง ๆ โดยศึกษาความหมายและรูปแบบของข้อมูลการขายที่เกี่ยวข้องกับธุรกิจทั้งหมด

1.2 สัมภาษณ์และเก็บรวบรวมข้อมูลจากเจ้าของกิจการและพนักงานของแต่ละฝ่าย เพื่อศึกษาปัญหาที่เกิดขึ้นและรวบรวมความต้องการ

1.3 วิเคราะห์ระบบและสรุปความต้องการของผู้ใช้งาน พร้อมทั้งกำหนดขอบเขตของระบบงานที่จะพัฒนา

2) การออกแบบระบบ (System Design)

2.1 ออกแบบรูปแบบของรายงานต่าง ๆ (Report Design) เพื่อให้ตรงกับวัตถุประสงค์และสามารถนำไปช่วยในการสนับสนุนการตัดสินใจของผู้บริหาร

2.2 ออกแบบโมเดลข้อมูลเชิงมิติ (Multi-dimensional Data Modeling Design) โดยโครงสร้างของข้อมูลจะอยู่ในรูปแบบ Star Schema เพื่อให้สามารถออกรายงานได้ตรงตามความต้องการ

2.3 ออกแบบวิธีการดึง แปลง และนำเข้าข้อมูล (Extract Transform Load - ETL) จากฐานข้อมูลเพื่อให้สามารถนำข้อมูลเข้าสู่ระบบได้อย่างถูกต้อง

3) การพัฒนาระบบ (System Development)

3.1 พัฒนาโปรแกรมและนำข้อมูลข้อมูลระบบงานย่อยเข้ามาในคลังข้อมูล (ETL)

3.2 พัฒนาคลังข้อมูลตามที่ได้มีการออกแบบไว้ (Data Warehouse Development)

3.3 พัฒนารูปแบบของรายงานที่ช่วยในการวิเคราะห์ตามที่ได้ออกแบบไว้ (Report Preparation)

3.4 พัฒนาตัวแบบเพื่อนำผลของสมการที่ได้ไปใช้สนับสนุนการตัดสินใจ

4) การทดสอบระบบ (System Testing)

4.1 ทดสอบการเชื่อมโยงของระบบความสอดคล้องของข้อมูลในแต่ละระบบงาน

4.2 ประเมินผลประสิทธิภาพของการพัฒนาตัวแบบ

4.3 ปรับปรุงและแก้ไขข้อผิดพลาดให้มีความสมบูรณ์

5) การจัดทำคู่มือการใช้งาน (User Document)

5.1 จัดทำคู่มือสำหรับการใช้งาน (User Manual) จะเป็นเอกสารที่บอกถึงขั้นตอนการใช้งานของระบบที่พัฒนาขึ้น เพื่อช่วยให้เกิดความเข้าใจในการใช้งานได้อย่างถูกต้อง โดยจะนำเสนอในรูปของหน้าจอแสดงผลพร้อมคำอธิบายประกอบรูปภาพ

1.5 เทคโนโลยีที่ใช้ในโครงการ

เทคโนโลยีที่ใช้สำหรับการพัฒนาโครงการ คลังข้อมูล ธุรกิจอัจฉริยะ และการวิเคราะห์ข้อมูลขั้นสูงของธุรกิจร้านอาหารและเครื่องดื่มมีรายละเอียดดังนี้

ตารางที่ 1: เทคโนโลยีที่ใช้ในการพัฒนาระบบ

ด้าน Software	
ระบบปฏิบัติการ	macOS
ระบบจัดการฐานข้อมูล	Google BigQuery
เครื่องมือที่ใช้ในการเตรียมข้อมูล	Python 3.8.0
เครื่องมือที่ใช้ในการพัฒนาธุรกิจอัจฉริยะ	Tableau Desktop Version 2021.3
เครื่องมือที่ใช้ในการแสดงผลธุรกิจอัจฉริยะ	Tableau Desktop Version 2021.3
เครื่องมือที่ใช้ในการวิเคราะห์ข้อมูลขั้นสูง	Python 3.8.0 - pandas - matplotlib - plotly - numpy - seaborn - scikit-learn - mlxtend

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

1.6 ประโยชน์ที่คาดว่าจะได้รับ

ประโยชน์ที่คาดว่าจะได้รับจากการพัฒนาโครงการ “คลังข้อมูล ธุรกิจอัจฉริยะ และการวิเคราะห์ข้อมูลขั้นสูงของธุรกิจร้านอาหารและเครื่องดื่ม” มีดังนี้

1) ทำให้เกิดการรวบรวมและการจัดเก็บข้อมูลด้านการดำเนินงานและข้อมูลด้านการให้บริการที่เกี่ยวข้องกัน และพัฒนาเป็นระบบสนับสนุนตัดสินใจของเจ้าของกิจการธุรกิจเพื่อให้ตัดสินใจวางแผนดำเนินธุรกิจได้อย่างมีประสิทธิผล
2) ระบบคลังข้อมูลทำให้การออกรายงานมีความยืดหยุ่นมากขึ้น เจ้าของกิจการสามารถเรียกดู และปรับเปลี่ยนมุมมองของรายงานได้อย่างทันท่วงทีตามความต้องการ ทำให้ได้รายงานที่เอื้อต่อการวิเคราะห์และการวางแผนกลยุทธ์การตลาด
3) ช่วยลดปริมาณการทำงานของพนักงานแต่ละฝ่ายในการจัดเตรียมสารสนเทศเพื่อเสนอแก่เจ้าของกิจการ โดยเจ้าของกิจการสามารถเข้าถึงข้อมูลที่ต้องการเพื่อประกอบการตัดสินใจได้โดยตรงและรวดเร็ว
4) ทำให้เจ้าของกิจการสามารถประเมินแนวโน้มความต้องการของผู้บริโภคได้ โดยพิจารณาจากปริมาณยอดขายที่เพิ่มขึ้น รวมถึงพฤติกรรมการใช้บริการของลูกค้า และเมนูอาหารและเครื่องดื่มที่เป็นที่นิยมของลูกค้าแต่ละกลุ่ม
5) สนับสนุนการวิเคราะห์เพื่อวางแผนกลยุทธ์ลดค่าใช้จ่ายที่ไม่จำเป็น ช่วยเพิ่มผลกำไรแก่ธุรกิจ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทที่ 2
เหตุผลและแนวคิด

แนวคิดที่สำคัญที่นำมาประยุกต์ใช้ในการพัฒนาระบบ ได้แก่ แนวคิดด้านคลังข้อมูล (Data Warehouse) แนวคิดด้านธุรกิจอัจฉริยะ (Business Intelligence) แนวคิดด้านการวิเคราะห์ข้อมูลเชิงทำนาย

2.1 แนวคิดด้านคลังข้อมูล (Data Warehouse)

การจัดทำคลังข้อมูลมีความสำคัญมากเพราะปัจจุบันนี้มีผู้ใช้และผู้บริหารของหน่วยงานเริ่มเข้าใจความสำคัญของข้อมูลมากขึ้น และเริ่มตระหนักว่า หากนำข้อมูลมาวิเคราะห์ให้เข้าใจสถานภาพหรือเหตุการณ์ที่เกิดขึ้นแล้วจะทำให้องค์กรสามารถตอบสนองต่อเหตุการณ์ที่เกิดขึ้นได้ดียิ่งขึ้น และจะทำให้องค์กรดำเนินธุรกิจตามบรรลุวัตถุประสงค์และเป้าหมายได้ดียิ่งขึ้นตามไปด้วย

2.1.1 นิยามคลังข้อมูล

คลังข้อมูล (Data Warehouse) หมายถึง ฐานข้อมูลขนาดใหญ่ของหน่วยงานหนึ่ง ๆ หรือองค์กร ซึ่งเก็บรวบรวมข้อมูลจากฐานข้อมูลระบบงานที่ใช้ดำเนินธุรกิจขององค์กรประจำวัน และฐานข้อมูลอื่นภายนอกองค์กร โดยข้อมูลที่ได้ถูกจัดเก็บในคลังข้อมูลนั้นมีวัตถุประสงค์ในการนำไปประยุกต์ และมีลักษณะของการจัดเก็บที่แตกต่างไป จากข้อมูลในฐานข้อมูลระบบงานอื่น โดยการย้ายข้อมูลจากฐานข้อมูลปกติเข้าไปไว้ในคลังข้อมูลมีข้อดีหลายอย่าง เช่น ทำให้องค์กรหรือเจ้าของข้อมูลมีรูปแบบการเก็บข้อมูลที่เหมาะสมกับการเรียกใช้มากยิ่งขึ้น และนำไปช่วยตัดสินใจหรือใช้ในงานวิเคราะห์ในแง่มุมต่างๆ นอกจากนั้นระบบคลังข้อมูลยังรวมข้อมูลที่ใช้อยู่ในปัจจุบันกับข้อมูลในอดีตเป็นฐานข้อมูลเดียวกัน ซึ่งสามารถเรียกใช้งานได้โดยง่าย พร้อมสำหรับการจัดการข้อมูลและไปวิเคราะห์ เหมาะกับผู้ใช้ระดับสูง หรือพนักงานทั่วไปที่มีหน้าที่ในการวิเคราะห์ข้อมูล เพราะสามารถเข้าถึงและเรียกใช้ได้ง่าย ทำงานได้สะดวกและมีประสิทธิภาพ (Panthong, Rattanawadee, 2560)

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

2.1.2 ลักษณะคลังข้อมูล

1) **Integrated**

การรวบรวมข้อมูลมาจากหลายแหล่งเพื่อใช้การวิเคราะห์ อาจทำให้เกิดปัญหาความไม่ถูกต้อง และความไม่สอดคล้องกันของข้อมูล ดังนั้น ระบบคลังข้อมูลจึงต้องมีคุณสมบัติที่มีการรวบรวมข้อมูลการดำเนินธุรกรรมจากหลาย ๆ แหล่ง เข้ามาไว้ภายใต้โครงสร้างเดียวกัน ทำให้ข้อมูลมีมาตรฐานเดียวกันในการจัดเก็บ สามารถนำไปใช้ดำเนินหรือวิเคราะห์ต่อได้อย่างถูกต้อง

2) **Subject-Oriented**

ข้อมูลในคลังข้อมูลจะเลือกเก็บแต่ข้อมูลที่สามารถมาใช้ในเชิงวิเคราะห์หรือ ตัดสินใจมากกว่าเก็บข้อมูลเพื่อตอบคำถามรายละเอียดปลีกย่อย เช่น หากจุดประสงค์ ต้องการทราบข้อมูลเกี่ยวกับยอดขาย คลังข้อมูลจะถูกออกแบบสร้างให้เน้นแต่เรื่องการขาย ประกอบด้วย ยอดขายราคาขาย สินค้าที่ขาย เป็นต้น

3) **Non-Volatile**

ข้อมูลในคลังข้อมูลจะไม่มีการแก้ไขเปลี่ยนแปลงอีกหลังจากที่ถูกโหลดเข้าสู่ฐานข้อมูลแล้วเว้นแต่ในกรณีที่ข้อมูลที่โหลดเข้าไปนั้นเป็นข้อมูลที่ไม่ถูกต้อง เมื่อข้อมูลถูกนำเข้าไปใส่ไว้ในระบบ data warehouse แล้ว การใช้งานโดยส่วนใหญ่มากกว่าร้อยละ 99 จะเป็นการ “อ่าน” ข้อมูลเพื่อใช้ในการวิเคราะห์และสนับสนุนการตัดสินใจในรูปแบบต่าง ๆ

4) **Time-Variant**

ลักษณะข้อมูลในคลังข้อมูลจะเป็นข้อมูลประวัติศาสตร์ (Historical Data) เก็บจากอดีตถึงปัจจุบันย้อนหลังเป็นเวลาหลายปี ทั้งนี้เพื่อที่จะได้นำไปวิเคราะห์เปรียบเทียบหาแนวโน้มของข้อมูล เพื่อช่วยสนับสนุน และเป็นฐานประกอบการตัดสินใจ ดังนั้น คลังข้อมูลจึงเน้นการจัดเก็บข้อมูลตามช่วงเวลา เช่น ข้อมูลการขายของสินค้ารายการหนึ่งจะถูกเก็บในคลังข้อมูลย้อนหลังไป 3 ถึง 5 ปี หรืออาจจะตั้งแต่เริ่มจำหน่ายสินค้า เป็นต้น เพื่อให้สามารถวิเคราะห์ถึงยอดขายในอดีตและพยากรณ์แนวโน้มในอนาคตต่อไปได้ (PanaEk, 2555)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2.1.3 ประโยชน์ของคลังข้อมูล

การจัดเก็บข้อมูลในรูปแบบคลังข้อมูลมีประโยชน์ (mindphp, 2563) ดังนี้

1) ช่วยเสริมสร้างความรู้และช่วยสนับสนุนการตัดสินใจของบุคลากรในองค์กร เนื่องจากข้อมูลที่เป็นข้อมูลจริงจากการดำเนินธุรกิจ จึงทำให้ทราบถึงผลการดำเนินการตามจริง ผู้ใช้งานเห็นภาพการดำเนินธุรกิจ และใช้หากลยุทธ์ในการดำเนินกิจการต่อไป
2) สามารถเข้าถึงข้อมูลที่สำคัญได้อย่างสะดวกและรวดเร็ว เพราะข้อมูลมีการจัดเก็บและรวบรวมอยู่ที่เดียวกัน รวมไปถึงรูปแบบการจัดเก็บมีความเหมือนกัน ไม่ว่าจะเข้าถึงข้อมูลในด้านใดก็มีการเข้าถึงในรูปแบบเดียวกัน ทำให้เปรียบเทียบได้ง่าย
3) แยกฐานข้อมูลที่ใช้ปฏิบัติงานออกจากฐานข้อมูลที่ใช้วิเคราะห์ ทำให้การดูข้อมูลเชิงลึกทำได้ง่ายขึ้น เพราะรูปแบบการจัดเก็บต่างจากการเก็บในรูปแบบผลการดำเนินการประจำวัน ซึ่งเน้นเนื้อหาที่สำคัญที่สามารถนำไปใช้การวิเคราะห์ต่อได้
4) ให้ข้อมูลสรุปในมุมมองระดับสูง และสามารถเจาะลึกลงไปได้ เพราะมีรูปแบบการมองในหลายมิติ ไม่ว่าจะเป็นการดูในภาพรวม หรือการเจาะลึกเจาะลึกลงไป ก็สามารถทำได้ในทีเดียว ทำให้การวิเคราะห์ข้อมูลเป็นไปได้ง่ายและมีประสิทธิภาพมากขึ้น

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

2.1.4 การออกแบบฐานข้อมูลสำหรับคลังข้อมูล

การออกแบบคลังข้อมูล (โกเมศ อัมพวัน, 2560) สามารถจัดทำเป็นขั้นตอนดังต่อไปนี้

1) กำหนดความต้องการ

ความต้องการประกอบด้วยความต้องการของผู้ใช้ข้อมูล และความต้องการทางด้านเทคนิค การกำหนดขอบเขตงาน และกำหนดแหล่งที่มาของข้อมูลที่ต้องใช้ สามารถศึกษาจากงานที่ใช้อยู่เดิม หรือเป็นความต้องการและรูปแบบของรายงานที่ต้องการใช้สำหรับการวิเคราะห์

2) ออกแบบและสร้างฐานข้อมูล (DW, 2559)

โครงสร้างของฐานข้อมูลสำหรับงานคลังข้อมูลจะแตกต่างจากการออกแบบงานที่เป็น ฐานข้อมูลเชิงสัมพันธ์ (OLTP) สำหรับคลังข้อมูลนั้น เน้นการออกแบบให้ดีนอร์มัลไลซ์ (Denormalized) มากที่สุด เพื่อสำหรับการเรียกข้อมูลแล้วได้ผลเร็วที่สุด จึงสามารถยอมให้เกิดการซ้ำซ้อนของข้อมูลในแต่ละตารางได้ โดยในการออกแบบฐานข้อมูลของคลังข้อมูล ประกอบด้วยตารางหลัก 2 อย่างคือ ตารางข้อเท็จจริง (Fact Table) และตารางมิติ (Dimension Table)

• ตารางข้อเท็จจริง (Fact Table) เป็นตารางหลักที่ใช้เก็บข้อมูลและใช้เพื่อตอบคำถามที่ต้องการ ซึ่งการออกแบบตารางข้อเท็จจริงจะเลือกเฉพาะคอลัมน์ที่ต้องการใช้งานเท่านั้น และพยายามลดขนาดของคอลัมน์ที่มีความยาวมากเกินไปโดยไม่จำเป็น ข้อมูลจะไม่มีการแก้ไขหรือเปลี่ยนแปลง ยกเว้นแต่การเพิ่มข้อมูลใหม่เข้าไปในตารางเท่านั้น

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

• ตารางมิติ (Dimension Table) เป็นตารางที่เก็บความหมายของข้อมูลที่ใช้ในตารางข้อเท็จจริงมีประโยชน์เพื่อช่วยให้การสอบถามแสดงคำอธิบายข้อมูลต่าง ๆ ได้ชัดเจนขึ้น

3) เลือกชนิดของข้อมูลที่เป็นตัวเลขสำหรับการวิเคราะห์

ในทางคลังข้อมูลจะเรียกตัวเลขที่ใช้ในการวิเคราะห์ว่าตัววัด (Measure) ได้แก่ การเลือกคอลัมน์ที่มีชนิดข้อมูลเป็นตัวเลข เช่น จำนวนชิ้นในการขาย และจำนวนคนทำงาน เป็นต้น

4) การเตรียมข้อมูล

เป็นขั้นตอนการนำข้อมูลที่ผ่านขบวนการตรวจสอบความถูกต้องของชุดข้อมูลมา ETL (Extraction, Transformation and Loading) ซึ่งเทคนิคการโหลดข้อมูลเข้าสู่คลังข้อมูล เป็นวิธีที่เร็วและคุ้มกับค่าใช้จ่ายที่ต้องเสียไป ซึ่งอาจมีประโยชน์สำหรับโครงการต่างๆ ที่ต้องการข้อมูลสำหรับตัดสินใจ หรือการทำงานอย่างรวดเร็ว

5) ออกแบบการเพิ่ม หรือปรับปรุงข้อมูลในคลังข้อมูล (Porpor, 2563)

การเพิ่มข้อมูลในคลังข้อมูลเป็นสิ่งสำคัญ เนื่องจากจะต้องมีการนำข้อมูลจากระบบ OLTP (Online Transaction Processing) มาเพิ่มที่คลังข้อมูลเพื่อให้มีข้อมูลทันสมัย นอกจากนี้ ยังมีการปรับเปลี่ยนโครงสร้างที่เก็บข้อมูล เพราะเปลี่ยนแนวความคิด หรือเพิ่มวิธีในการวิเคราะห์ข้อมูล เป็นต้น การเพิ่มและปรับปรุงข้อมูลสามารถทำได้หลายวิธี ได้แก่

• Incremental Update คือ การเพิ่มข้อมูลต่อจากข้อมูลที่มีอยู่เดิม และไม่ได้เปลี่ยนแปลงโครงสร้างของข้อมูล โดยข้อมูลส่วนนี้ยังไม่เคยถูกนำเข้าในคลังข้อมูล

• Refresh Data คือ การทำให้ข้อมูลทั้งหมดมีความถูกต้องทันสมัย โดยทำการประมวลผลข้อมูลใหม่ ซึ่งอาจมีผลกระทบกับการเปลี่ยนแปลงโครงสร้างของฐานข้อมูลในคลังข้อมูล

• Rebuild the Dimension Structure การปรับปรุงโครงสร้างที่ใช้สำหรับการวิเคราะห์ เช่น การเพิ่มมิติ และการเพิ่มตัววัด เป็นต้น

6) การบำรุงรักษาฐานข้อมูล

เนื่องจากข้อมูลที่ใช้ในการเก็บและวิเคราะห์เป็นข้อมูลที่มีขนาดใหญ่การเตรียมแผนและสำรองข้อมูลจากระบบคลังข้อมูลจึงมีความสำคัญ และควรทำอย่างสม่ำเสมอรวมทั้งการทดสอบและวางแผนการกู้ระบบเมื่อมีความเสียหายเกิดขึ้นด้วย

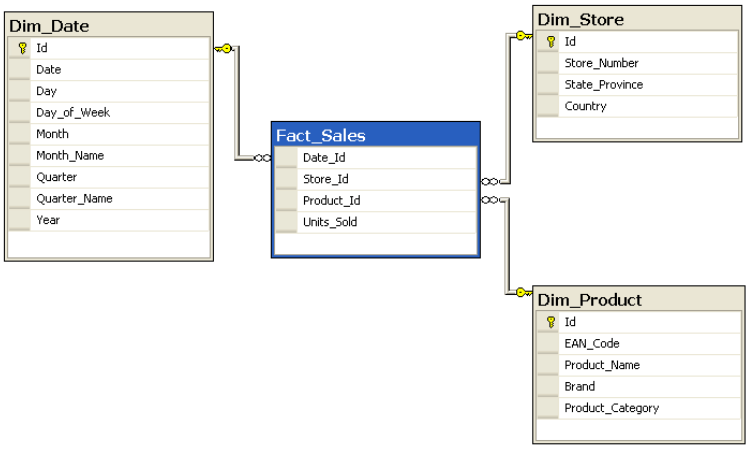
จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2.1.5 การออกแบบโครงสร้างของฐานข้อมูลหลายมิติ

การออกแบบโครงสร้างในการจัดเก็บข้อมูลของคลังข้อมูล เป็นการจัดรูปแบบของตารางเก็บข้อมูลแบบ Denormalization ซึ่งทำให้ข้อมูลที่ถูกจัดเก็บอยู่ในลักษณะของเชิงมิติ แทนที่จะเป็นการเก็บข้อมูลของ OLTP ตามแบบเชิงสัมพันธ์ปกติ การออกแบบโครงสร้างของฐานข้อมูลหลายมิติแบ่งออกเป็น 2 ประเภท (Smallcombe, 2563) ดังนี้

1) Star Schema (โครงสร้างแบบดาว)

เป็นโครงสร้างในการจัดเก็บข้อมูลที่มีรูปร่าง Diagram คล้ายรูปดาว เป็นความสัมพันธ์ระหว่างตารางที่ไม่ซับซ้อน ทำให้สามารถสืบค้นข้อมูลได้อย่างรวดเร็ว โดยองค์ประกอบของโครงสร้างประกอบไปด้วยตารางหลักที่อยู่ตรงกลางเรียกว่า Fact Table (ตารางค่าที่แท้จริง) เป็นเหมือนศูนย์รวมข้อมูล และรายล้อมไปด้วยตารางที่เรียกว่า Dimensional Table (ตารางมิติ) ซึ่งมีจำนวนเท่าไรก็ได้ แต่สิ่งสำคัญนั้นคือ ในแต่ละตารางนั้นต้องมี Key ที่สัมพันธ์ไปยัง Fact Table ตารางเดียวเท่านั้น ดังรูปที่ 5

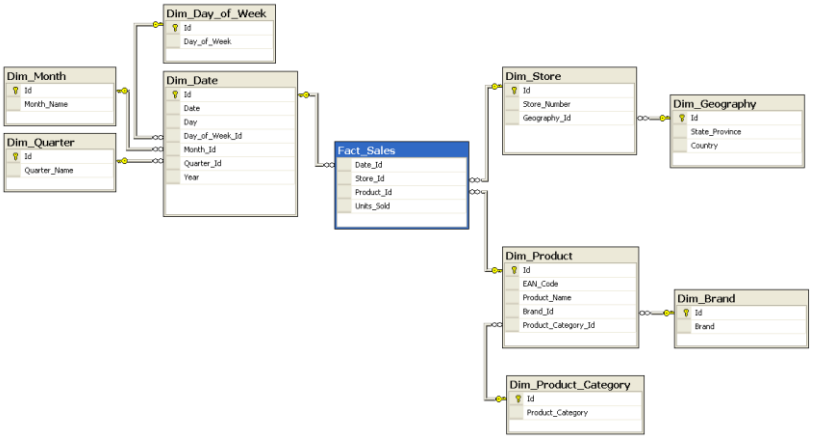


ที่มา : https://en.wikipedia.org/wiki/Star_schema

รูปที่ 5: ตัวอย่างการออกแบบโมเดลข้อมูลหลายมิติแบบ Star Schema

2) Snowflake Schema (โครงสร้างแบบเกล็ดหิมะ)

เป็นโครงสร้างที่แตกต่างจาก Star Schema ตรง Dimension Table สามารถมี หลายระดับ และ โดยระหว่าง Dimension Table มี Key ที่เชื่อมโยงกัน จึงเป็น โครงสร้างที่ซับซ้อนมาก และใช้ในการสืบค้นยาก ลักษณะของข้อมูลมีความเป็น Normalized ดังรูปที่ 6



ที่มา : https://en.wikipedia.org/wiki/Snowflake_schema

รูปที่ 6: ตัวอย่างการออกแบบโมเดลข้อมูลหลายมิติแบบ Snowflake Schema

2.2 แนวคิดด้านระบบธุรกิจอัจฉริยะ (Business Intelligence: BI)

ในปัจจุบันเทคโนโลยีมีการเปลี่ยนแปลงอย่างรวดเร็วเช่นเดียวกับระบบธุรกิจที่มีการแข่งขันกัน ดังนั้น องค์กรจึงต้องแข่งขันด้วยการใช้ข้อมูลสารสนเทศที่ทันสมัยเพื่อสนับสนุนการตัดสินใจอย่างรวดเร็วและสามารถนำไปวางแผน หรือแก้ไขปัญหาเชิงธุรกิจได้ การจะได้ข้อมูลและสารสนเทศเหล่านั้นจำเป็นต้องมีการหาวิธีการในการเก็บรวบรวมข้อมูลให้ได้มากที่สุด และมีการแสดงผลออกมาในรูปแบบที่ง่ายต่อการทำความเข้าใจ จึงจำเป็นต้องมีระบบที่สามารถช่วยเตรียมข้อมูลที่มีคุณค่าทางธุรกิจให้แก่องค์กรได้ (MBAToday, 2561)

2.2.1 นิยามระบบธุรกิจอัจฉริยะ Business Intelligence (BI) หรือ ระบบธุรกิจอัจฉริยะ

เป็นชุดเครื่องมือ หรือซอฟต์แวร์ (Software) ที่ใช้ในการทำงานทางด้านเทคโนโลยีสารสนเทศ ซึ่งประกอบด้วยชุดคำสั่งเพื่อทำการรวบรวมข้อมูลจากแหล่งข้อมูลที่มาจากระบบ สารสนเทศต่าง ๆ ทั้งในและนอกองค์กรที่เรียกว่า แหล่งข้อมูลขนาดใหญ่ (Big Data) ซึ่งจะมีการนำข้อมูลที่มีความหลากหลายมาวิเคราะห์จัดการตัวชุดคำสั่งงาน เพื่อให้เกิดเป็นสารสนทศตามที่ผู้ใช้ต้องการ และจัดทำการนำเสนอในรูปแบบต่าง ๆ ในหลายมิติให้เหมาะสมกับมุมมองในการวิเคราะห์ และแสดงความสัมพันธ์ที่เป็นรูปภาพหรือกราฟที่ดูง่ายยิ่งขึ้น เพื่อประโยชน์ในการวางแผนกลยุทธ์ด้านต่าง ๆ (Meeieis, 2560)

2.2.2 องค์ประกอบของระบบธุรกิจอัจฉริยะ

องค์ประกอบของระบบธุรกิจอัจฉริยะ (เกียรติพงษ์ อุดมธนะธีระ, 2561) มีดังนี้

1) ข้อมูลขนาดใหญ่ (Big Data) เริ่มจากการตั้งข้อปัญหาและจัดหาแหล่งข้อมูล ซึ่งอาจเป็นข้อมูลทั้งจากภายในองค์กรหรือจากภายนอกองค์กร โดยมีการจัดการข้อมูลจากแหล่งที่มาต่าง ๆ ที่หลากหลาย
2) ทำการจัดการข้อมูล (Data Sources) ให้เป็นชุดข้อมูลที่มีหมวดหมู่ และมีการตรวจสอบข้อมูล
3) ทำการปรับเปลี่ยนข้อมูล (Data Transformation) โดยจัดรูปแบบข้อมูลให้มีมาตรฐานที่ สามารถจัดการหรือนำมาใช้งานต่อได้ง่าย แล้วจึงนำไปเก็บที่คลังข้อมูล (Data Warehouse) ซึ่งการทำงานในส่วนนี้เรียกว่า การทำงานแบบ ETL ประกอบด้วยกระบวนการต่าง ๆ ดังนี้

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

• Extract การคัดแยกหรือจัดกลุ่มข้อมูล โดยดึงข้อมูลจากแหล่งข้อมูล และนำมาคัดแยกว่าข้อมูลตรงกับการใช้งานหรือไม่

• Transform การปรับเปลี่ยนรูปแบบของข้อมูลให้ตรงกับความต้องการ โดยตัดข้อมูลที่ไม่จำเป็นออกเพื่อเพิ่มคุณภาพข้อมูล

• Load การนำเข้าข้อมูลที่มีการจัดการแล้วไปจัดเก็บในพื้นที่ที่ได้จัดเตรียมไว้ ซึ่งคือคลังข้อมูล

เครื่องมือที่ใช้ในการกระบวนการ ETL จะทำการทำความสะอาดข้อมูล (Data Cleansing) เพื่อให้ข้อมูลมีความสม่ำเสมอ สอดคล้องกันทั้งหมด ก่อนจะนำบรรจุลงที่เก็บที่เรียกว่า คลังข้อมูล (Data Warehouse) ต่อไป

4) คลังข้อมูล (Data Warehouse) เป็นที่จัดเก็บข้อมูลนำมาจากแหล่งข้อมูลภายในองค์กร ซึ่งก็คือระบบสารสนเทศในระดับปฏิบัติการ และแหล่งข้อมูลภายนอกที่ผู้บริหารเห็นว่ามีความ จำเป็นต้องใช้ในการตัดสินใจของผู้บริหาร ซึ่งข้อมูลเหล่านั้นจะถูกนำมาจัดเตรียมให้อยู่ในรูปแบบที่พร้อมจะทำงานเชิงวิเคราะห์ (Analytical Data) ตามที่ผู้บริหารต้องการได้ คลังข้อมูลจะเป็นฐานข้อมูลสำหรับการวิเคราะห์ด้วยชุดคำสั่งงานต่าง ๆ ต่อไป ดังรูปที่ 7 เช่น การประมวลผลเชิงวิเคราะห์แบบออนไลน์หรือโอแลบ (On-Line Analytical Processing, OLAP) การทำเหมืองข้อมูล (Data Mining) และระบบสารสนเทศอื่น ๆ เป็นต้น
5) ดาต้ามาร์ท (Data Mart) คือ ข้อมูลในคลังที่มีการจัดแบ่งกลุ่มขนาดเล็ก การเก็บข้อมูลจะเลือกเอาที่จำเป็นและต้องการใช้ในงานเฉพาะเท่านั้น ซึ่งจะเป็นส่วนที่ย่อยจาก Data Warehouse มีลักษณะเฉพาะเพื่อให้มีขนาดของข้อมูลที่เล็กลงและค่าใช้จ่ายต่ำ
6) การแก้ปัญหาทางคณิตศาสตร์ (Operations Research & Numerical Methods) เป็นชุดคำสั่งงานเพื่อการวิเคราะห์ข้อมูล ประกอบด้วยชุดคำสั่งงานหลายแบบที่จะทำการวิเคราะห์ในจุดมุ่งหมายและประเด็นที่แตกต่างหลากหลายกันไป
7) เครื่องมือที่ใช้ในการวิเคราะห์ข้อมูลในหลายมิติ (Online Analytical Processing, OLAP) เป็นชุดคำสั่งงานที่ช่วยให้ผู้ใช้งานวิเคราะห์ข้อมูลที่มาจากคลังข้อมูล มีการประมวลผลเชิงวิเคราะห์แบบออนไลน์ การวิเคราะห์ข้อมูลที่เกิดขึ้นบ่อยจะเป็นการวิเคราะห์ข้อมูลหลายมิติ (Multidimensionality) เพื่อช่วยให้ผู้วิเคราะห์ได้มองเห็นข้อมูลในเชิงลึกในมิติต่าง ๆ เป็นการเสริมความเข้าใจในสถานการณ์ให้มากขึ้น
8) การทำเหมืองข้อมูล (Data Mining) เป็นชุดคำสั่งงานที่ใช้ในการวิเคราะห์ข้อมูลเพื่อค้นหา ความสัมพันธ์ในข้อมูลจากคลังข้อมูลที่ไม่เคยมีการค้นพบมาก่อน

เพราะการได้ค้นพบสิ่งใหม่ก่อนผู้อื่นอาจจะสร้างความได้เปรียบในการแข่งขัน ตัวอย่างผลการวิเคราะห์ที่นำเสนอจากการทำเหมืองข้อมูล เช่น การวิเคราะห์เพื่อจัดประเภทลูกค้า การค้นหากลุ่มของลูกค้า การค้นหา ลักษณะหรือพฤติกรรมของลูกค้าในแต่ละกลุ่ม การพยากรณ์พฤติกรรมของลูกค้าที่อาจจะพาไปสู่การกระทำที่ไม่ดี เช่น การฉ้อโกงองค์กร เป็นต้น ซึ่งในโครงการนี้จะสนใจการวิเคราะห์เพื่อจัดประเภทของกลุ่มใบเสร็จ เพื่อค้นหากลุ่มของลูกค้าตามพฤติกรรมการสั่งซื้ออาหารและเครื่องดื่มโดยใช้เทคนิคที่เรียกว่า K-means

• K-means คือการแบ่งกลุ่ม แบบ Clustering ซึ่งการแบ่งกลุ่มในลักษณะนี้จะใช้พื้นฐานทาง สถิติ ซึ่งหน้าที่ของ clustering คือการจับกลุ่มของข้อมูลที่มีลักษณะใกล้เคียงกันเป็นกลุ่ม เดียวกัน

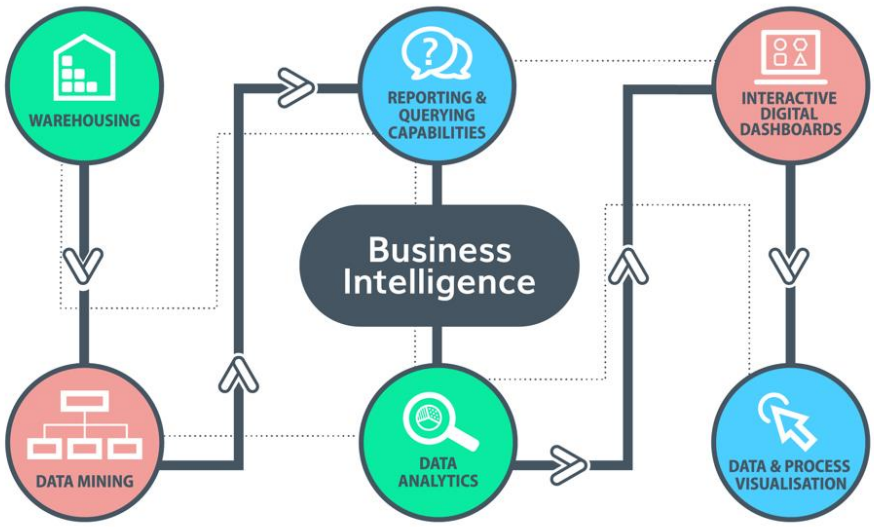
• การหาจำนวนกลุ่ม k ที่เหมาะสมที่สุดด้วยวิธี Elbow Method (Weerasak, 2560)

เป็นวิธีหนึ่งซึ่งใช้การวัดข้อผิดพลาด (Error measurement) ของผลรวมของระยะห่างระหว่าง Object กับ Centroid เมื่อความผิดพลาดลดน้อยลง เส้นโค้งที่มีความชันจะเริ่มโค้ง และราบเรียบ (Smooth) จนเกิดเป็นมุมลักษณะเหมือน Elbow ณ จุดนี้เป็นจุดที่ให้ค่า จำนวนกลุ่ม Cluster ที่ดีที่สุด

• การหาจำนวนกลุ่ม k ที่เหมาะสมที่สุดด้วยวิธี Silhouette (Weerasak, 2560)

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

Silhouette เป็นเทคนิคที่ใช้วัดว่า Instance นั้นมีความเหมือนกับ Cluster ที่มันอยู่มาก เพียงใด เมื่อเทียบกับ Cluster อื่นๆ ค่าของ Silhouette อยู่ในช่วง -1 ถึง 1 ยิ่งมีค่ามาก แสดงว่า Instance มีความคล้ายกับ cluster ของมันมากและมีความคล้ายกับ Cluster อื่นน้อย



ที่มา : https://www.4x-treme.com/business-intelligence/?lang=en

รูปที่ 7: Business Intelligence Model

2.3 แนวคิดด้านการวิเคราะห์ข้อมูลเชิงทำนาย

2.3.1 การวิเคราะห์ข้อมูลเชิงทำนาย

การวิเคราะห์เชิงทำนาย (Predictive analysis) เป็นการวิเคราะห์ข้อมูลในอดีตเพื่อนำข้อมูลมาหา รูปแบบความสัมพันธ์ในชุดข้อมูลที่สามารถนำมาเป็นต้นแบบในการทำนายการคาดการณ์ผล หรือสิ่งที่น่าจะเกิดขึ้นในอนาคต ซึ่งจะช่วยให้ผู้ใช้ข้อมูลสามารถตัดสินใจได้อย่างมีประสิทธิภาพ (Krooann, 2560)

เทคนิคการวิเคราะห์เชิงทำนายสามารถแบ่งได้หลายประเภท เช่น วิธีการวิเคราะห์การถดถอย (Regression Analysis) วิธีการเรียนรู้ของเครื่อง (Machine Learning) และวิธีการหากฎของความสัมพันธ์ (Association Rules Discovery) เป็นต้น ซึ่งในโครงการนี้จะสนใจการวิเคราะห์เชิงทำนายแบบวิธีการหากฎ ของความสัมพันธ์ และวิธีการการวิเคราะห์การถดถอย

2.3.2 กฎของความสัมพันธ์

กฎความสัมพันธ์ (Association Rule) คือการทำเหมืองข้อมูลโดยใช้กฎความสัมพันธ์ เพื่อหาความสัมพันธ์ของเหมืองข้อมูลมักใช้ในธุรกิจการค้าปลีก (retailing business) เช่น ร้านค้าสะดวกซื้อ หรือ ซุปเปอร์มาเก็ต เป็นการวิเคราะห์ตะกร้าตลาด (Market basket analysis) เพื่อศึกษาพฤติกรรมการซื้อสินค้าของลูกค้า และหา

ความสัมพันธ์ของสินค้าที่ลูกค้าซื้อ เพื่อนำผลลัพธ์ที่ได้จากความสัมพันธ์ (Mindphp, 2562)

2.3.3 การวิเคราะห์การถดถอย

การวิเคราะห์การถดถอย (Regression Analysis) เป็นวิธีการทางสถิติที่ใช้เพื่อศึกษาความสัมพันธ์ระหว่างตัวแปรตั้งแต่ 2 ตัวขึ้นไป ซึ่งประกอบด้วย ตัวแปรที่ทราบค่า หรือ ที่เรียกว่า ตัวแปรต้น (Independent variable) และตัวแปรที่เราต้องการจะทราบค่าเรียกว่าตัวตอบสนองหรือตัวแปรตาม (Dependent variable) ว่าเป็นเหตุผลของกันและกันหรือไม่ (สุทิน ชนะ บุญ, 2564)

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

บทที่ 3
โครงสร้างองค์กรและการดำเนินงาน

บทนี้จะกล่าวถึงประวัติขององค์กร โครงสร้างองค์กร ลักษณะการดำเนินงานขององค์กรในส่วนที่เกี่ยวข้องกับการพัฒนา และปัญหาที่เกิดขึ้นในปัจจุบันขององค์กรต้นแบบนี้

3.1 ข้อมูลเกี่ยวกับองค์กร

กิจการร้านอาหารและเครื่องดื่ม (Bluebird Jazz Bar) ตั้งอยู่บนถนนทองหล่อ กรุงเทพฯ บริหารโดย นายอรรถสิทธิ์ โคมินทร์ (คุณหมอแอนดี้) โดยจุดเริ่มต้นเกิดจากสมัยอดีตที่คุณหมอเรียนอยู่ต่างประเทศ และได้ใช้เวลาส่วนใหญ่หลังเลิกเรียนเพื่อเดินทางไปฟังเพลงแจ๊สคลาสสิคที่ร้านขายแผ่นเสียงเก่าเป็นประจำ จนกระทั่งเมื่อถึงเวลาเรียนจบ และกลับมาประเทศไทย นอกจากงานประกอบอาชีพแพทย์เป็นงานหลักแล้ว ยังได้ประกอบอาชีพที่สองควบคู่ไปด้วย ซึ่งก็คือการเปิดร้านอาหารและเครื่องดื่มเป็นของตนเอง ตั้งแต่ปี พ.ศ. 2563 เป็นต้นมา

บรรยากาศภายในร้านอาหารและเครื่องดื่ม Bluebird Jazz Bar ให้ความรู้สึกอบอุ่นเหมือนบ้านหลังที่สอง นอกจากจะมี โต๊ะ เก้าอี้ โซฟา ไว้สำหรับรองรับลูกค้าที่เข้ามาใช้บริการรับประทานอาหารและเครื่องดื่มเป็นประจำแล้ว ยังมีกิจกรรมต่าง ๆ ที่สร้างบรรยากาศ และความสุขให้กับลูกค้าที่เข้ามาใช้บริการอีกด้วย เช่น การเปิดเพลงดนตรีคลอเบาๆ ด้วยเครื่องเล่นแผ่นเสียง เพื่อช่วยสร้างบรรยากาศในระหว่างการรับประทานอาหารและเครื่องดื่ม การจัดจำหน่ายบัตรคอนเสิร์ต เพื่อให้ลูกค้ารับชมดนตรีหลากหลายประเภท ไม่ว่าจะเป็น Acoustic Folk, Piano Solo, Jazz-Trio และ Mini Orchestra เป็นต้น

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ปัจจุบันรูปแบบบริการขายอาหารและเครื่องดื่มของร้านแบ่งออกเป็น 2 ประเภท ได้แก่ รูปแบบ Café และ รูปแบบ Bar โดยมีตัวอย่างเมนูอาหารและเครื่องดื่มดังนี้

1. การบริการรูปแบบ Café

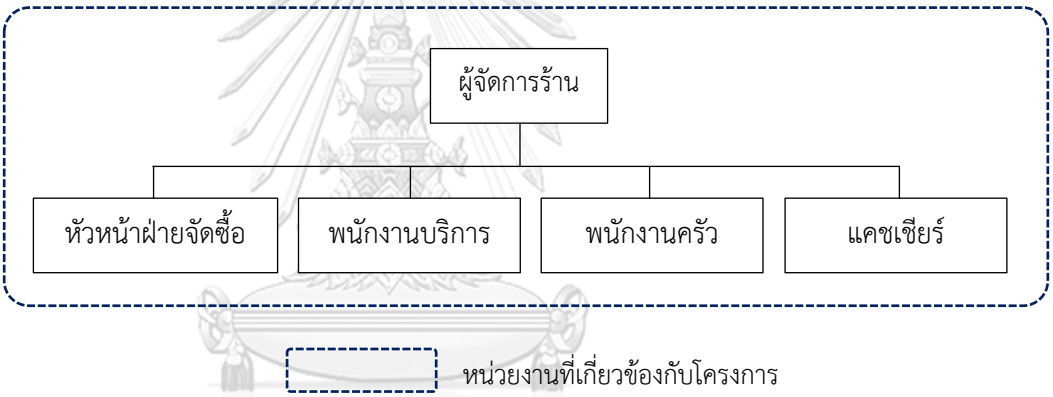
- ประเภท Dessert ได้แก่ Cheesecake, Brownie, Banoffee และ Tiramisu ฯลฯ
- ประเภท Coffee ได้แก่ Americano, Latte, Caramel Macchiato และ Mocha ฯลฯ
- ประเภท Tea ได้แก่ Matcha Latte, Tea pot and Earl Grey Fruity tea ฯลฯ
- ประเภท Other ได้แก่ Apple Juice, Baby Bluebird และ Lemon Honey Soda ฯลฯ

2. การบริการรูปแบบ Bar

- ประเภท Food & Snacks ได้แก่ Potato wedges และ Shumai Sausage Arabiki ฯลฯ
- ประเภท Beer ได้แก่ Singha, Asahi, Stella, Kirin, Chatri และ Hoegaarden ฯลฯ
- ประเภท Cocktails ได้แก่ Old fashioned, Bluebird, Negroni และ Margarita ฯลฯ
- ประเภท Gin & Tonic ได้แก่ Hendrick, Gordon, Roku และ Beefeater ฯลฯ
- ประเภท Wine ได้แก่ Red, White, Sparking rose และ Nua Brut Sparkling ฯลฯ

3.2 โครงสร้างองค์กร

โครงสร้างองค์กรของร้านอาหารและเครื่องดื่ม (Bluebird Jazz Bar) แสดงได้ดังรูปที่ 8



รูปที่ 8: โครงสร้างองค์กรของร้าน Bluebird Jazz Bar

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

กิจการร้านมีการแบ่งงานและหน้าที่ความรับผิดชอบ ดังนี้

1) **ผู้จัดการร้าน** มีหน้าที่ ดูแลรับผิดชอบการบริหารกิจการทั้งหมด
2) **หัวหน้าฝ่ายจัดซื้อ** มีหน้าที่ ประเมินยอดขาย และซื้อวัตถุดิบเข้าร้านให้เพียงพอสำหรับให้บริการลูกค้า
3) **พนักงานบริการ** มีหน้าที่ รับ Order และ เสิร์ฟเมนูอาหาร และเครื่องดื่ม
4) **พนักงานครัว** มีหน้าที่ จัดทำอาหารและเครื่องดื่มตามรายการที่ลูกค้าสั่ง
5) **แคชเชียร์** มีหน้าที่ รับชำระเงินจากลูกค้า และออกใบเสร็จรับเงิน

3.3 การดำเนินงานขององค์กร

1) **การบริการร้าน**

- แบ่งออกเป็น 2 รูปแบบ คือ Café และ Bar

2) **การรับ Order**

- พนักงานบริการจะทำการต้อนรับลูกค้า และยื่นเมนูอาหารและเครื่องดื่มเพื่อรับ Order จากนั้นจึงกลับมาบันทึกคำสั่งซื้อผ่านระบบ POS ที่บริเวณเคาน์เตอร์บาร์

3) **การออกใบเสร็จรับชำระเงิน**

- เมื่อลูกค้าต้องการเช็คบิล พนักงานบริการจะสั่งพิมพ์ใบเสร็จ และนำส่งให้กับลูกค้าชำระเงินตามจำนวนที่ได้ระบุไว้ในใบเสร็จ

4) **การชำระเงินของลูกค้า**

- สำหรับการชำระเงินสามารถชำระได้ 3 ช่องทาง คือ เงินสด เงินโอนและบัตรเครดิต

3.4 ปัญหาที่เกิดขึ้นในปัจจุบัน

1) ปัญหายอดขาย

- ธุรกิจเพิ่งจะเริ่มต้นได้ไม่นานมากนัก และประสบกับช่วงวิกฤตสถานการณ์ระบาดของโควิด-19 ทำให้ยังไม่ทราบสถานะแนวโน้มของยอดขายว่าเป็นอย่างไร
- ต้องการปรับปรุงยอดขายแต่ยังไม่มีแนวทางการส่งเสริมการขายที่ชัดเจน

2) ปัญหาต้นทุน

- มีอาหารหรือเครื่องดื่มในเมนูบางชนิดที่ไม่ได้ถูกขายออกไป ทำให้วัตถุดิบคงค้าง เกิดเป็นค่าใช้จ่ายที่ไม่จำเป็น และสูญเสียโดยเปล่าประโยชน์

3) ปัญหาลูกค้า

- ลูกค้าเก่าบางรายที่เคยเข้ามาใช้บริการหายไป
- ลูกค้าประจำยังคงมีสัดส่วนที่น้อยกว่าลูกค้าสัญจร

4) ปัญหาคู่แข่ง

- มีคู่แข่งมากในกลุ่มของร้านอาหารและเครื่องดื่มประเภทเดียวกันจึงจำเป็นต้องค้นหากลยุทธ์การขายที่แตกต่างจากคู่แข่ง

บทที่ 4
การพัฒนาระบบงาน

ในบทนี้กล่าวถึงขั้นตอนต่าง ๆ ของการพัฒนาโครงการพิเศษ “คลังข้อมูล ธุรกิจอัจฉริยะ และการวิเคราะห์ข้อมูลขั้นสูงของธุรกิจร้านอาหารและเครื่องดื่ม” โดยจะกล่าวถึงการวิเคราะห์ระบบ การออกแบบระบบ และการพัฒนาระบบ

4.1 การวิเคราะห์ระบบ

4.1.1 คุณสมบัติที่ต้องการของระบบ

โครงการพิเศษ “คลังข้อมูล ธุรกิจอัจฉริยะ และการวิเคราะห์ข้อมูลขั้นสูงของธุรกิจร้านอาหารและเครื่องดื่ม” มีคุณสมบัติที่ต้องการโดยรวมดังต่อไปนี้

1) การรวบรวมข้อมูลไว้ภายใต้มาตรฐานข้อมูลเดียวกัน (Integrated System) ระบบที่พัฒนามีการรวบรวมข้อมูลที่เกี่ยวข้องจากหน่วยงานต่าง ๆ ภายใน องค์กร ซึ่งข้อมูลที่ได้มานั้นอยู่ในหลากหลายฐานข้อมูล ได้นำมาจัดเก็บให้มี มาตรฐานและอยู่ในฐานข้อมูลเดียวกัน เพื่อให้เกิดความถูกต้อง สอดคล้องกัน ลด ความซ้ำซ้อน และมีความเหมาะสมสำหรับการนำไปวิเคราะห์ข้อมูล
2) การสร้างรูปแบบรายงานที่ยืดหยุ่นและหลากหลาย (Flexibility and Diversity) ระบบที่พัฒนาขึ้นมีการนำเสนอรายงานหลากหลายรูปแบบ ทั้งรูปแบบของ แผนภูมิชนิดต่างๆ และ ตารางข้อมูลในหลายมิติ เพื่อให้เหมาะสมกับข้อมูลและ ผู้ใช้งานในหลายรูปแบบ โดยผู้ใช้ระบบสามารถเจาะลึกข้อมูลลงไปในรายละเอียด (Drill Down) หรือเปลี่ยนจากรายละเอียดมาเป็นข้อมูลสรุป (Roll Up) เพื่อดู ข้อมูลในระดับต่าง ๆ ได้ เพื่อให้ผู้ใช้งานเข้าใจข้อมูลได้อย่างรวดเร็ว และยัง สามารถเลือกพิจารณาผลลัพธ์บางส่วนที่สนใจ โดยการเลือกเฉพาะค่าบางค่าของแต่ละมิติ (Slice) นอกจากนี้ผู้ใช้งานยังสามารถส่งออกข้อมูลเพื่อนำไปใช้งานใน รูปแบบของแฟ้มเอกสารได้หลากหลายรูปแบบ เช่น PDF XML หรือ Microsoft Excel เป็นต้น
3) ส่วนต่อประสานกับผู้ใช้งาน (User Interface) การติดต่อระหว่างผู้ใช้งานกับระบบผ่านรูปแบบ Graphic User Interface (GUI) ของ เครื่องมือ Tableau Desktop ง่ายต่อการใช้งาน และผู้ใช้งานสามารถทำความ เข้าใจระบบได้ง่ายยิ่งขึ้น ซึ่งเครื่องมือ Tableau Desktop เป็นซอฟต์แวร์ที่สามารถติดตั้งที่เครื่องของผู้ใช้งาน Tableau Desktop และสามารถเข้าถึงคลังข้อมูลผ่านเว็ปบราวเซอร์ได้ด้วย Tableau Server เช่นกัน

4) การควบคุมด้านความปลอดภัยในการเข้าใช้งานระบบเครื่องมือ Tableau Desktop มีการจัดการทางด้านความปลอดภัย โดยการควบคุมผู้ใช้งานผ่าน E-mail ที่ต้องลงชื่อเข้าใช้ผ่านการ Log in และมีการจัดการสิทธิในการเข้าถึงข้อมูลในองค์กรในหลายรูปแบบ เช่น การเปิดเผยข้อมูลแบบสาธารณะ การกำหนดการเข้าถึงรายบุคคลตาม E-mails หรือ แบบ Workspace ที่กลุ่มผู้ใช้งานที่จะต้องใช้งานร่วมกัน เป็นต้น

4.1.2 ความต้องการโดยละเอียดของระบบ

โครงการ “คลังข้อมูล ธุรกิจอัจฉริยะ และการวิเคราะห์ข้อมูลขั้นสูงของธุรกิจร้านอาหารและเครื่องดื่ม” ประกอบด้วย 5 ระบบย่อย โดยมีรายละเอียดของแต่ละระบบดังนี้

4.1.2.1 ระบบวิเคราะห์ยอดขาย (Sales Analysis System)

ภาพรวมของระบบ (System Overview)

ระบบนี้เป็นการแสดงผลประกอบการของธุรกิจร้านอาหารและเครื่องดื่มในรูปแบบ Dashboard เพื่อให้ผู้จัดการร้าน และหัวหน้าฝ่ายจัดซื้อใช้ตรวจสอบสถานะยอดขายเฉลี่ยต่อใบเสร็จ (Average sales per receipt) ในมิติต่าง ๆ ได้ทันที เช่น มิติเวลา มิติรูปแบบการให้บริการ และมิติสินค้า เป็นต้น

ผู้ใช้ (Users)

1) ผู้จัดการร้าน
2) หัวหน้าฝ่ายจัดซื้อ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

คำถามผู้บริหาร (Management Questions)

1) กิจการร้านมีแนวโน้มยอดขายเป็นอย่างไร
2) ตลอดระยะเวลาที่ผ่านมามียอดขายเฉลี่ยต่อใบเสร็จเป็นเท่าไหร่
3) กิจการสามารถสร้างยอดขายได้ดีในช่วงวันหยุดสุดสัปดาห์ใช่หรือไม่
4) ประเภทเมนูใดที่มียอดขายสูงที่สุด
5) การแสดงดนตรีสดมีผลต่อยอดขายใช่หรือไม่

แดชบอร์ดการวิเคราะห์ (Analytics Dashboard)

1) แดชบอร์ดการวิเคราะห์ทิศทางและแนวโน้มยอดขายของกิจการ
2) แดชบอร์ดการวิเคราะห์ยอดขายแยกตามประเภทรายการอาหารและเครื่องดื่ม

มิติ (Dimensions)

1) มิติเวลา (Time Dimensions)

มิติเวลามีลำดับชั้นในการวิเคราะห์ ดังนี้

- ปี (Year)
- เดือน (Month)
- วัน (Day)
- ชื่อวันในสัปดาห์ (Day of Week)
- การจัดแสดงดนตรี (Event Day)

2) มิติรูปแบบการให้บริการ (Service Dimensions)

มิติรูปแบบการให้บริการมีลำดับชั้นในการวิเคราะห์ ดังนี้

- ประเภทรูปแบบการให้บริการ (Type of Service)

3) มิติสินค้า (Product Dimensions)

มิติสินค้ามีลำดับชั้นในการวิเคราะห์ ดังนี้

- ประเภทสินค้า (Category Name)
- ชื่อสินค้า (Product Name)
- จำนวนการสั่งสินค้า (Quantity)
- ราคาต่อหน่วย (Price)

ค่าวัด (Measures)

ตารางที่ 2: ค่าวัดของระบบวิเคราะห์ยอดขาย (Sales Analysis System)

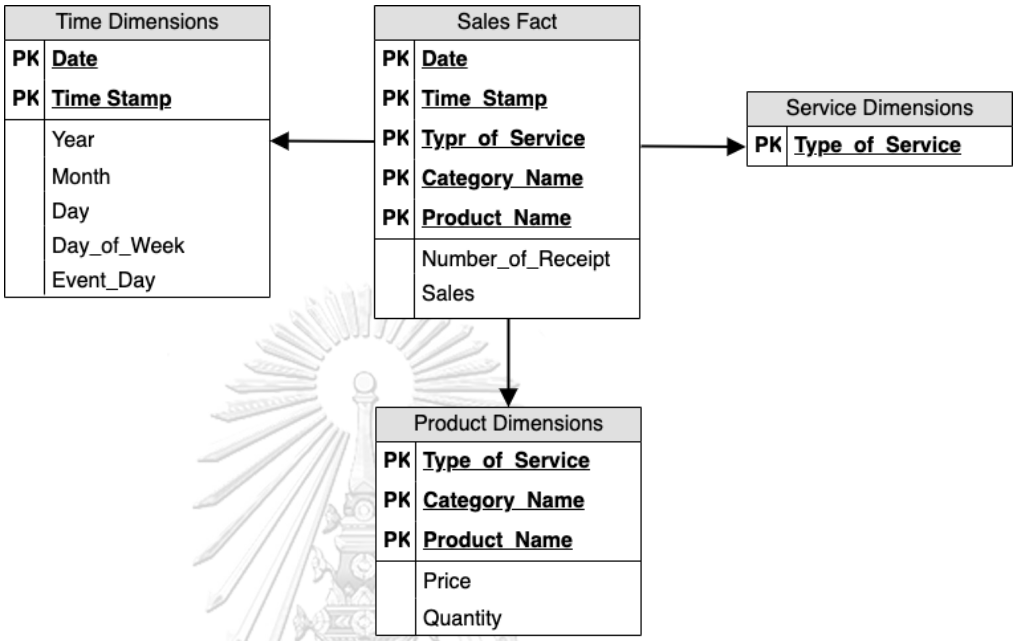
ลำดับ	ค่าวัด	ชื่อภาษาอังกฤษ (หน่วยวัด)	ชื่อภาษาไทย (หน่วยวัด)
1	Sales	Sales (Baht)	ยอดขาย (บาท)
2	Number of Receipt	Number of Receipt (Copies)	จำนวนใบเสร็จ (ใบ)

ตัวชี้วัดผลการดำเนินการหลัก (KPIs)

ตารางที่ 3: ตัวชี้วัดผลการดำเนินการหลักของระบบวิเคราะห์ยอดขาย (Sales Analysis System)

ลำดับ	ตัวชี้วัดผลการดำเนินการหลัก (ชื่อภาษาอังกฤษและภาษาไทย) (หน่วยวัด)	สูตรคำนวณ
1	Average sales per receipt (Baht) ยอดขายเฉลี่ยต่อใบเสร็จ (บาท)	$\frac{\text{ยอดขายต่อใบเสร็จทั้งหมด}}{\text{จำนวนใบเสร็จ}}$
2	Growth average sales per receipt (%) แนวโน้มอัตราการเติบโตของยอดขายเฉลี่ยต่อใบเสร็จ (%)	$\frac{(\text{ยอดขายต่อใบเสร็จของช่วงเวลาถัดไป} - \text{ยอดขายต่อใบเสร็จของช่วงเวลาก่อนหน้า})}{\text{ยอดขายต่อใบเสร็จของช่วงเวลาก่อนหน้า}} \times 100$
3	Average sales per receipt by Product Category (Baht) ยอดขายเฉลี่ยต่อใบเสร็จของแต่ละประเภทสินค้า (บาท)	$\frac{\text{ยอดขายต่อใบเสร็จของประเภทสินค้าที่สนใจ}}{\text{จำนวนใบเสร็จของประเภทสินค้าที่สนใจ}}$
4	Growth average sales per receipt by Product Category (%) แนวโน้มอัตราการเติบโตของยอดขายเฉลี่ยต่อใบเสร็จของแต่ละประเภทสินค้า (%)	$\frac{(\text{ยอดขายต่อใบเสร็จของประเภทสินค้าที่สนใจของช่วงเวลาถัดไป} - \text{ยอดขายต่อใบเสร็จของประเภทสินค้าที่สนใจของช่วงเวลาก่อนหน้า})}{\text{ยอดขายต่อใบเสร็จของประเภทสินค้าที่สนใจของช่วงเวลาก่อนหน้า}} \times 100$
5	Growth average sales per receipt by Event Day (%) แนวโน้มอัตราการเติบโตของยอดขายเฉลี่ยต่อใบเสร็จของวันที่มีการแสดงดนตรี (%)	$\frac{(\text{ยอดขายต่อใบเสร็จของวันที่มีการแสดงดนตรีของช่วงเวลาถัดไป} - \text{ยอดขายต่อใบเสร็จของวันที่มีการแสดงดนตรีของช่วงเวลาก่อนหน้า})}{\text{ยอดขายต่อใบเสร็จของวันที่มีการแสดงดนตรีของช่วงเวลาก่อนหน้า}} \times 100$

โมเดลข้อมูลหลายมิติ (Multidimensional Data Model)



รูปที่ 9: โมเดลข้อมูลหลายมิติของระบบวิเคราะห์ยอดขาย (Sales Analysis System)

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

คำถามของผู้บริหาร ผู้ใช้ และแดชบอร์ดการวิเคราะห์ (Management Questions, Users and Analytics Dashboard)

ตารางที่ 4: คำถามของผู้บริหาร ผู้ใช้ และแดชบอร์ดการวิเคราะห์ของระบบวิเคราะห์ยอดขาย (Sales Analysis System)

คำถามของผู้บริหาร (Management Questions)	ผู้ใช้ (Users)	แดชบอร์ดการวิเคราะห์ (Analytics Dashboard)
1) กิจการร้านมีแนวโน้มยอดขายเป็นอย่างไร	1) ผู้จัดการร้าน 2) หัวหน้าฝ่ายจัดซื้อ	1) แดชบอร์ดการวิเคราะห์ทิศทางและแนวโน้มยอดขายของกิจการ
2) ตลอดระยะเวลาที่ผ่านมามียอดขายเฉลี่ยต่อใบเสร็จเป็นเท่าไหร่	1) ผู้จัดการร้าน 2) หัวหน้าฝ่ายจัดซื้อ	1) แดชบอร์ดการวิเคราะห์ทิศทางและแนวโน้มยอดขายของกิจการ
3) กิจการร้านสามารถสร้างยอดขายได้ดีในช่วงวันหยุดสุดสัปดาห์ใช่หรือไม่	1) ผู้จัดการร้าน 2) หัวหน้าฝ่ายจัดซื้อ	1) แดชบอร์ดการวิเคราะห์ทิศทางและแนวโน้มของยอดขายกิจการ
4) ประเภทเมนูใดที่มียอดขายสูงที่สุด	1) ผู้จัดการร้าน 2) หัวหน้าฝ่ายจัดซื้อ	1) แดชบอร์ดการวิเคราะห์ยอดขายแยกตามรายการอาหารและเครื่องดื่ม
5) วันที่มีการแสดงดนตรีมีผลต่อยอดขายสูงใช่หรือไม่	1) ผู้จัดการร้าน 2) หัวหน้าฝ่ายจัดซื้อ	1) แดชบอร์ดการวิเคราะห์ทิศทางและแนวโน้มยอดขายของกิจการ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

แดชบอร์ดการวิเคราะห์ ค่าวัด ตัวชี้วัดผลการดำเนินงานหลัก และมิติ

(Analytics Dashboard, Measures, KPIs and Dimensions)

ตารางที่ 5: แดชบอร์ดการวิเคราะห์ ค่าวัด ตัวชี้วัดผลการดำเนินงานหลัก และมิติของระบบวิเคราะห์ยอดขาย (Sales Analysis System)

แดชบอร์ดการวิเคราะห์ (Analytics Dashboard)	ค่าวัด (Measures)	ตัวชี้วัดผลการดำเนินงานหลัก (KPIs)	มิติ (Dimensions)
1) แดชบอร์ดการวิเคราะห์ทิศทางและแนวโน้มยอดขายของกิจการ	1) ยอดขาย 2) จำนวนใบเสร็จ	1) ยอดขายเฉลี่ยต่อใบเสร็จ	1) มิติเวลา 2) มิติสินค้า
	1) ยอดขาย 2) จำนวนใบเสร็จ	1) แนวโน้มอัตราการเติบโตของยอดขายเฉลี่ยต่อใบเสร็จ	1) มิติเวลา 2) มิติสินค้า
	1) ยอดขาย 2) จำนวนใบเสร็จ	1) แนวโน้มอัตราการเติบโตของยอดขายเฉลี่ยต่อใบเสร็จของวันที่มีการแสดงดนตรี	1) มิติเวลา 2) มิติสินค้า
2) แดชบอร์ดการวิเคราะห์ยอดขายแยกตามรายการอาหารและเครื่องดื่ม	1) ยอดขาย 2) จำนวนใบเสร็จ	1) แนวโน้มอัตราการเติบโตของยอดขายเฉลี่ยต่อใบเสร็จของแต่ละประเภทสินค้า	1) มิติเวลา 2) มิติสินค้า

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

4.1.2.2 ระบบการวิเคราะห์การสั่งอาหารและเครื่องดื่ม (Food & Beverage Order Analysis System)

ภาพรวมของระบบ (System Overview)

ระบบนี้เป็นการวิเคราะห์การสั่งอาหารและเครื่องดื่มในรูปแบบของ Dashboard เพื่อให้ผู้จัดการร้าน ใช้ตรวจสอบสัดส่วนของวันหรือเวลาที่ลูกค้าเข้ามาใช้บริการ (Percentage of datetime in receipts) จำนวนลูกค้าเฉลี่ยต่อใบเสร็จ (Average number of customers per receipt) และสัดส่วนวิธีการชำระเงินของลูกค้าในใบเสร็จ (Percentage of payment method in receipts) ตามมิติต่าง ๆ ได้ทันที เช่น มิติใบเสร็จ มิติลูกค้า มิติเวลา และมิติการชำระเงิน

เป็นต้น

ผู้ใช้ (Users)

1) ผู้จัดการร้าน
2) หัวหน้าฝ่ายจัดซื้อ

คำถามผู้บริหาร (Management Questions)

1) ลูกค้าส่วนใหญ่มักมาช่วงวันใดในสัปดาห์
2) ระยะเวลาที่ลูกค้าใช้บริการในร้านเฉลี่ยต่อใบเสร็จเป็นเท่าไหร่
3) ช่วงเวลาใดที่มีจำนวนลูกค้าเข้ามาใช้บริการมากที่สุด
4) แนวโน้มพฤติกรรมการชำระเงินของลูกค้าในแต่ละช่องทางเป็นอย่างไร

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

แดชบอร์ดการวิเคราะห์ (Analytics Dashboard)

1) แดชบอร์ดการวิเคราะห์พฤติกรรมการเข้าใช้บริการร้าน
2) แดชบอร์ดการวิเคราะห์ใบเสร็จการสั่งอาหารและเครื่องดื่มที่พบ

มิติ (Dimensions)

1) มิติใบเสร็จ (Receipt Dimensions)

   มิติใบเสร็จมีลำดับชั้นในการวิเคราะห์ ดังนี้

   - เลขที่ใบเสร็จ (Receipt Id)

2) มิติลูกค้า (Customer Dimensions)

   มิติลูกค้ามีลำดับชั้นในการวิเคราะห์ ดังนี้

- ประเภทลูกค้า (Type of Customer)
- จำนวนลูกค้า (Total of Customer)

3) มิติเวลา (Time Dimensions)

มิติเวลามีลำดับชั้นในการวิเคราะห์ ดังนี้

- ปี (Year)
- เดือน (Month)
- วัน (Day)
- ชื่อวันในสัปดาห์ (Day of Week)
- ระยะเวลารวมที่ใช้บริการ (Spending Time)
- การจัดแสดงดนตรี (Event Day)

4) มิติการชำระเงิน (Payment Dimensions)

มิติการชำระเงินมีลำดับชั้นในการวิเคราะห์ ดังนี้

- วิธีการชำระเงิน (Payment Method)
- ประเภทของบัตรเครดิต (Type of Credit Card)

ค่าวัด (Measures)

ตารางที่ 5: ค่าวัดของระบบวิเคราะห์การสั่งอาหารและเครื่องดื่ม (Food & Beverage Order Analysis System)

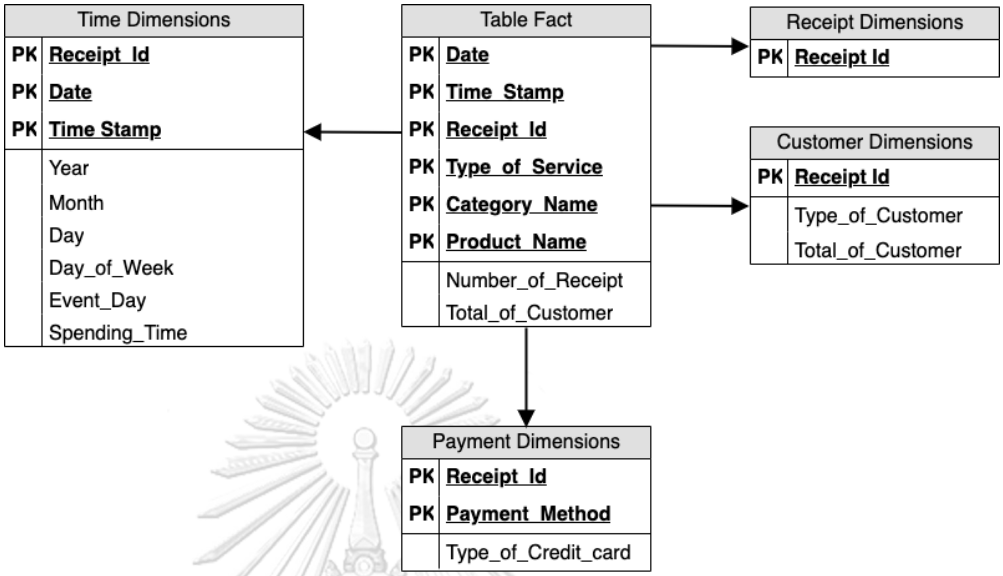
ลำดับ	ค่าวัด	ชื่อภาษาอังกฤษ (หน่วยวัด)	ชื่อภาษาไทย (หน่วยวัด)
1	Number of Receipt	Number of Receipt (Copies)	จำนวนใบเสร็จ (ใบ)
2	Total of Customer	Total of Customer (Person)	จำนวนลูกค้า (คน)

ตัวชี้วัดผลการดำเนินการหลัก (KPIs)

ตารางที่ 6: ตัวชี้วัดผลการดำเนินการหลักของระบบวิเคราะห์การสั่งอาหารและเครื่องดื่ม (Food & Beverage Order Analysis System)

ลำดับ	ตัวชี้วัดผลการดำเนินการหลัก (ชื่อภาษาอังกฤษและภาษาไทย) (หน่วยวัด)	สูตรคำนวณ
1	Percentage of datetime in receipts (%) สัดส่วนของวันหรือเวลาที่ลูกค้าเข้ามาใช้บริการ (%)	$\frac{\text{จำนวนใบเสร็จรวมเฉพาะวันหรือเวลาที่สนใจ}}{\text{จำนวนใบเสร็จรวมทั้งหมด}} \times 100$
2	Average number of customers per event day (Person) จำนวนลูกค้าเฉลี่ยต่อวันที่มีการแสดงดนตรี (คน)	$\frac{\text{จำนวนลูกค้ารวมทั้งหมดเฉพาะวันที่มีการแสดงดนตรี}}{\text{จำนวนวันทั้งหมดที่มีการแสดงดนตรี}}$
3	Average number of customers per non-event day (Person) จำนวนลูกค้าเฉลี่ยต่อวันที่ไม่มีการแสดงดนตรี (คน)	$\frac{\text{จำนวนลูกค้ารวมทั้งหมดเฉพาะวันที่ไม่มีการแสดงดนตรี}}{\text{จำนวนวันทั้งหมดที่ไม่มีการแสดงดนตรี}}$
4	Percentage of payment method in receipts (%) สัดส่วนวิธีการชำระเงินของลูกค้าในใบเสร็จ (%)	$\frac{\text{จำนวนใบเสร็จรวมเฉพาะวิธีการชำระเงินที่สนใจ}}{\text{จำนวนใบเสร็จรวมทั้งหมด}} \times 100$

โมเดลข้อมูลหลายมิติ (Multidimensional Data Model)



รูปที่ 10: โมเดลข้อมูลหลายมิติของระบบการวิเคราะห์การสั่งอาหารและเครื่องดื่ม (Food & Beverage Order Analysis System)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

คำถามของผู้บริหาร ผู้ใช้ และแดชบอร์ดการวิเคราะห์ (Management Questions, Users and Analytics Dashboard)

ตารางที่ 7: คำถามของผู้บริหาร ผู้ใช้ และแดชบอร์ดการวิเคราะห์ของระบบการวิเคราะห์การสั่งอาหารและเครื่องดื่ม (Food & Beverage Order Analysis System)

คำถามของผู้บริหาร (Management Questions)	ผู้ใช้ (Users)	แดชบอร์ดการวิเคราะห์ (Analytics Dashboard)
1) ลูกค้าส่วนใหญ่มักมาช่วงวันใดในสัปดาห์	1) ผู้จัดการร้าน 2) หัวหน้าฝ่ายจัดซื้อ	1) แดชบอร์ดการวิเคราะห์พฤติกรรมการเข้าใช้บริการร้าน
2) ระยะเวลาที่ลูกค้าใช้บริการในร้านเฉลี่ยต่อใบเสร็จเป็นเท่าไหร่	1) ผู้จัดการร้าน 2) หัวหน้าฝ่ายจัดซื้อ	1) แดชบอร์ดการวิเคราะห์พฤติกรรมการเข้าใช้บริการร้าน
3) ช่วงเวลาใดที่มีจำนวนลูกค้าเข้ามาใช้บริการมากที่สุด	1) ผู้จัดการร้าน 2) หัวหน้าฝ่ายจัดซื้อ	1) แดชบอร์ดการวิเคราะห์พฤติกรรมการเข้าใช้บริการร้าน
4) แนวโน้มพฤติกรรมการชำระเงินของลูกค้าในแต่ละช่องทางเป็นอย่างไร	1) ผู้จัดการร้าน 2) หัวหน้าฝ่ายจัดซื้อ	1) แดชบอร์ดการวิเคราะห์ใบเสร็จการสั่งอาหารและเครื่องดื่มที่พบ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

แดชบอร์ดการวิเคราะห์ ค่าวัด ตัวชี้วัดผลการดำเนินงานหลัก และมิติ

(Analytics Dashboard, Measures, KPIs and Dimensions)

ตารางที่ 8: แดชบอร์ดการวิเคราะห์ ค่าวัด ตัวชี้วัดผลการดำเนินงานหลัก และมิติของระบบการวิเคราะห์การสั่งอาหารและเครื่องดื่ม (Food & Beverage Order Analysis System)

แดชบอร์ดการวิเคราะห์ (Analytics Dashboard)	ค่าวัด (Measures)	ตัวชี้วัดผลการดำเนินงานหลัก (KPIs)	มิติ (Dimensions)
1) แดชบอร์ดการวิเคราะห์พฤติกรรมการเข้าใช้บริการร้าน	1) จำนวนใบเสร็จ	1) สัดส่วนของวันหรือเวลาที่ลูกค้าเข้ามาใช้บริการ	1) มิติใบเสร็จ 2) มิติเวลา
	1) จำนวนลูกค้า	1) จำนวนลูกค้าเฉลี่ยต่อวันที่มีการแสดงดนตรี	1) มิติลูกค้า 2) มิติเวลา
	1) จำนวนลูกค้า	1) จำนวนลูกค้าเฉลี่ยต่อวันที่ไม่มีการแสดงดนตรี	1) มิติลูกค้า 2) มิติเวลา
2) แดชบอร์ดการวิเคราะห์ใบเสร็จการสั่งอาหารและเครื่องดื่มที่พบ	1) จำนวนใบเสร็จ	1) สัดส่วนวิธีการชำระเงินของลูกค้าในใบเสร็จ	1) มิติใบเสร็จ 2) มิติการชำระเงิน

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

4.1.2.3 ระบบแบ่งกลุ่มใบเสร็จค่าอาหารและเครื่องดื่ม (Food & Beverage Receipt Segmentation System)

ภาพรวมของระบบ (System Overview)

ระบบนี้จะใช้เทคนิค Unsupervised Learning เพื่อจัดกลุ่มที่ได้จากใบเสร็จค่าอาหารและเครื่องดื่มโดยนำเสนอเป็นรายงานให้กับผู้จัดการร้าน เพื่อทำความเข้าใจพฤติกรรมการสั่งอาหารและเครื่องดื่มของลูกค้า

ผู้ใช้ (Users)

1) ผู้จัดการร้าน
2) หัวหน้าฝ่ายจัดซื้อ

คำถามผู้บริหาร (Management Questions)

1) ใบเสร็จค่าอาหารและเครื่องดื่มของลูกค้าสามารถจำแนกได้กี่กลุ่ม
2) ลักษณะเฉพาะใบเสร็จค่าอาหารและเครื่องดื่มของแต่ละกลุ่มเป็นอย่างไร

เทคนิคการวิเคราะห์ (Analytic Techniques)

1) K-means คือ วิธีการหนึ่งใน Data mining อยู่ในกลุ่มของ Unsupervised Learning คือการเรียนรู้แบบไม่ต้องสอน ไม่มีคำตอบตายตัว โดยการทำงานหลักของ K-means เริ่มจากการกำหนดจุดศูนย์กลางของกลุ่ม (centroid) มาจำนวน k จุดโดยที่ k คือจำนวนกลุ่มที่คาดว่าจะได้จากการแบ่งกลุ่ม จากนั้นคำนวณระยะห่างระหว่างข้อมูลในแต่ละแถวกับ centroid เพื่อใช้ในการจัดกลุ่มข้อมูลที่มีลักษณะใกล้เคียงกันเป็นกลุ่มเดียวกัน โดยระยะห่างนั้นสามารถคำนวณได้โดยใช้ Euclidean distance
2) ใช้หลักการ Elbow method เพื่อเลือกค่า k ที่เหมาะสม หรือการหาค่า Optimal cluster number

ผลการวิเคราะห์ (Analytic Results)

1) จำนวนกลุ่มใบเสร็จค่าอาหารและเครื่องดื่มที่มีลักษณะใกล้เคียงกัน
2) คุณลักษณะของแต่ละกลุ่ม

เครื่องมือที่ใช้ (Tools)

1) Python Programing Language
   - pandas
   - matplotlib

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

- numpy
- seaborn
- scikit-learn

ขั้นตอนการพัฒนา

ขั้นตอนที่ 1 ทำความเข้าใจข้อมูล (Data Understanding)

1. รวบรวมข้อมูลทั้งหมดที่จำเป็นสำหรับนำมาแบ่งกลุ่ม เช่น วันและเวลาระยะเวลาการเข้าใช้บริการร้าน จำนวนลูกค้าในใบเสร็จ ประเภทของลูกค้า ร้อยละมูลค่ายอดขายของประเภทอาหารและเครื่องดื่ม มูลค่ารวมของใบเสร็จ และวิธีการชำระเงิน เป็นต้น
2. ทำ Data Exploration เพื่อทำความเข้าใจกับข้อมูลต่างๆ
3. ออกแบบข้อมูลที่เหมาะสมในการจัดทำ Model ดังตารางที่ 9

ตารางที่ 9: ข้อมูลที่จะนำมาใช้แบ่งกลุ่มใบเสร็จค่าอาหารและเครื่องดื่ม

ชื่อตัวแปร	ชนิดข้อมูล	คำอธิบาย
Day of week	String	วันในสัปดาห์ที่เข้าใช้บริการ
Total of customer	Integer	จำนวนลูกค้าในใบเสร็จ
Type of customer	String	ประเภทของลูกค้า
Entry time	Integer	เวลาที่เข้าใช้บริการร้าน
Spending time	Integer	ระยะเวลาการเข้าใช้บริการรวม
Event day	String	เป็นวันมีการแสดงดนตรีเกิดขึ้นใช่หรือไม่
Percentage of category value	Float	ร้อยละมูลค่ายอดขายของประเภทอาหารและเครื่องดื่ม
Total of sales	Float	มูลค่ารวมของใบเสร็จ
Payment method	String	วิธีการชำระเงิน

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ขั้นตอนที่ 2 เตรียมข้อมูล (Data Preparation)

การเตรียมข้อมูลสำหรับการแบ่งกลุ่มใบเสร็จค่าอาหารและเครื่องดื่ม มีขั้นตอนดังต่อไปนี้

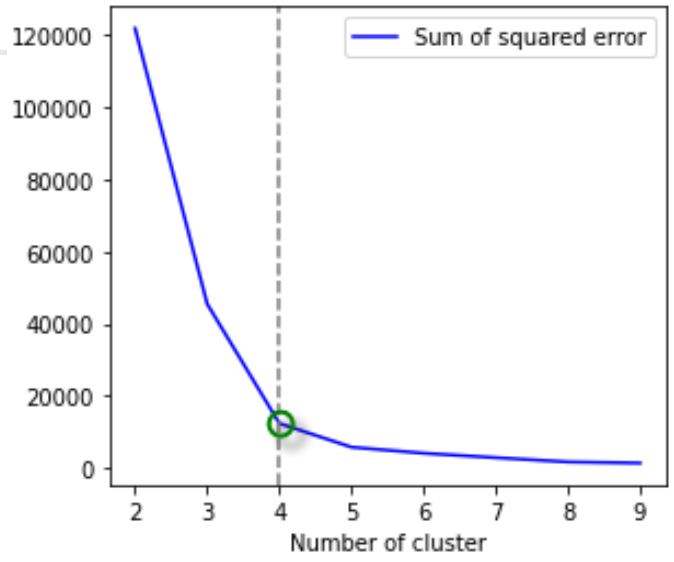
1. Data Integration รวมข้อมูลจากหลายตาราง เพื่อจัดทำให้เป็นตารางใหม่ ที่เหมาะสม
2. Data Cleansing โดยตรวจสอบ และจัดการ Missing Value พร้อมกับตัด Outliner หรือข้อมูลที่ Inconsistent ออก
3. Data Transformation ทำการปรับ Format ของ Feature ต่าง ๆ ให้เหมาะสม รวมถึงการทำ One-Hot encoding สร้างตัวแปร Dummies สำหรับเตรียมเข้าโมเดล

ขั้นตอนที่ 3 วิเคราะห์ข้อมูล (Data Analysis)

1) สร้างโมเดล KMeans Clustering และทดสอบประสิทธิภาพของโมเดลที่ได้ โดยหาค่า K จำนวน Cluster ที่เหมาะสมที่สุด ด้วยวิธี Elbow method

ขั้นตอนที่ 4 ประเมินผลการวิเคราะห์ (Result Evaluation)

1) วัด Performance ในการทำ Clustering ของ K-Mean จากค่าผลรวมของระยะห่างระหว่าง Object กับ Centroid เพื่อการเลือกค่า K ที่เหมาะสม โดยการหาจำนวนกลุ่มที่ดีที่สุดจากค่าผลรวมระยะห่างกำลังสอง ซึ่งลดลงอย่างมีนัยสำคัญ และผลของการคำนวณที่ได้ ก่อนบริเวณที่ความชันของเส้นโค้งจะเริ่มเรียบ (Smooth) คล้ายข้อศอก (Elbow) จะเป็นจุดที่แสดงค่าของจำนวนกลุ่ม (k) ที่เหมาะสมที่สุด



รูปที่ 11: กราฟแสดงตัวอย่างค่าผลรวมระยะห่างกำลังสอง[1-20]

4.1.2.4 ระบบวิเคราะห์กฎความสัมพันธ์ในใบเสร็จค่าอาหารและเครื่องดื่ม (Food & Beverage Receipt Association Analysis System)

ภาพรวมของระบบ (System Overview)

ระบบนี้เป็นการค้นหาความสัมพันธ์ที่เกิดขึ้นร่วมกันบ่อยครั้ง จากพฤติกรรมการสั่งอาหารและเครื่องดื่มจากใบเสร็จที่ผ่านมา โดยใช้เทคนิคการค้นหากฎความสัมพันธ์ และประเมินผลกฎความสัมพันธ์ด้วย Support Confidence และ Lift เพื่อนำกฎความสัมพันธ์ที่ได้ไปประกอบการตัดสินใจในการจัดทำกลยุทธ์ส่งเสริมการขายในอนาคต

ผู้ใช้ (Users)

1) ผู้จัดการร้าน
2) หัวหน้าฝ่ายจัดซื้อ

คำถามผู้บริหาร (Management Questions)

1) เมนูอาหารและเครื่องดื่มใดที่พบความสัมพันธ์อย่างมีนัยสำคัญมากที่สุดตั้งแต่เปิดกิจการ
2) ถ้าจะส่งเสริมการขายในช่วงเวลา Happy hour ควรเลือกชุดเมนูใด
3) ถ้าจะส่งเสริมการขายในวันที่มีการแสดงดนตรี ควรเลือกชุดเมนูใด

เทคนิคการวิเคราะห์ (Analytic Techniques)

1) ใช้กฎความสัมพันธ์ (Association Rule) เพื่อหาความสัมพันธ์ของสินค้าในรายการขาย ดังที่พบในการประยุกต์ในธุรกิจค้าปลีก (retailing business) เช่น ร้านค้าสะดวกซื้อ หรือ ซุปเปอร์มาร์เก็ต เพื่อศึกษาพฤติกรรมการซื้อสินค้าของลูกค้าผ่านการหาความสัมพันธ์ในรายการซื้อของลูกค้า

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ผลการวิเคราะห์ (Analytic Results)

1) ความสัมพันธ์ที่พบในรายการซื้อของลูกค้า

เครื่องมือที่ใช้ (Tools)

1) Python Programing Language
   - pandas
   - matplotlib
   - numpy
   - seaborn
   - mlxtend

ขั้นตอนการพัฒนา

ขั้นตอนที่ 1 ทำความเข้าใจข้อมูล (Data Understanding)

1) รวบรวมข้อมูลทั้งหมดที่จำเป็น เช่น เลขที่ใบเสร็จ, วันในสัปดาห์ที่เข้าใช้บริการ, ช่วงเวลาในการเข้าใช้บริการร้าน, ระยะเวลาการเข้าใช้บริการรวม, เป็นวันมีการแสดงดนตรีเกิดขึ้นใช่หรือไม่ จำนวนลูกค้าในใบเสร็จ ประเภทของลูกค้า ประเภทสินค้า และชื่อสินค้า เป็นต้น
2) ทำ Data Exploration เพื่อทำความเข้าใจกับข้อมูลต่างๆ
3) ออกแบบรูปแบบข้อมูลที่เหมาะสมในการจัดทำ Model ดังตารางที่ 10

ตารางที่ 10: ข้อมูลที่จะนำมาใช้วิเคราะห์กฎความสัมพันธ์ในใบเสร็จค่าอาหารและเครื่องดื่ม

ชื่อตัวแปร	ชนิดข้อมูล	คำอธิบาย
Receipt Id	String	เลขที่ใบเสร็จ
Day of week	String	วันในสัปดาห์ที่เข้าใช้บริการ
Entry time	Integer	เวลาที่เข้าใช้บริการร้าน
Spending time	Integer	ระยะเวลาการเข้าใช้บริการรวม
Event Day	String	เป็นวันมีการแสดงดนตรีเกิดขึ้นใช่หรือไม่
Total of customer	Integer	จำนวนลูกค้าในใบเสร็จ
Type of customer	String	ประเภทของลูกค้า
Category Name	String	ประเภทสินค้า
Product Name	String	ชื่อสินค้า

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ขั้นตอนที่ 2 เตรียมข้อมูล (Data Preparation)

การเตรียมข้อมูลสำหรับการวิเคราะห์กฎความสัมพันธ์ในใบเสร็จค่าอาหารและเครื่องดื่มมีขั้นตอนดังต่อไปนี้

1) Data Integration รวมข้อมูลจากหลายตาราง เพื่อจัดทำให้เป็นตารางใหม่ ที่เหมาะสม
2) Data Cleansing โดยตรวจสอบ และจัดการ Missing Value
3) Data Transformation ทำการปรับ Format ของข้อมูลทั้ง Discrete or Continuous ให้เป็นรูปแบบดังตารางที่ 10

ขั้นตอนที่ 3 วิเคราะห์ข้อมูล (Data Analysis)

1) เริ่มต้นจากการสร้าง itemsets ที่มีจำนวนสมาชิก 1 item ที่เป็นไปได้จาก items ทั้งหมด
2) ทำการคัดเลือก Frequent Itemsets โดยพิจารณาค่า Support ของแต่ละ itemsets เพื่อหา Frequent k-Itemsets (Fk) โดยจะเก็บ itemsets ที่มีค่า Support ถึงเกณฑ์ที่ตั้งไว้ (Support Threshold) เท่านั้น
3) นำ Frequent Itemsets ทั้งหมดที่หามาได้จาก Apriori Algorithm ไปเข้าสู่กระบวนการสร้างกฎความสัมพันธ์ หรือ Rules Generation ซึ่งจะมีการพิจารณาค่า Confidence หรือ Lift ในการคัดเลือกกฎความสัมพันธ์ที่เหมาะสม

ขั้นตอนที่ 4 ประเมินผลการวิเคราะห์ (Result Evaluation)

1) Support A => B เป็นตัววัดประสิทธิภาพสำหรับ itemset โดยเป็นตัวเลขที่แสดงสัดส่วนของจำนวนใบเสร็จการซื้อขายที่มี itemset ต่อจำนวนใบเสร็จการซื้อขายทั้งหมด หรือกล่าวได้ว่า “มีการซื้อขายที่มี itemset อยู่ด้วยเป็นสัดส่วนเท่าไหร่ในการซื้อขายทั้งหมด” หากเขียนในรูปแบบสมการทางคณิตศาสตร์จะได้ว่า

$$\frac{\text{จำนวนใบเสร็จการซื้อขายที่มี itemset}}{\text{จำนวนใบเสร็จการซื้อขายทั้งหมด}}$$

2) Confidence เป็นตัววัดประสิทธิภาพสำหรับ association rule โดยเป็นตัวเลขที่แสดงความน่าจะเป็นที่กลุ่มของสินค้า B จะถูกซื้อขายอยู่ในใบเสร็จเดียวกัน หลังจากที่กลุ่มของสินค้า A ถูกซื้อขายอยู่ในใบเสร็จเดียวกันแล้ว จากกฎของความสัมพันธ์ A => B ค่าของ confidence ถูกเขียนด้วยสมการทางคณิตศาสตร์ คือ

$$\frac{\text{จำนวนใบเสร็จการซื้อขายที่มีทั้ง A และ B}}{\text{จำนวนใบเสร็จการซื้อขายที่มี A}}$$

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

3) Lift A => B เป็นตัววัดประสิทธิภาพสำหรับ association rule ที่ทำการเปรียบเทียบ ความน่าจะเป็นที่จะพบกลุ่มของสินค้า B ในใบเสร็จที่มีการซื้อขายกลุ่มของสินค้า A กับ ความน่าจะเป็นที่จะพบกลุ่มของสินค้า B ในใบเสร็จทั้งหมด จากกฎของความสัมพันธ์ A => B ค่าของ Lift ถูกเขียนด้วยสมการทางคณิตศาสตร์

$$\frac{\text{Confident A => B}}{\text{Support B}}$$

จุฬาลงกรณ์มห
CHULALONGKORN

4.1.2.5 ระบบพยากรณ์ยอดขาย (Sales Forecasting System)

ภาพรวมของระบบ (System Overview)

ระบบนี้เป็นการพัฒนาตัวแบบพยากรณ์ยอดขายธุรกิจ โดยใช้สมการการถดถอยเชิงพหุคูณ (Multiple Linear Regression Analysis) ในการช่วยพัฒนาและเลือกตัวแบบที่ดีที่สุดจากตัวแปรต่าง ๆ ที่ได้จากใบเสร็จค่าอาหารและเครื่องดื่ม พร้อมนำเสนอเป็นรายงานให้กับผู้จัดการร้าน ในการคาดการณ์ยอดขายธุรกิจในอนาคต

ผู้ใช้ (Users)

1) ผู้จัดการร้าน
2) หัวหน้าฝ่ายจัดซื้อ

คำถามผู้บริหาร (Management Questions)

1) ยอดขายที่จะเกิดขึ้นในวันศุกร์-อาทิตย์ของสัปดาห์หน้าเป็นอย่างไร
2) ยอดขายที่จะเกิดขึ้นในเดือนหน้าทั้งเดือนเป็นอย่างไร

เทคนิคการวิเคราะห์ (Analytic Techniques)

1) นำข้อมูลที่มีอยู่ทั้งหมด มาค้นหาตัวแปรต้นที่ส่งผลต่อตัวแปรตาม (มูลค่าในใบเสร็จ หรือ Sales) โดยใช้ทฤษฎีความสัมพันธ์ของ Multilinear Correlation ที่เข้ามาช่วยในการคัดเลือกตัวแปรต้นดังกล่าว
2) นำตัวแปรต้นที่ได้มาสร้างเป็นสมการถดถอยเชิงพหุคูณ Multiple Linear Regression จากนั้นประเมินความคลาดเคลื่อนของสมการ
3) ปรับปรุงสมการ โดยการปรับจูนพารามิเตอร์เพื่อค้นหาสมการเส้นตรงที่ให้ค่าความแม่นยำสูงสุด
4) ทำนายยอดขายโดยการใช้สมการสมการถดถอยเชิงพหุคูณ Multiple Linear Regression ที่เลือก

ผลการวิเคราะห์ (Analytic Results)

1) ตัวแบบทำนายยอดขาย

เครื่องมือที่ใช้ (Tools)

1) Python Programing Language

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

- pandas
- matplotlib
- numpy
- seaborn
- scikit-learn

ขั้นตอนการพัฒนา

ขั้นตอนที่ 1 ทำความเข้าใจข้อมูล (Data Understanding)

1) รวบรวมข้อมูลทั้งหมดที่จำเป็นสำหรับนำมาสร้างโมเดล เช่น ปี เดือน วัน วันในสัปดาห์ที่เข้าใช้บริการ เป็นวันมีการแสดงดนตรีเกิดขึ้นใช่หรือไม่ เป็นวันหยุดพิเศษใช่หรือไม่ จำนวนวันหยุดพิเศษยาวต่อเนื่องเป็นต้น
2) ทำ Data Exploration เพื่อทำความเข้าใจกับข้อมูลต่างๆ
3) ออกแบบข้อมูลที่เหมาะสมในการจัดทำ Model ดังตารางที่ 11

ตารางที่ 11: ข้อมูลที่จะนำมาใช้สร้างสมการถดถอยเชิงพหุคูณ

ชื่อตัวแปร	ชนิดข้อมูล	คำอธิบาย
Year	Integer	ปี
Month	Integer	เดือน
Day	Integer	วัน
Day of Week	String	วันในสัปดาห์ที่เข้าใช้บริการ
Event Day	String	เป็นวันมีการแสดงดนตรีเกิดขึ้นใช่หรือไม่
Special Day Off	String	เป็นวันหยุดพิเศษใช่หรือไม่
Continue Day Off	Integer	จำนวนวันหยุดติดกันถัดไป
Sales	Float	ยอดค่าใช้จ่ายในใบเสร็จ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ขั้นตอนที่ 2 เตรียมข้อมูล (Data Preparation)

การเตรียมข้อมูลสำหรับระบบพยากรณ์ยอดขาย มีขั้นตอนดังต่อไปนี้

1) Data Integration รวมข้อมูลจากหลายตาราง เพื่อจัดทำให้เป็นตารางใหม่ ที่เหมาะสม
2) Data Cleansing โดยตรวจสอบ และจัดการ Missing Value พร้อมกับตัด Outliner หรือข้อมูลที่ Inconsistent ออก
3) Data Transformation ทำการปรับ Format ของข้อมูลทั้ง Discrete or Continuous รวมถึงการทำ One-Hot encoding สร้างตัวแปร Dummies สำหรับเตรียมเข้าโมเดล
4) Data Split เพื่อแบ่งข้อมูลออกเป็น Train, Validation และ Test Set

ขั้นตอนที่ 3 วิเคราะห์ข้อมูล (Data Analysis)

1) การวิเคราะห์การถดถอยเชิงพหุคูณ (Multiple Linear Regression Analysis)

ศึกษาความสัมพันธ์ระหว่างตัวแปรอิสระหลายตัวกับตัวแปรตาม 1 ตัว เพื่อศึกษาว่ามีตัวแปรอิสระตัวใดบ้างที่ร่วมกันทำนายหรือพยากรณ์ หรืออธิบายการผันแปรของตัวแปรตามได้โดยเขียนความสัมพันธ์ในรูปแบบของสมการได้ดังนี้

สมการในรูปของประชากร $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_k X_k + \varepsilon$

สมการในรูปของตัวอย่าง $y = b_0 + b_1 x_1 + b_2 x_2 + ... + b_k x_k + e$

สมการทำนายผล (สมการพยากรณ์) $y\ hat = b_0 + b_1 x_1 + b_2 x_2 + ... + b_k x_k$

โดยสัญลักษณ์ที่ใช้มีความหมายดังนี้

Xi คือ ค่าของตัวแปรอิสระแต่ละตัว (จะใช้สัญลักษณ์ xi สำหรับค่าที่ได้จากตัวอย่างและ สำหรับค่าประมาณหรือตัวทำนาย)

Y คือ ค่าของตัวแปรตาม (จะใช้สัญลักษณ์ y สำหรับค่าที่ได้จากตัวอย่าง และใช้ค่า y hat สำหรับค่าประมาณหรือตัวทำนาย)

k คือ จำนวนตัวแปรอิสระในสมการถดถอย

β_0 คือ ค่าคงที่ (Constant) ของสมการถดถอย (จะใช้สัญลักษณ์ b0 สาหรับค่าที่ได้จาก ตัวอย่าง และ สำหรับค่าประมาณหรือตัวทำนาย) โดยที่ β_0 หรือ b0 จะเป็นจุดตัด (Intercept) แกน y ของสมการ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

β_i คือ ค่าสัมประสิทธิ์การถดถอย (Regression Coefficient) ของตัวแปรอิสระ Xi แต่ละตัว (จะใช้สัญลักษณ์ bi สำหรับค่าที่ได้จากตัวอย่าง และ สำหรับค่าประมาณหรือตัวทำนาย) โดยที่ค่า β_i หรือ bi จะแสดงอัตราการเปลี่ยนแปลงของค่า xi ต่อค่า y ดังนี้ คือ ถ้าค่า xi เปลี่ยนไป 1 หน่วย จะทำ ให้ค่า y เปลี่ยนไป bi หน่วย

ε คือ ค่าความคลาดเคลื่อน (Error or Residual) ระหว่างค่า Y และค่า y hat (จะใช้สัญลักษณ์ e สำหรับค่าที่ได้จากตัวอย่าง)

2) ข้อตกลงเบื้องต้น (Assumption) ในการวิเคราะห์การถดถอย

2.1) ตัวแปรอิสระ X และตัวแปรตาม Y ต้องเป็นตัวแปรเชิงปริมาณ (Quantitative Variable) หรือ ตัวแปรต่อเนื่อง (Continuous Variable) หรือ มีระดับการวัดเป็น Interval หรือ Ratio Scale

2.2) ในกรณีที่ตัวแปรอิสระ X บางตัวมีระดับการวัดเป็น Nominal หรือ Ordinal Scale จะต้องแปลงข้อมูลให้ เป็นตัวแปรหุ่น (Dummy Variable) คือ มีค่า 0 กับ 1 ก่อนจึงจะนำไปวิเคราะห์ และตัวแปรหุ่นไม่ควรจะมีหลายตัว เพราะจะทำให้เกิดความคลาดเคลื่อนมากขึ้น

2.2) ตัวแปรอิสระแต่ละตัวมีความสัมพันธ์เชิงเส้นตรงตัวแปรตาม

2.3) ตัวแปรอิสระไม่ควรมีความสัมพันธ์กันหรือเป็นอิสระต่อกัน (ค่าสหสัมพันธ์ไม่ควรเกิน 0.7) ในกรณีการวิเคราะห์ถดถอยแบบพหุคูณ เพราะจะทำให้เกิด Multicollinearity คือ การที่ตัวแปรอิสระมีความสัมพันธ์กันมาก ซึ่งจะมีผลกระทบทำให้ค่าสัมประสิทธิ์การตัดสินใจ (R Square) สูงเกินความเป็นจริง

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2.4) การแจกแจงของตัวแปรตามเป็นแบบโค้งปกติ (Normal Distribution) ที่ทุกค่าของ X

2.5) ค่าของ Y มีความแปรปรวนเท่ากันทุกค่าของ X

2.6) ความแปรปรวนของค่าความคลาดเคลื่อนจากการพยากรณ์ (Residual) ที่ทุกจุดบนเส้นถดถอยมีค่าเท่ากัน

ขั้นตอนที่ 4 ประเมินผลการวิเคราะห์ (Result Evaluation)

1) R Square

R Square คือ ค่าสัมประสิทธิ์การตัดสินใจพหุคูณ (Coefficient of Multiple Determination) เป็นค่าที่แสดงอิทธิพลของตัวแปรอิสระทั้งหมดในสมการที่มีต่อตัวแปรตาม หรือ ตัวแปรอิสระทั้งหมดที่อยู่ในสมการสามารถอธิบายการผันแปรของตัวแปรตามได้มากน้อยเพียงใด โดยอธิบายว่าตัวแปรอิสระทั้งหมดมีอิทธิพลต่อตัวแปรตามอยู่ร้อยละ R Square ส่วนที่เหลือจะเป็นอิทธิพลจากตัวแปรอื่นที่ไม่ได้นำมาวิเคราะห์ หรือตัวแปรอิสระทั้งหมดสามารถทำนายตัวแปรตามได้ร้อยละ R Square

2) Adjusted R Square

ข้อมูลทุกตัวมีค่าความผันแปรที่ไม่สามารถอธิบายรวมอยู่ด้วยเสมอ ซึ่งค่า R-squared ที่ได้จากข้างต้น ไม่ได้บอกว่าค่าสูงสุดควรเป็นเท่าใด ในการปรับแต่งค่า R-squared ให้มีค่าเพิ่มขึ้นอาจเป็นเพียงการเพิ่มตัวแปรทำนายเพื่อให้อธิบายในสิ่งที่ไม่สามารถอธิบายจนทำให้เกิดผลที่ผิดพลาดได้ ดังนั้น ค่า adjusted R-squared จะช่วยทำให้สามารถประเมินได้ว่าตัวแปรทำนายกี่ตัวที่ควรจะอยู่ในตัวแบบ

3) Cost Function

Cost function ของ Linear regression คือค่าความผิดพลาดของค่าจริง และค่าที่เกิดขึ้นจากการทำนายของโมเดล ซึ่งสามารถหาได้จากความแตกต่างที่เกิดขึ้น (diff)

$$e_i = y_i - \hat{y}_i$$

เมื่อมองจากสมการแล้ว ค่า Error สามารถเกิดขึ้นได้ทั้งทางบวกและทางลบ ดังนั้นเพื่อการคำนวณผลรวมของ Error สุดท้ายจึงใช้ค่าผลรวมของเลขยกกำลัง หรือที่เรียกว่า Residual Sum of Squares (RSS)

$$RSS = \sum_{i}^{n} (y_i - \hat{y}_i)^2$$

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

เนื่องจากค่า RSS คือค่าผลรวม และมีแนวโน้มที่จะสูงขึ้นเรื่อย ๆ ดังนั้นเพื่อการวัดค่า Error ที่เกิดขึ้นของสมการเส้นตรง จึงใช้เทคนิคการคำนวณ Cost function แบบ MSE

$$MSE = \frac{1}{n}\sum_{i}^{n}(y_i - \hat{y}_i)^2$$

4.2 การออกแบบระบบ

ในการพัฒนา “คลังข้อมูล ธุรกิจอัจฉริยะ และการวิเคราะห์ข้อมูลขั้นสูงของธุรกิจร้านอาหารและเครื่องดื่ม” มีการออกแบบระบบโดยสามารถแบ่งออกเป็นส่วนต่าง ๆ ได้แก่ การออกแบบผลลัพธ์ การออกแบบข้อมูลนำเข้า การออกแบบส่วนต่อประสานผู้ใช้ และการออกแบบการรักษาความปลอดภัย

4.2.1 การออกแบบผลลัพธ์ (Output Design)

การออกแบบผลลัพธ์ (Output Design) สำหรับการพัฒนา“คลังข้อมูล ธุรกิจอัจฉริยะ และการวิเคราะห์ข้อมูลขั้นสูงของธุรกิจร้านอาหารและเครื่องดื่ม” จะนำเสนอในรูปแบบของตารางและกราฟประเภท ต่าง ๆ ซึ่งจะขึ้นอยู่กับวัตถุประสงค์ของระบบงานและข้อมูลที่ต้องการนำเสนอหรือเปรียบเทียบ โดย การออกแบบจึงเน้นรูปแบบที่เข้าใจง่าย ตรงตามวัตถุประสงค์ในการวิเคราะห์ โดยแบ่งผลลัพธ์เป็น 9 ประเภท ได้แก่

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

1) รายงานในรูปแบบแผนภูมิฮีทแมพ (Heat Map Report)

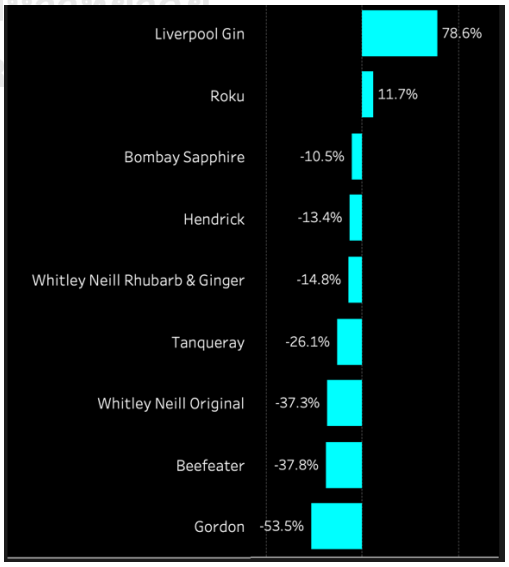
เหมาะสำหรับการแสดงผลข้อมูลตัวแปร 2 มิติ โดยสีเข้มหมายถึงสัดส่วนของข้อมูลที่มีค่ามากที่สุด และค่อย ๆ ไล่ระดับสีตามปริมาณที่ลดลง เช่น รายงานจำนวนนับใบเสร็จช่วงเวลาที่ขายดีที่สุดในแต่ละวันในสัปดาห์ ดังรูปที่ 12

Entry Time	Sun	Mon	Tue	Wed	Thu	Fri	Sat
Null	30	4		1	10	10	49
11 AM	1						
16 PM							3
17 PM	3			1		4	1
18 PM	6		1		7	1	9
19 PM	20		6	6	10	20	32
20 PM	49	3	12	23	42	36	98
21 PM	44	12	5	28	51	69	100
22 PM	27	20	2	22	77	43	67
23 PM	31	19	1	19	28	43	44
24 PM	21	4		3	5	8	47

รูปที่ 12: รายงานจำนวนนับใบเสร็จช่วงเวลาที่ขายดีที่สุดในแต่ละวันในสัปดาห์

2) รายงานในรูปแบบแผนภูมิแท่ง (Bar Chart)

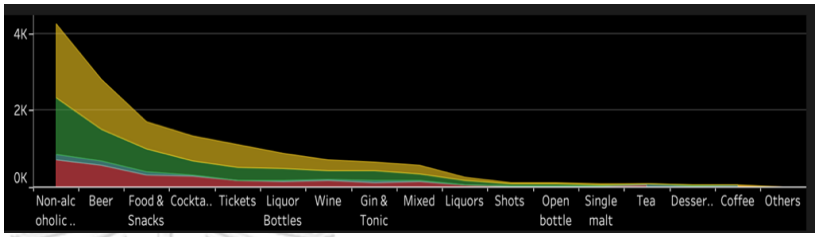
เหมาะสำหรับการแสดงผลข้อมูลตัวแปร 1 มิติ และสามารถลำดับข้อมูลเพื่อให้ง่ายต่อการวิเคราะห์ผล เช่น รายงานการแนวโน้มการเติบโตยอดของสินค้า ดังรูปที่ 13



รูปที่ 13: รายงานการแนวโน้มการเติบโตยอดของสินค้า

3) รายงานในรูปแบบแผนภูมิพื้นที่ (Area Chart)

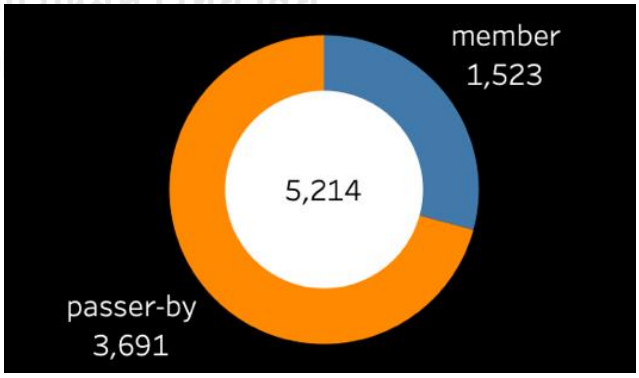
เหมาะสำหรับการแสดงผลข้อมูลตัวแปร 1 มิติ โดยสามารถแสดงผลตัวแปรในแกน x ให้มีความต่อเนื่องกันตามสัดส่วนปริมาณได้ พร้อมทั้งลำดับค่าตัวแปรในแกน y จากมากไปน้อย เช่น รายงานจำนวนใบเสร็จแยกตามรายการอาหารและเครื่องดื่มของแต่ละไตรมาส ดังรูปที่ 14



รูปที่ 14: รายงานจำนวนใบเสร็จแยกตามรายการอาหารและเครื่องดื่มของแต่ละไตรมาส

4) รายงานในรูปแบบแผนภูมิโดนัท (Donut Chart)

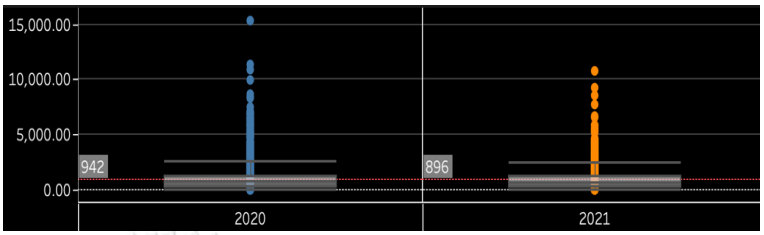
เหมาะสำหรับการแสดงผลข้อมูลตัวแปร 1 มิติ สามารถแสดงสัดส่วนปริมาณของแต่ละตัวแปรได้ โดยข้อดีของกราฟโดนัท คือ บริเวณรูตรงกลางจะแสดงถึงผลรวมข้อมูลเช่น รายงานจำนวนนับใบเสร็จแยกตามประเภทของลูกค้า ดังรูปที่ 15



รูปที่ 15: รายงานจำนวนนับใบเสร็จแยกตามประเภทของลูกค้า

5) รายงานในรูปแบบแผนภูมิบ็อกซ์ (Box plot)

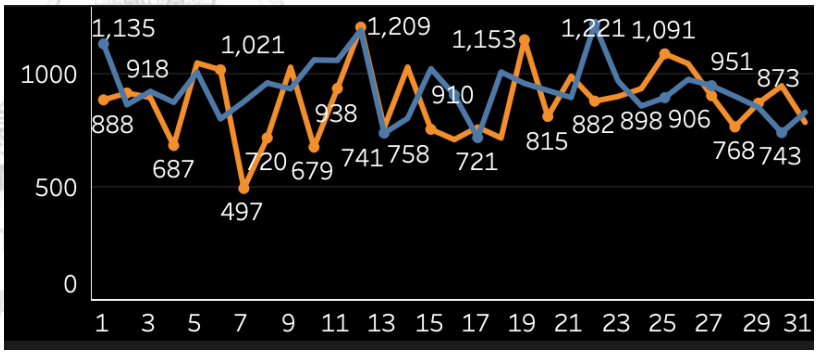
เหมาะสำหรับการแสดงผลข้อมูลตัวแปร 1 มิติ และสามารถแสดงสัดส่วนการกระจายตัวของตัวแปรที่ต้องการได้ด้วย เช่น รายงานยอดขายเฉลี่ยต่อใบเสร็จแยกตามรายปี ดังรูปที่ 16



รูปที่ 16: รายงานยอดขายเฉลี่ยต่อใบเสร็จแยกตามรายปี

6) รายงานในรูปแบบแผนภูมิเส้น (Line Chart)

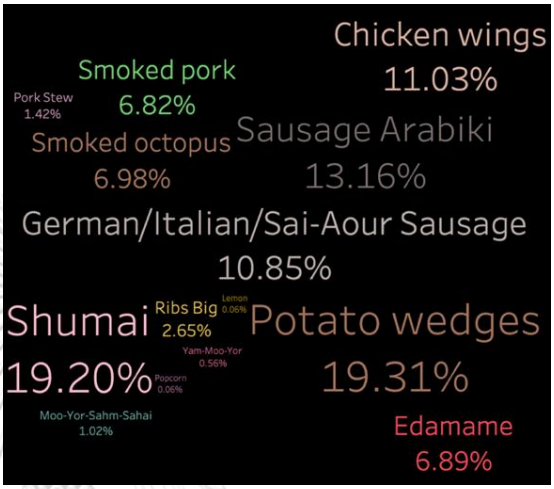
เหมาะสำหรับการแสดงผลข้อมูลตัวแปร 1 มิติ นิยมใช้ตัวแปรในแกน x ที่เกี่ยวข้องกับช่วงเวลาเพื่อแสดงผลแนวโน้มตัวแปร y ได้อย่างต่อเนื่อง เช่น รายงานยอดขายเฉลี่ยต่อใบเสร็จแยกตามรายวัน ดังรูปที่ 17



รูปที่ 17: รายงานยอดขายเฉลี่ยต่อใบเสร็จแยกตามรายวัน

7) รายงานในรูปแบบแผนภูมิกลุ่มคำ (Word Cloud)

เหมาะสำหรับการแสดงผลข้อมูลตัวแปร 1 มิติ โดยแสดงผลของกลุ่มข้อความที่พบเจอมากที่สุดตามสัดส่วนปริมาณของข้อมูลในแต่ละกลุ่มข้อความนั้น ๆ เช่น รายงานสัดส่วนเปอร์เซ็นต์ยอดขายของสินค้าต่าง ๆ ดังรูปที่ 18

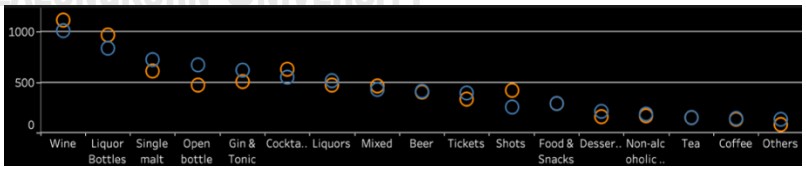


รูปที่ 18: รายงานสัดส่วนเปอร์เซ็นต์ยอดขายของสินค้าต่าง ๆ

8) รายงานในรูปแบบแผนภูมิเซอเคิลวิว (Circle View)

เหมาะสำหรับการแสดงผลข้อมูลตัวแปร 1 มิติ โดยแสดงปริมานสัดส่วนข้อมูลแบบจุดลงบนกราฟ และสามารถเรียงลำดับข้อมูลจากมากไปน้อย เช่น รายงานยอดขายเฉลี่ยต่อใบเสร็จแยกตามรายการอาหารและเครื่องดื่ม ดังรูปที่ 19

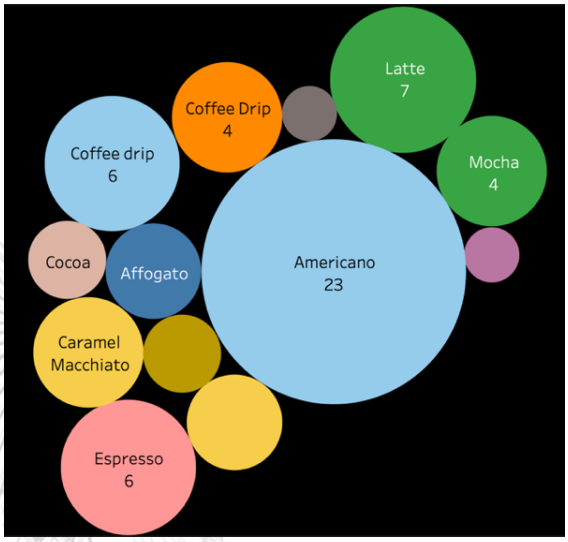
จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY



รูปที่ 19: รายงานยอดขายเฉลี่ยต่อใบเสร็จแยกตามรายการอาหารและเครื่องดื่ม

9) รายงานในรูปแบบแผนภูมิแพกเกจบับเบิล (Packed Bubble)

เหมาะสำหรับการแสดงผลข้อมูลตัวแปร 1 มิติ สามารถแสดงสัดส่วนของข้อมูลตามขนาดของบับเบิลได้ เช่น รายงานจำนวนนับใบเสร็จแยกตามชื่อเมนูในรายการอาหารและเครื่องดื่ม ดังรูปที่ 20



รูปที่ 20: รายงานจำนวนนับใบเสร็จแยกตามชื่อเมนูในรายการอาหารและเครื่องดื่ม

4.2.2 การออกแบบข้อมูลนำเข้า (Input Design)

การนำเข้าข้อมูลสำหรับโครงการ “คลังข้อมูล ธุรกิจอัจฉริยะ และการวิเคราะห์ข้อมูลขั้นสูงของธุรกิจร้านอาหารและเครื่องดื่ม” มีการนำเข้าข้อมูลสู่คลังข้อมูลแบบระบบ Manual โดย ขั้นตอนในการนำเข้าจะ แบ่งเป็น 3 ส่วน ได้แก่

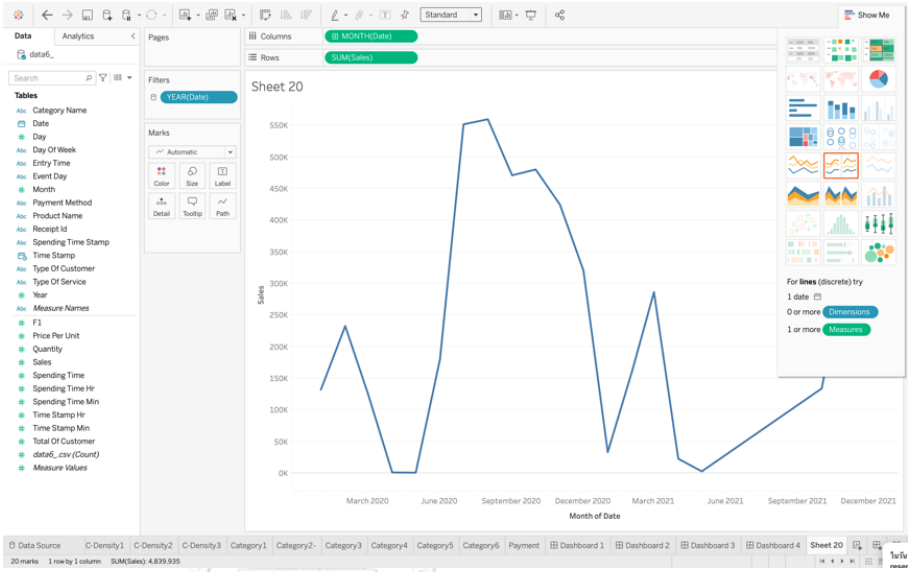
1) การรวบรวมข้อมูลจากหน่วยงานต่างๆ ในองค์กรในรูปแบบของ Excel เพื่อทำการแปลงข้อมูลให้อยู่ในรูปแบบมาตรฐานเดียวกันทั้งระบบการจัดการฐานข้อมูล
2) การเตรียมข้อมูล โดยการตรวจสอบความถูกต้องของข้อมูลก่อนนำเข้า รวมไปถึง การตัดข้อมูลส่วนเกินออก เพื่อให้ข้อมูลมีความเกี่ยวข้องและถูกต้องแม่นยำก่อน การนำเข้าสู่คลังข้อมูล
3) นำข้อมูลจาก Excel เข้าสู่ฐานข้อมูล Google BigQuery
4) เชื่อมต่อฐานข้อมูล Google BigQuery เข้ากับ Tableau Desktop แล้วเชื่อมโยงความสัมพันธ์ของข้อมูลตามที่ออกแบบไว้และจัดทำรายงานต่าง ๆ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

4.2.3 การออกแบบส่วนติดต่อผู้ใช้ (User Interface Design)

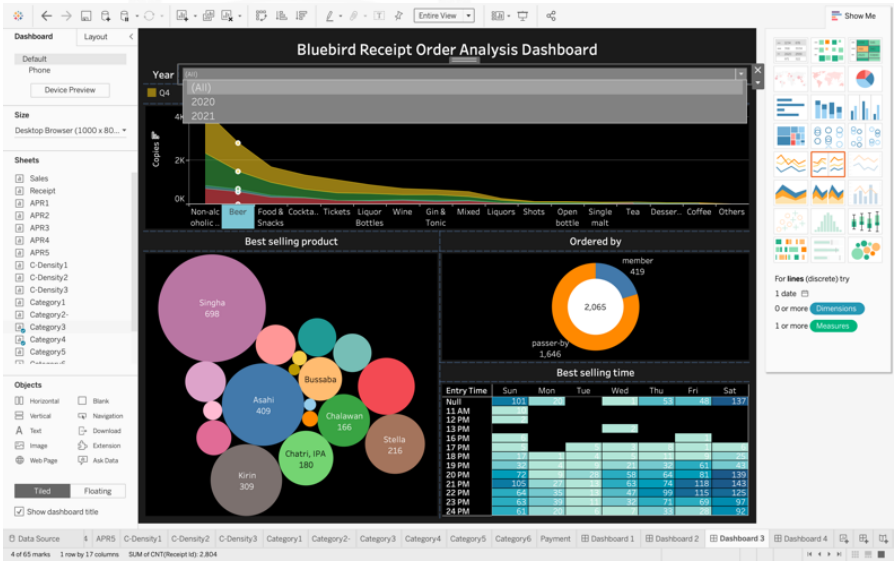
การออกแบบส่วนติดต่อผู้ใช้ (User Interface Design) จะนำเสนอบนโปรแกรม Tableau Desktopโดยแบ่งออกเป็นส่วนต่าง ๆ ได้ดังนี้

1) หน้าจอการสร้างรายงาน เป็นหน้าจอแสดงผลการสร้างกราฟที่ต้องการ ดังรูปที่ 21



รูปที่ 21: หน้าจอการสร้างรายงาน

2) หน้าจอ Dashboard เป็นหน้าจอแสดงผลรวบรวมรายงานต่างๆไว้หน้าจอเดียว โดยผู้ใช้สามารถเลือกปรับมุมมองได้ตามมิติที่สนใจได้ ดังรูปที่ 22



รูปที่ 22: หน้าจอ Dashboard

4.2.4 การออกแบบการรักษาความปลอดภัย

การพัฒนาระบบคลังข้อมูลเพื่อใช้ในการตัดสินใจการดำเนินธุรกิจนั้นเป็นการนำข้อมูลที่มีความสำคัญต่อการวางแผนกลยุทธ์ขององค์กรมาใช้ จึงควรมีการกำหนดสิทธิ์ในการเข้าถึงข้อมูลสำหรับผู้ใช้งานตามหน้าที่ความรับผิดชอบที่เหมาะสม ซึ่งการออกแบบระบบรักษาความปลอดภัย จะกำหนดสิทธิ์การเข้าถึงตามความจำเป็นในการใช้งานความเกี่ยวข้องกับหน้าที่ความรับผิดชอบ และการเข้าถึงข้อมูล โดยจะแบ่งกลุ่มกำหนดสิทธิ์สำหรับผู้ใช้งานแต่งต่างกันไป มีรายละเอียดของสิทธิ์ในการเข้าใช้งานของเจ้าหน้าที่และผู้บริหารของบริษัทต้นแบบ ถึงแม้ว่าโครงการนี้พัฒนาด้วย Tableau Desktop Version 2021.3 ที่ยังไม่สามารถจำกัดสิทธิ์ได้ แต่ผู้พัฒนาได้มีการออกแบบระบบรักษาความปลอดภัยไว้ได้หากการใช้งานเป็นในรูปแบบของ Server ดังนี้

ตารางที่ 12: สิทธิ์ในการเข้าถึงระบบ

สิทธิในการเข้าถึงระบบ	ผู้จัดการร้าน	หัวหน้าฝ่ายจัดซื้อ	พนักงานบริการ	พนักงานครัว	แคชเชียร์
ระบบวิเคราะห์ยอดขาย	✓	✓			
ระบบการวิเคราะห์การสั่งอาหารและเครื่องดื่ม	✓	✓			
ระบบแบ่งกลุ่มใบเสร็จค่าอาหารและเครื่องดื่ม	✓	✓			
ระบบวิเคราะห์กฎความสัมพันธ์ในใบเสร็จค่าอาหารและเครื่องดื่ม	✓	✓			
ระบบพยากรณ์ยอดขาย	✓	✓			

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

4.3 การติดตั้งและพัฒนาระบบ

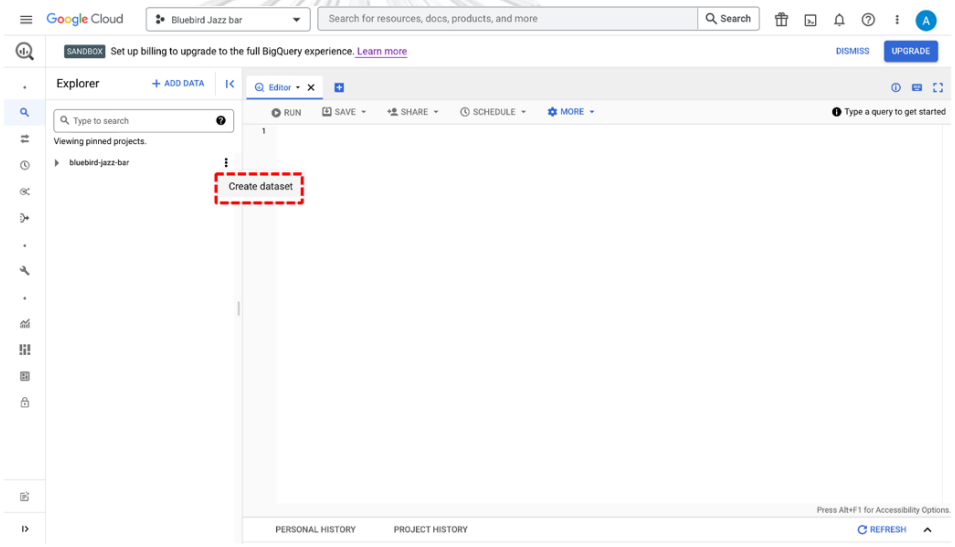
สำหรับการติดตั้งและพัฒนาระบบ “คลังข้อมูล ธุรกิจอัจฉริยะ และการวิเคราะห์ข้อมูลขั้นสูงของธุรกิจร้านอาหารและเครื่องดื่ม” มีขั้นตอนดังนี้

1) การติดตั้งซอฟต์แวร์ที่ใช้งานในระบบ โปรแกรมที่นำมาใช้งานของโครงการนี้ประกอบด้วย Google BigQuery และ Tableau Desktop ซึ่งทำงานบนระบบปฏิบัติการ mac OS

2) การจัดการและนำข้อมูลเข้าสู่ฐานข้อมูล

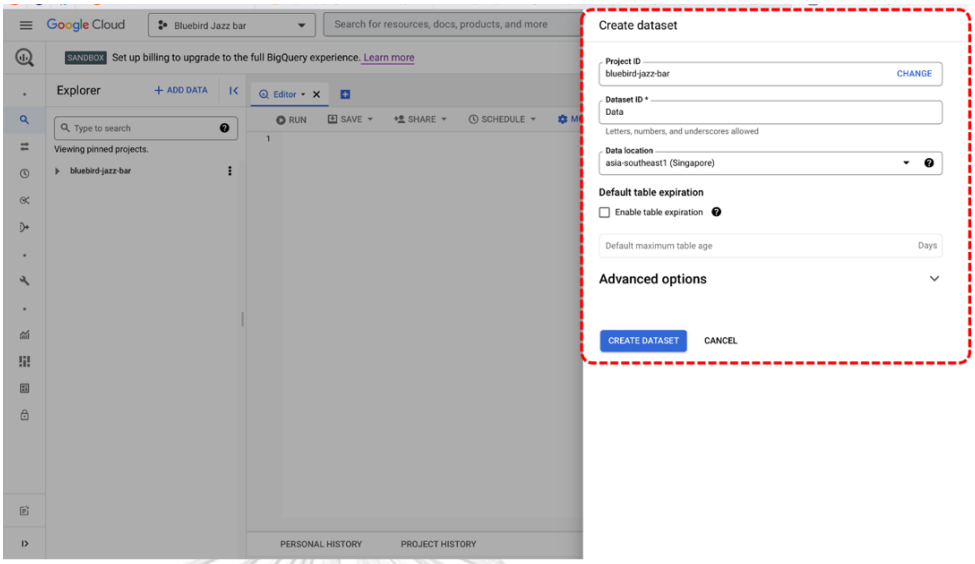
ผู้ดูแลระบบจะต้องดำเนินการ Sign in Google BigQuery เพื่อใช้เป็นเครื่องมือสำหรับนำเข้าข้อมูลเข้าสู่ฐานข้อมูล ซึ่งการนำเข้าข้อมูลเข้าสู่ฐานข้อมูลมีขั้นตอนดังต่อไปนี้

1) สร้างฐานข้อมูลใหม่ โดยคลิกขวาที่จุด 3 จุด เลือก Create dataset ดังรูปที่ 23



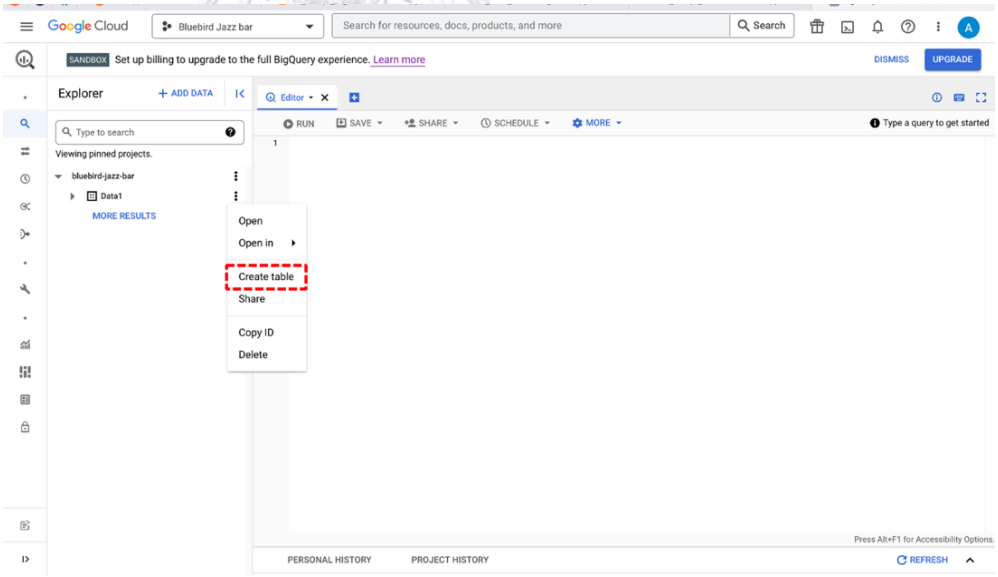
รูปที่ 23: การสร้างฐานข้อมูลใหม่

2) ตั้งชื่อ Dataset ID ที่ต้องการ และกด CREATE DATASET ดังรูปที่ 24



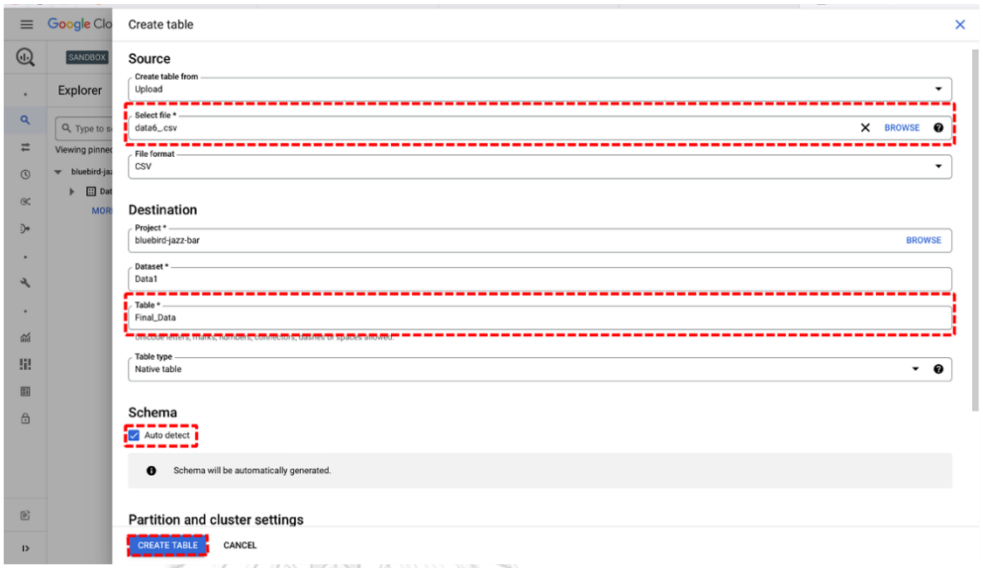
รูปที่ 24: หน้าจอการสร้างฐานข้อมูลใหม่

3) นำข้อมูลลงฐานข้อมูลโดยคลิกขวา จุด 3 จุด ที่ฐานข้อมูลที่สร้างจากข้อที่ 2 จากนั้นเลือก Create table ดังรูปที่ 25



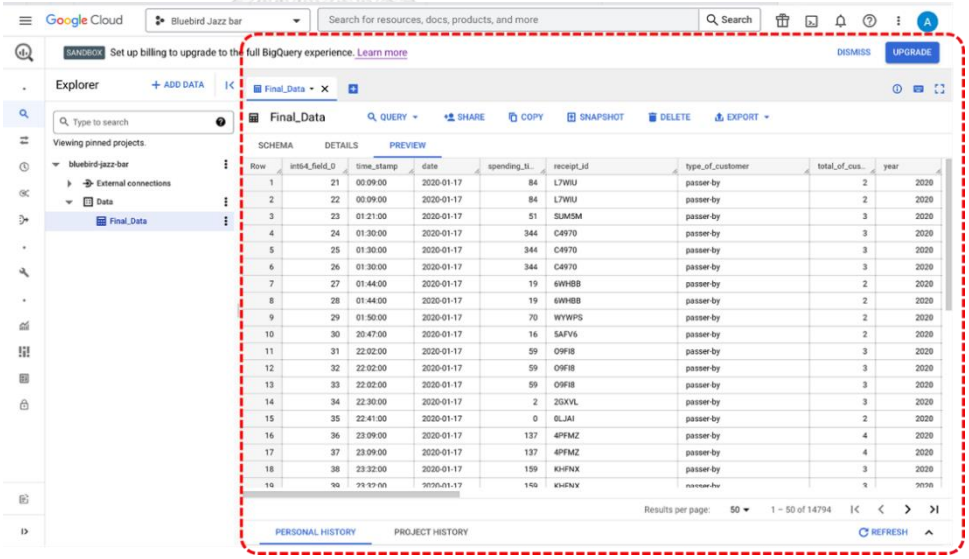
รูปที่ 25: การนำเข้าข้อมูลลงในฐานข้อมูล

4) โปรแกรมจะปรากฏหน้าต่าง Create table ขึ้นมาดังรูป 26 จากนั้น Browse เลือกไฟล์ Excel ที่ต้องการ พร้อมตั้งชื่อ Table และ เลือกโหมด Auto detect ที่หัวข้อ Schema



รูปที่ 26: หน้าจอการ Upload Data เพื่อสร้าง Table ในฐานข้อมูล

5) เมื่อกดเลือก CREATE TABLE จากข้อที่ 4 จะได้ผลลัพธ์ขึ้นมาดังรูป 27

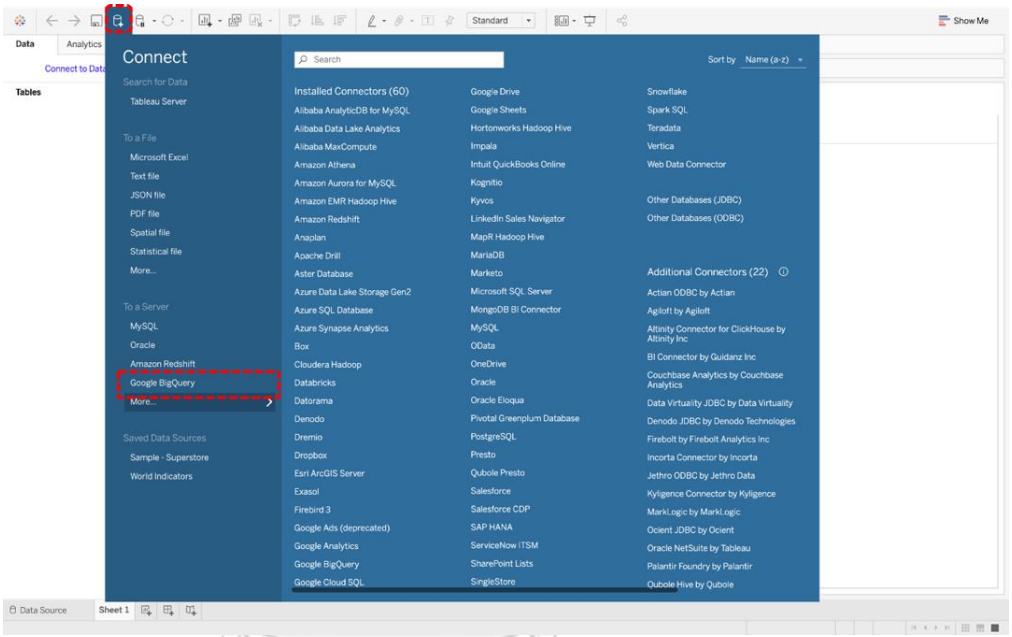


รูปที่ 27: หน้าจอการนำเข้าข้อมูลในฐานข้อมูลที่สำเร็จ

4.4 การเชื่อมต่อข้อมูลจากฐานข้อมูลเข้าสู่โปรแกรมที่ใช้สำหรับพัฒนาระบบคลังข้อมูล

การเชื่อมต่อข้อมูลระหว่างฐานข้อมูลกับระบบเครื่องมือที่ใช้ในการพัฒนาและแสดงผลธุรกิจอัจฉริยะ โดยนำฐานข้อมูลของแต่ละระบบที่สร้างใน Google Bigquery มาเชื่อมต่อกับโปรแกรม Tableau Desktop Version 2021.3 ซึ่งเป็นเครื่องมือที่ใช้ในการพัฒนาและแสดงผลธุรกิจอัจฉริยะ มีขั้นตอนดังนี้

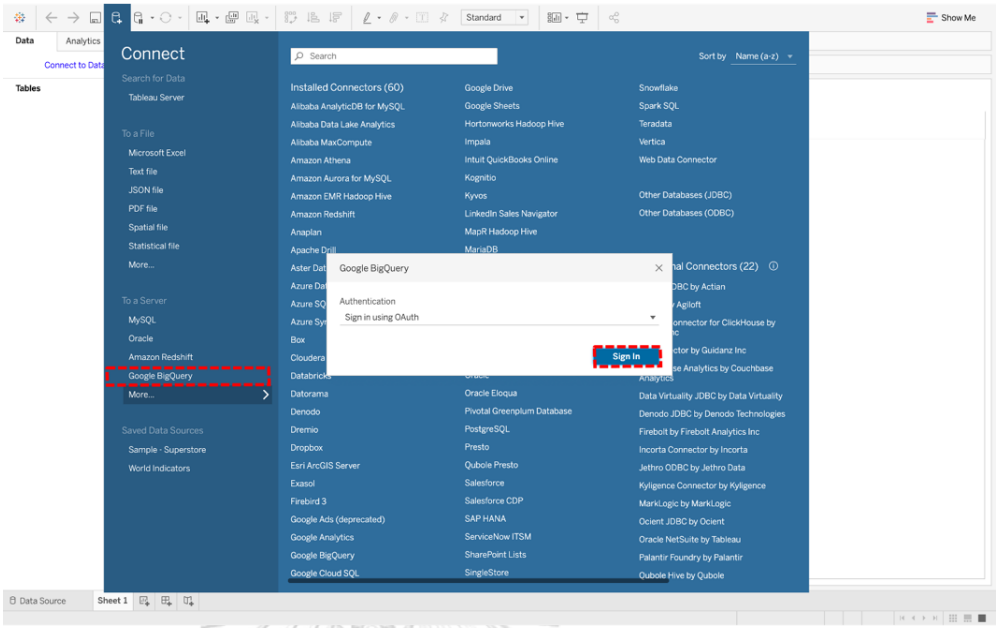
1) เชื่อมต่อกับฐานข้อมูลจากการนำเข้าข้อมูลจาก Google BigQuery ดังรูปที่ 28



รูปที่ 28: หน้าจอการนำเข้าข้อมูลจาก Google Bigquery

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

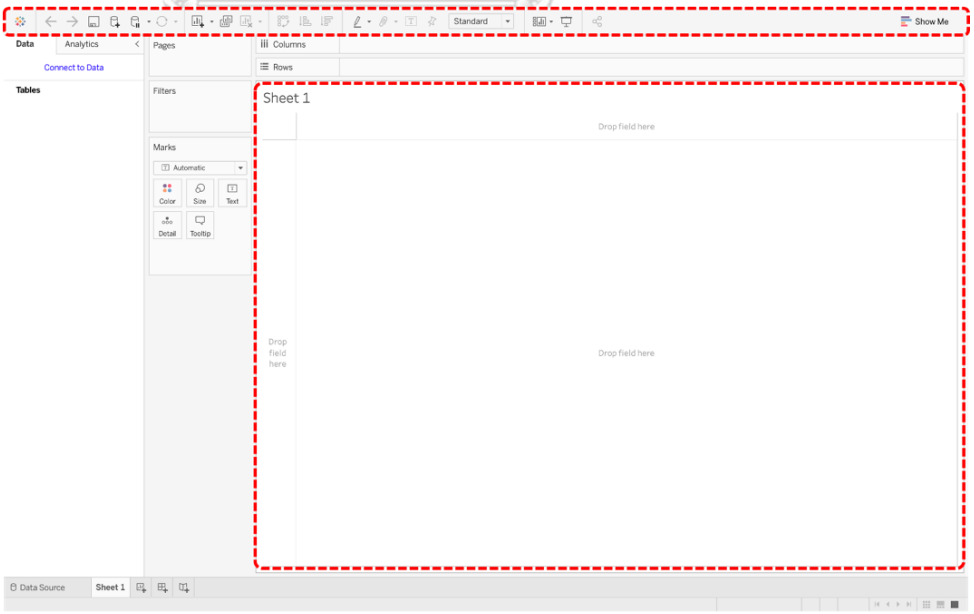
2) Sign In เพื่อเชื่อมต่อกับฐานข้อมูลจาก Google Bigquery ดังรูปที่ 29



รูปที่ 29: Sign In เพื่อเชื่อมต่อข้อมูลจาก Google Biquery

3) หลังจากการ Sign in และนำเข้ามูลเรียบร้อยแล้ว ผู้ใช้สามารถสร้าง กราฟจากหน้า Worksheet ได้

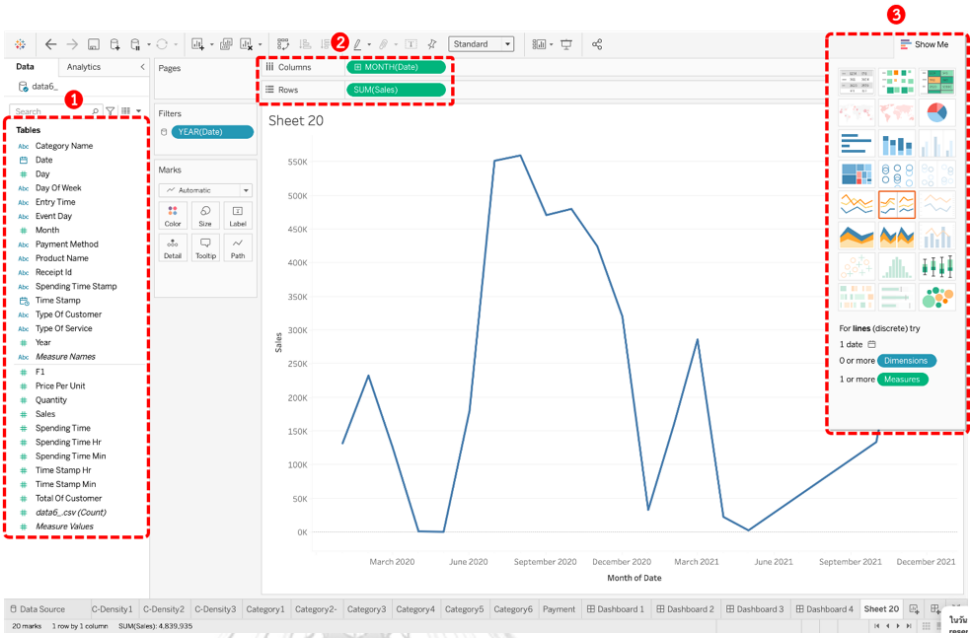
ดังรูปที่ 30



รูปที่ 30: หน้าจอการสร้างรายงาน

4.5 การสร้างรายงาน

4.5.1 การสร้างรายงานผ่าน Tableau Desktop มีขั้นตอนดังนี้



รูปที่ 31: หน้าจอการสร้างรายงาน Tableau Desktop

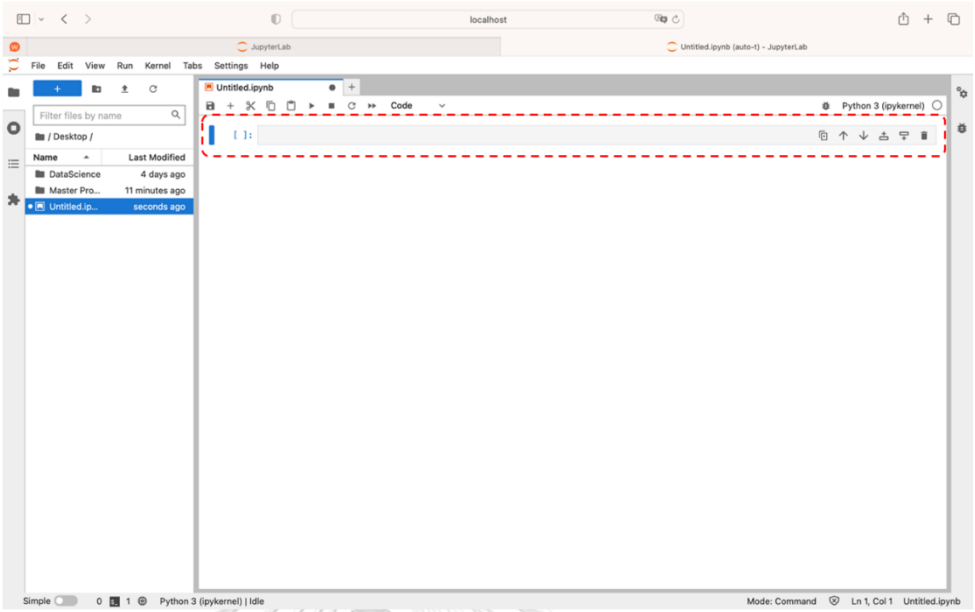
1) เลือกข้อมูลที่ต้องการนำมาใช้สร้างกราฟ

2) นำข้อมูลที่ถูกเลือก ลากมาไว้ในส่วนที่ต้องการ เช่น Columns หรือ Rows

3) Show me คือ การเลือกรูปแบบกราฟที่ต้องการให้โปรแกรมแสดงบนหน้าจอ

เมื่อกดเลือกเสร็จแล้วผลลัพธ์จะปรากฏบนหน้าจอดังรูปที่ 31

4.5.2 การสร้างรายงานผ่านโปรแกรม Jupyter Lab มีขั้นตอนดังนี้



รูปที่ 32: หน้าจอการสร้างรายงานผ่าน Jupyter Lab

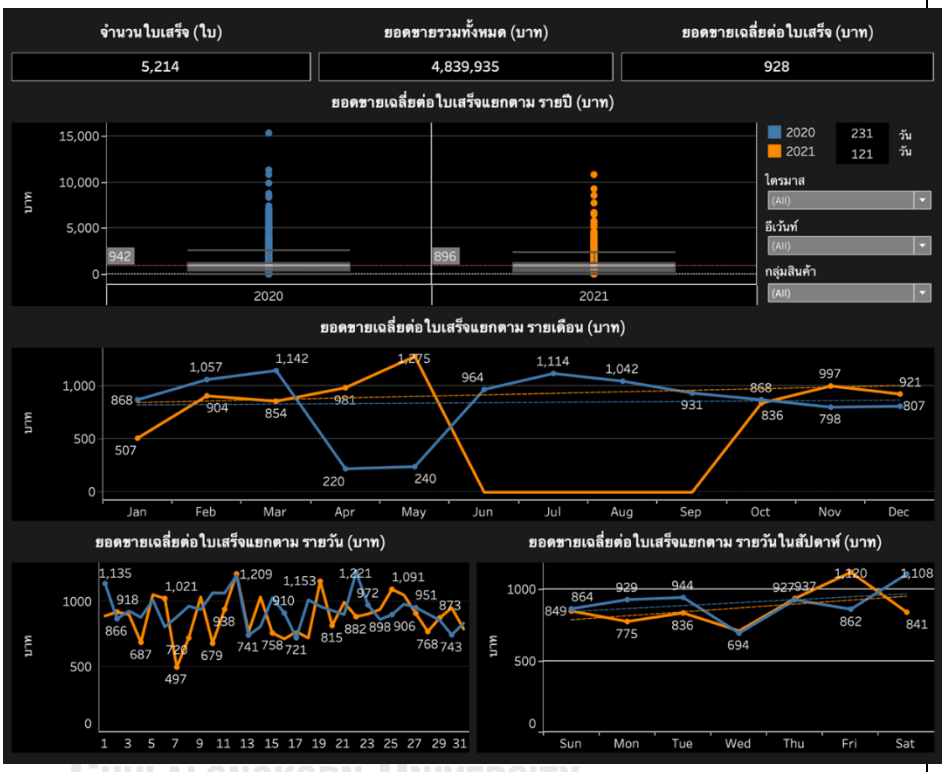
1) Import library ที่จำเป็นต้องใช้
2) นำเข้าข้อมูลลงระบบโดยการเขียนคำสั่ง
3) ทำความสะอาดข้อมูล
4) พัฒนาตัวแปรใหม่ที่เหมาะสม
5) ทดสอบโมเดล
6) วัดผลประสิทธิภาพโมเดล
7) สรุปผลที่ได้เพื่อเขียนรายงาน

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

4.5.3 ตัวอย่างการสร้างรายงานจาก Tableau Desktop มีดังนี้

1) ระบบวิเคราะห์ยอดขาย (Sales Analysis System)

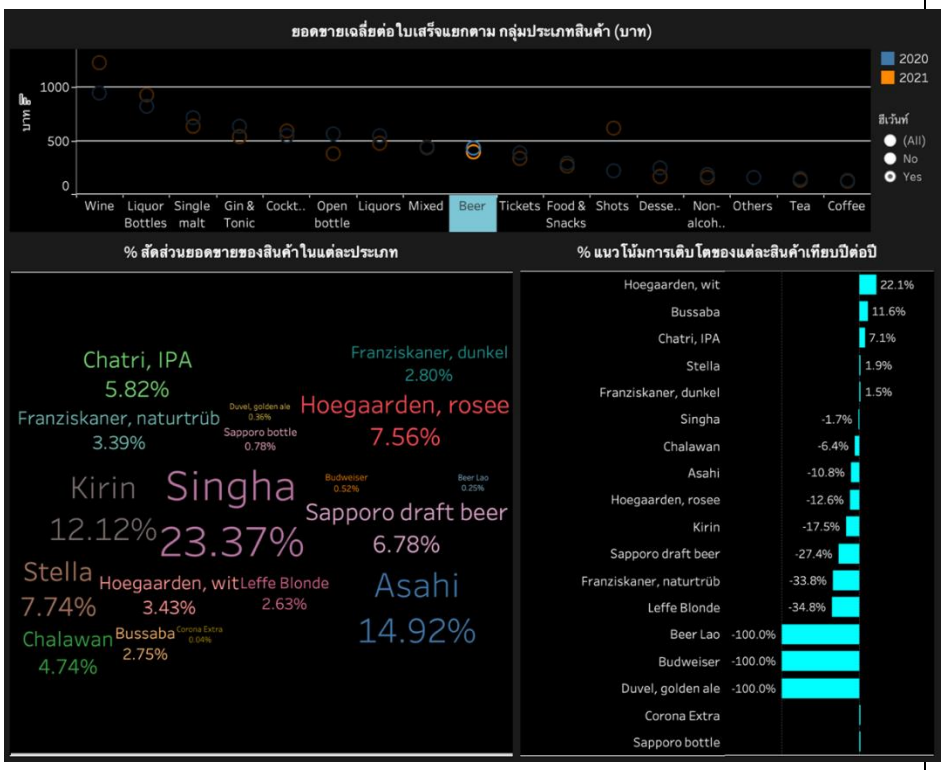
ชื่อรายงาน	แดชบอร์ดการวิเคราะห์ทิศทางและแนวโน้มยอดขายของกิจการ
ผู้ใช้งาน	1) ผู้จัดการร้าน 2) หัวหน้าฝ่ายจัดซื้อ
ประโยชน์	เพื่อใช้วิเคราะห์ข้อมูลยอดขายแยกตามช่วงเวลาที่สนใจ



CHULALONGKORN UNIVERSITY

รูปที่ 33: แดชบอร์ดการวิเคราะห์ทิศทางและแนวโน้มยอดขายของกิจการ

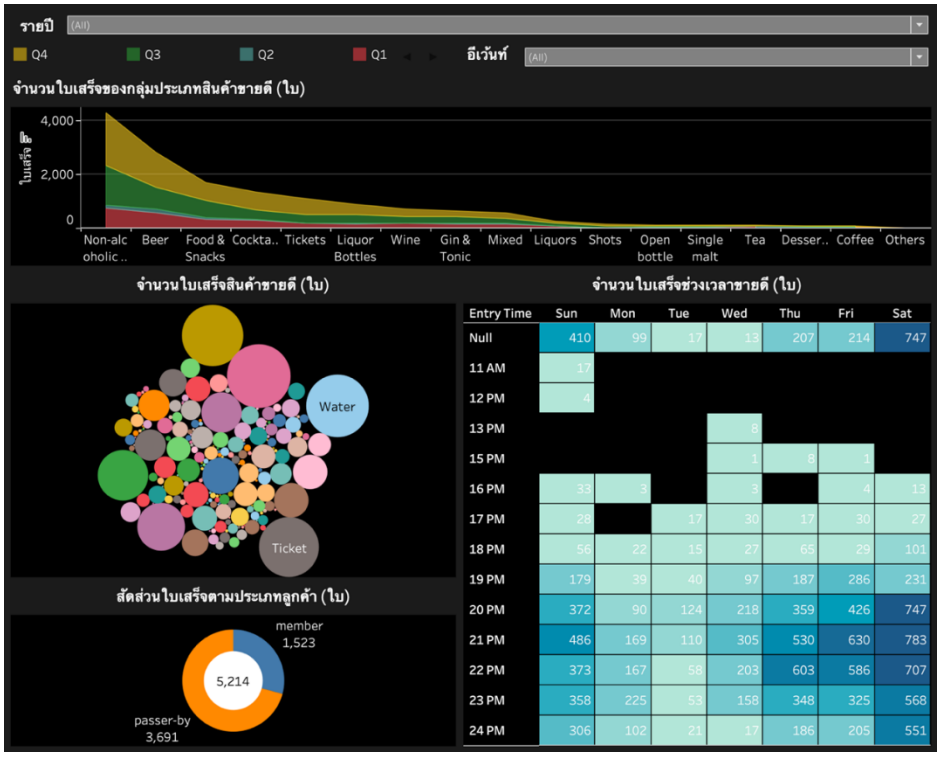
ชื่อรายงาน	แดชบอร์ดการวิเคราะห์ยอดขายแยกตามประเภทรายการอาหารและเครื่องดื่ม
ผู้ใช้งาน	1) ผู้จัดการร้าน 2) หัวหน้าฝ่ายจัดซื้อ
ประโยชน์	เพื่อใช้วิเคราะห์ข้อมูลยอดขายแยกตามประเภทรายการอาหารและเครื่องดื่ม



รูปที่ 34: แดชบอร์ดการวิเคราะห์ยอดขายแยกตามประเภทรายการอาหารและเครื่องดื่ม

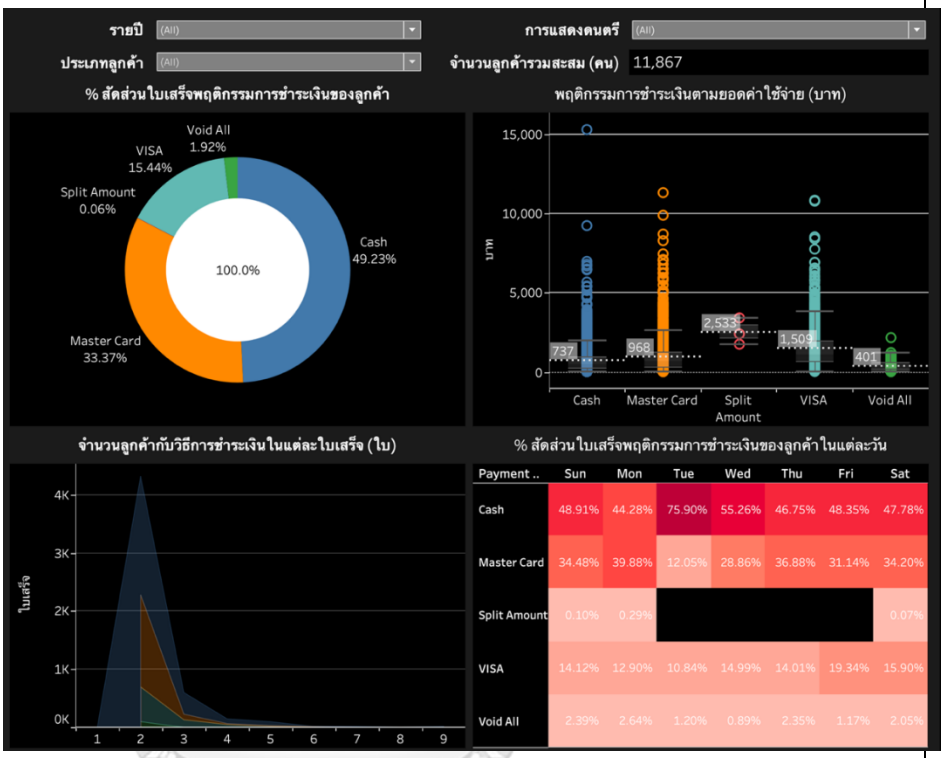
2) ระบบการวิเคราะห์การสั่งอาหารและเครื่องดื่ม (Food & Beverage Order Analysis System)

ชื่อรายงาน	แดชบอร์ดการวิเคราะห์ใบเสร็จการสั่งอาหารและเครื่องดื่มที่พบ
ผู้ใช้งาน	1) ผู้จัดการร้าน 2) หัวหน้าฝ่ายจัดซื้อ
ประโยชน์	เพื่อใช้วิเคราะห์จำนวนใบเสร็จแยกตามประเภทรายการอาหารและเครื่องดื่ม



รูปที่ 35: แดชบอร์ดการวิเคราะห์ใบเสร็จการสั่งอาหารและเครื่องดื่มที่พบ

ชื่อรายงาน	แดชบอร์ดการวิเคราะห์พฤติกรรมการเข้าใช้บริการร้าน
ผู้ใช้งาน	1) ผู้จัดการร้าน 2) หัวหน้าฝ่ายจัดซื้อ
ประโยชน์	เพื่อใช้วิเคราะห์พฤติกรรมการชำระเงินของลูกค้า



รูปที่ 36: แดชบอร์ดการวิเคราะห์ใบเสร็จการสั่งอาหารและเครื่องดื่มที่พบ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

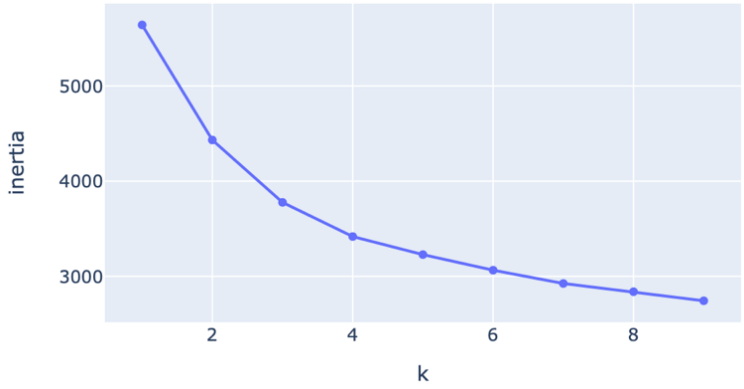
4.5.4 ตัวอย่างการสร้างรายงานจาก Jupyter Lab มีดังนี้

1) ระบบแบ่งกลุ่มใบเสร็จค่าอาหารและเครื่องดื่ม (Food & Beverage Receipt Segmentation System)

ชื่อรายงาน	รายงานการหาค่าเฉลี่ยและจำนวนนับที่ได้จากการโมเดลข้อมูล โดยแยกตามกลุ่มใบเสร็จค่าอาหารและเครื่องดื่ม
ผู้ใช้งาน	1) ผู้จัดการร้าน 2) หัวหน้าฝ่ายจัดซื้อ
ประโยชน์	ใช้วิเคราะห์เพื่อทำความเข้าใจกลุ่มใบเสร็จค่าอาหารและเครื่องดื่ม

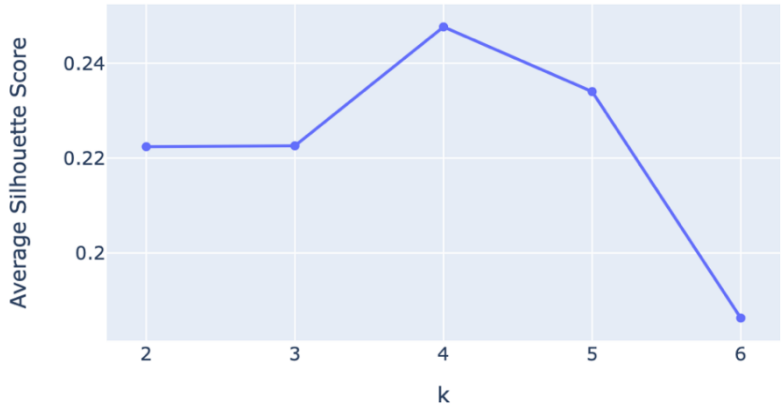
พารามิเตอร์ที่ใช้เป็นตัวชี้วัดในการแบ่งกลุ่ม ได้แก่

1. Elbow testing



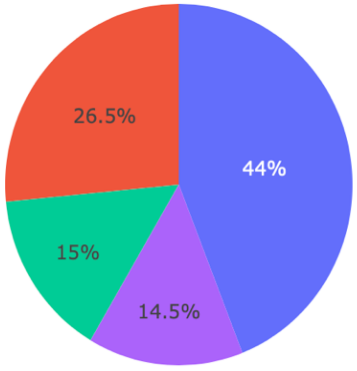
รูปที่ 37: กราฟเส้นแสดงผลลัพธ์การแบ่งกลุ่มที่เหมาะสมโดยใช้เทคนิค Elbow testing

2. Average Silhouette Score



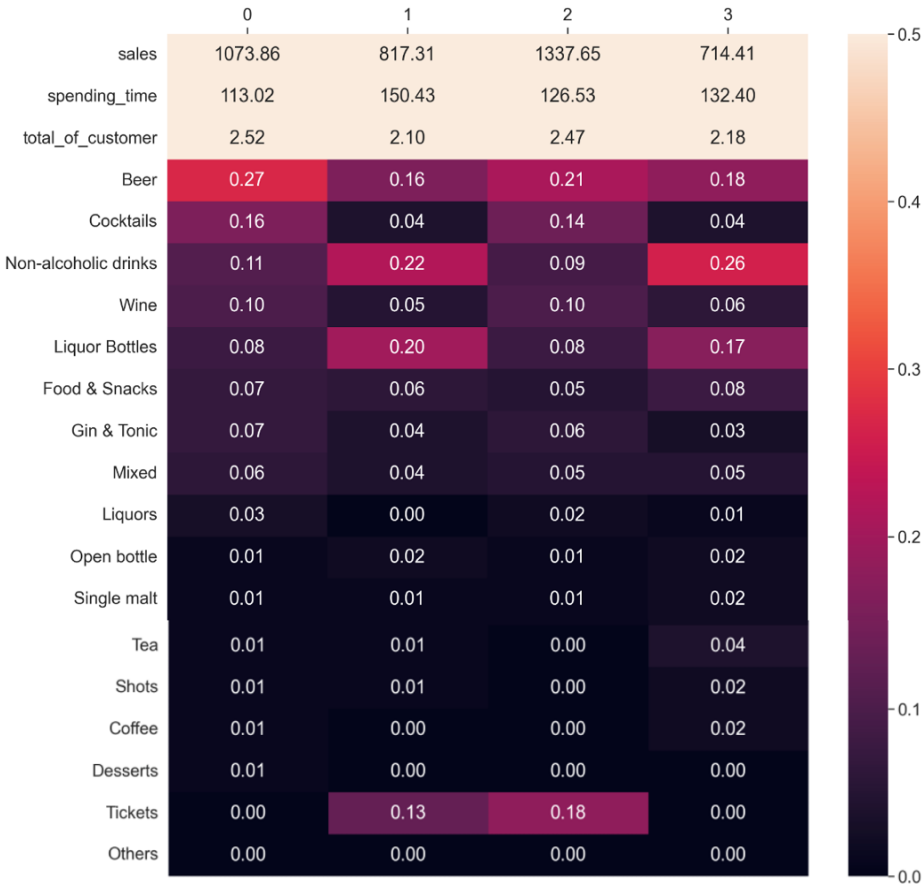
รูปที่ 38: กราฟเส้นแสดงผลลัพธ์การแบ่งกลุ่มที่เหมาะสมโดยใช้เทคนิค Average Silhouette

3. สัดส่วนของข้อมูลแยกตามจำนวนกลุ่ม Cluster



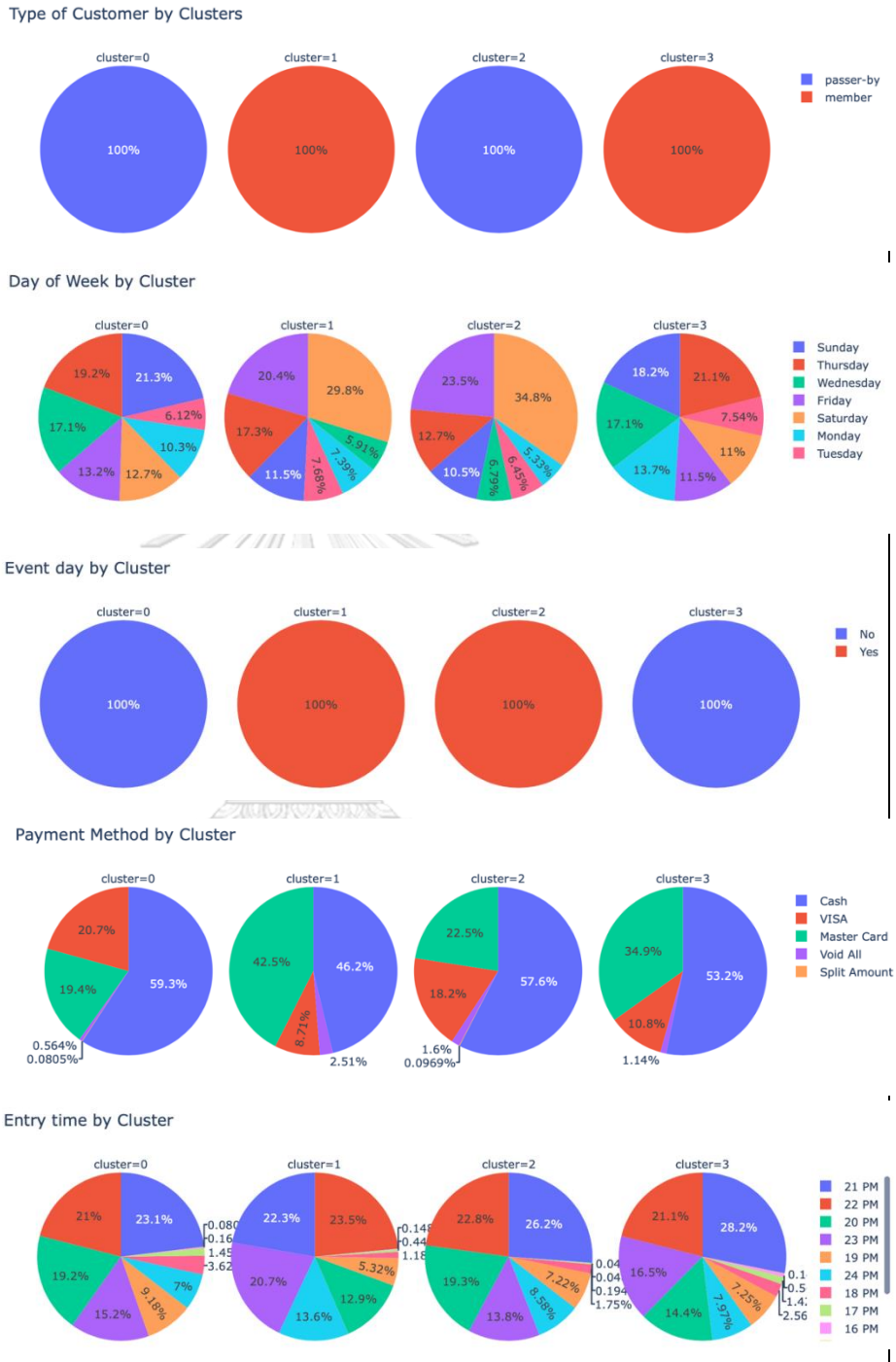
รูปที่ 39: กราฟวงกลมแสดงสัดส่วนของข้อมูลแยกตามจำนวนกลุ่ม Cluster

4. ค่าเฉลี่ยของแต่ละตัวแปรที่สนใจ โดยแยกตาม Cluster



รูปที่ 40: รายงานแสดงค่าเฉลี่ยของแต่ละตัวแปรที่สนใจ โดยแยกตาม Cluster

5. สัดส่วนของตัวแปรที่สนใจ โดยแยกตาม Cluster



รูปที่ 41: กราฟวงกลมแสดงสัดส่วนของตัวแปรที่สนใจ โดยแยกตาม Cluster

2) ระบบวิเคราะห์กฎความสัมพันธ์ในใบเสร็จค่าอาหารและเครื่องดื่ม (Food & Beverage Receipt Association Analysis System)

ชื่อรายงาน	รายงานเมนูอาหารและเครื่องดื่มที่พบความสัมพันธ์อย่างมีนัยสำคัญมากที่สุดตั้งแต่เปิดกิจการ
ผู้ใช้งาน	1) ผู้จัดการร้าน 2) หัวหน้าฝ่ายจัดซื้อ
ประโยชน์	ใช้วิเคราะห์เพื่อค้นหาความสัมพันธ์สองสิ่งที่เกิดขึ้นพร้อมกัน และนำไปใช้ในการตัดสินใจส่งเสริมการขาย

พารามิเตอร์ที่ใช้วัด ได้แก่ minimum support = 5%, confidence และ lift

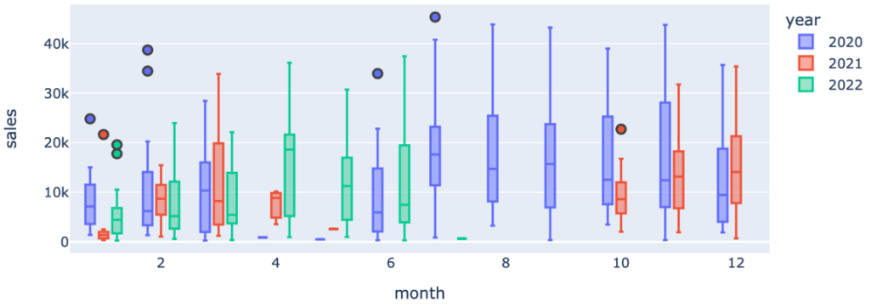
	antecedents	consequents	support	confidence	lift
0	(Non-alcoholic drinks)	(Liquor Bottles)	0.235857	0.414796	1.724730
1	(Liquor Bottles)	(Non-alcoholic drinks)	0.235857	0.980700	1.724730
2	(Non-alcoholic drinks)	(Liquor Bottles, Food & Snacks)	0.065854	0.115816	1.720771
3	(Liquor Bottles, Food & Snacks)	(Non-alcoholic drinks)	0.065854	0.978448	1.720771
4	(Non-alcoholic drinks)	(Liquor Bottles, Tickets)	0.064694	0.113776	1.712595
5	(Liquor Bottles, Tickets)	(Non-alcoholic drinks)	0.064694	0.973799	1.712595
6	(Liquor Bottles)	(Food & Snacks, Non-alcoholic drinks)	0.065854	0.273824	1.412980
7	(Food & Snacks, Non-alcoholic drinks)	(Liquor Bottles)	0.065854	0.339820	1.412980
8	(Gin & Tonic)	(Cocktails)	0.050769	0.369198	1.352419
9	(Cocktails)	(Gin & Tonic)	0.050769	0.185972	1.352419
10	(Tickets, Non-alcoholic drinks)	(Liquor Bottles)	0.064694	0.314972	1.309659
11	(Liquor Bottles)	(Tickets, Non-alcoholic drinks)	0.064694	0.268999	1.309659
12	(Food & Snacks)	(Wine)	0.077459	0.228792	1.290745
13	(Wine)	(Food & Snacks)	0.077459	0.436989	1.290745
14	(Food & Snacks)	(Beer, Non-alcoholic drinks)	0.065274	0.192802	1.167994
15	(Beer, Non-alcoholic drinks)	(Food & Snacks)	0.065274	0.395431	1.167994
16	(Beer)	(Tickets, Food & Snacks)	0.055991	0.135249	1.151117
17	(Tickets, Food & Snacks)	(Beer)	0.055991	0.476543	1.151117
18	(Beer)	(Tickets)	0.194082	0.468816	1.101573
19	(Tickets)	(Beer)	0.194082	0.456033	1.101573
20	(Food & Snacks)	(Beer)	0.151726	0.448158	1.082550
21	(Beer)	(Food & Snacks)	0.151726	0.366503	1.082550
22	(Food & Snacks)	(Cocktails)	0.099507	0.293916	1.076651
23	(Cocktails)	(Food & Snacks)	0.099507	0.364506	1.076651

รูปที่ 42: ตัวอย่างรายงานประเภทกลุ่มสินค้าคู่ที่พบความสัมพันธ์อย่างมีนัยสำคัญ

3) ระบบพยากรณ์ยอดขาย (Sales Forecasting System)

ชื่อรายงาน	รายงานผลลัพธ์การพยากรณ์ยอดขาย
ผู้ใช้งาน	1) ผู้จัดการร้าน 2) หัวหน้าฝ่ายจัดซื้อ
ประโยชน์	เพื่อนำสมการถดถอยแบบพหุคูณที่ได้ไปพยากรณ์ยอดขายที่จะเกิดขึ้นในอนาคต

1. ยอดค่าใช้จ่ายใบเสร็จรายวันแยกตามรายเดือนและรายปี



รูปที่ 43: กราฟแสดงยอดค่าใช้จ่ายใบเสร็จรายวันแยกตามรายเดือนและรายปี

2. การประมาณค่าพารามิเตอร์จากการเลือกช่วงเวลามาสร้างสมการพหูคูณด้วยวิธีกำลังสองน้อยที่สุด (Ordinary Least Squares: OLS)

OLS Regression Results

Dep. Variable:	sales	R-squared:	0.499
Model:	OLS	Adj. R-squared:	0.485
Method:	Least Squares	F-statistic:	34.98
Date:	Sun, 30 Oct 2022	Prob (F-statistic):	6.93e-56
Time:	02:56:22	Log-Likelihood:	-4474.2
No. Observations:	434	AIC:	8974.
Df Residuals:	421	BIC:	9027.
Df Model:	12		
Covariance Type:	nonrobust		

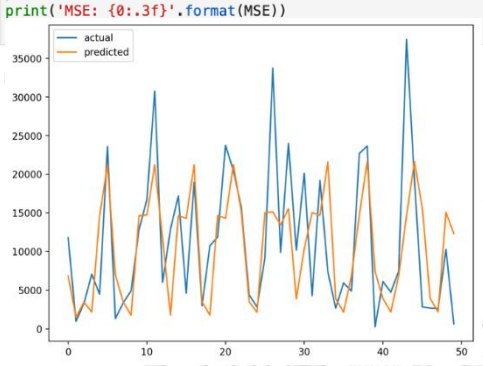
	coef	std err	t	P>|t|	[0.025	0.975]
const	7196.4653	3188.145	2.257	0.025	929.799	1.35e+04
quarter	5397.8372	4019.059	1.343	0.180	-2502.085	1.33e+04
month	828.5320	4807.043	0.172	0.863	-8620.262	1.03e+04
day	43.9801	1222.866	0.036	0.971	-2359.704	2447.664
continue_day_off	-1101.3211	5826.311	-0.189	0.850	-1.26e+04	1.04e+04
day_of_week_Monday	-6145.6178	3234.044	-1.900	0.058	-1.25e+04	211.267
day_of_week_Saturday	6372.3324	1885.553	3.380	0.001	2666.061	1.01e+04
day_of_week_Sunday	-231.1536	3075.825	-0.075	0.940	-6277.041	5814.734
day_of_week_Thursday	-762.1747	3082.636	-0.247	0.805	-6821.450	5297.101
day_of_week_Tuesday	-3741.4461	3446.522	-1.086	0.278	-1.05e+04	3033.089
day_of_week_Wednesday	-5464.7616	3086.942	-1.770	0.077	-1.15e+04	602.977
event_day_Yes	8168.0251	871.747	9.370	0.000	6454.505	9881.545
special_day_off_Yes	444.9494	1508.837	0.295	0.768	-2520.843	3410.741

Omnibus:	30.354	Durbin-Watson:	1.731
Prob(Omnibus):	0.000	Jarque-Bera (JB):	38.083
Skew:	0.582	Prob(JB):	5.38e-09
Kurtosis:	3.867	Cond. No.	39.1

รูปที่ 44: รายงานการประมาณค่าพารามิเตอร์จากการเลือกช่วงเวลามาสร้างสมการพหูคูณด้วยวิธีกำลังสองน้อยที่สุด (Ordinary Least Squares: OLS)

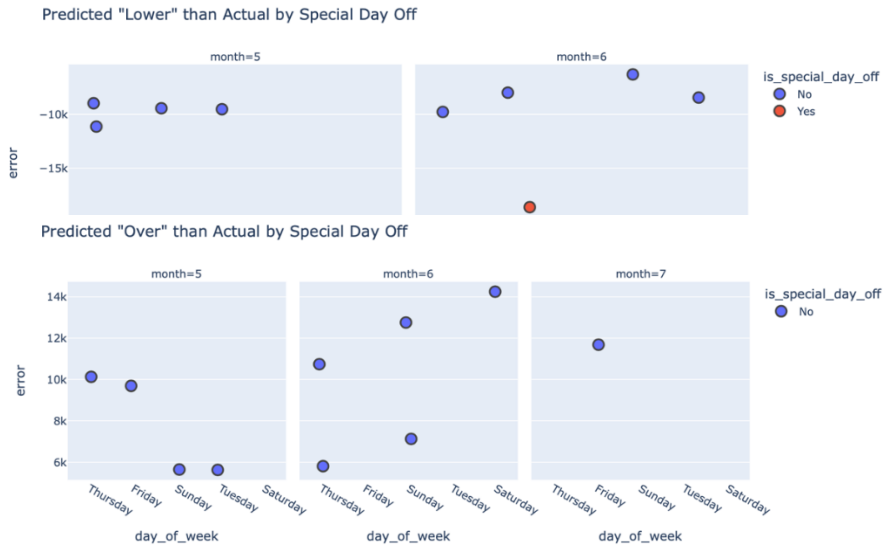
3. การทำนายผลของชุดทดสอบ test set โดยใช้เทคนิคการเรียนรู้เครื่องจักร (Machine Learning)

```
reg = LinearRegression()
reg.fit(X_trainplusvalidate_scaled,y_trainplusvalidate) # Y2020, Y2021 and Y2022 (Jan - Apr) >>> 428 Days

y_pred = reg.predict(X_test_scaled) # Y2022 (May - July) >>> 50 Days

R2 = r2_score(y_test, y_pred)
MSE = mean_squared_error(y_test, y_pred)
RMSE = np.sqrt(mean_squared_error(y_test, y_pred))

print('R2: {0:.3f}'.format(R2))
print('MSE: {0:.3f}'.format(MSE))
```



รูปที่ 45: รายงานการทำนายผลของชุดทดสอบ test set

4. การวิเคราะห์การทำนายผลผิดพลาดของชุดทดสอบผ่านตัวแปรที่สนใจ



รูปที่ 46: รายงานการวิเคราะห์การทำนายผลผิดพลาดของชุดทดสอบ

บทที่ 5
บทสรุปปัญหาและข้อเสนอแนะ

ในบทนี้จะกล่าวถึงบทสรุปปัญหา และข้อเสนอแนะ ของการพัฒนาโครงการ “คลังข้อมูลธุรกิจอัจฉริยะ และการวิเคราะห์ข้อมูลขั้นสูงของธุรกิจร้านอาหารและเครื่องดื่ม” สำหรับใช้เป็นแนวทางในการ พัฒนาระบบสารสนเทศอื่นๆ ต่อไป

5.1 บทสรุป

การพัฒนาโครงการ “คลังข้อมูล ธุรกิจอัจฉริยะ และการวิเคราะห์ข้อมูลขั้นสูงของธุรกิจร้านอาหารและเครื่องดื่ม” เริ่มต้นจากการสัมภาษณ์เจ้าของกิจการ เพื่อศึกษาการดำเนินงานและปัญหาที่เกิดขึ้นในปัจจุบัน โดยรวบรวมไฟล์ใบเสร็จทั้งหมดของกิจการร้านอาหารและเครื่องดื่ม และนำมาจัดรูปแบบข้อมูลให้อยู่ในรูปแบบเดียวกัน เพื่อเตรียมพร้อมสำหรับการนำมาพัฒนาเป็นคลังข้อมูลและระบบสนับสนุนการตัดสินใจสำหรับเจ้าของกิจการร้าน เพื่อใช้ในการตัดสินใจในการดำเนินธุรกิจ ให้ทันเวลากับสภาพแวดล้อมที่เปลี่ยนแปลงไปในการดำเนินธุรกิจในปัจจุบันได้อย่างมีประสิทธิภาพ

โครงการ “คลังข้อมูล ธุรกิจอัจฉริยะ และการวิเคราะห์ข้อมูลขั้นสูงของธุรกิจร้านอาหารและเครื่องดื่ม” ดังกล่าว ประกอบด้วย 5 ระบบย่อย ดังต่อไปนี้

1) ระบบวิเคราะห์ยอดขาย (Sales Analysis System)
2) ระบบการวิเคราะห์การสั่งอาหารและเครื่องดื่ม (Food & Beverage Order Analysis System)ระบบวิเคราะห์ลูกค้า (Customer Analysis System)
3) ระบบแบ่งกลุ่มใบเสร็จค่าอาหารและเครื่องดื่ม (Food & Beverage Receipt Segmentation System)
4) ระบบวิเคราะห์กฎความสัมพันธ์ในใบเสร็จค่าอาหารและเครื่องดื่ม (Food & Beverage Receipt Association Analysis System)
5) ระบบพยากรณ์ยอดขาย (Sales Forecasting System)

เทคโนโลยีที่ใช้ในโครงการนี้ทั้งหมดเป็นชุดโปรแกรมสำหรับพัฒนา Business Intelligence ซึ่งเป็นซอฟต์แวร์ที่มีฟังก์ชันการใช้งานที่หลากหลายที่ช่วยให้ผู้ใช้สามารถ ออกแบบและพัฒนาระบบสารสนเทศได้อย่างหลากหลาย สามารถนำข้อมูลจำนวนมากที่มีอยู่ในองค์กรมาทำการวิเคราะห์ได้อย่างรวดเร็ว และยังรองรับการเข้าถึงข้อมูลจากหลายฐานข้อมูล ได้เช่นกัน นอกจากนั้นยังความ

ยืดหยุ่นให้สามารถปรับเปลี่ยนมุมมองในการวิเคราะห์เพื่อให้ผู้ใช้งานได้เห็นข้อมูลในมิติต่าง ๆ เพื่อค้นพบข้อมูลเชิงลึกที่อาจเป็นประโยชน์แก่องค์กรได้อีกด้วย

กล่าวโดยสรุป โครงการ “คลังข้อมูล ธุรกิจอัจฉริยะ และการวิเคราะห์ข้อมูลขั้นสูงของธุรกิจร้านอาหารและเครื่องดื่ม” นี้ สามารถบรรลุวัตถุประสงค์ที่กำหนดไว้ได้อย่างครบถ้วน ดังนี้

1) เพื่อพัฒนาคลังข้อมูล โดยรวบรวมข้อมูลที่เก็บในรูปแบบ Excel และข้อมูลที่เก็บแยกกันมาเชื่อมโยงและรวมไว้ที่เดียวกัน เพื่อลดความซ้ำซ้อนของข้อมูล สะดวกต่อการนำข้อมูลมาใช้ในการวิเคราะห์
2) เพื่อสร้างรายงานเชิงวิเคราะห์ที่มีลักษณะ Interactive เพื่อให้เจ้าของกิจการสามารถเรียกดูรายงานในมุมมองที่หลากหลายและใช้สนับสนุนการตัดสินใจและวางแผนกลยุทธ์ต่าง ๆ ตามความต้องการ
3) เพื่อแก้ปัญหาการนำเสนอข้อมูลที่ล่าช้า เนื่องจากในปัจจุบันเทคโนโลยีพัฒนาขึ้นมาก ทำให้ธุรกิจและพฤติกรรมของผู้บริโภคเปลี่ยนแปลงอย่างรวดเร็ว ดังนั้น การจะก้าวนำคู่แข่งจึงต้องมีการตัดสินใจอย่ารวดเร็วเพื่อให้ทันต่อความต้องการของผู้บริโภค

5.2 ปัญหา

ปัญหาที่พบในการพัฒนาโครงการ “คลังข้อมูล ธุรกิจอัจฉริยะ และการวิเคราะห์ข้อมูลขั้นสูงของธุรกิจร้านอาหารและเครื่องดื่ม” สามารถแบ่งออกเป็น 2 ส่วนหลัก ดังต่อไปนี้

1) ปัญหาด้านข้อมูล

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ปัญหาที่พบ: เนื่องจากข้อมูลที่ได้มีความตรงไปตรงมา และจำเป็นต้องสร้างใหม่เพื่อแปรผลตัวแปรให้มีความเข้าใจได้ง่าย และเหมาะสมกับแนวทางในการใช้แก้ปัญหาธุรกิจมากขึ้น

แนวทางในการแก้ไข: ทำความเข้าใจข้อมูล ศึกษาวิธีการแปรผลตัวแปรจากแหล่งศึกษาหาข้อมูลทั่วไป จากนั้นฝึกฝนและนำมาปรับใจในการพัฒนาโครงการ

2) ปัญหาการสร้างรายงาน

ปัญหาที่พบ: การสร้าง Measure ใน Tableau Desktop ใหม่ จากข้อมูลจำนวนมากทำให้ใช้เวลาค่อนข้างนาน รวมไปถึงการแสดงผลของแต่ละรายงานต้องใช้เวลาเช่นกัน

แนวทางการแก้ไข : หากในขั้นตอนการออกแบบมี Measure ที่ต้องการคำนวณอยู่แล้ว สามารถทำการคำนวณก่อนการนำเข้า Tableau Desktop ก่อนได้เพื่อลดเวลาการสร้างใหม่ในโปรแกรม

5.3 ข้อเสนอแนะ

จากการพัฒนาโครงการ “คลังข้อมูล ธุรกิจอัจฉริยะ และการวิเคราะห์ข้อมูลขั้นสูงของธุรกิจร้านอาหารและเครื่องดื่ม” ผู้พัฒนาพบปัญหาต่าง ๆ ดังที่กล่าวไว้ในข้างต้น ดังนั้นจึงข้อเสนอแนวทางในการพัฒนาโครงการ สำหรับผู้ที่สนใจทำโครงการลักษณะเดียวกัน โดยการแก้ไขการพยากรณ์ ในระบบพยากรณ์ยอดขาย เนื่องจากในปี ค.ศ. 2021 ที่ผ่านมา มีสถานการณ์ โควิด-19 อาจทำให้ยอดขายลดลงจากปกติเป็นอย่างมาก ทำให้การพยากรณ์ มีการคลาดเคลื่อนได้ หากสถานการณ์ได้กลับไปสู่ปกติ และการดำเนินการกิจการร้านอาหารและเครื่องดื่มสามารถดำเนินได้เหมือนเดิม ควรนำข้อมูลมาทำการพยากรณ์ใหม่ เพื่อให้การพยากรณ์นั้น คลาดเคลื่อนน้อยที่สุด

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บรรณานุกรม

1. Chaiyadecha, S. *Multiple linear regression* ด้วย *TensorFlow 2.0* และ *Keras*. 2020; Available from: https://lengyi.medium.com/multiple-linear-regression-2-tensorflow-2-0-keras-e1da62e873ff.
2. Coraline. สถิติเบื้องต้น. 2562; Available from: https://www.coraline.co.th/single-post/basic-statistic-2.
3. DW. การพัฒนาคลังข้อมูล (*Data Warehouse)*. 2559; Available from: http://chilchil-learning.blogspot.com/.
4. fillgoods. รู้จักทฤษฎีพฤติกรรมผู้บริโภค *6W1H* ตัวช่วยตามหาลูกค้าของธุรกิจออนไลน์. 2021; Available from: https://fillgoods.co/payments/no-shop-consumer-behavior-6w1h/.
5. Krooann. การวิเคราะห์เชิงทำนาย. 2560; Available from: https://sites.google.com/site/tecnologysci5/3-2-kar-wikheraah-cheing-thanay.
6. MBAToday. ระบบธุรกิจอัจฉริยะกับการจัดการ *Big Data*. 2561; Available from: https://mbamagazine.net/index.php/intelligent/nidawisdom/item/626-business-intelligence-big-data.
7. Meeieis. *Business Intelligence (BI)*. 2560; Available from: http://www.mindphp.com: http://www.mindphp.com/forums/viewtopic.php?f=79&t=41083.
8. mindphp. รู้จักกับคลังข้อมูล (*Data Warehouse)* คืออะไร มีประโยชน์อะไรบ้าง. 2563; Available from: www.mindphp.com: https://www.mindphp.com/forums/viewtopic.php?f=28&t=60570.
9. PanaEk. ลักษณะสำคัญของ *Data Warehouse*. 2555; Available from: https://bzinsight.wordpress.com/2012/06/03/ ลักษณะสำคัญของ-data-warehouse/.
10. Panthong, R. *Data Warehousing and Application*. 2560; Available from: www.ict.up.ac.th: https://www.ict.up.ac.th/rattanawadeep/dw/chapter1.pdf.
11. Porpor. การออกแบบคลังข้อมูล. 2563; Available from: www.thai-pdpa.com: https://thai-pdpa.com/ออกแบบคลังข้อมูล/.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

12. Smallcombe, M. *Star Schema vs Snowflake Schema*. 2563; Available from: https://www.xplenty.com/blog/snowflake-schemas-vs-starschemas-what-are-they-and-how-are-they-different/.
13. Thachai, W. การหาจำนวน *k* ที่เหมาะสมที่สุดด้วยวิธี *Silhouette*. 2017; Available from: https://medium.com/espressofx-notebook/การหาจำนวน-k-ที่เหมาะสมที่สุดด้วยวิธี-silhouette-b367fdae24d4.
14. Thachai, W. การหาจำนวน *k* ที่เหมาะสมที่สุดด้วยวิธี *Elbow Method*. 2017; Available from: https://medium.com/espressofx-notebook/การหาจำนวน-k-ที่เหมาะสมที่สุดด้วยวิธี-elbow-method-79b9a75f934.
15. Thankittikoon, J. *What is K-Means Clustering* 2019; Available from: https://medium.com/@Jutharath.Thankittikoon/what-is-k-means-clustering-ee36ab6f7638.
16. ชนะบุญ, ส. สถิติและการวิเคราะห์ข้อมูลในงานวิจัยเบื้องต้น. 2564; Available from: http://www.kkpho.go.th/index/component/attachments/download/1933.htm.
17. เศรษฐเสถียร, ณ. มาทำความรู้จัก *Association Rule:* เครื่องมือเพื่อการวิเคราะห์ตะกร้าตลาด กัน! 2021; Available from: https://bigdata.go.th/big-data-101/data-science/what-is-association-rule/.
18. สสว. บทวิเคราะห์ธุรกิจ *SMEs (Quick Study)* สาขาธุรกิจร้านอาหาร. 2020; Available from: https://www.sme.go.th/upload/mod_download/download-20210909160132.pdf.
19. อัมพวัน, โ. *Data Warehouse Design*. 2560; Available from: https://staff.informatics.buu.ac.th/~komate/886452/data%20warehouse-ch2.pdf.
20. อุดมธนะธีระ, เ. *BI* องค์ประกอบของ *Business Intelligence (BI)*. 2561; Available from: www.iok2u.com/index.php/article/information-technology/1047-bi-business-intelligence-bi.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ประวัติผู้เขียน

ชื่อ-สกุล	นาย อานนท์ ศิลานันท์
วัน เดือน ปี เกิด	5 ตุลาคม 2536
สถานที่เกิด	โรงพยาบาลตากสิน กรุงเทพฯ
ที่อยู่ปัจจุบัน	35/197 หมู่ที่ 5 ต.บางบ่อ อ.บางบ่อ จ.สมุทรปราการ

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY