



การสร้างรายงานอัตโนมัติด้วยชื่อของรายงานภาษาไทยและรูปแบบข้อมูล

กรณีศึกษา ธุรกิจบริการ

Automatic Report Generation based on Report's Title in Thai and

Data Schema: A Case Study on Service Business

ไตรรัตน์ พัฒน์เอี่ยม

Trirat Pat-iam

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญา

วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ

มหาวิทยาลัยสงขลานครินทร์

A Thesis Submitted in Partial Fulfillment of the Requirements for the

Degree of Master of Science in Information Technology

Prince of Songkla University

2561

ลิขสิทธิ์ของมหาวิทยาลัยสงขลานครินทร์



การสร้างรายงานอัตโนมัติด้วยชื่อของรายงานภาษาไทยและรูปแบบข้อมูล

กรณีศึกษา ธุรกิจบริการ

Automatic Report Generation based on Report's Title in Thai and

Data Schema: A Case Study on Service Business

นายไตรรัตน์ พัฒน์เอี่ยม

Trirat Pat-iam

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญา

วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ

มหาวิทยาลัยสงขลานครินทร์

A Thesis Submitted in Partial Fulfillment of the Requirements for the

Degree of Master of Science in Information Technology

Prince of Songkla University

2561

ลิขสิทธิ์ของมหาวิทยาลัยสงขลานครินทร์



(3)

ขอรับรองว่า ผลงานวิจัยนี้มาจากการศึกษาวิจัยของนักศึกษาเอง และได้แสดงความขอบคุณบุคคลที่มี  
ส่วนช่วยเหลือแล้ว

ลงชื่อ.....

(ดร.ณัฐพงศ์ ทองเทพ)

อาจารย์ที่ปรึกษาวิทยานิพนธ์

ลงชื่อ.....

(นายไตรรัตน์ พัฒน์เอี่ยม)

นักศึกษา



(4)

ข้าพเจ้าขอรับรองว่า ผลงานวิจัยนี้ไม่เคยเป็นส่วนหนึ่งในการอนุมัติปริญญาในระดับใดมาก่อน และ  
ไม่ได้ถูกใช้ในการยื่นขออนุมัติปริญญาในขณะนี้

ลงชื่อ.....

(นายไตรรัตน์ พัฒน์เอี่ยม)

นักศึกษา

ชื่อวิทยานิพนธ์	การสร้างรายงานอัตโนมัติด้วยชื่อของรายงานภาษาไทยและรูปแบบข้อมูล กรณีศึกษา ธุรกิจบริการ
ผู้เขียน	นายไตรรัตน์ พัฒน์เอี่ยม
สาขาวิชา	เทคโนโลยีสารสนเทศ
ปีการศึกษา	2560

### บทคัดย่อ

สำหรับธุรกิจนั้น สิ่งสำคัญประการหนึ่งคือ การวิเคราะห์ ออกแบบ และสร้างรายงานเพื่อนำเสนอสารสนเทศ ซึ่งประกอบด้วยตาราง และแผนภูมิ อย่างไรก็ตาม การสร้างรายงานอย่างมีประสิทธิภาพนั้น ต้องใช้เวลาและความพยายาม ดังนั้น ระบบหรือกลไกอัตโนมัติที่จะช่วยสร้างรายงานที่สามารถตั้งชื่อรายงาน เลือกตาราง หรือแผนภูมิที่เหมาะสม จึงเป็นสิ่งจำเป็น งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อ นำเสนอแบบจำลองในการสร้างรายงาน ด้วยการศึกษาวิเคราะห์โครงสร้างของชื่อรายงานที่เกิดขึ้นในภาษาไทย โดยรวบรวมตัวอย่างรายงานจำนวน 159 ฉบับจากเว็บไซต์บริการข้อมูลอิเล็กทรอนิกส์ของภาครัฐในประเทศไทย ผลการวิเคราะห์พบว่า ชื่อของรายงานสามารถแบ่งออกเป็น 3 ส่วน ได้แก่ ส่วนเริ่มต้น ส่วนกลาง และส่วนท้าย โดยแต่ละส่วนจะมีคุณลักษณะของคำสำคัญที่แตกต่างกัน นำไปสู่การพัฒนาแบบจำลอง ที่สามารถสร้างรายงานจากชื่อรายงาน ซึ่งใช้เพียงข้อมูลดิบที่จัดเตรียมโดยผู้ใช้ แบบจำลองจะดึงชนิดของข้อมูลของแต่ละคอลัมน์ร่วมกับคำสำคัญ จากนั้นแบบจำลองจะสร้างชื่อรายงานแบบต่าง ๆ ที่สอดคล้องกับข้อมูลดังกล่าว สำหรับการทดสอบแบบจำลอง ผู้วิจัยได้เลือกรายงานจำนวน 64 ฉบับ จากตัวอย่างรายงานที่ใช้ในกลุ่มธุรกิจบริการ ซึ่งผลการทดสอบพบว่า ค่าความคล้ายคลึง Cosine Similarity ระหว่างชื่อรายงานที่สร้างจากแบบจำลองกับชื่อรายงานที่ปรากฏจริง มีคะแนนความคล้ายคลึงเฉลี่ยร้อยละ 60 ผลการวิจัยพบว่า แบบจำลองที่นำเสนอนี้ สามารถช่วยลดระยะเวลาในการสร้างรายงานสำหรับธุรกิจได้

**คำสำคัญ:** การสร้างรายงานโดยอัตโนมัติ, ระบบแนะนำ, โครงสร้างข้อมูล, ธุรกิจ

<b>Thesis Title</b>	Automatic Report Generation based on Report's Title in Thai and Data Schema: A Case Study on Service Business
<b>Author</b>	Mr.Trirat Pat-iam
<b>Major Program</b>	Information Technology
<b>Academic</b>	2017

### ABSTRACT

In business, there is a high necessity of creating a variety of reports whose main contents include tables and graphs. Effective report writing takes time and effort, and an automatic report generation system to create simple reports with an appropriate title, a table and a graph is desirable. The goal of this research is to propose a report generation model based on the investigation of title structures in the Thai language. We collected 159 reports as report samples from an open government data website of Thailand. Our analysis on titles revealed that a title can be divided into three parts: the beginning, middle, and ending phrases, and each part has its own preference on the choice of keywords. Based on this, the proposed model creates a report title candidate list based on the table of data that is uploaded by the user. In order to create a title candidate list, the model first tries to obtain the data type of each column names and search for the relevant keywords in the keyword list, which are then combined together to form a title. For the model evaluation, we selected 64 reports from samples in services businesses. The cosine similarity between the title proposed by the model and the actual title is approximately 0.6 on average. This model is expected to reduce the burden of the user who needs to write many short reports in business.

**Keywords:** Report generation, Recommendation system, Data structure, Business

### กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จได้จากความช่วยเหลือของ ดร.ณัฐพงศ์ ทองเทพ อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่ได้ให้คำแนะนำเกี่ยวกับกระบวนการวิจัยและกระบวนการด้านเอกสาร ข้าพเจ้าขอขอบพระคุณเป็นอย่างสูง

ขอขอบพระคุณคณะกรรมการสอบ ประกอบด้วย ดร.กาญจนา ทองกลิ่น Prof. Dr. Kiyota Hashimoto และ ดร.พุทธิพร ธนธรรมเมธี ที่แนะนำเกี่ยวกับกระบวนการวิจัยมาตลอด

ขอขอบพระคุณคณาจารย์ รวมทั้งบุคลากรประจำวิทยาลัยการคอมพิวเตอร์ ที่ถ่ายทอดความรู้ในทุกด้าน และขอขอบคุณ นายบุรฮาน หวังเหล็ม และนายณัฐ กุลพัฒนโสภณ สำหรับความช่วยเหลือตลอดการศึกษาในหลักสูตรของข้าพเจ้า

ขอขอบพระคุณมารดา น้องชาย และบุคคลสำคัญในชีวิต ที่คอยสนับสนุนและเป็นกำลังใจให้กับข้าพเจ้าเสมอมา

ไตรรัตน์ พัฒน์เยี่ยม

## สารบัญ

	หน้า
บทคัดย่อ	(5)
ABSTRACT	(6)
กิตติกรรมประกาศ	(7)
สารบัญ	(8)
สารบัญตาราง	(10)
สารบัญภาพประกอบ	(11)
<b>บทที่ 1 บทนำ</b>	<b>1</b>
1.1 ความเป็นมาของปัญหาและปัญหา	1
1.2 คำถามวิจัย	2
1.3 วัตถุประสงค์	2
1.4 สมมติฐาน	2
1.5 ความสำคัญและประโยชน์ของงานวิจัย	3
1.6 ขอบเขตของงานวิจัย	3
<b>บทที่ 2 การทบทวนวรรณกรรมและทฤษฎีที่เกี่ยวข้อง</b>	<b>4</b>
2.1 รายงาน	4
2.2 ข้อมูลและการเก็บข้อมูล	6
2.3 แผนภูมิ	9
2.4 เครื่องมือในการสร้างรายงาน	17
2.5 ภาษาไทย	19
2.6 หลักการทางสถิติเบื้องต้น	28
2.7 กลุ่มอุตสาหกรรมและหมวดธุรกิจ	32
2.8 การหาความคล้ายกันของข้อความ (Sentence Similarity Measurement)	39
2.9 การตัดคำ	40
<b>บทที่ 3 ขั้นตอนการดำเนินการวิจัย</b>	<b>42</b>
3.1 การเก็บข้อมูล (Data collection)	44
3.2 การออกแบบแบบจำลอง (Report generation model)	52

## สารบัญ (ต่อ)

	หน้า
<b>บทที่ 4 ผลการดำเนินงาน</b>	60
4.1 การทดสอบแบบจำลอง	60
4.2 การพัฒนาระบบ	79
<b>บทที่ 5 สรุปผลการทดลอง</b>	89
5.1 สรุปและข้อเสนอแนะ	89
5.2 สรุปผลการทดลอง	89
5.3 ปัญหาและอุปสรรค	91
5.4 ข้อเสนอแนะ	91
5.5 ลักษณะการนำไปใช้ประโยชน์	92
<b>ภาคผนวก ก คำจำกัดความคำสำคัญในงานวิจัย</b>	97
<b>ภาคผนวก ข รายชื่อรายงานต้นฉบับ</b>	108
<b>ภาคผนวก ค คำหยุดที่ปรากฏในชื่อของรายงาน</b>	120
<b>ภาคผนวก ง สถิติคำและชนิดของคำที่ปรากฏในรายงาน</b>	124
<b>ภาคผนวก จ ตัวอย่างชื่อรายงานที่แบบจำลองสร้างขึ้นจากชื่อรายงาน</b>	138
<b>ภาคผนวก ฉ คำสำคัญที่ปรากฏในชื่อรายงาน</b>	144
<b>ประวัติผู้เขียน</b>	147

## สารบัญตาราง

	หน้า
ตารางที่ 2.1 ตารางแสดงสัญลักษณ์ทางสถิติ	29
ตารางที่ 2.2 ตารางกลุ่มและหมวดธุรกิจแบ่งตามตลาดหลักทรัพย์แห่งประเทศไทย	36
ตารางที่ 3.1 ตารางรายละเอียดรายงานจาก WWW.DATA.GO.TH โดยแยกตามหมวดธุรกิจ	44
ตารางที่ 3.2 ค่าทางสถิติของจำนวนคำที่ปรากฏในชื่อรายงานแยกตามหมวดหมู่	49
ตารางที่ 3.3 ตัวอย่างการบันทึกประเภทข้อมูลตามคำสำคัญ	50
ตารางที่ 3.4 ตัวอย่างของคำสำคัญที่ปรากฏในแต่ละตำแหน่งของชื่อรายงาน	51
ตารางที่ 3.5 ตารางแสดงผลลัพธ์จากการผสมคำสำคัญเพื่อสร้างเป็นชื่อของรายงาน	57
ตารางที่ 4.1 ค่าสถิติเบื้องต้นของกลุ่มตัวอย่างการทดลองที่ 1.1	64
ตารางที่ 4.2 ค่าสถิติเบื้องต้นของกลุ่มตัวอย่างการทดลองที่ 1.2	67
ตารางที่ 4.3 ค่าสถิติเบื้องต้นของกลุ่มตัวอย่างการทดลองที่ 2.1	71
ตารางที่ 4.4 ลำดับคะแนน COSINE SIMILARITY ในอันดับ 1 - 5	72
ตารางที่ 4.5 ลำดับคะแนน COSINE SIMILARITY ในอันดับ 6 - 10	73
ตารางที่ 4.6 ค่าสถิติเบื้องต้นของกลุ่มตัวอย่างการทดลองที่ 3	73
ตารางที่ 4.7 ตารางเปรียบเทียบงานวิจัยที่มีแนวทางการประยุกต์ใช้ขั้นตอนวิธีคล้ายคลึงกัน	77

## สารบัญภาพประกอบ

	หน้า
รูปที่ 2.1 ตัวอย่างแผนภูมิแท่งแนวตั้ง จากเครื่องมือ R STUDIO	10
รูปที่ 2.2 ตัวอย่างแผนภูมิแท่งแนวนอน จากเครื่องมือ R STUDIO	10
รูปที่ 2.3 ตัวอย่างฮิสโตแกรมจากเว็บไซต์ WWW.SOFTSCHOOLS.COM	11
รูปที่ 2.4 ตัวอย่างแผนภูมิจุดจากเครื่องมือ R STUDIO	12
รูปที่ 2.5 ตัวอย่างแผนภูมิกล่องจากเครื่องมือ FUSION CHART	13
รูปที่ 2.6 ตัวอย่างแผนภูมิวงกลม จากเครื่องมือ R STUDIO	14
รูปที่ 2.7 ตัวอย่างแผนภูมิกระจายจากเว็บไซต์ WWW.STAT.YALE.EDU	14
รูปที่ 2.8 ตัวอย่างแผนภูมิเส้น จากเครื่องมือ R STUDIO	15
รูปที่ 3.1 ภาพแบบขั้นตอนการศึกษาและออกแบบแบบจำลอง	43
รูปที่ 3.2 ตัวอย่างชุดชื่อรายงานที่ดาวนโหลดมา	45
รูปที่ 3.3 การวิเคราะห์คุณลักษณะของการตั้งชื่อรายงาน (REPORT'S TITLE ANALYSIS)	46
รูปที่ 3.4 ภาพหน้าจอเครื่องมือ LEXTO สำหรับการตัดคำ	46
รูปที่ 3.5 ตัวอย่างผลการตัดคำด้วยเครื่องมือ LEXTO	47
รูปที่ 3.6 ตัวอย่างผลจากการตัดคำหยุด	48
รูปที่ 3.7 การจัดกลุ่มประเภทข้อมูลของแถวข้อมูล (COLUMN'S NAME CATEGORIZATION)	51
รูปที่ 3.8 แบบจำลองการสร้างรายงาน (REPORT GENERATION MODEL)	53
รูปที่ 3.9 ตัวอย่างชื่อแถวจากรายงานที่ผู้สร้างรายงานอัปโหลดเข้าสู่ระบบ	55
รูปที่ 3.10 ตัวอย่างชื่อแถวที่ผ่านการตรวจสอบประเภทข้อมูลโดยใช้หัวตาราง	55
รูปที่ 3.11 ตัวอย่างคำสำคัญที่ดึงจากตารางคำสำคัญ	56
รูปที่ 3.12 การสร้างชื่อรายงานด้วยการผสมคำสำคัญ (REPORT TITLE GENERATION)	57
รูปที่ 3.13 การคำนวณข้อมูลจากรูปแบบแผนภูมิที่ถูกเลือก (DATA COMPUTATION)	59
รูปที่ 3.14 การสร้างรายงานตามรูปแบบรายงานที่ผู้สร้างรายงานสนใจ (REPORT GENERATION)	59
รูปที่ 4.1 แผนภูมิแสดงค่าความคล้ายกันของข้อมูล	63
รูปที่ 4.2 แผนภูมิแสดงค่าความคล้ายกันของข้อมูลของคำสำคัญในการทดลองที่ 1 สมมติฐานที่ 2	66
รูปที่ 4.3 แผนภูมิแสดงค่าความคล้ายกันของข้อมูลของคำสำคัญในการทดลองที่ 2	70



## บทที่ 1

### บทนำ

#### 1.1 ความเป็นมาของปัญหาและปัญหา

รายงาน คือ เอกสารที่สร้างขึ้นเพื่อสรุปรวมข้อมูลสำคัญ โดยมีจุดประสงค์เพื่อให้ผู้อ่านรายงานสามารถเห็นภาพรวมของข้อมูลและนำไปใช้ได้ ในด้านธุรกิจนั้น รายงานเป็นหนึ่งในเครื่องมือที่ใช้ในการพิจารณาทิศทางของธุรกิจที่ดำเนินมาแล้ว กำลังดำเนิน และรอดำเนินการ ดังนั้น รายงานจึงถือเป็นสิ่งสำคัญที่จะช่วยผลักดันธุรกิจให้เข้าไปในทิศทางที่ถูกต้อง ในความเป็นจริง การสร้างรายงานขึ้นมาจำเป็นต้องรวบรวมข้อมูล ออกแบบรูปแบบรายงาน วิเคราะห์ สรุปผล และนำเสนอออกมาตามที่ได้ออกแบบไว้ ซึ่งต้องอาศัยผู้เชี่ยวชาญในการสร้างรายงาน ส่งผลให้กลุ่มธุรกิจบางกลุ่มอาจมองข้ามความสำคัญของรายงาน ในงานวิจัยนี้ผู้วิจัยได้ตระหนักถึงความสำคัญของรายงาน จึงได้นำเสนอแบบจำลองที่ช่วยอำนวยความสะดวก และลดความยุ่งยากในการสร้างรายงานให้กับผู้สร้างรายงาน โดยผู้วิจัยมองว่า “ชื่อ” เป็นสิ่งที่สื่อถึงคุณลักษณะของข้อมูลได้ดี จึงได้ออกแบบขั้นตอนวิธีที่จะนำคำสำคัญมาผสมกันเพื่อสร้างเป็นชื่อรายงาน และเรียงลำดับของชื่อที่น่าจะเหมาะสมกับข้อมูล โดยการเรียงลำดับค่าความใกล้เคียงของชื่อรายงานที่เกิดจากการผสมคำสำคัญกับชื่อรายงานต้นฉบับ โดยแบบจำลองจะแนะนำชื่อของรายงานที่เกี่ยวข้องกับโครงสร้างของข้อมูลต้นฉบับเหล่านั้นเพื่อเป็นแนวทางในการสร้างรายงานให้กับผู้สร้างรายงาน

## 1.2 คำถามวิจัย

- 1.2.1 โครงสร้างของข้อมูลมีความสัมพันธ์กับชื่อของรายงานหรือไม่?
- 1.2.2 แบบจำลองที่พัฒนาขึ้นสามารถใช้ในการสร้างรายงานได้หรือไม่?

## 1.3 วัตถุประสงค์

- 1.3.1 เพื่อศึกษาความสัมพันธ์ระหว่างโครงสร้างข้อมูลและชื่อรายงาน
- 1.3.2 เพื่อออกแบบแบบจำลองในการสร้างรายงานโดยใช้โครงสร้างของข้อมูลและชื่อของรายงานเป็นที่ตั้ง

## 1.4 สมมติฐาน

- 1.4.1 โครงสร้างของข้อมูลมีความสัมพันธ์กับสารสนเทศในรายงาน
- 1.4.2 แบบจำลองที่พัฒนาขึ้นสามารถใช้ในการสร้างรายงานได้

## 1.5 ความสำคัญและประโยชน์ของงานวิจัย

แบบจำลองอำนวยความสะดวกให้กับผู้สร้างรายงาน ซึ่งผู้สร้างรายงานสามารถสร้างรายงานภายใน 4 ขั้นตอน คือ

- 1.5.1 ผู้สร้างรายงานเลือกข้อมูลต้นฉบับเพื่อใช้ในการสร้างรายงาน
- 1.5.2 ผู้สร้างรายงานเลือกแถวที่ใช้ในการสร้างรายงาน
- 1.5.3 ผู้สร้างรายงานเลือกแผนภูมิที่ใช้แสดงข้อมูลในรายงาน
- 1.5.4 แบบจำลองสร้างรายงานตามที่คุณสร้างรายงานต้องการ

## 1.6 ขอบเขตของงานวิจัย

1.6.1 ข้อมูลสำคัญที่ใช้ในการสร้างชื่อรายงานรวบรวมจากเว็บไซต์ [www.data.go.th](http://www.data.go.th) ซึ่งเป็นรายงานของภาครัฐในประเทศไทย โดยรายงานเหล่านั้นมีจุดประสงค์ในการกระจายข้อมูลให้กับประชาชน ภาคธุรกิจ และหน่วยงานของรัฐ เพื่อลดความเหลื่อมล้ำทางข้อมูล

1.6.2 รายงานต้นฉบับที่ใช้ในการทดสอบนำมาจากหมวดหมู่สังคมและสวัสดิการ เนื่องจากมีคลังรายงานขนาดใหญ่ที่สุดในหมวดหมู่ทั้ง 8 หมวดหมู่ และชื่อของรายงานมีคำศัพท์เฉพาะสูงสุด 3 คำ โดยมักเป็นชื่อหน่วยงาน หรือชื่อสถานที่

- 1.6.3 รูปแบบแผนภูมิในรายงานมี 3 รูปแบบ คือ
  - 1.6.3.1 แผนภูมิมวงกลม
  - 1.6.3.2 แผนภูมิเส้น
  - 1.6.3.3 แผนภูมิแท่ง

## บทที่ 2

### การทบทวนวรรณกรรมและทฤษฎีที่เกี่ยวข้อง

#### 2.1 รายงาน

##### 2.1.1 ความหมายของรายงาน

รายงาน คือ งานเขียนที่มีจุดประสงค์ และมีกลุ่มเป้าหมายชัดเจน เป็นเอกสารใช้สำหรับแสดงหลักฐานที่ผ่านการวิเคราะห์ เพื่อใช้สำหรับแก้ปัญหาหรือวางแผน โดยรายงานจะแสดงข้อมูลที่สื่อถึงจุดประสงค์ และสามารถทำความเข้าใจได้ง่าย การสร้างรายงานนั้น จำเป็นต้องมีความชัดเจนในเรื่องของขั้นตอนการวิเคราะห์ข้อมูลและลำดับการแสดงผล ผู้สร้างรายงานต้องทราบถึงข้อมูลที่ต้องการสื่อ กลุ่มเป้าหมายของรายงาน จึงจะสามารถสร้างรายงานได้ โดยรายงานมีจุดประสงค์หลัก 2 ข้อ คือ

- 1) เพื่อสรุปข้อมูลที่ผู้สร้างรายงานได้รับจากการสังเกตทดลอง
- 2) เพื่อสร้างสื่อที่ใช้อธิบายข้อมูลที่ผู้สร้างรายงานได้ค้นพบ

##### 2.1.2 หัวใจหลักของการสร้างรายงาน

การสร้างรายงานที่มีประสิทธิภาพนั้น ผู้สร้างรายงานจะต้องเข้าใจหัวใจของการสร้างรายงาน ซึ่งประกอบด้วย 6 ข้อ ดังนี้

2.1.2.1 เข้าใจจุดประสงค์และค่านึงถึงจุดประสงค์ของรายงาน

2.1.2.2 รวบรวม ประเมิน และวิเคราะห์ทุกข้อมูลที่เกี่ยวข้อง

2.1.2.3 จัดองค์ประกอบของรายงานอย่างมีเหตุผล

2.1.2.4 ออกแบบรูปแบบของรายงานในทิศทางเดียวกัน และเป็นรูปแบบที่

สอดคล้องกับจุดประสงค์ของรายงาน

2.1.2.5 เขียนอภิปรายข้อมูล โดยอ้างอิงจากข้อมูลที่ได้จากการสร้างรายงาน

2.1.2.6 สร้างรายงานด้วยความรอบคอบ ในกรณีที่มีหลักเกณฑ์กำหนดไว้  
จะต้องสร้างรายงานให้ถูกต้องตามหลักเกณฑ์นั้น

### 2.1.3 ลำดับการสร้างรายงาน

การสร้างรายงานเบื้องต้นประกอบด้วยขั้นตอนดังนี้

2.1.3.1 ทำความเข้าใจจุดประสงค์ของรายฉบับที่กำลังจะสร้าง คือ สิ่งที่สำคัญที่สุดในการสร้างรายงาน ผู้สร้างรายงานต้องมั่นใจว่าเข้าใจจุดประสงค์ที่แท้จริงของการสร้างรายงาน รวมทั้งทราบแน่ชัดว่าใครคือผู้อ่านรายงานฉบับนี้

2.1.3.2 รวบรวมข้อมูลสำคัญที่ใช้ในการสร้างรายงาน เป็นขั้นตอนของการรวบรวมข้อมูลที่สัมพันธ์กับรายงานที่จะสร้าง ซึ่งข้อมูลอาจจะมาจากหลายแหล่ง หรือแหล่งเดียวก็ได้ ขึ้นอยู่กับระดับความลึกซึ้งของข้อมูลของรายงาน

2.1.3.3 จัดการข้อมูลที่มีอยู่ เป็นการเลือกข้อมูลว่าข้อมูลใดเป็นข้อมูลที่สามารถหาข้อมูลที่ต้องการได้ โดยเริ่มจากการจัดกลุ่มข้อมูลเป็นกลุ่ม ๆ แล้วดูความเชื่อมโยงระหว่างข้อมูลกับจุดประสงค์ของรายงาน การเรียงลำดับความสำคัญจะช่วยให้การเลือกข้อมูลเพื่อใช้ในการสร้างรายงานง่ายขึ้น

2.1.3.4 วิเคราะห์ข้อมูลที่มีอยู่ เป็นการวิเคราะห์ข้อมูลที่ได้รวบรวมไว้เพื่อหาข้อมูลสำคัญที่จะตอบจุดประสงค์ของรายงานได้ รวมทั้งเป็นการหาข้อมูลที่มีความขัดแย้งต่อผลลัพธ์ที่ควรจะเป็น

2.1.3.5 สร้างรายงาน เป็นการสร้างรายงาน โดยหากเป็นการเขียนรายงานเชิงวิชาการ รายงานควรประกอบด้วย 5 ส่วน ดังนี้

- 1) บทนำ เป็นส่วนที่ใช้แสดงถึงจุดประสงค์ของรายงาน
- 2) ส่วนขยายความ เป็นส่วนที่ใช้ขยายความจุดประสงค์แต่ละข้อ รวมถึงให้คำจำกัดความเกี่ยวกับส่วนสำคัญที่จะเกิดขึ้นในรายงาน
- 3) ส่วนข้อมูล เป็นส่วนที่ใช้แสดงข้อมูลสำคัญที่ช่วยยืนยันผลลัพธ์
- 4) ส่วนอธิบายข้อมูล เป็นส่วนที่ใช้อธิบายความสำคัญของข้อมูล
- 5) ส่วนสรุปข้อมูล เป็นส่วนที่ใช้แสดงผลลัพธ์จากการศึกษา

2.1.3.6 พิจารณารายงาน เป็นการพิจารณารายงานที่สร้าง กรณีที่เป็นแบบร่างให้ตรวจสอบข้อความให้ถูกต้อง โดยการพิจารณานั้นจะต้องพิจารณาจากหลายมุมมอง

2.1.3.7 พิจารณารูปแบบรายงาน เป็นการพิจารณารูปแบบรายงาน เพื่อให้รายงานมีข้อมูลที่ถูกต้องที่สุด การพิจารณาในขั้นสุดท้าย ต้องตรวจสอบทุกแง่มุม ทั้งความถูกต้องของข้อมูล การแสดงผลข้อมูล รวมถึงการสะกดคำให้ถูกต้อง

## 2.2 ข้อมูลและการเก็บข้อมูล

### 2.2.1 ข้อมูล

ข้อมูล หมายถึง สิ่งที่เกี่ยวข้องกับทั้งสิ่งมีชีวิตและไม่มีชีวิต อาจถูกบันทึกอยู่ในรูปแบบของตัวเลข ตัวอักษร เช่น ข้อมูลปริมาณพนักงานราชการในแต่ละหน่วย ข้อความบันทึกความคิดเห็นของผู้ร่วมงาน โดยข้อมูลนั้นสามารถแบ่งได้หลายวิธี ยกตัวอย่างเช่น

#### 2.2.1.1 การแบ่งประเภทข้อมูลตามแหล่งที่มา

1) ข้อมูลปฐมภูมิ (Primary Data) เป็นข้อมูลที่เกิดจากการเก็บรวบรวมข้อมูลด้วยตนเองจากแหล่งข้อมูล เช่น การสัมภาษณ์และบันทึกข้อมูล การสังเกต

2) ข้อมูลทุติยภูมิ (Secondary Data) เป็นข้อมูลที่ไม่ได้เกิดจากการรวบรวมเอง แต่เป็นการนำข้อมูลที่ถูกรวบรวมไว้แล้วมาใช้งาน

#### 2.2.1.2 การแบ่งประเภทข้อมูลตามวิธีการวัด

1) มาตรวัดนามบัญญัติ (Nominal Scale) เป็นการแบ่งกลุ่มข้อมูลเป็นกลุ่มย่อย ตามลักษณะที่ปรากฏ เช่น เพศ ศาสนา เขตการปกครอง

2) มาตรวัดอันดับ (Ordinal Scale) เป็นการแบ่งกลุ่มข้อมูลที่แสดงถึงระดับคะแนน สามารถเปรียบเทียบระหว่างกลุ่ม ว่ากลุ่มไหนดีกว่าได้ เช่น ลำดับที่ 1 คนชอบกินไก่ ลำดับที่ 2 คนชอบกินหมู ลำดับที่ 3 คนชอบกินปลา

3) มาตรวัดแบบช่วง (Interval Scale) เป็นการแบ่งกลุ่มเป็นช่วง ๆ โดยแต่ละช่วงมีความห่างของข้อมูลเท่ากัน ค่าของข้อมูลประเภทนี้ เป็นค่าสมมุติ ใช้ในการบ่งบอกปริมาณเพื่อการเปรียบเทียบ โดยค่าที่เป็น 0 ไม่ได้หมายความว่า ไม่มีข้อมูล เช่น อุณหภูมิ

4) มาตรการอัตราส่วน (Ratio Scale) เป็นการแบ่งกลุ่มเป็นช่วง ๆ เช่นเดียวกับมาตรวัดแบบช่วง แต่ข้อมูลประเภทนี้จะบันทึกค่าของข้อมูลจริง ๆ โดยค่า 0 หมายความว่า ไม่มีค่า เช่น ความสูง เวลา ระยะทาง

#### 2.2.1.3 การแบ่งประเภทข้อมูลตามลักษณะของข้อมูล

1) ข้อมูลเชิงปริมาณ (Quantitative Data) เป็นข้อมูลที่สามารถพิจารณาได้ว่า ค่ามีมากหรือน้อย โดยแบ่งเป็น 2 ประเภท คือ

1.1) ประเภทต่อเนื่อง (Continuous Data) เป็นข้อมูลตัวเลขที่มีค่าเป็นจำนวนจริง โดยค่าของข้อมูลจะเป็นไปได้ทุกค่า เช่น เงินเดือน ส่วนสูง

1.2) ประเภทไม่ต่อเนื่อง (Discrete Data) เป็นข้อมูลตัวเลขที่มีค่าเป็นจำนวนเต็ม เช่น จำนวนคน จำนวนสุนัข

2) ข้อมูลเชิงคุณภาพ (Qualitative Data) เป็นข้อมูลที่ใช้แบ่งข้อมูลเป็นกลุ่ม โดยไม่สามารถพิจารณาค่าของข้อมูลได้ชัดเจน เช่น คุณภาพสินค้า ความยากง่ายของข้อสอบ

#### 2.2.2 การเก็บรวบรวมข้อมูล

การเก็บข้อมูล แบ่งต้นแบ่งได้เป็น 3 วิธี ดังนี้

2.2.2.1 การเก็บรวบรวมข้อมูลจากการรายงาน การเก็บข้อมูลวิธีนี้ มักเป็นการเก็บข้อมูลแบบทุติยภูมิ โดยการเก็บข้อมูลจากรายงานที่ออกมาจากระบบหรือเอกสารประกอบการทำงาน เพื่อนำข้อมูลเหล่านั้นมาประมวลผล เพื่อต่อยอดความรู้

2.2.2.2 การเก็บรวบรวมข้อมูลด้วยการสำรวจและแจกจ่าย การเก็บข้อมูลแบบนี้ มักเป็นการเก็บข้อมูลแบบปฐมภูมิ โดยแบ่งออกเป็น 2 ประเภท คือ

1) การสำรวจประชากร เป็นการเก็บข้อมูลแบบละเอียด โดยการเก็บข้อมูลจากทุกส่วนที่ผู้เก็บข้อมูลสนใจ เช่น การสัมภาษณ์ผู้เข้าสัมภาษณ์งานทั้งหมด เพื่อเก็บข้อมูลประจำตัว การเก็บข้อมูลจากหน่วยงานทั้งหมดในเทศบาลเมืองภูเก็ต

2) การสำรวจตัวอย่าง เป็นการเก็บข้อมูลจากตัวอย่างกลุ่มที่สนใจ โดยสุ่มเลือกจากกลุ่มประชากร หรือเลือกโดยมีเงื่อนไข เช่น สุ่มคะแนนสอบนักศึกษา 15 คนจากนักศึกษาสาขาวิชาเทคโนโลยีสารสนเทศ เพื่อหาแนวโน้มคะแนนสอบของนักศึกษาสาขาวิชาดังกล่าว

2.2.2.3 การเก็บรวบรวมข้อมูลจากการทดลอง เป็นการเก็บข้อมูลที่ต้องมีกลุ่มเป้าหมายที่ต้องการเก็บ อาจจะมีหนึ่งกลุ่ม หรือมากกว่า มักใช้กับการเก็บข้อมูลเพื่อนำไปคำนวณทางสถิติ เช่น การเก็บข้อมูลจากผู้ป่วยที่มีภาวะเสี่ยงต่อโรคหัวใจ

### 2.2.3 ประเภทข้อมูล

งานวิจัยนี้อ้างอิงประเภทข้อมูลจากข้อกำหนดประเภทข้อมูล ฉบับปี ค.ศ. 2018 ที่ใช้งานในระบบฐานข้อมูล MySQL 8.0 (ORACLE, 2018) โดยในข้อกำหนดประเภทข้อมูลได้กำหนดประเภทข้อมูลไว้ 5 ประเภท ดังนี้

#### 2.2.3.1 ข้อมูลชนิดตัวเลข (Numeric Types) แบ่งเป็น 3 ชนิด

- 1) จำนวนเต็ม (Integer Value) เช่น 1 2 3
- 2) ทศนิยมแบบค่าจริง (Fixed – Point Exact Value) เช่น 1.005
- 3) ทศนิยมแบบประมาณค่า (Approximate Value) เป็นค่าที่เกิดจากการประมาณ เช่น ค่าพาย ประมาณ 3.14

#### 2.2.3.2 ข้อมูลชนิดวันและเวลา (Date and Time Types) แบ่งเป็น 2 ชนิด

- 1) วัน วันที่ และ แสตมป์เวลา (Date, Datetime and Timestamp Value) เช่น 2018-06-22 1970-01-01 00:00:01
- 2) เวลา (Time Value) เช่น 00:00:01

#### 2.2.3.3 ข้อมูลอักขระ (String Types) แบ่งเป็น 3 ชนิด

- 1) อักขระ (Characters Value) เช่น ชื่อ นามสกุล ตัวอักษร
- 2) ข้อความ (Text Value) เป็นชุดของอักขระหลายชุดมาประกอบกัน เช่น รายละเอียด คำอธิบาย
- 3) ข้อความแจกแจงกลุ่ม (Enumerated Value) เป็นอักขระที่มีการแจกแจงเป็นกลุ่มได้ เช่น เล็ก กลาง ใหญ่

2.2.3.4 ข้อมูลเชิงพื้นที่ (Spatial Data Types) เป็นข้อมูลเกี่ยวกับภูมิศาสตร์ โดยรองรับคุณสมบัติต่าง ๆ ตามที่ OGC (The Open Geospatial Consortium) กำหนด เช่น การบันทึกข้อมูลภูเขา เมือง ถนน สี่แยก



2.2.3.5 ข้อมูลสัญกรณ์วัตถุจาวาสคริปต์ (JSON : Javascript Object Notation) เป็นข้อมูลตามมาตรฐาน RFC 4627<sup>1</sup> ซึ่งเก็บข้อมูลเป็นข้อมูลอักษรที่ถูกจัดเป็นหมวดหมู่ มีขนาดเล็ก

## 2.3 แผนภูมิ

แผนภูมิ เป็นหนึ่งในเครื่องมือสำคัญที่ใช้สำหรับอธิบาย หรือ สรุปข้อมูลที่มีความซับซ้อนให้อยู่ในรูปแบบที่อ่านได้ง่ายมากขึ้น สามารถใช้แทนตารางเพื่อนำเสนอข้อมูลในลักษณะรูปภาพขนาดเล็ก โดยแผนภูมิมียุคหลายประเภท ซึ่งแต่ละประเภทมีความเหมาะสมกับรูปแบบข้อมูลแตกต่างกัน งานวิจัยนี้ ผู้วิจัยได้ทำการศึกษาแผนภูมิในมุมมองต่าง ๆ ตามหัวข้อ ดังนี้

### 2.3.1 ประเภทของแผนภูมิ

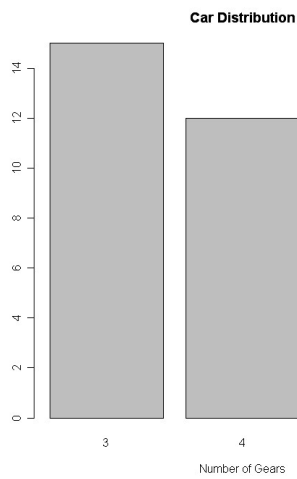
ส่วนนี้ผู้วิจัยจะอธิบายเกี่ยวกับแผนภูมิ ซึ่งเป็นเครื่องมือในการนำเสนอข้อมูลที่เป็นส่วนสำคัญในรายงาน โดยแผนภูมิที่ใช้งานทั่วไปนั้นมี 7 ประเภท (DataCamp, 2017; IBM Knowledge Center, 2018; ReportLab, 2018) แต่ละประเภทมีจุดประสงค์ในการแสดงข้อมูลต่างกัน ดังนี้

#### 2.3.1.1 แผนภูมิแท่ง (Bar chart)

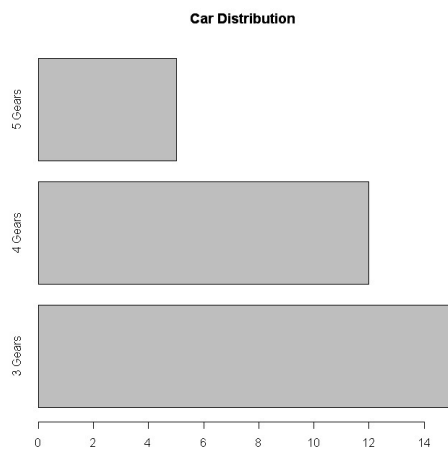
แผนภูมิแท่ง คือ แผนภูมิที่มีจุดประสงค์เพื่อให้ผู้อ่านรายงานสามารถเห็นความแตกต่างของข้อมูล ทั้งในรูปแบบของค่าตัวเลข ความถี่ รวมทั้งรูปแบบอื่น ๆ ที่เป็นข้อมูลแบบไม่ต่อเนื่อง ความสูงของแท่งกราฟแต่ละแท่งแสดงให้เห็นถึงปริมาณของข้อมูลแต่ละกลุ่ม ในพื้นฐานแกนนอนจะเป็นข้อมูลที่สื่อถึงกลุ่มของข้อมูล เช่น ชื่อกลุ่ม ราคา ส่วนแกนตั้งเป็นค่าสำหรับบ่งชี้ปริมาณข้อมูล เช่น ตัวเลข ลำดับชั้น ทั้ง 2 แกนสามารถสลับกันได้ขึ้นอยู่กับความเหมาะสมระหว่างจำนวนกลุ่ม และตัวบ่งชี้ปริมาณข้อมูล (DataCamp, 2017)

<sup>1</sup> <https://tools.ietf.org/html/rfc4627>

โดยรูปที่ 2.1 แสดงให้เห็นปริมาณของรถยนต์ แยกตามปริมาณเกียร์ที่รถคนนั้นมี โดยมีแกนนอนเป็นจำนวนเกียร์ ซึ่งเป็นข้อมูลที่อยู่ถึงกลุ่มของรถยนต์ และมีแกนตั้งเป็นจำนวนรถยนต์ในแต่ละกลุ่ม ทั้งนี้ยังสามารถแสดงแผนภูมิในแนวนอน ดังเช่นรูปที่ 2.2 โดยขึ้นอยู่กับขนาดของข้อมูลและความแตกต่างกันของข้อมูล



รูปที่ 2.1 ตัวอย่างแผนภูมิแท่งแนวตั้ง จากเครื่องมือ R Studio

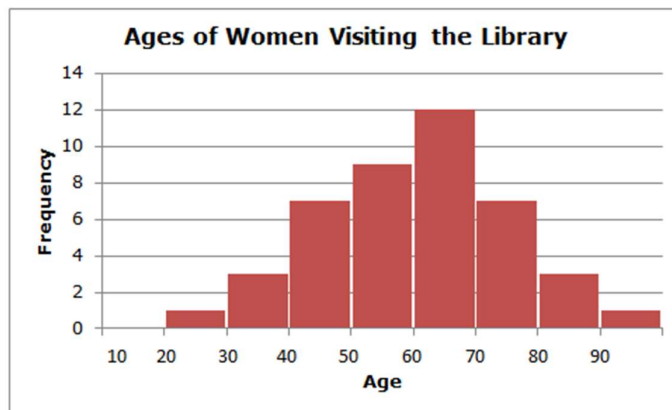


รูปที่ 2.2 ตัวอย่างแผนภูมิแท่งแนวนอน จากเครื่องมือ R Studio

### 2.3.1.2 ฮิสโตแกรม (Histogram chart)

ฮิสโตแกรม คือ แผนภูมิที่แสดงข้อมูลเป็นรูปแบบแท่ง เช่นเดียวกับแผนภูมิแท่ง โดยมีสิ่งที่แตกต่างกัน คือ ฮิสโตแกรมจะใช้นำเสนอข้อมูลประเภทไม่ต่อเนื่อง เช่น แผนภูมิแสดงช่วงอายุของคนที่ชอบดูรายการโทรทัศน์ในช่วงเก้าโมงถึงเที่ยง จะเห็นว่าข้อมูลอายุนั้นมีความแตกต่างกันอย่างมากเพราะเป็นไปได้หลายค่า หากต้องแสดงทั้งหมดอาจจะส่งผลให้แผนภูมิอ่านได้ยาก ฉะนั้นจึงจำเป็นต้องจัดข้อมูลให้อยู่ในรูปแบบของช่วงชั้นอายุ (DataCamp, 2017)

ซึ่งรูปที่ 2.3 แสดงตัวอย่างของฮิสโตแกรม โดยมีข้อมูล คือ จำนวนผู้หญิงที่ใช้บริการห้องสมุด โดยแบ่งตามช่วงอายุ ความสูงของแท่งข้อมูลแสดงถึงความถี่ของจำนวนคน

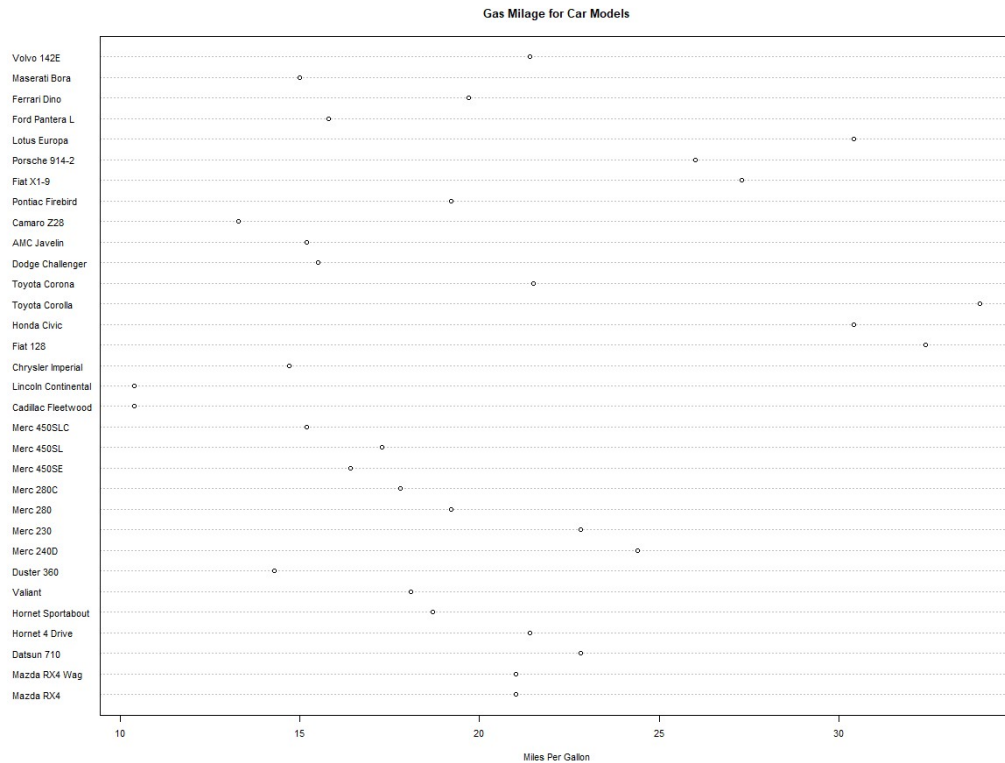


รูปที่ 2.3 ตัวอย่างฮิสโตแกรมจากเว็บไซต์ [www.softschools.com](http://www.softschools.com)

### 2.3.1.3 แผนภูมิจุด (Dot chart)

แผนภูมิจุด คือ แผนภูมิที่แสดงความหนาแน่นของข้อมูลเป็นจุดบนแผนภูมิ โดยข้อมูลที่เหมาะสมกับแผนภูมิจุด คือ ข้อมูลต่อเนื่อง ข้อมูลเชิงปริมาณ และข้อมูลตัวแปรเดียว ซึ่งขนาดของข้อมูลทั้งหมดไม่ควรเกิน 30 ข้อมูล เนื่องจากจุดที่แสดงบนแผนภูมิ หากมีปริมาณเยอะ จะทำให้การอ่านผลเกิดความผิดพลาดได้ จุดเด่นของแผนภูมิจุด คือ ทำให้ผู้อ่านแผนภูมิสามารถเห็นค่าผิดปกติได้ชัดเจน

ซึ่งรูปที่ 2.4 แสดงตัวอย่างแผนภูมิจุด โดยมีข้อมูล คือ รุ่นของรถยนต์และระยะทางในหน่วยไมล์ที่ปริมาณน้ำมัน 1 แกลลอน สามารถพาไปได้

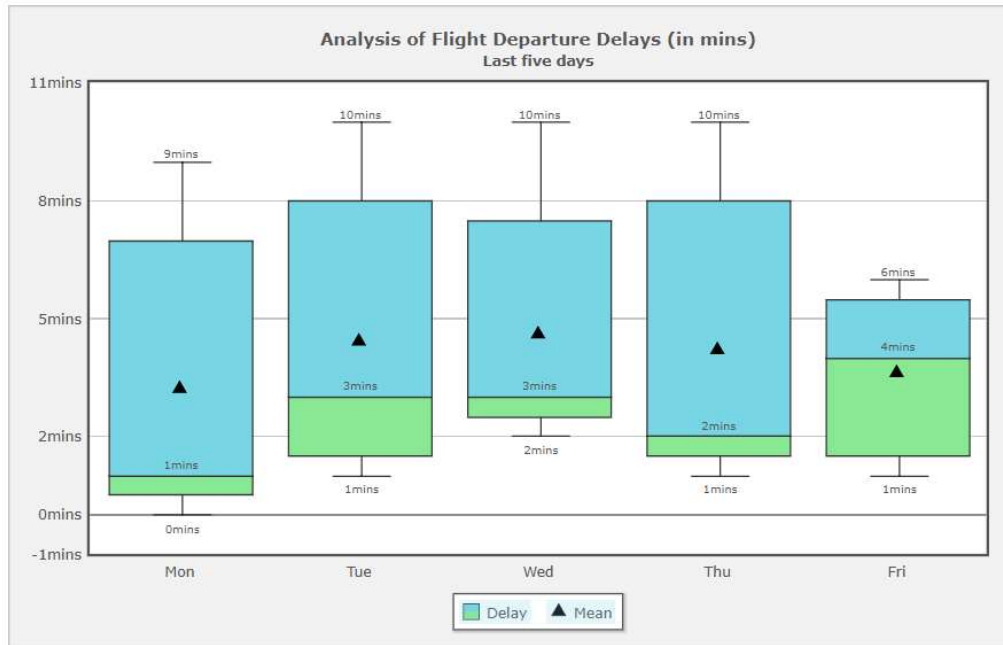


รูปที่ 2.4 ตัวอย่างแผนภูมิจุดจากเครื่องมือ R Studio

#### 2.3.1.4 แผนภูมิกำลัง (Box plot chart)

แผนภูมิกำลัง คือ แผนภูมิที่แสดงข้อมูล 5 ส่วน ประกอบด้วย ค่าต่ำสุด ค่าสูงสุด ค่ามัธยฐาน ค่าควอร์ไทล์ส่วนบน ค่าควอร์ไทล์ส่วนล่าง โดยจุดประสงค์ของแผนภูมิกำลัง คือ ทำให้ผู้อ่านแผนภูมิสามารถเห็นการกระจายและปัญหาของข้อมูลได้ง่าย เหมาะสำหรับการเปรียบเทียบข้อมูลระหว่างกลุ่ม

โดยรูปที่ 2.5 คือ ตัวอย่างของแผนภูมิกำลังที่แสดงให้เห็นความแตกต่างของข้อมูลการมาสายของเครื่องบินในช่วงวันจันทร์ถึงวันศุกร์ ซึ่งแผนภูมิแสดงให้เห็นปริมาณข้อมูล ค่าต่ำสุด ค่าสูงสุด มัธยฐาน ได้ชัดเจน

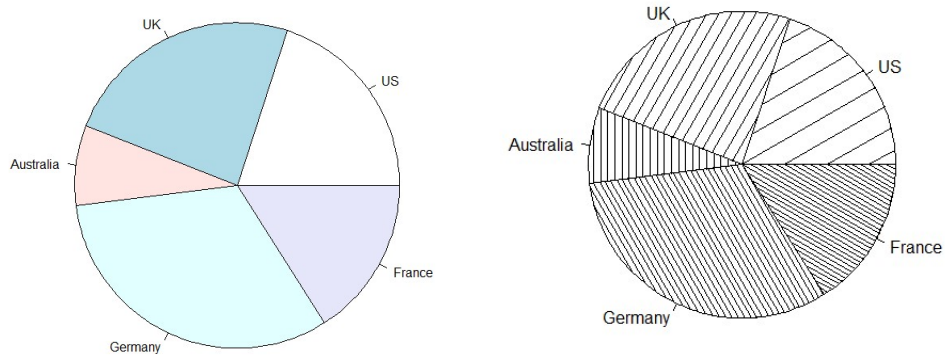


รูปที่ 2.5 ตัวอย่างแผนภูมิกำลังจากเครื่องมือ Fusion chart

### 2.3.1.5 แผนภูมิวงกลม (Pie chart)

แผนภูมิวงกลม คือ แผนภูมิที่มีจุดประสงค์เพื่อให้ผู้อ่านได้เห็นถึงความกระจายของข้อมูล เช่น อัตราส่วนของจำนวนพนักงานแต่ละแผนกของมหาวิทยาลัย โดยแผนภูมิวงกลมนั้นใช้สำหรับแสดงข้อมูลที่สามารถแบ่งแยกเป็นกลุ่มได้อย่างชัดเจนเท่านั้น (Nominal Data) และไม่เหมาะกับข้อมูลที่มีจำนวนกลุ่มมาก (DataCamp, 2017)

ซึ่งรูปที่ 2.6 คือ ตัวอย่างแผนภูมิวงกลม แสดงให้เห็นอัตราส่วนประชากรของประเทศทั้ง 5 ประเทศ โดยใช้สีหรือลายเพื่อให้ผู้อ่านรายงานสามารถเห็นความแตกต่างของข้อมูลได้อย่างชัดเจน

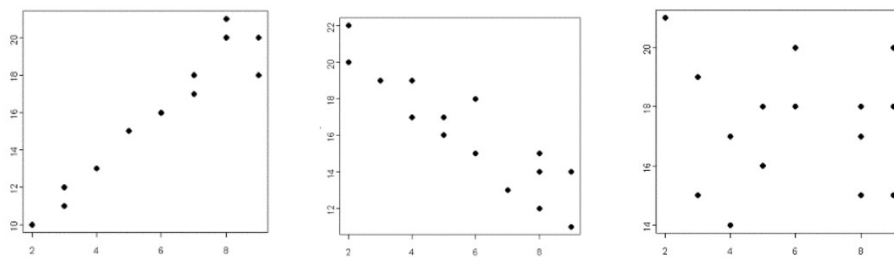


รูปที่ 2.6 ตัวอย่างแผนภูมิวงกลม จากเครื่องมือ R Studio

### 2.3.1.6 แผนภูมิการกระจาย (Scatter plot chart)

แผนภูมิการกระจายมีจุดประสงค์เพื่อแสดงข้อมูลที่ประกอบจาก 2 ตัวแปร มักใช้ในการเปรียบเทียบสหสัมพันธ์เชิงเส้น เพื่อให้เห็นความสัมพันธ์ระหว่างข้อมูล เช่น หากความสัมพันธ์ระหว่างจำนวนปีของประสบการณ์ในการทำงานกับปริมาณงานที่สามารถทำได้ในช่วงเวลาหนึ่ง

โดยรูปที่ 2.7 แสดงให้เห็นตัวอย่างของแผนภูมิกระจาย ซึ่งมีผลลัพธ์ 3 แบบ คือ แบบที่ 1 ข้อมูลมีความสัมพันธ์กันในทางบวก (Positive association) แบบที่ 2 ข้อมูลมีความสัมพันธ์กันในทางลบ (Negative association) และแบบที่ 3 ข้อมูลไม่มีความสัมพันธ์กัน

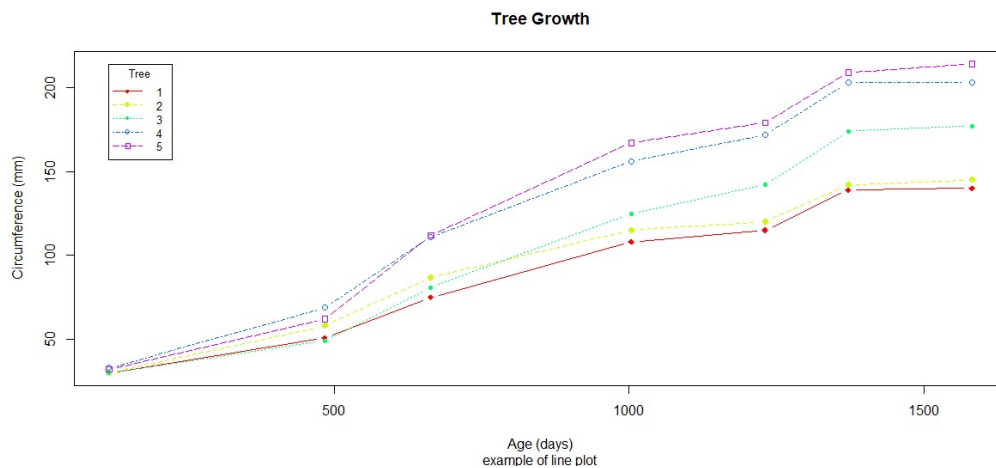


รูปที่ 2.7 ตัวอย่างแผนภูมิกระจายจากเว็บไซต์ [www.stat.yale.edu](http://www.stat.yale.edu)

### 2.3.1.7 แผนภูมิเส้น (Line chart)

แผนภูมิเส้น คือ แผนภาพที่ใช้แสดงข้อมูล ที่มีจะจัดเป็นช่วงเวลา โดยอาจมีข้อมูลตั้งแต่หนึ่งกลุ่มเป็นต้นไป เช่น ปริมาณน้ำฝนในเดือนมกราคม กุมภาพันธ์ และ มีนาคม แผนภูมิเส้นจะช่วยให้ผู้อ่านแผนภูมิสามารถเปรียบเทียบข้อมูลประเภทเดียวกัน ที่เกิดในช่วงเวลาที่ต่างกันได้ง่าย ผู้อ่านสามารถเห็นความเป็นไปของข้อมูลที่เกิดขึ้นได้ ช่วยให้การทำนายสิ่งอาจที่เกิดขึ้นง่ายขึ้น รวมทั้งทำให้เห็นจุดที่ข้อมูลมีการเปลี่ยนแปลง โดยปกติแกนนอนจะสื่อถึงช่วงเวลา และแกนตั้งจะสื่อถึงจำนวนหน่วยของข้อมูลที่เกิดขึ้น แต่ทั้งนี้สามารถสลับกันได้โดยพิจารณาตามความเหมาะสมของข้อมูล (DataCamp, 2017)

ซึ่งรูปที่ 2.8 คือ ตัวอย่างแผนภูมิเส้น แสดงให้เห็นถึงการเจริญเติบโตของต้นไม้ 5 ต้น โดยแกนนอนเป็นค่าช่วงเวลาในหน่วยวัน ส่วนในแกนตั้งเป็นขนาดวงปี มีหน่วยเป็นมิลลิเมตร การให้สีหรือลายทำให้ผู้อ่านแผนภูมิสามารถแยกข้อมูลได้ชัดเจนมากขึ้น



รูปที่ 2.8 ตัวอย่างแผนภูมิเส้น จากเครื่องมือ R Studio

### 2.3.2 ส่วนประกอบของแผนภูมิ

แผนภูมิมี่ส่วนประกอบ 5 ส่วน คือ

#### 2.3.2.1 Plot area

ส่วน Plot area เป็นส่วนหลักที่ใช้ในการแสดงข้อมูล โดยข้อมูลเหล่านั้นบรรจุอยู่ในกรอบข้อมูลของแกนตั้งและแกนนอน

#### 2.3.2.2 The x - axis

The x - axis คือ แกนนอนที่ใช้กำหนดขอบเขตของข้อมูลในแนวนอน โดยพื้นฐานมักจะเป็นค่าของหน่วยในการวัดข้อมูล หน่วยที่ใช้จำแนกข้อมูล หรืออาจจะเป็นกลุ่มของข้อมูล เช่น ตัวเลข อุณหภูมิ หรือ ปี ทั้งนี้สามารถสลับกับแกนตั้งได้ โดยพิจารณาตามความเหมาะสมของชุดข้อมูล

#### 2.3.2.3 The y - axis

The y - axis คือ แกนตั้งที่ใช้กำหนดขอบเขตของข้อมูลในแนวตั้ง โดยปกติจะอยู่ทางด้านซ้ายของแผนภูมิ ยกเว้นในกรณีที่จำนวนข้อมูลมีสองค่าจะอยู่ทั้งด้านซ้าย และด้านขวาของแผนภูมิ ค่าของข้อมูลของแกนตั้งจะเป็นตัวเลข ใช้สำหรับแสดงปริมาณข้อมูลที่เกิดขึ้น เช่น ความถี่ ปริมาณ

#### 2.3.2.4 Gridlines

Gridlines คือ เส้นที่สร้างขึ้นมาเพื่อใช้อ้างอิงตำแหน่งของข้อมูล เพื่อให้ผู้อ่านแผนภูมิใช้เวลาน้อยลงในการอ่านข้อมูล

#### 2.3.2.5 Title

Title คือ ชื่อของแผนภูมิ ซึ่งในการตั้งชื่อเบื้องต้นจะต้องตั้งชื่อให้สามารถสรุปใจความได้



## 2.4 เครื่องมือในการสร้างรายงาน

เครื่องมือในการสร้างรายงาน หมายถึง เครื่องมือที่ใช้สำหรับสกัด และแสดงข้อมูล ในรูปแบบแผนภูมิ ตาราง และอื่น ๆ โดยเครื่องมือสร้างรายงานจะช่วยอำนวยความสะดวกในด้านการเรียงข้อมูล การคัดกรองข้อมูล รวมทั้งการสรุปข้อมูลให้อยู่ในรูปแบบรายงานหน้าเดียวหรือมากกว่า หน้าหนึ่งเพื่อใช้ในงานพิมพ์ เพื่อตอบสนองรูปแบบการใช้งานของผู้อ่านรายงาน ในงานวิจัยนี้ ผู้วิจัยได้ ทำการศึกษาเกี่ยวกับเครื่องมือในการสร้างรายงาน ดังนี้

### 2.4.1 วิธีการทำงานของเครื่องมือในการสร้างรายงาน

การสร้างรายงานจำเป็นต้องมีผู้สร้างรายงาน โดยผู้สร้างรายงานจะต้องทำการ ออกแบบรูปแบบของรายงาน โดยใช้เครื่องมือตามที่ระบบแต่ละระบบได้เตรียมไว้ให้ ยกตัวอย่างเช่น เครื่องมือสร้างรายงาน JReport designer หลังจากออกแบบรูปแบบแล้ว ผู้สร้างรายงานจะต้อง กำหนดข้อมูลสำคัญที่ใช้ในการสร้างรายงาน กำหนดข้อมูลที่จะต้องใช้ ลักษณะการแสดงผลข้อมูล การกรองข้อมูล และอื่น ๆ ที่สำคัญต่อการสร้างรายงาน เครื่องมือสำหรับสร้างรายงานระดับองค์กร (Enterprise Reporting Tools) มักจะมาพร้อมเครื่องมือในการทำงานผ่านเว็บเบราว์เซอร์ ที่มีระบบ แดชบอร์ด (Dashboards) สำหรับการดูข้อมูลที่อยู่ในระบบ สามารถสร้างส่วนประกอบของรายงานไว้ ล่วงหน้า เมื่อต้องการสร้างรายงานสามารถดึงมาใช้งานได้ทันที ซึ่งทำให้การสร้างรายงานที่มีรูปแบบ ซ้ำ ๆ เดิม ๆ สามารถทำได้อย่างรวดเร็ว รวมทั้งในบางเครื่องมือยังมีความสามารถในการสร้างรายงาน โดยตั้งเวลาไว้ล่วงหน้า เช่น ต้องการให้สร้างรายงานการอัปเดตรายการข้อมูลประจำวัน ทุกหกโมงเย็น โดยใช้รูปแบบรายงานแบบตาราง

### 2.4.2 ตัวอย่างเครื่องมือในการสร้างรายงาน

#### 2.4.2.1 SAP Crystal Reports

SAP Crystal Reports เป็นเครื่องมือสร้างรายงาน ที่จะทำให้ผู้สร้างรายงาน สร้างรายงานทั้งในรูปแบบสำหรับงานพิมพ์ รวมทั้งรูปแบบที่สามารถโต้ตอบกับผู้ใช้งานได้ เครื่องมือ รองรับภาษามากถึง 24 ภาษา และรองรับรูปแบบข้อมูลที่หลากหลาย เช่น Native, ODBC,

JDBC connectivity และอื่น ๆ ตัวเครื่องมือมีลักษณะแบบ “What You See Is What You Get” (WYSIWYG) ทำให้ผู้ใช้งานสามารถมองเห็นภาพรวมของรายงานที่กำลังสร้าง ซึ่งเครื่องมือรองรับการนำเข้าการแสดงผลแบบ dynamic charts, Adobe Flash, Adobe Flex และ SAP Crystal Dashboard Design

#### 2.4.2.2 Jasper Reports

Jasper Reports คือ เครื่องมือที่เปิดให้ใช้งานแบบเสรี ตัวเครื่องมือถูกพัฒนาขึ้นด้วยภาษาจาวาซึ่งได้รับความนิยม (Carlos Alonso Vega, 2001) โดยรองรับข้อมูลขาเข้าจากทุกแหล่งข้อมูล และรายงานที่ถูกสร้างจะคมชัดภาพไม่แตกเมื่อขยาย สามารถใช้ทั้งในงานพิมพ์และงานเอกสารอิเล็กทรอนิกส์ เครื่องมือรองรับรูปแบบรายงานหลากหลายรูปแบบ เช่น HTML, PDF, Excel, OpenOffice, Word โดย Jasper Reports ประกอบด้วยชุดเครื่องมือหลายเครื่องมือ เช่น iReport Designer, Jaspersoft Studio, Jasper Reports Server

iReport Designer เป็นเครื่องมือที่อำนวยความสะดวกในการสร้างรายงานที่ซับซ้อน รายงานที่ประกอบด้วยแผนภูมิรูปภาพ โดยเครื่องมือจะอำนวยความสะดวกในการเชื่อมต่อข้อมูลผ่าน JDBC JavaBeans XML CSV เครื่องมือสามารถสร้างรายงานทั้งในรูปแบบ PDF RTF XML CSV HTML XHTML Text DOCX และอีกมากมาย ในปัจจุบัน iReport Designer ได้หยุดพัฒนาความสามารถไป โดยทางทีมพัฒนาได้สร้างเครื่องมือในการสร้างรายงานที่มีชื่อว่า Jaspersoft Studio

Jaspersoft Studio เป็นชุดเครื่องมือที่ใช้ในการสร้างรายงาน โดยมีความสามารถในการสร้างแม่แบบสำหรับรายงาน สร้างคำสั่งในการสร้างรายงาน สร้างคำสั่งในการสร้างรายงานที่ซับซ้อน โดย Jaspersoft Studio มีแผนภูมิ แผนผัง ตาราง และรูปแบบการแสดงผลให้เลือกใช้งานมากกว่า 50 รูปแบบ ช่วยให้ผู้สร้างรายงานสามารถสร้างรายงานทั้งในแบบที่มีความซับซ้อนน้อย เช่น รายงานที่มีรูปแบบสำหรับการพิมพ์สำเร็จรูป ไปจนถึงรายงานที่มีความซับซ้อน เช่น รายงานที่โต้ตอบกับผู้ใช้งานได้ ผู้ใช้งานสามารถเลือกสิ่งที่ต้องการจากหน้าเว็บ และสามารถสร้างรายงานออกมาในรูปแบบของ Open Office, Power Point, RTF, Word, Spread Sheet รวมทั้งสามารถสร้างรายงานในรูปแบบชุดข้อมูล เช่น Text, CSV, JSON, XML เครื่องมือ Jaspersoft Studio รองรับข้อมูลจากหลายแหล่ง เช่น CSV, Hibernate, Jaspersoft, Java, Beans, JDBC, JSON, NoSQL, TIBCO, Spotfire, Information Links, XML รวมทั้งข้อมูลที่ถูกสร้างขึ้นโดยผู้ใช้งาน

เอง โดยที่ผู้ใช้งานสามารถเลือกดาวน์โหลดโปรแกรมในรูปแบบส่วนเสริมสำหรับใช้งานกับ Eclipse หรือ จะดาวน์โหลดโปรแกรมในรูปแบบโปรแกรมสแตนด์ออล (Standalone)

#### 2.4.2.3 Pentaho Business Analytics

Pentaho Business Analytics คือ ชุดเครื่องมือที่ใช้ประมวลผลข้อมูล และสร้างรายงานที่พัฒนาด้วยภาษาจาวา เป็นโอเพนซอร์ซที่รองรับการนำโปรแกรมสำเร็จรูปอื่นมาประยุกต์เข้าด้วยกัน เครื่องมือมีความสามารถในการรวบรวมข้อมูลและวิเคราะห์ข้อมูล โดยสามารถทำงานร่วมกับข้อมูลที่มีขนาดใหญ่ และมีความหลากหลายของข้อมูล Pentaho Business Analytics มาพร้อมชุดเครื่องมือที่อำนวยความสะดวกให้กับผู้ใช้งานในการเตรียมข้อมูล จำแนกข้อมูล ผสมข้อมูลจากหลายแหล่ง รวมถึงเครื่องมือช่วยอำนวยความสะดวกในการเลือกรูปแบบในการแสดงผลของข้อมูล ซึ่งสามารถสร้างรายงานได้หลายรูปแบบ เช่น HTML, Excel, CSV, PDF และ RTF ทั้งยังสามารถใช้งานร่วมกับเครื่องมือเก็บข้อมูลอื่น ๆ เช่น เซ็นเซอร์ และเครื่องมือที่สามารถเชื่อมต่อกับคอมพิวเตอร์ได้

## 2.5 ภาษาไทย

### 2.5.1 คำในภาษาไทย

คำ หมายถึง กลุ่มของพยัญชนะ ที่สามารถออกเสียงเป็นเสียงพยัญชนะ เสียงสระ และเสียงวรรณยุกต์ โดยคำแต่ละคำจะต้องมีความหมาย (ราเนา พลอยขาว, 2549)

ประเภทของคำในหลักภาษาไทยมีทั้งหมด 7 ประเภท คือ

#### 2.5.1.1 คำนาม

เป็นคำที่ใช้เรียกแทนคน สัตว์ สิ่งของ เช่น โต๊ะ หรือนามธรรม เช่น ความไว้วางใจ (สุรางค์ จันทราภรณ์, 2547) โดยที่คำนามสามารถแบ่งได้เป็น 5 ประเภท ดังนี้

1) สามานยนาม หรือค่านามสามัญ เป็นคำที่ใช้แทนสิ่งต่าง ๆ ทั่วไป เช่น หมู เห็ด เป็ด

2) วิสามานยนาม หรือคำชื่อเฉพาะ เป็นชื่อเฉพาะของสิ่งต่าง ๆ ทั้งคน สัตว์ สิ่งของ สถานที่ เช่น ภูเก็ต พังงา

3) สมุหนาม หรือค่านามรวมหมู่ เป็นคำที่ใช้เรียกกลุ่มหรือหมวดหมู่ของสิ่งต่าง ๆ เช่น กอง กลุ่ม หมู่

4) อาการนาม หรือค่านิยมธรรม เป็นการประกอบคำ เช่น ความ หรือ การ เข้ากับค่านาม หรือ คำวิเศษณ์ เพื่อแสดงถึงกิจกรรม หรือสิ่งต่าง ๆ เช่น การเดิน การวิ่ง ความคิด

5) คำลักษณนาม แบ่งเป็น 6 ประเภท ดังนี้

5.1) ลักษณนามบอกชนิด เช่น หล่ง ใช้กับเปียโน ตัว ใช้กับกีตาร์

5.2) ลักษณนามบอกหมวดหมู่ เช่น โขลง ใช้กับช้างเมือง

5.3) ลักษณนามบอกสัณฐาน เช่น แผ่น ใช้กับกระดาษ

5.4) ลักษณนามบอกจำนวนและมาตรา เช่น คู่ ใช้กับรองเท้า

5.5) ลักษณนามบอกอาการ เช่น มวน ใช้กับบุหรี่

5.6) ลักษณนามซ้ำชื่อ เช่น จังหวัด ตำบล

#### 2.5.1.2 คำสรรพนาม

เป็นคำที่ใช้เรียกแทนคน สัตว์ สิ่งของ สามารถใช้ได้หลากหลายบริบท ขึ้นอยู่กับตำแหน่งที่ปรากฏในประโยค (สุรางค์ จันทราภรณ์, 2547) โดยที่คำสรรพนามสามารถแบ่งได้เป็น 7 ประเภท คือ

1) บุรุษสรรพนาม หรือสรรพนามแทนบุคคล ใช้เรียกแทนผู้ส่งสาร ผู้รับสาร หรือผู้ถูกกล่าวถึง เช่น ผม คุณ ท่าน เธอ

2) นิยมสรรพนาม หรือสรรพนามใช้ชี้ระยะ ใช้กล่าวถึงสถานที่ ทั้งในระยะใกล้ กลาง หรือไกล เช่น นี่ นั่น โนน

3) ปฤจฉาสรรพนาม หรือสรรพนามใช้ถาม ใช้ประกอบประโยค คำถาม เช่น ใคร อะไร ไหน

4) อนิยมสรรพนาม หรือสรรพนามบอกความไม่เจาะจง มักใช้ในการกล่าวถึงบุคคล สิ่งของ หรือ สิ่งต่าง ๆ โดยไม่เจาะจง อาจจะเป็นพหูพจน์หรือเอกพจน์ เช่น ใคร ๆ อะไร ๆ

5) วิภาคสรรพนาม หรือสรรพนามบอกความเป็นพรรคพวกเป็นส่วน ใช้ในการแสดงการเกิดเหตุการณ์ต่าง ๆ ที่เกิดเป็นกลุ่ม เช่น เขาวิงกัน

6) ประพันธสรรพนาม ใช้สำหรับกล่าวถึงข้อความ หรือประโยคก่อนหน้า มี 3 คำ คือ ที่ ซึ่ง อัน เช่น คุณครูตินักเรียนที่ไม่ได้ส่งการบ้าน

7) สรรพนามที่ใช้แทนนามที่อยู่ข้างหน้า ใช้สำหรับกล่าวถึงบุคคล สัตว์ สิ่งของ ที่อยู่ข้างหน้า เช่น สมชายเขาไม่ชอบแมว

#### 2.5.1.3 คำกริยา

คำกริยา คือ คำที่ใช้บ่งบอกถึงการแสดงอาการของคำนาม เช่น กิน วิ่ง เดิน นอน โดยคำกริยาสามารถแบ่งได้เป็น 4 ประเภท คือ

1) อกรรมกริยา เป็นคำกริยาที่ไม่ต้องมีกรรมมารับ เช่น ประตูเปิด  
 2) สกรรมกริยา เป็นคำกริยาที่ต้องมีกรรมมารับ เช่น เขาเปิดประตู  
 3) วิกตรรถกริยา เป็นคำกริยาที่ต้องมีส่วนขยายมาขยาย เพื่อให้ได้ใจความที่สมบูรณ์ เช่น หน้าตางั้นคล้ายประตู

4) กริยานุเคราะห์ เป็นคำกริยาที่ใช้ประกอบกับคำกริยาคำอื่น เช่น นักเรียนถูกตี

#### 2.5.1.4 คำวิเศษณ์

เป็นคำที่ใช้อธิบายเสริมคำนามสรรพนาม กริยา เพื่อให้ผู้อ่านข้อความได้เข้าใจถึงสถานการณ์ หรือลักษณะต่าง ๆ ที่เกิดขึ้น (สุรางค์ จันทราภรณ์, 2547) โดยคำวิเศษณ์มีทั้งหมด 9 ประเภท ดังนี้

1) ลักษณะวิเศษณ์ ใช้สำหรับบอกลักษณะของคำ อาจหมายถึงขนาด สี กลิ่น รส ความรู้สึก เช่น เล็ก เร็ว เมื่อประกอบกับคำในรูปประโยค จะได้ตัวอย่างประโยคดังนี้ เขาวิ่งเร็ว เขามีแขนเล็ก

2) กาลวิเศษณ์ ใช้สำหรับขยายความหมายในเรื่องของเวลาว่าเกิดขึ้นในเวลาใด อาจเป็นอดีต ปัจจุบัน อนาคต หรือเป็นเรื่องที่เกิดขึ้นซ้ำ ๆ ยกตัวอย่างเช่น คำว่า เช้า สาย ภายหลัง เมื่ออยู่ในรูปประโยคจะมีลักษณะดังตัวอย่างเช่น เขากินข้าวเช้าสาย

3) สถานวิเศษณ์ ใช้สำหรับแสดงให้เห็นถึงสถานที่ที่เกี่ยวข้องกับคำนั้น ๆ เช่น เหนือ ใต้ ใกล้ ไกล เมื่อนำสถานวิเศษณ์มาประกอบกับคำในรูปประโยค จะได้ดังตัวอย่างเช่น เขามีบ้านใกล้

4) ปริมาณวิเศษณ์ ใช้สำหรับแสดงสถานะปริมาณอันดับ หรือจำนวนของคำนั้น ๆ เช่น มาก น้อย อันดับต้น ๆ หมด เมื่อใช้ในรูปประโยคจะได้ดังตัวอย่างเช่น พนักงานห้าคนกำลังกินข้าว ข้าวแต่ละจานเยอะมาก

5) นิยมวิเศษณ์ เป็นคำวิเศษณ์ที่ใช้เพื่อยืนยันสถานะของคำ เช่น นี่ นั่น เมื่อนำมาใช้ในรูปประโยคจะได้ดังตัวอย่างเช่น ชายคนนี้เป็นคนดี

6) อนิยมวิเศษณ์ ใช้สำหรับการแสดงความหลากหลาย หรือความไม่แน่นอนของคำ เช่น อื่น ๆ ใด ๆ เมื่อนำมาใช้ในรูปประโยคจะได้ดังตัวอย่างเช่น ภูเก็ตเป็นเมืองที่ร้อนมากกว่าจังหวัดอื่น ๆ

7) ปฤจฉาวิเศษณ์ ใช้สำหรับกรณีที่ต้องการคำตอบ เช่น ใครคนไหน อะไร ทำไม ซึ่งมักจะตามหลังคำนามสรรพนามหรือกริยา ตัวอย่างรูปประโยค เช่น นักศึกษาคนใดทำข้อสอบไม่ผ่าน

8) ประติชญาวิเศษณ์ ใช้สำหรับแสดงคำขานรับ เช่น ครับ มักพบในประโยคเกี่ยวกับบทสนทนา ตัวอย่างรูปประโยค เช่น คุณครับ

9) ประติเชวีเศษณ์ ใช้สำหรับการปฏิเสธ หรือใช้เปลี่ยนลักษณะคำให้ตรงกันข้าม เช่น ไม่ อย่า ตัวอย่างรูปประโยค เช่น เดินในที่ร่มแล้วจะไม่ร้อน

#### 2.5.1.5 คำบุพบท

ใช้ประกอบคำนาม คำสรรพนาม หรือคำกริยา เพื่ออธิบายเกี่ยวกับลักษณะตำแหน่ง หรือการทำงานเพิ่มเติม เช่น บน ของ มากกว่า น้อยกว่า ตั้งแต่ โดยหน้าที่ของคำบุพบทมีดังนี้

- 1) อธิบายสถานที่ที่ตั้งของคำนาม ๆ เช่น คนชนบท
- 2) อธิบายเวลาที่เกี่ยวกับคำนาม ๆ เช่น เขาซ้อมดนตรีจนเช้า
- 3) อธิบายความเป็นเจ้าของ เช่น หนังสือเล่มนั้นเป็นของผม
- 4) อธิบายเกี่ยวกับลักษณะอาการ เช่น เราเดินบนริมทะเลที่ทอดยาว

คำบุพบทแบ่งเป็น 2 ประเภท คือ

1) คำบุพบทใช้สำหรับแสดงความสัมพันธ์ระหว่างคำ อาจอยู่ในรูปแบบของคำนามกับคำนาม คำนามกับคำกริยา และอื่น ๆ เช่น ผมซื้อกีตาร์ของเพื่อน

2) คำบุพบทที่ไม่มีความสัมพันธ์กับคำอื่น มักพบในบทประพันธ์ ตัวอย่างคำบุพบทในกรณีต่าง ๆ

1) คำว่า กับ ใช้อธิบายลักษณะของคำที่กำชับว่าเกิดร่วมกัน เช่น ฉันไต่ยืนกับหู

2) คำว่า แก่ ใช้นำหน้ากรรมที่ขยายการกระทำ เช่น นายกมอปโล่แก่นักเรียนเรียนดี

3) คำว่า แต่ ใช้นำหน้ากรรมเช่นเดียวกับ แก่ แต่ใช้ในกรณีที่แสดงความเคารพนับถือ เช่น เหล่านักศึกษาต่างมอบพวงมาลัยแต่อาจารย์ที่เขาเคารพในวันไหว้ครู

4) คำว่า ต่อ ใช้นำหน้าแสดงความเกี่ยวข้องกันหรือต่อเนื่องกัน เช่น นักศึกษายื่นคำร้องต่อหน่วยงานของมหาวิทยาลัย

### ตำแหน่งของคำบุพบทในประโยค

- 1) นำหน้าคำนาม เช่น เขาเขียนเอกสารด้วยปากกา
- 2) นำหน้าคำสรรพนาม เช่น เธอจะอยู่กับฉันตลอดไป
- 3) นำหน้าคำกริยา เช่น เอกสารชุดนี้ใช้ในการประชุมครั้งนี้
- 4) นำหน้าคำวิเศษณ์ เช่น เอกสารเขียนโดยประณีต

### 2.5.1.6 คำสันธาน

เป็นคำที่ใช้สำหรับเชื่อมคำด้วยกัน เช่น และ หรือ เมื่ออยู่ในรูปประโยค จะได้ดังตัวอย่าง เช่น หมาและแมว คำหรือขาว สามารถใช้สำหรับเชื่อมประโยคด้วยกัน เช่น ถ้าเขาขยันเขาคงรวย เขาเป็นผู้ชายและเธอเป็นผู้หญิง คำสันธานยังใช้งานในการแต่งเติมข้อความให้สละสลวย เช่น คำว่า ก็ เมื่ออยู่ในรูปประโยค จะได้ดังนี้ เราทุกคนต่างก็เป็นคนเหมือนกัน โดยคำสันธานแบ่งได้เป็น 4 ประเภท คือ

- 1) ใช้สำหรับเชื่อมข้อความที่มีความหมายคล้ายตามกัน เช่น กับ
- 2) ใช้สำหรับเชื่อมข้อความที่เป็นเหตุเป็นผล เช่น ถ้า...แล้ว ดังนั้นจึง
- 3) ใช้สำหรับเชื่อมข้อความที่ขัดแย้งกัน เช่น แต่ ทว่า แม้
- 4) ใช้สำหรับเชื่อมข้อความที่เป็นไปได้ได้อย่างใดอย่างหนึ่ง เช่น หรือ

### ลักษณะของคำสันธานที่ปรากฏในประโยค

- 1) คำสันธานบางคำมักเกิดคู่กัน เช่น ไม่...ก็ กว่า...ก็
- 2) เกิดระหว่างคำ เช่น สีขาวและสีดำ
- 3) เกิดหลังคำ เช่น ขาวก็ดี ดำก็ดี
- 4) เกิดคร่อมคำ เช่น ถึงเป็นหญิงก็อย่าวางใจ
- 5) เกิดระหว่างประโยค เช่น สัตว์ตัวนั้นคือหมาหรือสัตว์นั้นไม่ใช่หมา
- 6) เกิดหลังประโยค เช่น คนก็ดี สัตว์ก็ดี ล้วนต้องการน้ำ
- 7) เกิดคร่อมประโยค เช่น เขากินเยอะมาก ดังนั้นเขาจึงอ้วนมาก
- 8) คำสันธานแบ่งประโยคออกเป็นสองส่วน แบ่งคำเป็นสองคำ



9) คำบางคำ สามารถเป็นได้ทั้งคำสันธาน และ คำบุพบท เช่น เมื่อ เมื่ออยู่ในรูปประโยค ดังตัวอย่าง เมื่อเวลาเช้าเขาได้กินข้าวต้ม คำว่า เมื่อ ในกรณีนี้ จะเป็นคำบุพบท และ ในตัวอย่าง เราจะได้กินข้าวเมื่อเราเข้าถึงเมืองกรุงเทพ คำว่า เมื่อ ในกรณีนี้ คือ คำสันธาน

10) คำสรรพนาม คำกริยา บางคำ ถือเป็นคำบุพบท โดยสังเกตจาก บริบทของประโยค เช่น ให้ เมื่อใช้ในประโยค เขาร้องเพลงให้ผู้ป่วยคลายเครียด

#### 2.5.1.7 คำอุทาน

ใช้สำหรับสื่ออารมณ์ความรู้สึกหรือตอบสนองต่อการกระทำ เช่น โอ๊ย

### 2.5.2 วลี

วลี คือ กลุ่มคำที่ประกอบด้วยคำตั้งแต่สองคำขึ้นไปมาต่อเรียงกัน เพื่อให้เกิดความหมายมากขึ้น โดยวลีไม่สามารถใช้ได้ถ้าฟังจะต้องทำหน้าที่ใดหน้าที่หนึ่งในประโยค เช่น ประธาน กริยา กรรม เพื่อให้เกิดความหมายที่สมบูรณ์ โดยวลีแบ่งเป็น 7 ประเภท คือ

2.5.2.1 นามวลี เป็นวลีที่มีค่านามเป็นคำขึ้นต้น เช่น ปากกาดำ

2.5.2.2 สรรพนามวลี เป็นวลีที่มีคำสรรพนามเป็นคำขึ้นต้น เช่น ท่านอาจารย์

2.5.2.3 กริยาวลี เป็นวลีที่มีคำกริยาเป็นคำขึ้นต้น เช่น ต้องพยายาม

2.5.2.4 วิเศษณ์วลี เป็นวลีที่มีคำวิเศษณ์เป็นคำขึ้นต้น เช่น ยิ่งใหญ่อลังการ

2.5.2.5 บุพบทวลี เป็นวลีที่มีคำบุพบทเป็นคำขึ้นต้น เช่น ในห้องนั่งเล่น

2.5.2.6 สันธานวลี เป็นวลีที่ใช้เป็นคำสันธาน อาจขึ้นต้นด้วยคำสันธาน หรือไม่ก็ได้ เช่น ถึงแม้ว่า...

2.5.2.7 อุทานวลี เป็นวลีที่ใช้เป็นคำอุทาน อาจขึ้นต้นด้วยคำอุทานหรือไม่ก็ได้ เช่น โอ๊ย!

### 2.5.3 ประโยค

ประโยค คือ การนำคำหลายคำ มาเรียงกันแล้วทำให้เกิดความหมาย ประโยคนั้น ประกอบด้วยภาคประธานและภาคแสดง ใช้สื่อสารได้ทั้งการเขียน การพูด แต่จะแตกต่างกันในเรื่อง การเลือกใช้คำในสถานการณ์ต่าง ๆ โดยประโยคสามารถจำแนกเป็น 3 ประเภทได้ ดังนี้

### 2.5.3.1 เอกภรตประโยชน์

เอกภรตประโยชน์หรือประโยชน์ความเดียว หมายถึง ประโยคที่สื่อถึงความหมายเพียงอย่างเดียว โดยรูปประโยคจะประกอบด้วยภาคประธานและภาคแสดง เช่น นักบินออกหาหินสามารถแบ่งออกเป็น 4 รูปแบบ คือ

- 1) ประโยคขึ้นต้นด้วยผู้กระทำ เช่น เขาทำงานตอนเช้า
- 2) ประโยคขึ้นต้นด้วยผู้ถูกกระทำ เช่น เอกสารถูกเขียนโดยเขา
- 3) ประโยคขึ้นต้นด้วยกริยา เช่น เกิดการก่อการร้ายขึ้นเมื่อคืนนี้
- 4) ประโยคคำสั่งและขอร้อง เช่น กรุณางดสูบบุหรี่

โดยลักษณะของประโยคความเดียวสามารถแบ่งได้หลายลักษณะ เช่น

- 1) ประโยคบอกเล่า เป็นประโยคสำหรับอธิบายเรื่องราวที่เกิดขึ้น เช่น ผมเล่นไวโอลิน
- 2) ประโยคปฏิเสธ เป็นประโยคที่แสดงการไม่ยอมรับ เช่น ผมไม่กินข้าว
- 3) ประโยคคำสั่งและขอร้อง เป็นประโยคที่แสดงสิ่งที่ต้องการให้ผู้อื่นปฏิบัติ เช่น กรุณาถอดรองเท้าก่อนเข้าห้องเรียน
- 4) ประโยคคำถาม เป็นประโยคที่แสดงสิ่งที่ต้องการทราบ เช่น คุณมีน้ำหนักเท่าไร

### 2.5.3.2 อเนกภรตประโยชน์

อเนกภรตประโยชน์หรือประโยคความรวม เป็นประโยคที่มีมากกว่าหนึ่งประโยคในชุดเดียวกันโดยจะเชื่อมด้วยคำสันธาน ซึ่งแบ่งเป็น 3 ประเภท ดังนี้

- 1) ประโยคความรวมแบบคล้ายตาม ลักษณะประโยคที่นำมาเขียนร่วมกันจะมีความเชื่อมโยงกัน เช่น เมื่อเขาโดนเจ้านายดุเขาจึงตั้งใจทำงาน
- 2) ประโยคความรวมแบบขัดแย้ง ลักษณะประโยคจะมีความขัดแย้งกัน เช่น เธอหน้าตาดีแต่เธอนิสัยแย่มาก

3) ประโยคความรวมแบบเลือกอย่างใดอย่างหนึ่ง ลักษณะประโยค จะมีความขัดแย้งกันจะเป็นจริงอย่างใดอย่างหนึ่ง เช่น เธอจะดื่มชาหรือโกโก้

#### 2.5.3.3 สักรประโยค

เป็นประโยคที่มีใจความหลักหนึ่งประโยค แล้วมีประโยคย่อยเสริมอีกหนึ่ง ประโยค ซึ่งประโยคย่อยนั้น เป็นได้ทั้งคำนาม สรรพนาม ประโยคความซ้อนแบ่งเป็น 3 ประเภท คือ

1) นามานุประโยค คือ ประโยคที่ทำหน้าที่แทนคำนาม อาจเป็นส่วน ประธาน กริยา กรรม ก็ได้ เช่น คนคดโกงเป็นคนที่สังกรรม์เกียด

2) คุณานุประโยค คือ ประโยคที่ทำหน้าที่ขยายคำนามหรือ สรรพนาม โดยมีประพันธสรรพนามเป็นตัวเชื่อมประโยค เช่น ผมได้รับไวโอลินราคาแพงที่ดูเก่าแก่ตัว นี้มาจากปู่

3) วิเศษณานุประโยค คือ ประโยคที่ทำหน้าที่ขยายกริยา โดยใช้ ร่วมกับคำประพันธวิเศษณ์ เช่น เขาออกกำลังกายทุกวันเพื่อให้ร่างกายแข็งแรง

#### 2.5.4 โครงสร้างของประโยคในภาษาไทย

โครงสร้างของประโยคสามารถแยกย่อยได้เป็น 7 ส่วน

2.5.4.1 ส่วนประธาน ทำหน้าที่เป็นตัวหลักของประโยคในการแสดงถึงกริยา อาการ ที่เกิดขึ้น โดยประธานสามารถเป็นคำทั้ง 7 ประเภท อาจเป็นวลี หรือประโยค ก็ได้ เช่น เขา สุนัข ไก่ หนังสือ

2.5.4.2 ส่วนขยายประธาน ทำหน้าที่ขยายประธาน เพื่อให้เกิดความชัดเจน มากยิ่งขึ้น เช่น เขาคอนันคนตัวเล็ก ประชาชนจังหวัดภูเก็ต

2.5.4.3 ส่วนกริยา ทำหน้าที่แสดงถึงอาการที่กำลังเกิดขึ้น อาจเป็นการ กระทำของประธาน หรือสิ่งที่เกิดกับประธาน ก็ได้ เช่น เขาคอนันกิน

2.5.4.4 ส่วนขยายกริยา ทำหน้าที่ขยายกริยา เพื่อให้ชัดเจนมากยิ่งขึ้น เช่น เขาคอนันกำลังกิน

2.5.4.5 ส่วนกรรม ทำหน้าที่รองรับการกระทำของประธาน เช่น ช้างเชือก  
นั้นกินอ้อย

2.5.4.6 ส่วนขยายกรรม ทำหน้าที่ขยายกรรม เพื่อให้เกิดความชัดเจน เช่น  
อ้อยแก่

2.5.4.7 ส่วนเชื่อม ทำหน้าที่เสริมความให้กับประโยคให้ชัดเจนยิ่งขึ้น อาจใช้  
เพื่อเชื่อมประโยคสองประโยคด้วยกัน หรือใช้เพื่อเริ่มต้นประโยค เช่น ท่านสุภาพบุรุษและสุภาพสตรี

## 2.6 หลักการทางสถิติเบื้องต้น

ในงานวิจัยนี้ ผู้วิจัยได้เลือกหลักการทางสถิติ (ปริตถารณ์ กาญจนสำราญวงศ์,  
2560) มาเป็นวิธีการในการออกแบบ และทดสอบแบบจำลอง ดังนั้นในส่วนนี้จะเป็นการอธิบาย  
พื้นฐานที่สำคัญในการคำนวณสถิติ

### 2.6.1 คำศัพท์สำคัญทางสถิติ

2.6.1.1 ประชากร (Population) คือ กลุ่มเป้าหมายที่เราสนใจศึกษาทั้งหมด  
โดยแบ่งเป็น 2 ประเภท คือ

1) ประชากรจำกัด คือ จำนวนประชากรที่สามารถสำมะโนได้  
ครบถ้วน เช่น จำนวนพนักงานเงินรายได้ จำนวนสัตว์เลี้ยงในมหาวิทยาลัย

2) ประชากรอนันต์ คือ จำนวนประชากรที่ไม่สามารถสำมะโนได้  
เช่น จำนวนเม็ดทราย

2.6.1.2 กลุ่มตัวอย่าง (Sample) คือ ตัวแทนประชากรขนาดเล็กของกลุ่ม  
ประชากรขนาดใหญ่ ที่ถูกเลือกมาทำการทดลอง ทั้งแบบสุ่ม และแบบมีเงื่อนไข เช่น นักศึกษาถูกสุ่ม  
เพื่อวัดผลการเรียน โดยมาจากสาขา A จำนวน 15 คน สาขา B จำนวน 17 คน

2.6.1.3 ค่าพารามิเตอร์ (Parameter) คือ ค่าต่าง ๆ ที่ได้จากการนับ การ  
คำนวณ ของประชากร

2.6.1.4 ค่าสถิติ (Statistic) คือ ค่าที่ได้จากการนับ การคำนวณ ของกลุ่มตัวอย่าง

2.6.1.5 สถิติพรรณนา คือ การคำนวณทางสถิติ ที่ทำการคำนวณจากประชากร หรือกลุ่มตัวอย่าง โดยค่าสถิติที่ได้เป็นค่าจริง

2.6.1.6 สถิติอนุมาน คือ การคำนวณทางสถิติ ที่ทำการคำนวณจากกลุ่มตัวอย่าง แล้วพิจารณาผลลัพธ์อ้างอิงไปสู่ประชากร

### ตารางที่ 2.1 ตารางแสดงสัญลักษณ์ทางสถิติ

ภาษาไทย	สัญลักษณ์	
	ค่าพารามิเตอร์	ค่าสถิติ
ค่าเฉลี่ย	$\mu$	X
ส่วนเบี่ยงเบนมาตรฐาน	$\Sigma$	S
ความแปรปรวน	$\sigma^2$	S
จำนวน	N	N

จากตารางที่ 2.1 แสดงสัญลักษณ์ทางสถิติ 4 สัญลักษณ์ที่สำคัญต่องานวิจัย ซึ่งใช้ในการพิสูจน์สมมติฐานในบทที่ 4

### 2.6.2 การเตรียมพารามิเตอร์หรือค่าสถิติ

#### 2.6.2.1 ค่าเฉลี่ย

ค่าเฉลี่ย คือ ค่าที่เกิดจากการหาค่าที่ปรากฏมากที่สุดในชุดข้อมูล การคำนวณค่าเฉลี่ยนั้นมี 3 วิธี

1) ค่าเฉลี่ยเลขคณิต (Arithmetic Mean) เป็นวิธีการวัดแนวโน้มเข้าสู่ส่วนกลางของข้อมูล โดยค่าเฉลี่ยเลขคณิตนั้นเกิดจากการรวมค่าทุกค่าในชุดข้อมูล แล้วนำมาหาค่าแทนชุดข้อมูล ซึ่งมีสมการดังนี้

ค่าเฉลี่ยประชากร

$$\mu = \frac{\sum_{i=1}^N x_i}{N}$$

ค่าเฉลี่ยกลุ่มตัวอย่าง

$$\bar{X} = \frac{\sum_{i=1}^n x_i}{n}$$

2) ค่าเฉลี่ยเรขาคณิต (Geometric Mean) เป็นการหาค่าเฉลี่ยให้กับข้อมูลที่มีอัตราส่วนต่างกันมาก เช่น การหาค่าเฉลี่ยของปริมาณออกซิเจนจากถังออกซิเจน 11.5 ลิตร ที่ต้องใช้ในการดำน้ำ 3 ช่วง โดยช่วงแรกมีระยะทางร้อยละ 50 ของระยะทางทั้งหมด ช่วงที่สองมีระยะทางร้อยละ 10 ของระยะทางทั้งหมด และช่วงสุดท้ายมีระยะทางร้อยละ 30 ของระยะทางทั้งหมด จากตัวอย่างจะเห็นว่าระยะทางแต่ละช่วงต่างกันมาก จึงจำเป็นต้องใช้วิธีการหาค่าเฉลี่ยเรขาคณิตในการหาค่าเฉลี่ย โดยมีสมการดังนี้

ค่าเฉลี่ยเรขาคณิต

$$\text{G.M.} = \sqrt[n]{\sum_{i=1}^N x_i}$$

3) ค่าเฉลี่ยฮาร์โมนิก (Harmonic Mean) เป็นการหาค่าเฉลี่ยที่เป็นอัตราส่วนระหว่างข้อมูลแต่ละชุด เช่น การหาค่าเฉลี่ยในการทำงานของพนักงาน 3 คน โดยคนแรกใช้เวลา 5 นาทีในการทำงาน 1 ชิ้น คนที่สองใช้เวลา 6 นาทีในการทำงานชิ้นเดียวกัน โดยคนสุดท้ายใช้เวลา 10 นาทีในการทำงานชิ้นเดียวกัน การหาค่าเฉลี่ยจากอัตราส่วนเหล่านี้ต้องใช้ค่าเฉลี่ยฮาร์โมนิก ซึ่งมีสมการดังนี้

ค่าเฉลี่ยฮาร์โมนิก

$$H = \frac{n}{\sum_{i=1}^N \frac{1}{x_i}}$$

#### 2.6.2.2 ส่วนเบี่ยงเบนมาตรฐานและความแปรปรวน

ส่วนเบี่ยงเบนมาตรฐานและความแปรปรวน คือ ค่าที่ใช้อธิบายลักษณะการกระจายของข้อมูลระหว่างค่าเฉลี่ย และสมาชิกทุกตัวในชุดข้อมูล โดยมีสมการดังนี้

ส่วนเบี่ยงเบนมาตรฐานและความแปรปรวนประชากร

$$\sigma = \sqrt{\frac{\sum_{i=1}^N (x_i - \mu)^2}{N}}$$

$$\sigma^2 = \frac{\sum_{i=1}^N (x_i - \mu)^2}{N}$$

ส่วนเบี่ยงเบนมาตรฐานและความแปรปรวนกลุ่มตัวอย่าง

$$S = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{X})^2}{n - 1}}$$

$$S^2 = \frac{\sum_{i=1}^n (x_i - \bar{X})^2}{n - 1}$$

### 2.6.3 การพิสูจน์สมมติฐาน

สมมติฐาน คือ สิ่งที่ทำการศึกษาสถิติได้คาดไว้ว่าน่าจะเกิดขึ้น โดยอาจจะเกิดขึ้นหรือไม่ก็ได้ โดยในทางสถิติแล้วมีขั้นตอนวิธีการหลายขั้นตอนหลายวิธีที่ใช้ในการพิสูจน์ว่าสิ่งที่คาดไว้เป็นจริงหรือไม่ โดยแต่ละขั้นตอนวิธีมีจุดประสงค์การใช้งานที่แตกต่างกันขึ้นอยู่กับลักษณะของประชากรหรือกลุ่มตัวอย่าง

### 2.6.4 การตั้งสมมติฐานทางสถิติ

สมมติฐานแบ่งออกเป็น 2 ประเภท คือ

2.6.4.1 สมมติฐานว่าง หรือสมมติฐานหลัก เป็นสมมติฐานที่กำหนดไว้ให้เป็นจริงขณะที่ทำการทดสอบ

2.6.4.2 สมมติฐานแย้ง หรือสมมติฐานทางเลือก เป็นสมมติฐานที่ผู้ทดสอบสนใจ คำอธิบายของสมมติฐานแย้ง คือ จุดมุ่งหมายที่คาดว่าประชากรหรือกลุ่มตัวอย่างจะเป็น และเมื่อพิสูจน์สมมติฐานแล้วผลลัพธ์เป็นการปฏิเสธสมมติฐานว่าง นั้นหมายความว่า สิ่งที่คาดไว้ว่าประชากรจะเป็นจริง

### 2.6.5 การเลือกใช้วิธีการในการทดสอบสมมติฐาน

วิธีการในการทดสอบสมมติฐานที่เกี่ยวข้องกับค่าเฉลี่ยเบื้องต้นนั้น ประกอบด้วย Z – test T – test และการวิเคราะห์ความแปรปรวน ANOVA ซึ่งแต่ละวิธีมีจุดประสงค์ในการทดสอบที่ต่างกัน โดย Z – test และ T – test ใช้สำหรับการทดสอบกลุ่มตัวอย่างหรือประชากร 1 ถึง 2 กลุ่ม ส่วน ANOVA ใช้สำหรับทดสอบกลุ่มตัวอย่างหรือประชากรมากกว่า 2 กลุ่ม (สมพงษ์ พันธุ์รัตน์, 2561)

## 2.7 กลุ่มอุตสาหกรรมและหมวดธุรกิจ

ตลาดหลักทรัพย์แห่งประเทศไทย ได้แบ่งกลุ่มธุรกิจออกเป็น 8 กลุ่ม โดยแต่ละกลุ่มมีนิยามของกลุ่มรองรับรูปแบบธุรกิจที่มีในประเทศ สามารถเปรียบเทียบธุรกิจในกลุ่มเดียวกันได้ง่าย และใช้เป็นข้อมูลการลงทุนได้ (ตลาดหลักทรัพย์แห่งประเทศไทย, 2557) โดยแต่ละกลุ่มธุรกิจนั้นจะแบ่งเป็นหมวดย่อย รวมทั้งหมด 28 หมวด ดังนี้

### 2.7.1 กลุ่มเกษตรและอุตสาหกรรมอาหาร

กลุ่มอุตสาหกรรมเกษตรและอุตสาหกรรมอาหาร คือ กลุ่มธุรกิจที่เกี่ยวข้องกับการเพาะปลูก ปศุสัตว์ รวมทั้งด้านการผลิตอาหาร โดยแบ่งออกเป็น 2 หมวด ดังนี้

2.7.1.1 หมวดธุรกิจการเกษตร ดำเนินงานด้านการเกษตร เช่น การเพาะปลูก การปศุสัตว์ การประมง ป่าไม้ รวมทั้งการเก็บรักษาผลิตผลทางการเกษตร

2.7.1.2 หมวดอาหารและเครื่องดื่ม ดำเนินงานด้านการผลิตอาหาร กิจการร้านอาหารและเครื่องดื่ม

### 2.7.2 สินค้าอุปโภคบริโภค

กลุ่มอุตสาหกรรมสินค้าอุปโภคบริโภค คือ กลุ่มที่ดำเนินธุรกิจด้านการผลิต จำหน่าย สินค้าอุปโภคบริโภค แบ่งเป็น 3 หมวด ดังนี้



2.7.2.1 หมวดแฟชั่น ดำเนินงานด้านออกแบบ ผลิต และจำหน่ายสินค้าที่เกี่ยวข้องกับแฟชั่น เช่น เสื้อผ้า เครื่องแต่งกาย เครื่องประดับ

2.7.2.2 หมวดของใช้ในครัวเรือนและสำนักงาน ดำเนินงานด้านการผลิต และจำหน่ายสินค้าที่ใช้งานในครัวเรือน เช่น โต๊ะ เก้าอี้ ปากกา อุปกรณ์ดนตรี

2.7.2.3 หมวดของใช้ส่วนตัวและเวชภัณฑ์ ดำเนินงานด้านการผลิต และจำหน่ายสินค้าอุปโภค เช่น เครื่องสำอาง น้ำหอม อุปกรณ์ทางการแพทย์

### 2.7.3 ธุรกิจการเงิน

กลุ่มธุรกิจการเงิน คือ กลุ่มที่ดำเนินธุรกิจเกี่ยวกับการบริการทางการเงิน โดยแบ่งเป็น 3 หมวด คือ

2.7.3.1 หมวดธนาคาร ดำเนินธุรกิจตาม พ.ร.บ.การธนาคารพาณิชย์และกฎหมายที่เกี่ยวข้อง โดยธุรกิจหมวดธนาคารต้องดำเนินงานตามกฎหมายเฉพาะด้าน

2.7.3.2 หมวดเงินทุนและหลักทรัพย์ ดำเนินธุรกิจเกี่ยวกับการเงิน การลงทุน การกู้ยืม การเช่าซื้อ รวมทั้งผู้ให้บริการด้านหลักทรัพย์

2.7.3.3 หมวดประกันภัยและประกันชีวิต ดำเนินงานด้านประกันภัย ภายใต้ พ.ร.บ.ประกันภัย และพ.ร.บ.ประกันชีวิต รวมทั้งกฎหมายที่เกี่ยวข้อง

### 2.7.4 สินค้าอุตสาหกรรม

กลุ่มสินค้าอุตสาหกรรม คือ กลุ่มที่ดำเนินธุรกิจเกี่ยวกับการผลิต และจำหน่ายวัตถุดิบที่ใช้ในงานอุตสาหกรรมสินค้า รวมทั้งจำหน่ายเครื่องจักรที่ใช้ในงานด้านการผลิต โดยแบ่งออกเป็น 6 หมวด คือ

2.7.4.1 หมวดยานยนต์ ดำเนินธุรกิจเกี่ยวกับการผลิต ประกอบ ซ่อมบำรุง และจำหน่ายรถยนต์

2.7.4.2 หมวดวัสดุอุตสาหกรรมและเครื่องจักร ดำเนินงานด้านการผลิตและจำหน่ายเครื่องจักรที่ใช้งานด้านอุตสาหกรรม รวมถึงอะไหล่ และวัสดุที่ใช้ในการอุตสาหกรรม

2.7.4.3 หมวดบรรจุภัณฑ์ ดำเนินงานด้านการผลิตบรรจุภัณฑ์

2.7.4.4 หมวดกระดาษและวัสดุการพิมพ์ ดำเนินงานด้านการผลิตและจำหน่ายผลิตภัณฑ์ที่เกี่ยวข้องกับการพิมพ์ เช่น กระดาษ หมึก

2.7.4.5 หมวดปิโตรเคมีและเคมีภัณฑ์ ดำเนินธุรกิจเกี่ยวกับการผลิตและจำหน่ายสินค้าประเภทสารเคมีที่ใช้ในงานผลิต รวมทั้งสารเคมีที่ใช้ในงานต่าง ๆ เช่น ปุ๋ยที่ใช้ทางด้านเกษตร

2.7.4.6 หมวดเหล็ก ทำหน้าที่เกี่ยวกับการผลิต แปรรูป และจำหน่ายเหล็ก ที่ใช้เป็นวัตถุดิบในผลิต เช่น อลูมิเนียม

## 2.7.5 อสังหาริมทรัพย์และก่อสร้าง

กลุ่มอสังหาริมทรัพย์และก่อสร้าง คือ กลุ่มที่ดำเนินธุรกิจเกี่ยวกับการก่อสร้าง วิศวกรรม และบริหารงานด้านอสังหาริมทรัพย์ แบ่งออกเป็น 4 หมวด คือ

2.7.5.1 หมวดวัสดุก่อสร้าง ดำเนินงานในฐานะผู้ผลิต และตัวแทนจำหน่ายวัสดุก่อสร้าง รวมทั้งวัสดุตกแต่งที่ไม่ใช่เหล็ก

2.7.5.2 หมวดบริการรับเหมาก่อสร้าง ให้บริการในด้านการสร้างสิ่งปลูกสร้าง ทั้งภายใน และ ภายนอก รวมทั้งให้คำแนะนำกับผู้รับบริการในด้านโครงสร้าง และการตกแต่ง

2.7.5.3 หมวดพัฒนาอสังหาริมทรัพย์ ทำหน้าที่บริการเกี่ยวกับ การขาย เช่าซื้อ และบริหารงานเกี่ยวกับอสังหาริมทรัพย์

2.7.5.4 หมวดกองทุนรวมอสังหาริมทรัพย์และกองทรัสต์เพื่อการลงทุนในอสังหาริมทรัพย์ ดำเนินการเกี่ยวกับการลงทุนด้านอสังหาริมทรัพย์ เช่น การหากำไรจากลงทุน โดยการลงทุนมาจากกองทุน หรือกำไรจากการซื้อของอสังหาริมทรัพย์

## 2.7.6 ทรัพยากร

กลุ่มทรัพยากร คือ กลุ่มที่ดำเนินธุรกิจเกี่ยวกับการจัดการทรัพยากรต่าง ๆ ทั้งด้านเชื้อเพลิง การทำเหมืองแร่ แบ่งออกเป็น 2 หมวด คือ

2.7.6.1 หมวดพลังงานและสาธารณูปโภค ดำเนินงานเกี่ยวกับการผลิตสำรวจ ขุดเจาะ กลั่น และจำหน่ายพลังงานธรรมชาติ เช่น น้ำมัน ไฟฟ้า

2.7.6.2 เหมืองแร่ ดำเนินงานด้านการถลุง จำหน่ายแร่ เฉพาะแร่ที่ไม่ให้พลังงาน ทั้งโลหะ และอโลหะ

### 2.7.7 บริการ

กลุ่มบริการ คือ กลุ่มที่ดำเนินธุรกิจด้านการให้บริการ ยกเว้นการบริการด้านการเงิน และข้อมูลสารสนเทศ

2.7.7.1 หมวดยานยนต์ ดำเนินธุรกิจเกี่ยวกับการบริการจำหน่ายสินค้า ทั้งแบบค้าปลีก และค้าส่ง โดยมีหน้าร้าน เช่น ห้างสรรพสินค้า หรือไม่มีหน้าร้าน เช่น ร้านค้าออนไลน์ สินค้าที่จำหน่าย เป็นสินค้าที่ไม่ได้เป็นวัตถุดิบ

2.7.7.2 หมวดยานยนต์ ให้บริการด้านการแพทย์ รวมทั้งงานบริการด้านความงาม

2.7.7.3 หมวดยานยนต์และสิ่งพิมพ์ ให้บริการด้านการผลิต และจำหน่ายสื่อต่าง ๆ เช่น โรงพิมพ์ โรงละคร

2.7.7.4 หมวดยานยนต์เฉพาะกิจ ให้บริการโดยมีจุดประสงค์เฉพาะ เช่น หน่วยงานด้านการศึกษา หน่วยงานด้านการบินของเสีย

2.7.7.5 หมวดยานยนต์ท่องเที่ยวและสันทนาการ ให้บริการเกี่ยวกับโรงแรมและที่พักชั่วคราว บริษัทนำเที่ยว รวมทั้งให้บริการเช่าสถานที่เพื่อสันทนาการต่าง ๆ

2.7.7.6 ขนส่งและโลจิสติกส์ ให้บริการด้านการขนส่งทุกช่องทาง เช่น การขนส่งทางอากาศ การขนส่งสินค้าทางเรือ

### 2.7.8 เทคโนโลยี

กลุ่มเทคโนโลยี คือ กลุ่มที่ดำเนินธุรกิจทั้งการผลิตวัตถุดิบที่ใช้ในการผลิตอุปกรณ์ด้านเทคโนโลยี รวมถึงการให้บริการเกี่ยวกับการสื่อสาร

2.7.8.1 ชิ้นส่วนอิเล็กทรอนิกส์ ดำเนินการผลิตชิ้นส่วนอิเล็กทรอนิกส์เพื่อใช้กับเครื่องใช้ไฟฟ้า

2.7.8.2 เทคโนโลยีสารสนเทศและการสื่อสาร ดำเนินงานด้านการให้บริการ เทคโนโลยีสำหรับจัดการข้อมูลและสื่อสาร เช่น ผู้ให้บริการเครือข่ายโทรคมนาคม รวมทั้งดำเนินการผลิต หรือจำหน่ายอุปกรณ์สำหรับเทคโนโลยี

ตารางที่ 2.2 ตารางกลุ่มและหมวดธุรกิจแบ่งตามตลาดหลักทรัพย์แห่งประเทศไทย

ลำดับ	กลุ่มธุรกิจ	หมวดธุรกิจ
1	เกษตรและอุตสาหกรรม อาหาร (Agro & Food Industry)	1. ธุรกิจการเกษตร (Agribusiness)
		2. อาหารและเครื่องดื่ม (Food & Beverage)
2	สินค้าอุปโภคบริโภค (Consumer Products)	1. แฟชั่น (Fashion)
		2. ของใช้ในครัวเรือนและสำนักงาน (Home & Office Products)
		3. ของใช้ส่วนตัวและเวชภัณฑ์ (Personal Products & Pharmaceuticals)
3	ธุรกิจการเงิน (Financials)	1. ธนาคาร (Banking)
		2. เงินทุนและหลักทรัพย์ (Finance & Securities)
		3. ประกันภัยและประกันชีวิต (Insurance)

ตารางที่ 2.2 ตารางกลุ่มและหมวดธุรกิจแบ่งตามตลาดหลักทรัพย์แห่งประเทศไทย (ต่อ)

ลำดับ	กลุ่มธุรกิจ	หมวดธุรกิจ
4	สินค้าอุตสาหกรรม (Industrials)	1. ยานยนต์ (Automotive)
		2. วัสดุอุตสาหกรรมและเครื่องจักร (Industrial Materials & Machine)
		3. บรรจุภัณฑ์ (Packaging)
		4. กระดาษและวัสดุการพิมพ์ (Paper & Printing Materials)
		5. ปิโตรเคมีและเคมีภัณฑ์ (Petrochemicals & Chemicals)
		6. เหล็ก (Steel)
5	อสังหาริมทรัพย์และก่อสร้าง (Property & Construction)	1. วัสดุก่อสร้าง (Construction Materials)
		2. บริการรับเหมาก่อสร้าง (Construction Services)
		3. พัฒนาอสังหาริมทรัพย์ (Property Development)
		4. กองทุนรวมอสังหาริมทรัพย์ และกองทรัสต์เพื่อการลงทุนในอสังหาริมทรัพย์ (Property Fund & Real Estate Investment Trusts)

ตารางที่ 2.2 ตารางกลุ่มและหมวดธุรกิจแบ่งตามตลาดหลักทรัพย์แห่งประเทศไทย (ต่อ)

ลำดับ	กลุ่มธุรกิจ	หมวดธุรกิจ
6	ทรัพยากร (Resources)	1. พลังงานและสาธารณูปโภค (Energy & Utilities)
		2.เหมืองแร่ (Mining)
7	บริการ (Services)	1. พาณิชยกรรม (Commerce)
		2. การแพทย์ (Health Care Services)
		3. สื่อและสิ่งพิมพ์ (Media & Publishing)
		4. บริการเฉพาะกิจ (Professional Services)
		5. การท่องเที่ยวและสันทนาการ (Tourisms & Leisure)
		6. ขนส่งและโลจิสติกส์ (Transportation & Logistics)
8	เทคโนโลยี (Technology)	1. ชิ้นส่วนอิเล็กทรอนิกส์ (Electronic Components)
		2. เทคโนโลยีสารสนเทศและการสื่อสาร (Information & Communication Technology)

## 2.8 การหาความคล้ายกันของข้อความ (Sentence Similarity Measurement)

การหาความคล้ายกันของข้อความ เป็นวิธีในการเปรียบเทียบชุดข้อมูล 2 ชุดที่มีลักษณะเป็นเวกเตอร์ เพื่อพิสูจน์ว่าข้อความที่ต้องการเปรียบเทียบกับมีความคล้ายกันหรือไม่ โดยวิธีการวัดความคล้ายกันของข้อความนั้นมีหลายวิธี เช่น Jaccard similarity, Overlap similarity, Dice similarity และอื่น ๆ โดยวิธีที่ให้ผลลัพธ์ที่ดีที่สุดและนิยมที่สุด คือ Cosine similarity (Zadeh and Goel 2012; Chien and Immorlica, 2005; Chuang and Chien, 2005; Sahami and Heilman, 2006; Spertus et al., 2005) โดย Cosine similarity เป็นวิธีการคำนวณค่าแบบจุด (Dot product หรือ Scalar Product) ที่ได้มาจากการแปลงข้อความให้เป็นเวกเตอร์ โดยมีสมการดังนี้

$$\frac{\#(x,y)}{\sqrt{\#(x)}\sqrt{\#(y)}}$$

กำหนดให้  $\#(x) = \sum_{i=1}^n x[i]$  ,  $\#(y) = \sum_{i=1}^n y[i]$  และ  $\#(x,y) = \sum_{i=1}^n (x[i])(y[i])$  โดย X คือ เวกเตอร์ของข้อความที่ 1 และ Y คือ เวกเตอร์ของข้อความที่ 2 ผลลัพธ์ที่ได้ คือ ค่าองศาความคล้ายกันซึ่งมีค่าตั้งแต่ 0 ถึง 1 ยิ่งค่าเข้าใกล้ 1 หมายความว่า ข้อความทั้ง 2 มีความคล้ายกันสูง

2.8.1 การแปลงข้อความให้อยู่ในรูปเวกเตอร์ด้วย TF – IDF (Term Frequency - Inverse Document Frequency)

TF – IDF เป็นหนึ่งในวิธีการทางสถิติที่ใช้ในการค้นคืนข้อมูลและทำเหมืองข้อมูล (Wu, et al., 2008; Capelle, et al., 2012) โดยมีจุดประสงค์เพื่อหาคำสำคัญที่เกิดขึ้นในข้อความหรือเอกสาร โดยคำใดที่ปรากฏในข้อความหลายครั้งหรือเกิดร่วมกันในข้อความหลายชุดอาจมีความสำคัญต่อข้อความหลายข้อความ TF – IDF ยังสามารถใช้ในการคัดเลือกคำหยุดหรือจัดกลุ่มคำได้ด้วย การคำนวณค่า TF – IDF แบ่งเป็น 2 ขั้นตอน ดังนี้

2.8.1.1 หาค่า TF (Term Frequency) โดยการหาความถี่ของคำที่เกิดขึ้นในข้อความหรือประโยคทุกประโยค โดยกรณีที่ข้อความมีขนาดต่างกันมากอาจส่งผลให้ค่าบางคำเกิดขึ้นบ่อยยิ่งขึ้นจึงควรจะนำความถี่ของคำหารด้วยจำนวนคำของแต่ละข้อความ เพื่อช่วยลดความแตกต่างกันของคำ

2.8.1.2 หาค่า IDF (Inverse Document Frequency) เพื่อพิสูจน์ว่าคำสำคัญนั้นมีความสำคัญต่อข้อความจริง โดยการคำนวณจากจำนวนข้อความทั้งหมดหารด้วยจำนวนข้อความที่มีคำสำคัญอยู่ เนื่องจากการหาค่า TF เพียงลำพังนั้น คำทุกคำ คือ คำสำคัญหมด ไม่ได้คำนึงถึงระดับความสำคัญ ซึ่งคำบางคำอาจพบบ่อยในข้อความ แต่ไม่ได้มีความสำคัญ เช่น คำเชื่อม

## 2.9 การตัดคำ

การตัดคำ คือ การแบ่งข้อความที่ต่อเนื่องกับให้เป็นกลายเป็นคำตามขอบเขตของแต่ละคำ (Morpheme) โดยในภาษาไทยนั้น ลักษณะการเขียนจะมีวรรคตอนภายในประโยคไม่ชัดเจน แต่จะมีวรรคตอนระหว่างประโยค เพื่อให้ผู้อ่านทำความเข้าใจประโยคได้ (วิรัช ตรีเลิศล้ำ วาณิช, 2557)

### 2.9.1 ขั้นตอนวิธีในการตัดคำ

2.9.1.1 การตัดคำโดยใช้อักษรวิธี เป็นการตัดคำโดยอาศัยกฎ 5 ข้อ ดังนี้

- 1) พิจารณาจากสระไม่สามารถอยู่เดี่ยว ๆ ได้ จะต้องมียัญชนะมารองรับเสมอ เช่น อ อี อี อี อู อู อ๋ อ๋ อ๋ อ๋
- 2) พิจารณาจากสระที่ต้องมียัญชนะตามเสมอ เช่น เอ แอ โอ โอ
- 3) พิจารณาจากสระที่ต้องมียัญชนะนำเสมอ เช่น อะ อา อำ
- 4) พิจารณาจากตัวการันต์ เนื่องจากตัวการันต์ไม่มีโอกาสเป็นยัญชนะตัวแรกของคำได้ เช่น น์ ตร์
- 5) พิจารณาตัวอักษรที่เหลือประกอบกับกลุ่มอักษรที่ 2 3 และ 4



2.9.1.2 การตัดคำโดยการตรวจสอบคำในพจนานุกรมแล้วแยกเป็นคำจากขวาไปซ้าย (Longest Matching) เป็นการตัดคำที่ละพยางค์โดยตัดตัวขวาสุดของประโยคแล้วนำส่วนที่เหลือไปค้นหาในพจนานุกรมเช่น “ความก้าวหน้าทางวิทยาศาสตร์มีบทบาทสำคัญ” จะเริ่มค้นหา จากประโยคเต็มก่อน เมื่อไม่พบในพจนานุกรมก็จะตัดพยางค์ขวาสุดออกไปเป็น “ความก้าวหน้าทางวิทยาศาสตร์มีบทบาทสำคัญ” “ความก้าวหน้าทางวิทยาศาสตร์มีบทบาทสำคัญ” จนกระทั่งได้คำว่า “ความก้าวหน้า” จึงได้คำที่หนึ่ง หลังจากนั้นจึงค้นหาคำในตำแหน่งอื่นต่อไป (วิรัช ศรีเลิศล้ำวาณิช, 2557) โดยมีตัวอย่างเครื่องมือในการตัดคำคือ LexTo<sup>2</sup> ที่ใช้ Lexitron เป็นพจนานุกรมในการตัดคำ

2.9.1.3 การตัดคำโดยเลือกคำที่ไม่พบในพจนานุกรม (Unknown Word) การทำงานเหมือนกับ Longest Matching แต่เริ่มจากซ้ายไปขวา โดยพิจารณาทีละตัวอักษร เมื่อได้คำที่ตรงกับพจนานุกรมแล้วจะพิจารณาคำนั้นโดยเพิ่มพยางค์ตัวถัดไป โดยในรอบแรกจะได้คำที่มีความยาวมากที่สุดก่อน รอบต่อมาเมื่อได้คำที่ตรงกับพจนานุกรมแล้วจะข้ามไปพิจารณาคำต่อไป โดยกำหนดพยางค์ตัวถัดไปเป็นพยางค์แรกของคำใหม่ (วิรัช ศรีเลิศล้ำวาณิช, 2557)

---

<sup>2</sup> <http://www.sansarn.com/lexto/>

## บทที่ 3

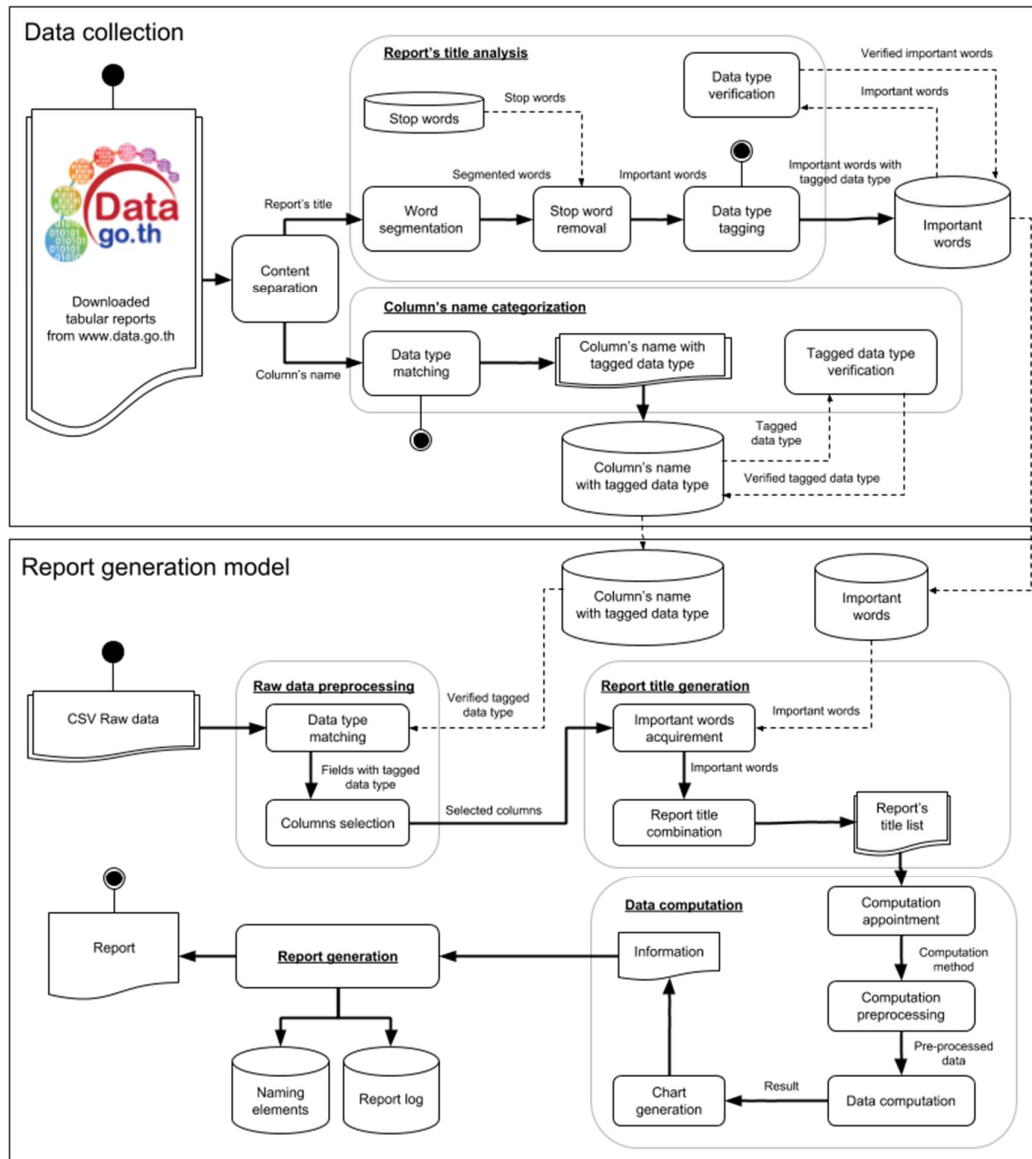
### ขั้นตอนการดำเนินการวิจัย

จากการศึกษาขั้นตอนการสร้างรายงาน ผู้วิจัยได้ออกแบบแบบจำลองที่สามารถแนะนำชื่อของรายงานเพื่อใช้ในการสร้างรายงานให้กับผู้สร้างรายงาน ซึ่งในการพัฒนานั้นจำเป็นต้องรวบรวมข้อมูลเกี่ยวกับชื่อรายงานที่ใช้ในหน่วยงานภายในประเทศไทย เพื่อนำมาศึกษารูปแบบและวิธีการในการตั้งชื่อ ในส่วนนี้ผู้วิจัยจะได้อธิบายขั้นตอนการออกแบบแบบจำลอง ซึ่งแบ่งออกเป็น 2 ขั้นตอน ดังนี้

3.1 การเก็บข้อมูล (Data collection)

3.2 การออกแบบแบบจำลองในการสร้างรายงาน (Report generation model)

ซึ่งรูปที่ 3.1 แสดงกระบวนการในการศึกษาและออกแบบแบบจำลองในการสร้างรายงานอัตโนมัติ โดยแบ่งเป็น 2 ขั้นตอน ตามกรอบการดำเนินงานที่แสดงดังรูป



รูปที่ 3.1 ภาพแบบขั้นตอนการศึกษาและออกแบบแบบจำลอง

### 3.1 การเก็บข้อมูล (Data collection)

จุดประสงค์ของขั้นตอนการเก็บข้อมูล คือ ศึกษาลักษณะการตั้งชื่อรายงาน และเก็บข้อมูลประเภทข้อมูลที่ปรากฏในแถวข้อมูลของรายงาน เพื่อนำมาออกแบบแบบจำลองในการสร้างรายงาน

ผู้วิจัยได้สำรวจและเก็บข้อมูลจากเว็บไซต์ <http://www.data.go.th> ซึ่งเป็นเว็บไซต์ที่ให้บริการข้อมูลในรูปแบบรายงานของภาครัฐในประเทศไทย โดยได้ดาวน์โหลดรายงาน จำนวน 158 ฉบับ ซึ่งรายงานที่ดาวน์โหลดมานั้นจะบันทึกเป็นลักษณะของไฟล์ข้อความ (Text file) โดยบันทึกแยกตามหมวดหมู่ 8 หมวดหมู่ และแยกเป็นกลุ่มธุรกิจอ้างอิงตามหลักเกณฑ์การแบ่งกลุ่มธุรกิจของหลักทรัพย์แห่งประเทศไทยเป็น 4 กลุ่ม ดังตารางที่ 3.1 โดยการเก็บข้อมูลแบ่งออกเป็น 3 ขั้นตอน ดังนี้

ตารางที่ 3.1 ตารางรายละเอียดรายงานจาก [www.data.go.th](http://www.data.go.th) โดยแยกตามหมวดหมู่

ลำดับ	ชื่อหมวดหมู่	กลุ่มธุรกิจอ้างอิงตามตลาดหลักทรัพย์แห่งประเทศไทย	จำนวน (ฉบับ)
1	สังคมและสวัสดิการ	บริการ	64
2	คมนาคมและโลจิสติกส์	บริการ	15
3	เศรษฐกิจ การเงิน และอุตสาหกรรม	ธุรกิจการเงิน	15
4	สาธารณสุข	บริการ	14
5	งบประมาณและการใช้จ่ายของภาครัฐ	-	9
6	พลังงาน ทรัพยากรธรรมชาติ และสิ่งแวดล้อม	ทรัพยากร	22
7	กฎหมาย ศาล และอาชญากรรม	-	16
8	เกษตรกรรมและชลประทาน	เกษตรและอุตสาหกรรม อาหาร	4

### 3.1.1 การแยกข้อมูลสำหรับการวิเคราะห์การตั้งชื่อรายงาน (Content Separation)

ขั้นตอนนี้มีจุดประสงค์ คือ แบ่งข้อมูลจากรายงานที่ดาวน์โหลดมาเป็น 2 ส่วน ประกอบด้วย ส่วนชื่อรายงาน ซึ่งนำไปใช้ในขั้นตอนการวิเคราะห์คุณลักษณะของการตั้งชื่อรายงาน (Report's title analysis) โดยมีลักษณะดังรูปที่ 3.2 และส่วนข้อมูลแถวรายงาน ที่จะนำไปใช้ในขั้นตอนการจัดกลุ่มประเภทข้อมูลของแถวข้อมูล (Column's name categorization)

รายได้เฉลี่ยต่อเดือนต่อครัวเรือน จำแนกตามภาค และจังหวัด พ.ศ. 2541 - 2558

รายได้เฉลี่ยต่อปีของคนและครัวเรือน (รายจังหวัด) จากข้อมูล จปฐ ปี 2558

จำนวนครัวเรือนที่คนในครัวเรือนมีรายได้เฉลี่ยน้อยกว่า 30,000 บาทต่อปี

รายจ่ายเฉลี่ยต่อปีของคนและครัวเรือน (รายจังหวัด) จากข้อมูล จปฐ ปี 2558

อัตราการว่างงาน จำแนกตามเพศ พ.ศ. 2549 - 2558

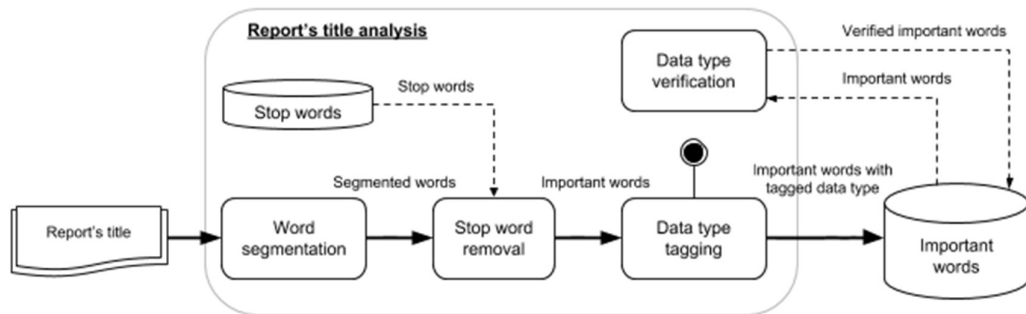
จำนวนประชากรอายุ 15 ปีขึ้นไป จำแนกตามสถานภาพแรงงาน เพศ ทัวราชอาณาจักร

ตำแหน่งงานว่าง จำแนกตามอาชีพ และวุฒิการศึกษา ทัวราชอาณาจักร ปี 2558

**รูปที่ 3.2** ตัวอย่างชุดชื่อรายงานที่ดาวน์โหลดมา

### 3.1.2 การวิเคราะห์คุณลักษณะของการตั้งชื่อรายงาน (Report's title analysis)

ขั้นตอนนี้เป็นขั้นตอนการศึกษาด้วยวิธีการทางสถิติ โดยมีจุดประสงค์ในการศึกษาองค์ประกอบของชื่อรายงานที่ดาวน์โหลดมา เพื่อให้เห็นแนวโน้มการประกอบคำเป็นชื่อของรายงานส่วนใหญ่ ว่าเกิดจากคำชนิดใดบ้างและมีการเรียงตำแหน่งคำอย่างไร โดยมีลักษณะดังรูปที่ 3.3 ซึ่งขั้นตอนนี้แบ่งออกเป็น 4 ขั้นตอน ดังนี้



รูปที่ 3.3 การวิเคราะห์คุณลักษณะของการตั้งชื่อรายงาน (Report's title analysis)

### 3.1.2.1 การตัดคำ (Word segmentation)

การตัดคำในงานวิจัยนี้ เป็นการเตรียมข้อมูลเบื้องต้นสำหรับการศึกษาลักษณะการตั้งชื่อ โดยใช้เครื่องมือ LexTo (Thai Lexeme Tokenizer) พัฒนาโดยศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ (National Electronics and Computer Technology Center: NECTEC) (ศูนย์เทคโนโลยีอิเล็กทรอนิกส์ และคอมพิวเตอร์แห่งชาติ, 2561) ในการตัดคำ ซึ่ง LexTo นั้นพัฒนาบนทฤษฎี Longest matching โดยการดึงคลังคำจากพจนานุกรมมาใช้ในการตัดคำ มีหน้าจอกการใช้งานดัง

โดยรูปที่ 3.4 เป็นผลลัพธ์ที่ได้จะเป็นชุดคำของแต่ละชื่อรายงาน โดยมีสัญลักษณ์ “|” ขึ้นระหว่างแต่ละคำ ดังรูปที่ 3.5



รูปที่ 3.4 ภาพหน้าจอเครื่องมือ LexTo สำหรับการตัดคำ

รายได้เฉลี่ยต่อเดือนต่อครัวเรือน | จำแนกตามภาค | และจังหวัด | พ.ศ. | 2541 | - | 2558 |  
 รายได้เฉลี่ยต่อปีของ|คน|และ|ครัวเรือน| |(รายได้จังหวัด)| |จาก|ข้อมูล| |จปฐ| |ปี| |2558|  
 จำนวนครัวเรือนที่|คน|ในครัวเรือน|มี|รายได้เฉลี่ย|น้อยกว่า| |30,000| |บาท|ต่อปี|  
 |รายจ่ายเฉลี่ยต่อปี|ของ|คน|และ|ครัวเรือน| |(รายได้จังหวัด)| |จาก|ข้อมูล| |จปฐ| |ปี| |2558|  
 อัตราการว่างงาน | จำแนกตามเพศ | พ.ศ. | 2549 | - | 2558 |  
 จำนวนประชากรอายุ | 15 | ปีขึ้นไป | จำแนกตามสถานภาพแรงงาน | เพศ | ทั่วราชอาณาจักร |  
 ตำแหน่งงานว่าง | จำแนกตามอาชีพ | และวุฒิการศึกษา | ทั่วราชอาณาจักร | ปี | 2558 |

รูปที่ 3.5 ตัวอย่างผลการตัดคำด้วยเครื่องมือ LexTo

### 3.1.2.2 การตัดคำหยุด (Stop word removal)

การตัดคำหยุดมีจุดประสงค์เพื่อลดขนาดของข้อความลงให้เหลือเฉพาะคำสำคัญ โดยผู้วิจัยรวบรวมคำหยุดที่ปรากฏในชื่อรายงานจำนวน 84 คำ ที่ได้ระบุใน ภาคผนวก ค ซึ่งขั้นตอนนี้มีข้อมูลขาเข้าเป็นชุดของคำที่ผ่านกระบวนการตัดคำโดย LexTo และมีผลลัพธ์ที่ได้เป็นชุดของคำที่เหลือเฉพาะคำสำคัญ

คำหยุดในงานวิจัยนี้ คือ คำที่สามารถแทนด้วยคำอื่นได้โดยไม่เปลี่ยนแปลงความหมาย เช่น “ของ” “แห่ง” “ทั้งสิ้น” หรือคำเชื่อมระหว่างคำหรือประโยค เช่น “และ” “หรือ” “แบบ” รวมถึงคำที่สามารถแทนด้วยคำอื่นเพื่อจำกัดขอบเขตของข้อมูลได้ เช่น “ระยะเวลา” “เดือน” “ปี”

#### ขั้นตอนวิธีที่ 3.1 การตัดคำหยุด (Stop word removal)

```

Function: Stop word removal
Input: -
Output: Set of important words
Begin:
  stop_word[] = get_stop_word_from_database()
  for contents in content_in_each_business_types:
    while not end of contents:
      read current_line from contents:
        split current_line to words by ` `
        for word in words:
          remove word if word exist in stop_words[]
End

```

ขั้นตอนวิธีที่ 3.1 ชุดคำสั่งจะดึงข้อมูลคำหยุดจากฐานข้อมูลคำหยุด และทำการวนรอบเพื่ออ่านไฟล์ที่ผ่านขั้นตอนในการตัดคำหยุดแล้ว โดยอ่านเป็นบรรทัด ซึ่งแต่ละบรรทัดจะแปลงค่าให้เป็นอาร์เรย์โดยการแบ่งแยกคำด้วยตัวคั่นคำ `|` หลังจากนั้นชุดคำสั่งจะคัดเลือกเฉพาะคำที่เป็นคำหยุดออกจากอาร์เรย์ เพื่อให้เหลือเพียงคำสำคัญที่ประกอบเป็นชื่อของรายงาน แล้วบันทึกข้อมูลเป็นไฟล์ใหม่ขึ้นมา ในรูปแบบ CSV ดังตัวอย่างในรูปที่ 3.6

รายได้ เฉลี่ย ภาค จังหวัด
รายได้ เฉลี่ย คน จังหวัด
คนใน รายได้ เฉลี่ย น้อยกว่า
รายจ่าย เฉลี่ย คน จังหวัด
อัตรา ว่างงาน เพศ
ประชากร อายุ ขึ้นไป สถานภาพ แรงงาน เพศ
ตำแหน่งงาน ว่าง อาชีพ วุฒิการศึกษา

รูปที่ 3.6 ตัวอย่างผลจากการตัดคำหยุด

### 3.1.2.3 การกำหนดชนิดข้อมูลให้กับคำสำคัญ (Data type tagging)

ขั้นตอนนี้มีจุดประสงค์เพื่อกำหนดประเภทข้อมูลให้กับคำสำคัญ โดยการใช้การเปรียบเทียบคำสำคัญกับคำสำคัญที่อยู่ในฐานข้อมูล หากสามารถเปรียบเทียบได้ว่ามีคำสำคัญนี้ในฐานข้อมูลจะทำการบันทึกประเภทข้อมูล และตำแหน่งของคำสำคัญเข้าสู่ฐานข้อมูล หากการเปรียบเทียบพบว่าคำสำคัญเป็นคำใหม่จะทำการบันทึกสถานะของประเภทข้อมูลเป็น UNKNOWN เพื่อทำการตรวจสอบประเภทข้อมูล



### 3.1.2.4 การตรวจสอบความถูกต้องของชนิดข้อมูลของแต่ละแถวข้อมูล (Data type verification)

ขั้นตอนนี้มีจุดประสงค์เพื่อตรวจสอบความถูกต้องของชนิดข้อมูลที่ได้นำมาบันทึกเข้าสู่ฐานข้อมูล เนื่องจากการกำหนดชนิดข้อมูลแบบอัตโนมัติมีโอกาสผิดพลาดได้ ทั้งจากการกำหนดประเภทข้อมูล UNKNOWN ให้คำสำคัญ หรือคำสำคัญคำนั้นเป็นได้หลายประเภทข้อมูล จึงจำเป็นที่จะต้องมีการตรวจสอบโดยอาศัยผู้ดูแลระบบในการตรวจสอบโดยใช้ประสบการณ์ควบคู่กับพจนานุกรม

ตารางที่ 3.2 ค่าทางสถิติของจำนวนคำที่ปรากฏในชื่อรายงานแยกตามหมวดหมู่

ชื่อหมวดธุรกิจ	จำนวนคำสำคัญ ที่ปรากฏในชื่อของรายงาน			ส่วนเบี่ยงเบน มาตรฐาน
	ต่ำสุด	สูงสุด	เฉลี่ย	
สังคมและสวัสดิการ	1	15	5.677	3.052
คมนาคมและโลจิสติกส์	1	7	3.562	1.711
เศรษฐกิจ การเงิน และอุตสาหกรรม	1	6	3.312	1.078
สาธารณสุข	1	3	2.667	0.617
งบประมาณและการใช้จ่ายของภาครัฐ	1	4	2.600	1.075
พลังงาน ทรัพยากรธรรมชาติ และ สิ่งแวดล้อม	1	8	4.522	2.042
กฎหมาย ศาล และอาชญากรรม	1	10	5.706	2.443
เกษตรกรรมและชลประทาน	1	3	2.200	0.837
ค่าเฉลี่ยคำสำคัญรวม	-	-	3.780	-

จากตารางที่ 3.2 แสดงให้เห็นถึงจำนวนค่าเฉลี่ยที่ใช้ประกอบเป็นชื่อของรายงานพบว่าจำนวนค่าเฉลี่ยอยู่ที่ประมาณ 3 คำ ดังนั้นในงานวิจัยนี้ผู้วิจัยจึงได้ออกแบบขั้นตอนวิธีการผสมชื่อรายงานจากคำสำคัญ โดยชื่อของรายงานนั้น มีโครงสร้างที่แบ่งเป็น 3 ส่วนด้วยกัน ดังนี้

1) ส่วนเริ่มต้นชื่อรายงาน (Beginning Phrase) เป็นส่วนเริ่มต้นชื่อรายงานที่มักเป็นบริบทสรุปภาพรวมของรายงานว่าต้องการสื่อสิ่งใด โดยอาจทำหน้าที่เป็นประธานของชื่อ มักเป็นคำนาม เช่น รายได้ มูลค่าสินทรัพย์ การบรรจุน้ำมัน ตำแหน่งงานว่าง

2) ส่วนกลางของชื่อรายงาน (Middle Phrase) เป็นส่วนที่กล่าวถึงวัตถุประสงค์หรือกลุ่มเป้าหมายของรายงาน โดยแบ่งออกเป็น 2 ส่วนย่อย คือ

2.1) ส่วนต้นของส่วนกลาง (Preceding of Middle Phrase) ส่วนนี้ทำหน้าที่ขยายประธาน เป็นได้ทั้งสรรพนาม กริยา เช่น ที่ลงทะเลเปียน รายชั่วโมง แต่ละอุตสาหกรรม ในประเทศ

2.2) ส่วนท้ายของส่วนกลาง (Succeeding of Middle Phrase) เป็นส่วนขยายเพื่อความสละสลวยหรือครบถ้วนของชื่อรายงานในส่วนเริ่มต้นและส่วนกลางของชื่อรายงาน โดยจะมีหรือไม่มีก็ได้ (Optional) เช่น ที่มีการเปลี่ยนแปลง ตามวัตถุประสงค์ ภายในระยะเวลาที่กำหนด

3) ส่วนท้ายชื่อรายงาน (Ending Phrase) สำหรับเป็นคำปิดท้ายของชื่อรายงาน ส่วนนี้เป็นส่วนที่สำคัญต่อการคำนวณข้อมูล มักเป็นส่วนขยายชื่อของรายงาน เพื่อความชัดเจน เช่น จำแนกตามระดับการศึกษา จำแนกตามภาค จำแนกตามเพศ

### ตารางที่ 3.3 ตัวอย่างการบันทึกประเภทข้อมูลตามคำสำคัญ

ลำดับ	ประเภทข้อมูล	คำสำคัญ	ตำแหน่งคำ
1	Numeric	รายได้	ส่วนเริ่มต้นชื่อรายงาน
2	Numeric	รายรับ	ส่วนเริ่มต้นชื่อรายงาน
3	Text	การยืม	ส่วนเริ่มต้นชื่อรายงาน

จากตารางที่ 3.3 แสดงตัวอย่างการบันทึกข้อมูลคำสำคัญเข้าสู่ฐานข้อมูล โดยข้อมูลที่ถูกรับบันทึกแบ่งออกเป็น 3 ส่วน คือ

- 1) ประเภทข้อมูล บันทึกประเภทข้อมูลของคำสำคัญนั้น
- 2) คำสำคัญ บันทึกคำสำคัญที่ใช้ในการประกอบเพื่อเป็นชื่อรายงาน
- 3) ตำแหน่งคำ บันทึกตำแหน่งของคำที่ปรากฏในชื่อของรายงาน

ตารางที่ 3.4 ตัวอย่างของคำสำคัญที่ปรากฏในแต่ละตำแหน่งของชื่อรายงาน

ส่วนเริ่มต้นชื่อรายงาน	ส่วนกลางของชื่อรายงาน	ส่วนท้ายชื่อรายงาน
จำนวนประชากร	ในแต่ละอุตสาหกรรม	จำแนกตามระดับการศึกษา
จำนวนผู้ลงทะเบียน	สมัครงาน	จำแนกตามอาชีพ
จำนวนคนต่างด้าว	ที่ลงทะเบียน	จำแนกตามไตรมาส

จากตารางที่ 3.4 แสดงตัวอย่างของคำสำคัญที่ปรากฏในชื่อรายงาน โดยแสดงตัวอย่างตามโครงสร้างของชื่อรายงานทั้ง 3 ส่วน

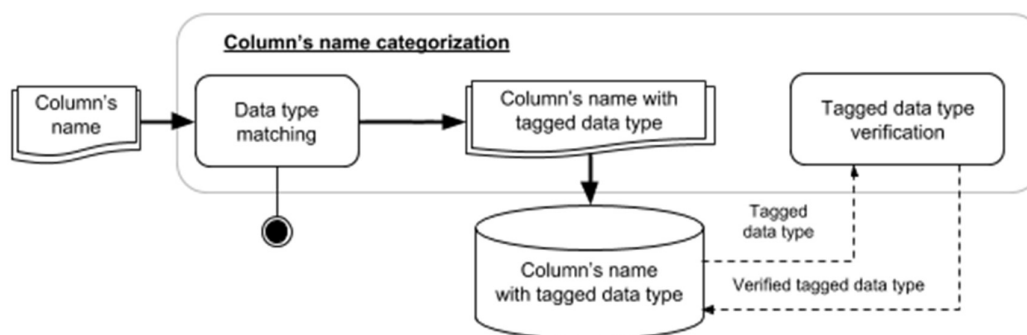
3.1.3 การจัดกลุ่มประเภทข้อมูลของแถวข้อมูล (Column's name categorization)

ขั้นตอนนี้มีจุดประสงค์เพื่อรวบรวมชื่อและศึกษาประเภทข้อมูลของแถวข้อมูลที่ปรากฏในรายงานที่ได้ดาวน์โหลดมา ซึ่งประกอบด้วยขั้นตอน 2 ขั้นตอน คือ

3.1.3.1 การกำหนดประเภทข้อมูลของแถวข้อมูล (Data type matching)

3.1.3.2 การตรวจสอบความถูกต้องของประเภทข้อมูล (Tagged data type verification)

โดยรูปที่ 3.7 แสดงให้เห็นถึงขั้นตอนในการจัดกลุ่มข้อมูลชื่อแถว และบันทึกเข้าสู่ฐานข้อมูล ซึ่งมีข้อมูลขาเข้า คือ ชุดของชื่อแถวที่ปรากฏในรายงานที่ได้ดาวน์โหลดมา โดยชื่อของแถวทุกชื่อจะถูกบันทึกเข้าสู่ฐานข้อมูลชื่อแถว



รูปที่ 3.7 การจัดกลุ่มประเภทข้อมูลของแถวข้อมูล (Column's name categorization)

### 3.1.3.1 การกำหนดประเภทข้อมูลของแถวข้อมูล (Data type matching)

ขั้นตอนนี้มีจุดประสงค์เพื่อกำหนดประเภทข้อมูลให้กับชื่อของแถวข้อมูล โดยการเปรียบเทียบชื่อของแถวแต่ละแถวกับชื่อแถวที่บันทึกในฐานข้อมูลชื่อแถว หากเปรียบเทียบแล้วพบชื่อของแถวที่ตรงกันนำชื่อแถวและประเภทข้อมูลที่ตรงกันมาบันทึกลงในฐานข้อมูล ในกรณีที่ไม่มีพบประเภทข้อมูลจะบันทึกชื่อแถวและสถานะข้อมูล UNKNOWN ลงในฐานข้อมูลเพื่อทำการตรวจสอบข้อมูลต่อไป

### 3.1.3.2 การตรวจสอบความถูกต้องของประเภทข้อมูล (Tagged data type verification)

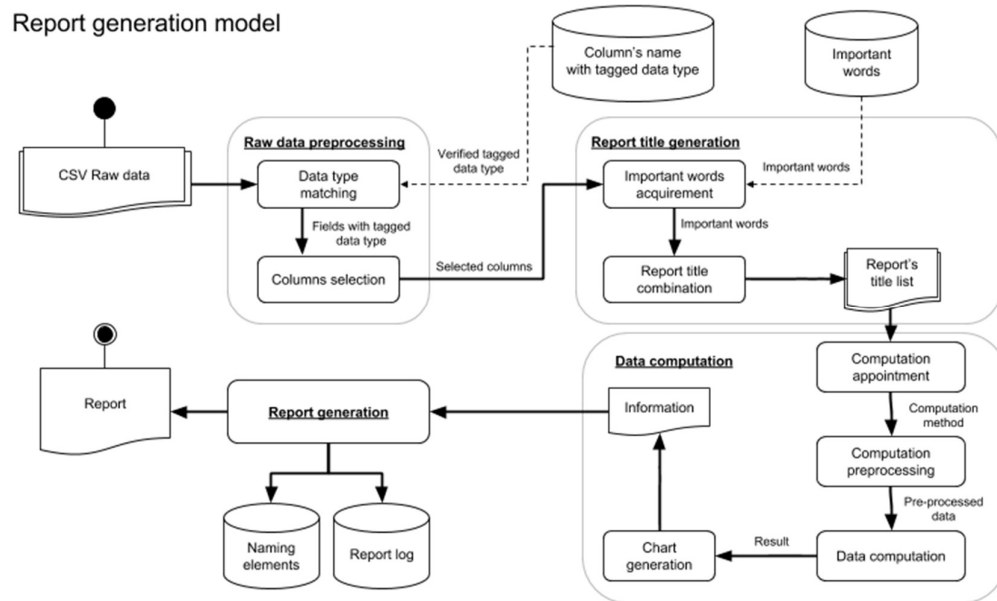
ขั้นตอนนี้มีจุดประสงค์เพื่อตรวจสอบความถูกต้องของประเภทข้อมูลและชื่อของแถวที่ถูกบันทึกในฐานข้อมูล เนื่องจากการกำหนดประเภทข้อมูลให้กับชื่อแถวโดยอัตโนมัติมีโอกาสผิดพลาดจากการกำหนดประเภทข้อมูล UNKNOWN ให้กับชื่อแถวข้อมูล หรืออาจเกิดจากกรณีที่ชื่อแถวข้อมูลมีประเภทข้อมูลที่ไม่ชัดเจน ซึ่งขั้นตอนนี้อาศัยผู้ดูแลระบบในการตรวจสอบความถูกต้อง โดยใช้ประสบการณ์ร่วมกับพจนานุกรม

## 3.2 การออกแบบแบบจำลอง (Report generation model)

งานวิจัยนี้ ผู้วิจัยได้ออกแบบแบบจำลองเพื่อใช้ในการสร้างรายงานขึ้นมา โดยแบบจำลองนั้นประกอบด้วยขั้นตอนวิธี 4 ขั้นตอน คือ

- 3.2.1 การเตรียมข้อมูลเบื้องต้นโดยแบบจำลอง (Raw data preprocessing)
- 3.2.2 การสร้างชื่อรายงานด้วยการผสมคำสำคัญ (Report title generation)
- 3.2.3 การคำนวณข้อมูล (Data computation)
- 3.2.4 การสร้างรายงาน (Report generation)

ซึ่งรูปที่ 3.8 แสดงถึงส่วนประกอบและขั้นตอนการทำงานของแบบจำลองในการสร้างรายงาน ประกอบด้วย 4 ขั้นตอน โดยมีข้อมูลขาเข้าเป็นไฟล์เอกสารชนิดคั่นด้วยสัญลักษณ์จุลภาค (Comma - Separated Values) และมีผลลัพธ์เป็นรายงานพร้อมกับแผนภูมิแสดงข้อมูล



รูปที่ 3.8 แบบจำลองการสร้างรายงาน (Report generation model)

### 3.2.1 การเตรียมข้อมูลเบื้องต้นโดยแบบจำลอง (Raw data preprocessing)

ขั้นตอนนี้ออกแบบเพื่อให้แบบจำลองทราบว่า ผู้สร้างรายงานใช้ข้อมูลต้นฉบับที่ประกอบด้วยข้อมูลชนิดใดบ้าง และผู้สร้างรายงานเลือกแถวข้อมูลใดในการคำนวณข้อมูล โดยประกอบด้วย 2 ขั้นตอน คือ

#### 3.2.1.1 การตรวจสอบประเภทข้อมูลโดยใช้หัวตาราง (Data type matching)

ในขั้นต้นแบบจำลองต้องทำการเตรียมข้อมูลเพื่อใช้ในการสร้างรายชื่อของรายงาน ข้อมูลที่จำเป็นต่อการสร้างรายงานแบบอัตโนมัติ คือ ประเภทของข้อมูล และชื่อของข้อมูล แต่ละแถวของข้อมูลตั้งต้น ขั้นตอนวิธีนี้จะทำการวิเคราะห์ว่า ข้อมูลตั้งต้นที่ผู้สร้างรายงานสนใจนั้น ประกอบด้วยข้อมูลประเภทใดบ้าง จุดประสงค์ของการทราบประเภทข้อมูลของแต่ละคอลัมน์ คือ ทำให้แบบจำลองสามารถนำประเภทของข้อมูลไปสร้างเป็นชุดรายชื่อของรายงานที่เกี่ยวข้องกับข้อมูลตั้งต้น

### ขั้นตอนวิธีที่ 3.2 การตรวจสอบประเภทข้อมูลโดยใช้หัวตาราง (Data type matching)

```

Function: Data type matching
Input: CSV file
Output: Set of column's id, column's name and column's type
Begin:
    columns_name_list[] = get_column_names_from_database()
    read column_names from csv_file
    for column_name in column_names:
        set data_type of current column_name to 'UNKNOWN'
        if column_name exist in columns_name_list:
            set data_type of current column_name to found
data_type
    append current columns_name to result
    return result
End

```

ขั้นตอนวิธีที่ 3.2 เป็นกระบวนการในการหาชนิดของข้อมูลในแต่ละแถวข้อมูล โดยเริ่มต้นจากชุดคำสั่งดึงชื่อแถวพร้อมด้วยประเภทข้อมูลจากฐานข้อมูลชื่อแถว เพื่อเตรียมสำหรับเปรียบเทียบระหว่างชื่อแถวที่ผู้สร้างรายงานเลือกใช้ในการสร้างรายงานกับชื่อแถวจากฐานข้อมูล ต่อมาชุดคำสั่งจะอ่านชื่อแถวจากไฟล์ตั้งต้นเพื่อเก็บในอาร์เรย์ หลังจากนั้นจะวนรอบชื่อแถวทุกชื่อเพื่อเปรียบเทียบชื่อแถวทีละชื่อกับชื่อจากฐานข้อมูล เบื้องต้นประเภทข้อมูลของชื่อแถวจะตั้งเป็น UNKNOWN และหากชื่อแถวใดตรงกับชื่อที่อยู่ในฐานข้อมูลหมายความว่าชื่อแถวนั้นตรงกับประเภทข้อมูลประเภทนั้น ยกตัวอย่างเช่น คำว่า “รายได้” ตรงกับคำว่า “รายได้” จากฐานข้อมูล ซึ่งตรงกับประเภทข้อมูล “ตัวเลข” ส่งผลให้คอลัมน์นี้มีสถานะเป็นตัวเลข

ซึ่งรูปที่ 3.9 แสดงตัวอย่างของชื่อแถวก่อนที่จะตรวจสอบประเภทข้อมูลโดยใช้หัวตาราง โดยประเภทข้อมูลของแต่ละแถวข้อมูลจะถูกตั้งให้เป็น UNKNOWN หลังจากนั้นชื่อแถวแต่ละชื่อจะถูกนำไปค้นหาประเภทข้อมูลจากฐานข้อมูล โดยได้ผลลัพธ์ดังรูปที่ 3.10

COLUMNS:	
1:	
	COLUMN_NAME: วัน
	DATA_TYPE: UNKNOWN
2:	
	COLUMN_NAME: ชื่อ
	DATA_TYPE: UNKNOWN
3:	
	COLUMN_NAME: ราคา
	DATA_TYPE: UNKNOWN

รูปที่ 3.9 ตัวอย่างชื่อแถวจากรายงานที่ผู้สร้างรายงานอัปโหลดเข้าสู่ระบบ

COLUMNS:	
1:	
	COLUMN_NAME: วัน
	DATA_TYPE: DATE
2:	
	COLUMN_NAME: ชื่อ
	DATA_TYPE: TEXT
3:	
	COLUMN_NAME: ราคา
	DATA_TYPE: AMOUNT

รูปที่ 3.10 ตัวอย่างชื่อแถวที่ผ่านการตรวจสอบประเภทข้อมูลโดยใช้หัวตาราง

### 3.2.2 การสร้างชื่อรายงานด้วยการผสมคำสำคัญ (Report title generation)

ขั้นตอนวิธีนี้ออกแบบเพื่อผสมคำสำคัญจากฐานข้อมูลเพื่อสร้างเป็นชื่อของรายงาน โดยจะเลือกคำสำคัญจากฐานข้อมูลที่ตรงกับประเภทข้อมูลที่ถูกเลือก จำนวนของคำสำคัญที่นำมาสร้างเป็นชื่อรายงานนั้นมีตั้งแต่ 1 – 4 คำ โดยแบ่งออกเป็น ส่วนเริ่มต้นชื่อรายงาน จำนวน 1 คำ ส่วนกลางของชื่อรายงาน จำนวน 1 – 2 คำ และส่วนท้ายชื่อรายงาน จำนวน 1 คำ ขั้นตอนวิธีในการผสมชื่อเริ่มต้นจากการค้นหาประเภทข้อมูลของแถวที่ถูกเลือกในการสร้างรายงาน แล้วนำประเภทข้อมูลที่ได้ไปค้นหาคำสำคัญตามประเภทข้อมูลนั้น หลังจากนั้นคำสำคัญที่ถูกดึงจากฐานข้อมูลจะถูกนำมาผสมกันในรูปแบบของเมทริกซ์คูณเมทริกซ์ ผลคูณ คือ เซ็ตของชื่อที่เกิดขึ้นทั้งหมด เซ็ตของคำเหล่านี้จะถูกเรียงลำดับตามชื่อที่นิยมใช้ในการสร้างรายงานจากมากไปน้อยโดยชื่อใดที่ไม่เคยถูกใช้งานจะถูกเรียงตามลำดับตัวอักษร

ซึ่งรูปที่ 3.11 แสดงตัวอย่างของการดึงคำสำคัญจากตารางคำสำคัญ โดยเป็นการดึงคำสำคัญจากส่วนเริ่มต้นชื่อรายงาน จำนวน 2 คำ ส่วนกลางของชื่อรายงาน จำนวน 3 คำ และส่วนท้ายของส่วนกลาง จำนวน 2 คำ ซึ่งคำสำคัญเหล่านี้มีประเภทข้อมูลเป็น TEXT เฉพาะฉะนั้นขนาดของเมทริกซ์คูณเมทริกซ์ จะมีค่าเท่ากับ  $2 * 3 * 2 = 12$  โดยผลลัพธ์ของเมทริกซ์คูณเมทริกซ์ แสดงดังตารางที่ 3.5

COLUMN_NAME: วัน	DATA_TYPE: TEXT
DATA_TYPE: TEXT	
ส่วนเริ่มต้นชื่อรายงาน:	['การยิ้ม', 'การคืน']
ส่วนกลางของชื่อรายงาน:	['พัสดุ', 'หนังสือ', 'เอกสาร']
ส่วนท้ายชื่อรายงาน:	['ของนักศึกษา', 'ของบุคลากร']

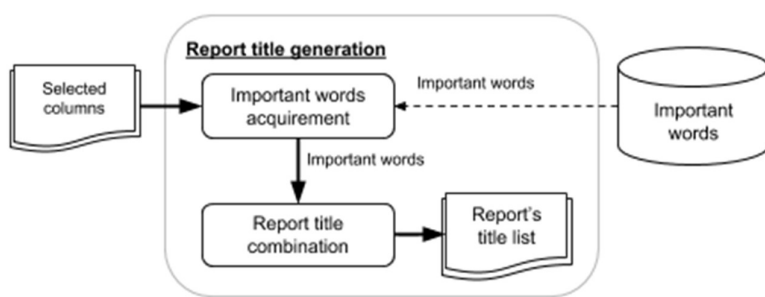
รูปที่ 3.11 ตัวอย่างคำสำคัญที่ดึงจากตารางคำสำคัญ



ตารางที่ 3.5 ตารางแสดงผลลัพธ์จากการผสมคำสำคัญเพื่อสร้างเป็นชื่อของรายงาน

ลำดับ	ส่วนเริ่มต้น ชื่อรายงาน	ส่วนกลางของ ชื่อรายงาน	ส่วนท้ายชื่อ รายงาน	ชื่อรายงาน
1	การยืม	พัสดุ	ของนักศึกษา	การยืมพัสดุของนักศึกษา
2	การยืม	พัสดุ	ของบุคลากร	การยืมพัสดุของบุคลากร
3	การยืม	หนังสือ	ของนักศึกษา	การยืมหนังสือของนักศึกษา
4	การยืม	หนังสือ	ของบุคลากร	การยืมหนังสือของบุคลากร
5	การยืม	เอกสาร	ของนักศึกษา	การยืมเอกสารของนักศึกษา
6	การยืม	เอกสาร	ของบุคลากร	การยืมเอกสารของบุคลากร
7	การคืน	พัสดุ	ของนักศึกษา	การคืนพัสดุของนักศึกษา
8	การคืน	พัสดุ	ของบุคลากร	การคืนพัสดุของบุคลากร
9	การคืน	หนังสือ	ของนักศึกษา	การคืนหนังสือของนักศึกษา
10	การคืน	หนังสือ	ของบุคลากร	การคืนหนังสือของบุคลากร
11	การคืน	เอกสาร	ของนักศึกษา	การคืนเอกสารของนักศึกษา
12	การคืน	เอกสาร	ของบุคลากร	การคืนเอกสารของบุคลากร

โดยรูปที่ 3.12 แสดงลักษณะการทำงานของกระบวนการสร้างชื่อรายงานด้วยการผสมคำสำคัญ ซึ่งประกอบด้วย 2 ส่วน คือ ส่วนดึงคำสำคัญจากฐานข้อมูล และส่วนสำหรับผสมคำสำคัญ โดยมีข้อมูลขาเข้าเป็นแถวที่ผู้สร้างรายงานเลือกเพื่อใช้ในการสร้างรายงาน และผลลัพธ์ คือ ชื่อของรายงานที่เป็นไปได้



รูปที่ 3.12 การสร้างชื่อรายงานด้วยการผสมคำสำคัญ (Report title generation)

### 3.2.3 การคำนวณข้อมูล (Data computation)

ขั้นตอนวิธีการนี้มีจุดประสงค์สำหรับกำหนดวิธีในการคำนวณข้อมูล คำนวณ และเตรียมสารสนเทศทั้งรูปแบบตารางและแผนภูมิสำหรับใช้ในการสร้างรายงาน โดยแบ่งการทำงานเป็น 4 ขั้นตอน ดังนี้

3.2.3.1 การกำหนดวิธีการในการคำนวณข้อมูล (Computation appointment) เป็นการกำหนดวิธีการที่จะใช้ในการคำนวณข้อมูลเพื่อสร้างเป็นแผนภูมิ โดยในงานวิจัยนี้ประกอบด้วยแผนภูมิ 3 ชนิด คือ Bar chart, Line chart และ Pie chart ทั้ง 3 ชนิด เป็นแผนภูมิเบื้องต้นที่นิยมใช้งาน (DataCamp, 2017; IBM Knowledge Center, 2018; ReportLab, 2018) ซึ่งแต่ละชนิดมีวิธีการในการคำนวณต่างกัน โดย Bar chart และ Line chart ใช้วิธีการนับความถี่ของข้อมูล ส่วน Pie chart ใช้วิธีคำนวณอัตราส่วนของความถี่ข้อมูล

3.2.3.2 การเตรียมข้อมูลสำหรับการคำนวณ (Computation preprocessing) เป็นการเตรียมข้อมูลสำหรับใช้คำนวณ เช่น

1) ข้อมูล DATE ชนิด Y/m/d ซึ่งจะถูกแปลงเป็นข้อมูลสามชนิด คือ ชนิดปี ชนิดปีเดือน และ ชนิดปีเดือนวัน เช่น จากข้อมูล 2018/01/01 จะถูกแปลงเป็นข้อมูล 3 ชุด คือ 2018 2018/01 และ 2018/01/01

2) ข้อมูล DATE ชนิด Full year Full month และ Full day จะถูกมองเป็น TEXT และใช้วิธีนับความถี่ในการคำนวณข้อมูล เช่น 2018 มกราคม วันจันทร์

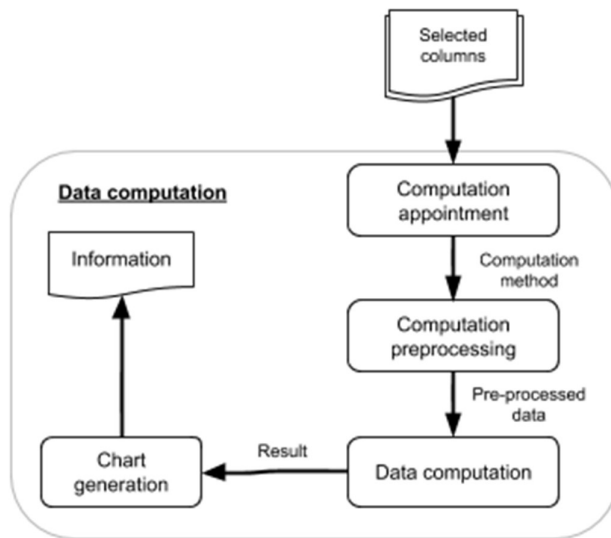
3) ข้อมูล TEXT จะถูกเรียงตามตัวอักษร เช่น กาญจนบุรี ขอนแก่น

4) ข้อมูล NUMBER จะถูกมองเป็น TEXT และเรียงลำดับโดยขนาดของข้อความ

3.2.3.3 การคำนวณข้อมูล (Data computation) เป็นขั้นตอนในการคำนวณข้อมูล โดยจะคำนวณข้อมูลตามรูปแบบแผนภูมิที่ได้เลือกไว้ โดยเบื้องต้นจะนับความถี่ของข้อมูลแต่ละแถวที่ถูกเลือก ในกรณีที่เลือกเป็น Pie chart จะคำนวณอัตราส่วน

3.2.3.4 การสร้างแผนภูมิ (Chart generation) เป็นการสร้างกราฟในรูปแบบไฟล์ PNG (Portable Network Graphics) แล้วแสดงผลในหน้าจอแสดงผล

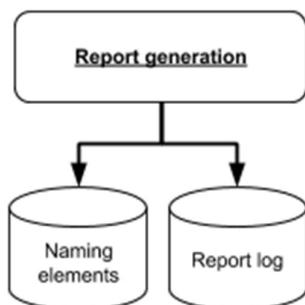
โดยรูปที่ 3.13 แสดงให้เห็นการคำนวณข้อมูลจากรูปแบบแผนภูมิที่ถูกเลือก ประกอบด้วยขั้นตอน 4 ขั้นตอน ซึ่งมีแถวที่ผู้สร้างรายงานเลือกเป็นข้อมูลขาเข้า



รูปที่ 3.13 การคำนวณข้อมูลจากรูปแบบแผนภูมิที่ถูกเลือก (Data computation)

### 3.2.4 การสร้างรายงาน (Report generation)

ขั้นตอนวิธีนี้มีจุดประสงค์เพื่อสร้างรายงานตามที่ต้องการ และเก็บข้อมูลการสร้างรายงาน เมื่อผู้สร้างรายงานเลือกและสร้างรายงาน แบบจำลองจะบันทึกประวัติในการสร้างรายงานครั้งนั้นเข้าสู่ฐานข้อมูล โดยประวัติในการสร้างรายงานนั้นประกอบด้วย ชื่อของรายงาน ชื่อแถวพร้อมกับประเภทข้อมูลของแถวนั้น และประเภทของแผนภูมิที่ผู้สร้างรายงานเลือก การบันทึกประวัตินั้นช่วยให้ผู้ดูแลระบบสามารถทำข้อมูลทางสถิติเพื่อนำข้อมูลทางสถิติเหล่านั้นมาใช้อ้างอิงในกระบวนการตรวจสอบความถูกต้องของการกำหนดประเภทข้อมูลให้กับชื่อแถวได้ ผลลัพธ์จากขั้นตอนนี้คือ ไฟล์รายงานในรูปแบบไฟล์ PDF (Portable Document Format) ดังรูปที่ 3.14



รูปที่ 3.14 การสร้างรายงานตามรูปแบบรายงานที่ผู้สร้างรายงานสนใจ (Report Generation)

## บทที่ 4

### ผลการดำเนินงาน

#### 4.1 การทดสอบแบบจำลอง

การทดสอบแบบจำลองเป็นการพิสูจน์การทำงานของแบบจำลองเพื่อวัดผลว่าสามารถสร้างชื่อของรายงานได้ใกล้เคียงกับความเป็นจริงตามที่ตั้งเกณฑ์ไว้หรือไม่ โดยกลุ่มทดสอบที่เลือกมาใช้ในการทดสอบคือ “หมวดหมู่สังคมและสวัสดิการ” เนื่องจากการดำเนินงานมีลักษณะเช่นเดียวกับกลุ่มธุรกิจบริการ ซึ่งครอบคลุมหน่วยงานที่เกี่ยวข้องกับสถานศึกษาที่ผู้วิจัยสนใจ โดยคลังรายงานของหมวดหมู่สังคมและสวัสดิการมีขนาดใหญ่ และมีลักษณะของชื่อรายงานที่ประกอบจากคำทั่วไป โดยมีคำเฉพาะเฉลี่ยน้อยกว่า 3 คำ ซึ่งการทดสอบแบบจำลองแบ่งออกเป็น 3 การทดสอบ ดังนี้

- 1) การทดสอบความคล้ายกันของชื่อรายงานด้วยวิธีการ Cosine similarity
- 2) การทดสอบการตั้งชื่อรายงานด้วยคำแต่ละประเภทและทดสอบความคล้ายกันด้วยวิธีการ Cosine similarity
- 3) การทดสอบการตั้งชื่อรายงานด้วยคำแต่ละประเภทโดยแยกตามประเภทข้อมูลของแต่ละแถวข้อมูลและทดสอบความคล้ายกันด้วยวิธีการ Cosine similarity

#### 4.1.1 การทดสอบความคล้ายกันของชื่อรายงานด้วยวิธีการ Cosine similarity

##### 4.1.1.1 จุดประสงค์ของการทดสอบ

- 1) เพื่อทดสอบขั้นตอนการสร้างชื่อรายงานด้วยการผสมคำสำคัญ
- 2) เพื่อทดสอบความคล้ายกันของชื่อรายงานต้นฉบับกับชื่อรายงานที่ผ่านการสร้างชื่อรายงานด้วยการผสมคำสำคัญ
- 3) เพื่อทดสอบความคล้ายกันของชื่อรายงานต้นฉบับกับชื่อรายงานที่ผ่านการสร้างชื่อรายงานด้วยการผสมคำสำคัญและการตรวจสอบความถูกต้องของคำสำคัญ

##### 4.1.1.2 สมมติฐาน

- 1) ชื่อของรายงานที่มีความคล้ายกับชื่อรายงานต้นฉบับมากที่สุด 10 อันดับแรกมีค่าความคล้ายกันเฉลี่ยที่ 0.5 ที่ระดับความเชื่อมั่น 0.95
- 2) ค่าเฉลี่ยความคล้ายกันของชื่อของรายงานที่ผ่านการตรวจสอบความถูกต้องของคำสำคัญให้ผลลัพธ์ที่ดีกว่าชื่อรายงานที่ไม่ผ่านการตรวจสอบความถูกต้องของคำสำคัญ

การทดสอบนี้มีจุดประสงค์เพื่อทดสอบกระบวนการในการบันทึกคำสำคัญ การผสมคำสำคัญเป็นชื่อรายงาน และการตรวจสอบความถูกต้องของคำสำคัญ โดยการทดลองแบ่งเป็น 2 ส่วน คือ

- 1) การทดสอบความคล้ายกันของชื่อรายงานต้นฉบับ และชื่อรายงานที่ผ่านกระบวนการผสมคำสำคัญโดยไม่ผ่านการตรวจสอบความถูกต้องของคำสำคัญ
- 2) การทดสอบความแตกต่างของค่าเฉลี่ยระหว่างความคล้ายกันของชื่อรายงานที่ผสมจากคำสำคัญที่ไม่ผ่านการตรวจสอบความถูกต้องของคำสำคัญและชื่อรายงานที่ผสมจากคำสำคัญที่ผ่านการตรวจสอบความถูกต้องของคำสำคัญ

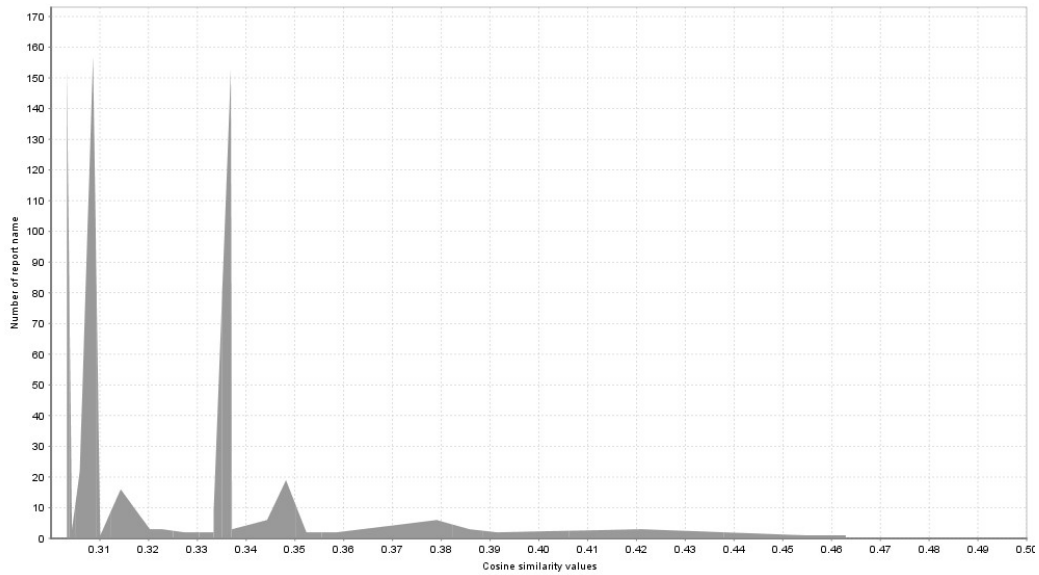
การทดสอบส่วนที่ 1 เป็นการทดสอบเพื่อแสดงให้เห็นผลลัพธ์จากการตัดคำ การบันทึกคำสำคัญ และการผสมคำสำคัญเป็นชื่อรายงาน โดยทั้งหมดเป็นกระบวนการแบบอัตโนมัติ การทดสอบนี้จะแสดงให้เห็นผลลัพธ์ความใกล้เคียงของชื่อรายงานที่สร้างผ่านแบบจำลองกับชื่อรายงานต้นฉบับ โดยใช้การทดสอบ Cosine Similarity ในการทดสอบและใช้ Z - test ในการทดสอบสมมติฐาน การทดสอบด้วย Cosine Similarity นั้น เป็นการทดสอบโดยการหาความคล้ายกันของข้อมูลสองชุด โดยข้อมูลสองชุดในที่นี้คือ ชื่อรายงานต้นฉบับ และ ชื่อรายงานที่สร้างจากแบบจำลอง

$$\frac{\sum_{n=0}^{n-1} X_n Y_n}{\sqrt{\sum_{n=0}^{n-1} (X)^2} \sqrt{\sum_{n=0}^{n-1} (Y)^2}} \quad (4.1)$$

จากสมการที่ 4.1 กำหนดให้  $x$  คือ เซตของความถี่คำสำคัญที่ปรากฏในชื่อรายงานต้นฉบับ  $y$  คือ เซตของความถี่คำสำคัญที่ปรากฏในชื่อรายงานที่ผ่านการสร้างชื่อรายงานด้วยการผสมคำสำคัญ โดย  $x$  และ  $y$  นั้น เป็นคำสำคัญที่เกิดขึ้นทั้งในชื่อรายงานต้นฉบับและชื่อรายงานที่ผ่านการสร้างชื่อรายงานด้วยการผสมคำสำคัญ  $\text{word}(X) \cap \text{word}(Y)$

#### 4.1.1.3 ขั้นตอนการทดลองเพื่อการพิสูจน์สมมติฐานที่ 1

- 1) รวบรวมชื่อของรายงานจากหมวดหมู่สังคมและสวัสดิการ ทั้งหมด 64 ฉบับ ไว้ในไฟล์เอกสาร
- 2) นำชื่อของรายงานเข้าสู่กระบวนการตัดคำและตัดคำหยุด เพื่อให้เหลือเฉพาะคำสำคัญ
- 3) บันทึกคำสำคัญ โดยบันทึกคำแรกและคำสุดท้ายลงในส่วนเริ่มต้นชื่อรายงาน (Beginning Phrase) และส่วนท้ายชื่อรายงาน (Ending Phrase) ส่วนคำสำคัญที่เหลือจะบันทึกในส่วนกลางของชื่อรายงาน (Middle Phrase)
- 4) สร้างรายชื่อของรายงานด้วยขั้นตอนวิธีการสร้างชื่อรายงานด้วยการสร้างชื่อรายงานด้วยการผสมคำสำคัญ (Report title generation)
- 5) คำนวณความคล้ายกันของชื่อรายงานด้วย Cosine Similarity โดยการแปลงชื่อแต่ละชื่อเป็นเวกเตอร์ (Vector) โดยการนับความถี่ของคำสำคัญที่เกิดร่วมกันระหว่างชื่อที่ผ่านการสร้างโดยแบบจำลองและชื่อรายงานจากหมวดหมู่สังคมและสวัสดิการ



รูปที่ 4.1 แผนภูมิแสดงค่าความคล้ายกันของข้อมูล

#### 4.1.1.4 ขั้นตอนการพิสูจน์สมมติฐานที่ 1

1) เรียงลำดับคะแนน Cosine Similarity จากมากไปหาน้อย  
 2) เลือกข้อมูลตัวอย่าง 10 อันดับแรกจากคะแนนที่เรียงไว้ เพื่อเป็นตัวอย่างในการอนุมานค่าสถิติ

3) หาค่าเฉลี่ยของคะแนน Cosine Similarity ของตัวอย่าง 10 ค่า

4) ทดสอบสมมติฐานด้วย Z-test

$$Z = \frac{\bar{X} - \mu}{\frac{\delta}{\sqrt{n}}} \quad (4.2)$$

จากสมการที่ 4.2 กำหนดให้  $\bar{X}$  คือ ค่าเฉลี่ยของกลุ่มตัวอย่าง  $\mu$  คือ ค่าเฉลี่ยตามสมมติฐาน  $\delta$  คือ ส่วนเบี่ยงเบนมาตรฐาน และ  $n$  คือ จำนวนกลุ่มตัวอย่าง

## 4.1.1.5 ผลการพิสูจน์สมมติฐานที่ 1

กำหนดให้

$$H_0: \mu = 0.5$$

$$H_1: \mu \neq 0.5$$

ค่าสถิติเบื้องต้นจากกลุ่มตัวอย่างมีดังนี้

ตารางที่ 4.1 ค่าสถิติเบื้องต้นของกลุ่มตัวอย่างการทดลองที่ 1.1

ค่าเฉลี่ยกลุ่มตัวอย่าง ( $\bar{X}$ )	0.5151
ส่วนเบี่ยงเบนมาตรฐาน ( $\delta$ )	0.0315
จำนวนสมาชิกกลุ่มตัวอย่าง ( $n$ )	10

แทนค่าลงในสมการ (4.2) โดยกำหนดค่าเฉลี่ยตามสมมติฐานที่ 1 ให้

$\mu = 0.50$  จะได้ดังนี้

$$Z = \frac{\bar{X} - \mu}{\frac{\delta}{\sqrt{n}}}$$

$$Z = \frac{0.5151 - 0.50}{\frac{0.0315}{\sqrt{10}}}$$

$$Z = 1.5164$$

ที่ระดับความเชื่อมั่น 0.05 ค่าวิกฤตที่ได้ คือ 1.96 สรุปได้ว่า  $1.5164 < 1.96$  จึงยอมรับ  $H_0$  ค่าเฉลี่ยของกลุ่มตัวอย่าง เท่ากับ 0.5 อย่างมีนัยสำคัญ



การทดสอบส่วนที่ 2 เป็นการทดสอบเพื่อแสดงให้เห็นผลลัพธ์ที่มีแนวโน้มดีขึ้นจากขั้นตอนการตรวจสอบความถูกต้องของคำสำคัญ

#### 4.1.1.6 ขั้นตอนการทดลองเพื่อการพิสูจน์สมมติฐานที่ 2

1) รวบรวมชื่อของรายงานจากหมวดหมู่สังคมและสวัสดิการ ทั้งหมด 64 ฉบับ ไว้ในไฟล์เอกสาร

2) นำชื่อของรายงานเข้าสู่กระบวนการตัดคำและตัดคำหยุด

3) บันทึกลำดับคำสำคัญ โดยบันทึกคำแรกและคำสุดท้ายลงในส่วนเริ่มต้นชื่อรายงาน (Beginning Phrase) และส่วนท้ายชื่อรายงาน (Ending Phrase) นอกเหนือจากนั้นบันทึกในส่วนของชื่อรายงาน (Middle Phrase)

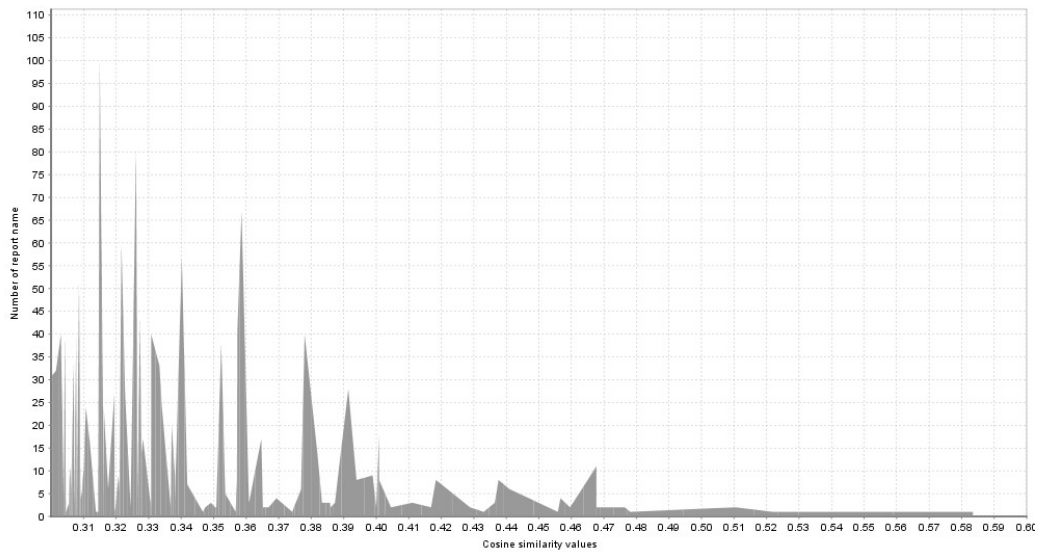
4) สร้างรายชื่อของรายงานด้วยขั้นตอนวิธีการสร้างชื่อรายงานด้วยการสร้างชื่อรายงานด้วยการผสมคำสำคัญ (Report title generation)

5) คำนวณความคล้ายกันของชื่อรายงานด้วย Cosine Similarity บันทึกค่า Cosine Similarity ไว้ในไฟล์เอกสารชุดที่ 1 โดยกำหนดให้กลุ่มนี้ คือ กลุ่มที่ไม่ผ่านการตรวจสอบความถูกต้องของคำสำคัญ

6) ตรวจสอบความเหมาะสมของคำสำคัญโดยขั้นตอนการตรวจสอบความถูกต้องของคำสำคัญ

7) สร้างรายชื่อของรายงานด้วยขั้นตอนวิธีการสร้างชื่อรายงานด้วยการผสมคำสำคัญ (Report title generation)

8) คำนวณความคล้ายกันของชื่อรายงานด้วย Cosine Similarity บันทึกค่า Cosine Similarity ไว้ในไฟล์เอกสารชุดที่ 2 โดยกำหนดให้กลุ่มนี้ คือ กลุ่มที่ผ่านการตรวจสอบความถูกต้องของคำสำคัญ



รูปที่ 4.2 แผนภูมิแสดงค่าความคล้ายกันของข้อมูลของคำสำคัญในการทดลองที่ 1 สมมติฐานที่ 2

#### 4.1.1.7 ขั้นตอนการพิสูจน์สมมติฐานที่ 2

- 1) เรียงลำดับคะแนน Cosine Similarity ของไฟล์เอกสารชุดที่ 1 และชุดที่ 2 จากมากไปหาน้อย
- 2) เลือกกลุ่มตัวอย่าง 100 อันดับแรกจากคะแนน Cosine Similarity ที่เรียงไว้ โดยแยกเป็นตัวอย่างข้อมูลจากไฟล์เอกสารชุดที่ 1 และชุดที่ 2
- 3) หาค่าเฉลี่ยของคะแนน Cosine Similarity ของกลุ่มตัวอย่างทั้งสองกลุ่ม
- 4) ทดสอบสมมติฐานด้วย Z-test

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\delta_1^2}{n_1} + \frac{\delta_2^2}{n_2}}} \quad (4.3)$$

จากสมการที่ 4.3 กำหนดให้  $\bar{X}_1$  คือ ค่าเฉลี่ยของกลุ่มตัวอย่างที่ 1  
 สัญลักษณ์  $\bar{X}_2$  คือ ค่าเฉลี่ยของกลุ่มตัวอย่างที่ 2 สัญลักษณ์  $\delta_1$  คือ ส่วนเบี่ยงเบนมาตรฐานของกลุ่ม  
 ตัวอย่างที่ 1 สัญลักษณ์  $\delta_2$  คือ ส่วนเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่างที่ 2 สัญลักษณ์  $n_1$  คือ  
 จำนวนกลุ่มตัวอย่างที่ 1 และสัญลักษณ์  $n_2$  คือ จำนวนกลุ่มตัวอย่างที่ 2

#### 4.1.1.8 ผลการพิสูจน์สมมติฐานที่ 2

กำหนดให้

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 > \mu_2$$

เนื่องจากสมมติฐานเป็นการทดสอบค่าเฉลี่ยว่ามีผลลัพธ์ที่เพิ่มขึ้นหรือไม่ จึง  
 กำหนดให้กลุ่มตัวอย่างที่ 1 คือ กลุ่มตัวอย่างที่ผ่านการตรวจสอบความถูกต้องของคำสั่ง และกลุ่ม  
 ตัวอย่างที่ 2 คือ กลุ่มตัวอย่างที่ไม่ผ่านการตรวจสอบความถูกต้องของคำสั่ง  
 ค่าสถิติเบื้องต้นจากกลุ่มตัวอย่างมีดังนี้

ตารางที่ 4.2 ค่าสถิติเบื้องต้นของกลุ่มตัวอย่างการทดลองที่ 1.2

	กลุ่มตัวอย่างที่ 1	กลุ่มตัวอย่างที่ 2
ค่าเฉลี่ยกลุ่มตัวอย่าง ( $\bar{X}$ )	0.4355	0.3519
ส่วนเบี่ยงเบนมาตรฐาน ( $\delta$ )	0.0374	0.0291
จำนวนสมาชิกกลุ่มตัวอย่าง ( $n$ )	100	100

แทนค่าลงในสมการ

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\delta_1^2}{n_1} + \frac{\delta_2^2}{n_2}}}$$

$$Z = \frac{0.4355 - 0.3519}{\sqrt{\frac{0.0374^2}{100} + \frac{0.0291^2}{100}}}$$

$$Z = 17.6446$$

ที่ระดับความเชื่อมั่น 0.05 ค่าวิกฤตที่ได้ คือ 1.96 สรุปได้ว่า  $17.6446 > 1.96$  จึงปฏิเสธ  $H_0$  ยอมรับ  $H_1$  ค่าเฉลี่ยของกลุ่มตัวอย่างที่ 1 มากกว่าค่าเฉลี่ยของกลุ่มตัวอย่างที่ 2 อย่างมีนัยสำคัญ

4.1.2 การทดสอบการตั้งชื่อรายงานด้วยคำแต่ละประเภทและทดสอบความคล้ายกันด้วยวิธีการ Cosine Similarity

#### 4.1.2.1 จุดประสงค์ของการทดสอบ

- 1) เพื่อเปรียบเทียบการผสมชื่อของรายงานจากประเภทคำสำคัญที่ต่างกัน
- 2) เพื่อนำผลไปใช้ในกระบวนการตรวจสอบความถูกต้องของประเภทข้อมูล

#### 4.1.2.2 สมมติฐาน

- 1) คำสำคัญที่ผ่านการปรับปรุงประเภทคำสำคัญให้ตรงตามสถิติประเภทคำสำคัญที่ใช้ในชื่อรายงานให้คะแนนความคล้ายกันมากกว่าคำสำคัญที่ไม่ผ่านการปรับปรุงประเภทคำสำคัญ

#### 4.1.2.3 ขั้นตอนการทดลองเพื่อพิสูจน์สมมติฐาน

- 1) รวบรวมชื่อของรายงานจากหมวดหมู่สังคมและสวัสดิการ ทั้งหมด 64 ฉบับ ไว้ในไฟล์เอกสาร
- 2) นำชื่อของรายงานเข้าสู่กระบวนการตัดคำ
- 3) ทำการกำหนดชนิดของคำสำคัญในแต่ละตำแหน่ง (Part of speech tagging)
- 4) เรียงลำดับคำสำคัญที่นิยมใช้ในชื่อรายงาน โดยนับความถี่ของคำที่ปรากฏในรายงาน
- 5) เลือกคำสำคัญ 10 อันดับแรกของแต่ละตำแหน่ง จากทั้งหมด 28 ตำแหน่ง

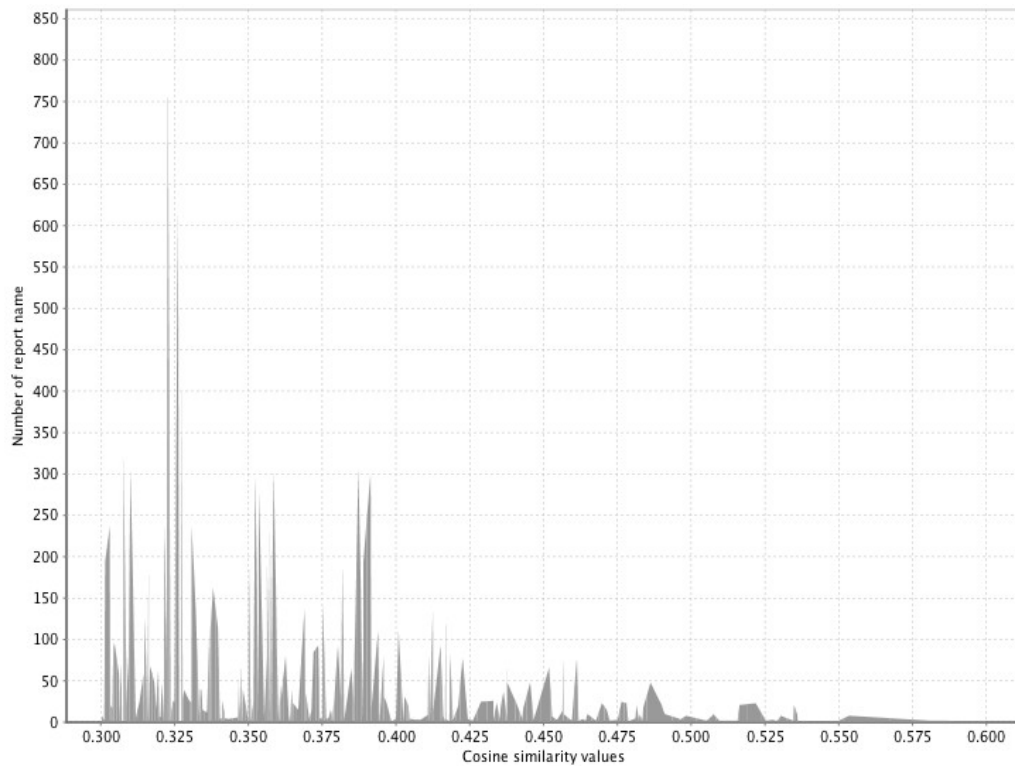
- 6) รวบรวมคำสำคัญทั้งหมดไว้ในไฟล์เอกสารเดียวกัน และลดความซ้ำซ้อนของคำสำคัญที่เกิดขึ้นซ้ำกัน
- 7) บันทึกลำดับคำสำคัญโดยให้แบ่งออกเป็น 4 กลุ่ม
- 8) ตรวจสอบความเหมาะสมของคำสำคัญโดยขั้นตอนการตรวจสอบความถูกต้องของคำสำคัญ
- 9) สร้างรายชื่อของรายงานด้วยขั้นตอนวิธีการสร้างชื่อรายงานด้วยการผสมคำสำคัญ (Report title generation)
- 10) คำนวณความคล้ายกันของชื่อรายงานด้วย Cosine Similarity บันทึกค่า Cosine Similarity ไว้ในไฟล์เอกสาร โดยกำหนดให้กลุ่มนี้คือกลุ่มที่ผ่านการการปรับปรุงประเภทคำสำคัญ
- 11) เปรียบเทียบค่าเฉลี่ยของกลุ่มที่ผ่านการการปรับปรุงประเภทคำสำคัญกับกลุ่มที่ผ่านการตรวจสอบความถูกต้องของคำสำคัญ

#### 4.1.2.4 ขั้นตอนการพิสูจน์สมมติฐาน

- 1) เรียงลำดับคะแนน Cosine Similarity ของไฟล์เอกสารจากมากไปหาน้อย
- 2) เลือกกลุ่มตัวอย่าง 10 อันดับแรกจากคะแนนที่เรียงไว้
- 3) หาค่าเฉลี่ยของคะแนนของกลุ่มที่ผ่านการการปรับปรุงประเภทคำสำคัญ
- 4) ทดสอบสมมติฐานด้วย Z-test โดยเปรียบเทียบค่าเฉลี่ยของกลุ่มที่ผ่านการปรับปรุงประเภทคำสำคัญกับกลุ่มที่ผ่านการตรวจสอบความถูกต้องของคำสำคัญ

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\delta_1^2}{n_1} + \frac{\delta_2^2}{n_2}}} \quad (4.4)$$

จากสมการที่ 4.4 กำหนดให้สัญลักษณ์  $\bar{X}_1$  คือ ค่าเฉลี่ยของกลุ่มตัวอย่างที่ 1 สัญลักษณ์  $\bar{X}_2$  คือ ค่าเฉลี่ยของกลุ่มตัวอย่างที่ 2 สัญลักษณ์  $\delta_1$  คือ ส่วนเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่างที่ 1 สัญลักษณ์  $\delta_2$  คือ ส่วนเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่างที่ 2 สัญลักษณ์  $n_1$  คือ จำนวนกลุ่มตัวอย่างที่ 1 สัญลักษณ์  $n_2$  คือ จำนวนกลุ่มตัวอย่างที่ 2



รูปที่ 4.3 แผนภูมิแสดงค่าความคล้ายกันของข้อมูลของคำสำคัญในการทดลองที่ 2

#### 4.1.2.5 ผลการพิสูจน์สมมติฐานที่ 2

กำหนดให้

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 > \mu_2$$

เนื่องจากสมมติฐานเป็นการทดสอบค่าเฉลี่ยว่ามีผลลัพธ์ที่เพิ่มขึ้นหรือไม่ จึง

กำหนดให้กลุ่มตัวอย่างที่ 1 คือ กลุ่มตัวอย่างที่ผ่านการปรับปรุงประเภทคำสำคัญ และกลุ่มตัวอย่างที่

2 คือ กลุ่มตัวอย่างที่การตรวจสอบความถูกต้องของคำสำคัญ

ค่าสถิติเบื้องต้นจากกลุ่มตัวอย่างมีดังนี้

ตารางที่ 4.3 ค่าสถิติเบื้องต้นของกลุ่มตัวอย่างการทดลองที่ 2.1

	กลุ่มตัวอย่างที่ 1	กลุ่มตัวอย่างที่ 2
ค่าเฉลี่ยกลุ่มตัวอย่าง ( $\bar{X}$ )	0.5840	0.5151
ส่วนเบี่ยงเบนมาตรฐาน ( $\delta$ )	0.0117	0.031
จำนวนสมาชิกกลุ่มตัวอย่าง ( $n$ )	10	10

แทนค่าลงในสมการจะได้

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\delta_1^2}{n_1} + \frac{\delta_2^2}{n_2}}}$$

$$Z = \frac{0.5840 - 0.5151}{\sqrt{\frac{0.0117^2}{10} + \frac{0.031^2}{10}}}$$

$$Z = 6.4693$$

ที่ระดับความเชื่อมั่น 0.05 ค่าวิกฤตที่ได้ คือ 1.96 สรุปได้ว่า  $6.4693 > 1.96$  จึงปฏิเสธ  $H_0$  ยอมรับ  $H_1$  ค่าเฉลี่ยของกลุ่มตัวอย่างที่ 1 มากกว่าค่าเฉลี่ยของกลุ่มตัวอย่างที่ 2 อย่างมีนัยสำคัญ

4.1.3 การทดสอบการตั้งชื่อรายงานด้วยคำแต่ละประเภทโดยแยกตามประเภทข้อมูลของแต่ละแถวข้อมูลและทดสอบความคล้ายกันด้วยวิธีการ Cosine similarity

#### 4.1.3.1 จุดประสงค์ของการทดสอบ

1) เพื่อเปรียบเทียบการผสมชื่อของรายงานโดยการเลือกแถวในการสร้างรายงานที่ต่างกัน

#### 4.1.3.2 สมมติฐาน

1) การสร้างชื่อรายงานจากการเลือกแถวในการสร้างรายงานที่ต่างกัน ให้ผลความใกล้เคียงไม่ต่างกัน

#### 4.1.3.3 ขั้นตอนการทดลองเพื่อพิสูจน์สมมติฐาน

- 1) รวบรวมชื่อของรายงานจากหมวดหมู่สังคมและสวัสดิการทั้งหมด 64 ฉบับ ไว้ในไฟล์เอกสาร
- 2) นำชื่อของรายงานเข้าสู่กระบวนการตัดคำ
- 3) นำชื่อของรายงานเข้าสู่กระบวนการตัดคำหยุด
- 4) บันทึกลำดับคำสำคัญโดยให้คำแรกและคำสุดท้ายเป็นประธาน และส่วนขยายชื่อรายงาน นอกเหนือจากนั้นให้เป็น ส่วนขยายประธานและส่วนขยายกริยา
- 5) สร้างรายชื่อของรายงานด้วยขั้นตอนวิธีการสร้างชื่อรายงานด้วยการผสมคำสำคัญ (Report title generation) โดยทำการสร้างชื่อรายงานจากการเลือกข้อมูล 3 ชนิด
  - 5.1) ชนิดข้อความ
  - 5.2) ชนิดตัวเลข
  - 5.3) ชนิดวันเวลา
- 6) คำนวณความคล้ายกันของชื่อรายงานด้วย Cosine Similarity โดยเปรียบเทียบกับรายงานต้นฉบับ และบันทึกผลลัพธ์เป็นไฟล์เอกสาร 3 ไฟล์ แยกตามประเภทข้อมูลของแถวข้อมูลที่ใช้ในการสร้างรายงาน

ตารางที่ 4.4 ลำดับคะแนน Cosine Similarity ในอันดับ 1 - 5

ประเภทข้อมูล	ลำดับคะแนน Cosine Similarity				
	1	2	3	4	5
ข้อความ	0.5996	0.5954	0.5892	0.5892	0.5892
ตัวเลข	0.5303	0.5303	0.5303	0.5303	0.5303
วันเวลา	0.5892	0.5892	0.5809	0.5809	0.5809



ตารางที่ 4.5 ลำดับคะแนน Cosine Similarity ในอันดับ 6 - 10

ประเภทข้อมูล	ลำดับคะแนน Cosine Similarity				
	6	7	8	9	10
ข้อความ	0.5870	0.5870	0.5834	0.5834	0.5809
ตัวเลข	0.5303	0.5303	0.5163	0.5163	0.5163
วันเวลา	0.5809	0.5809	0.5809	0.5359	0.5359

## 4.1.3.4 ขั้นตอนการพิสูจน์สมมติฐาน

1) เรียงลำดับคะแนน Cosine Similarity ของไฟล์เอกสารชุดที่ 1 ชุดที่ 2 และชุดที่ 3 จากมากไปหาน้อย

2) เลือกกลุ่มตัวอย่าง 10 อันดับแรกจากคะแนนที่เรียงไว้ โดยแยกเป็นตัวอย่างข้อมูลจากไฟล์เอกสารชุดที่ 1 ชุดที่ 2 และชุดที่ 3

3) หาค่าเฉลี่ยของคะแนนของกลุ่มตัวอย่างทั้งสองกลุ่ม

4) ทดสอบสมมติฐานด้วย ANOVA

## 4.1.3.5 ผลการพิสูจน์สมมติฐาน

กำหนดให้

$$H_0: \mu_1 = \mu_2 = \mu_3$$

$$H_1: \mu_i \neq \mu_j$$

เนื่องจากสมมติฐานเป็นการทดสอบค่าเฉลี่ยว่ามีผลลัพธ์ที่ต่างกันหรือไม่ จึงกำหนดให้กลุ่มตัวอย่างที่ 1 คือ กลุ่มตัวอย่างที่ผ่านการปรับปรุงประเภทคำสำคัญ และกลุ่มตัวอย่างที่ 2 คือ กลุ่มตัวอย่างที่การตรวจสอบความถูกต้องของคำสำคัญ ค่าสถิติเบื้องต้นจากกลุ่มตัวอย่างมีดังนี้

ตารางที่ 4.6 ค่าสถิติเบื้องต้นของกลุ่มตัวอย่างการทดลองที่ 3

	กลุ่มตัวอย่างที่ 1	กลุ่มตัวอย่างที่ 2	กลุ่มตัวอย่างที่ 3
ค่าเฉลี่ยกลุ่มตัวอย่าง ( $\bar{X}$ )	0.5736	0.5884	0.5261
ส่วนเบี่ยงเบนมาตรฐาน ( $\delta$ )	0.0191	0.0053	0.0063
จำนวนสมาชิกกลุ่มตัวอย่าง ( $n$ )	10	10	10

$$SST = \sum \sum X_{ij}^2 - \frac{T^2}{n} \quad (4.5)$$

จากสมการที่ 4.5 กำหนดให้สัญลักษณ์  $i$  คือ กลุ่มของประเภทข้อมูลซึ่งแบ่งเป็น 3 กลุ่ม คือ ข้อความ ตัวเลข และ วันที่ สัญลักษณ์  $j$  คือ สมาชิกแต่ละตัวในกลุ่ม  $i$  สัญลักษณ์  $X_{ij}$  คือ ค่า Cosine Similarity ของสมาชิก  $j$  ในกลุ่ม  $i$  สัญลักษณ์  $T$  คือ ผลรวมของคะแนน Cosine Similarity ทั้งหมด และสัญลักษณ์  $n$  คือ จำนวนสมาชิกทั้งหมด แทนค่าต่าง ๆ ลงในสมการจะได้ดังนี้

$$SST = (0.5996^2 + 0.5954^2 + \dots + 0.5359^2 + 0.5359^2) - \frac{16.8825^2}{30}$$

$$SST = 9.526226 - 9.500626$$

$$SST = 0.0256$$

$$SSB = \frac{\sum T_j^2}{n_j} - \frac{T^2}{n} \quad (4.6)$$

จากสมการ 4.6 กำหนดให้สัญลักษณ์  $T_j$  คือ ผลรวมของกลุ่มข้อมูล 3 กลุ่ม คือ ข้อความ ตัวเลข และ วันที่ สัญลักษณ์  $n_j$  คือ จำนวนสมาชิกของแต่ละกลุ่ม สัญลักษณ์  $T$  คือ ผลรวมทั้งหมดของสมาชิกจากทั้ง 3 กลุ่ม และสัญลักษณ์  $n$  คือ จำนวนสมาชิกทั้งหมด แทนค่าต่าง ๆ ลงในสมการจะได้ดังนี้

$$SSB = \left( \frac{5.7360^2}{10} + \frac{5.8849^2}{10} + \frac{5.2615^2}{10} \right) - \frac{16.8825}{30}$$

$$SSB = (3.2902 + 3.4632 + 2.7683) - 9.5006$$

$$SSB = 0.0211$$

หาค่า Sum Square of Error (SSE)

$$SSE = SST - SSB \quad (4.7)$$

แทนค่าต่าง ๆ ลงในสมการจะได้ดังนี้

$$SSE = 0.0256 - 0.0211$$

$$SSE = 0.0045$$

หาค่า Mean Square Between (MSB)

$$MSB = \frac{SSB}{df} \quad (4.8)$$

จากสมการ 4.8 กำหนดให้สัญลักษณ์  $df$  คือ ค่าองศาเสรีของกลุ่ม โดยมีค่าเป็นจำนวนกลุ่มลบด้วย 1 ดังนั้นจะได้สมการดังนี้

$$MSB = \frac{SSB}{k - 1} \quad (4.9)$$

แทนค่าต่าง ๆ ลงในสมการจะได้ดังนี้

$$MSB = \frac{0.0211}{3 - 1}$$

$$MSB = 0.0106$$

หาค่า Mean Square Error (MSE)

$$MSE = \frac{SSE}{n - k} \quad (4.10)$$

จากสมการ 4.10 กำหนดให้สัญลักษณ์  $n$  คือ จำนวนสมาชิกทั้งหมด และสัญลักษณ์  $k$  คือ จำนวนกลุ่ม แทนค่าต่าง ๆ ลงในสมการจะได้ดังนี้

$$MSE = \frac{0.0045}{30 - 3}$$

$$MSE = 0.00016$$

หาค่า F

$$F = \frac{MSB}{MSE} \quad (4.11)$$

แทนค่าต่าง ๆ ลงในสมการจะได้ดังนี้

$$F = \frac{0.0106}{0.00016}$$

$$F = 66.25$$

ที่ระดับความเชื่อมั่น 0.05 ค่าวิกฤตที่ได้ คือ 3.3541 สรุปได้ว่า  $66.25 > 3.3541$  จึงปฏิเสธ  $H_0$  ยอมรับ  $H_1$  ค่าเฉลี่ยของกลุ่มตัวอย่างมีอย่างน้อย 1 ค่าที่มีค่าเฉลี่ยแตกต่างจากกลุ่มอื่น

ตารางที่ 4.7 ตารางเปรียบเทียบงานวิจัยที่ประยุกต์ใช้ขั้นตอนวิธีคล้ายคลึงกัน

ผู้วิจัย	จุดประสงค์	ขั้นตอนวิธีที่ใช้	ผลลัพธ์
Steve Chien and Nicole Immorlica (2005)	สร้างคำค้นให้กับผู้ใช้งาน จากการค้นหาความคล้ายกันของข้อความที่ผ่านการปรับขนาดให้เหมาะสม	ค้นหาความสัมพันธ์ระหว่างประโยค (Correlations Efficiently) โดยเลือกคำค้นจากคำค้นที่มีคะแนนไม่น้อยกว่าที่กำหนด (Threshold)	ที่คะแนนความสัมพันธ์ที่ 0.9 แบบจำลองให้ผลลัพธ์ที่ถูกต้อง 19,330 จาก คำค้นที่เป็นไปได้ทั้งหมด 15,000,000
Anna Huang (2008)	เปรียบเทียบขั้นตอนวิธีที่เหมาะสมกับการจัดกลุ่มข้อความ	เปรียบเทียบขั้นตอนวิธีระหว่าง Jaccard Coefficient Pearson Correlation Coefficient Averaged Kullback-Leibler Divergence Euclidean Distance และ Cosine Similarity โดยข้อมูลถูกเตรียมโดยขั้นตอนวิธี K-mean algorithm และเลือก Purity และ Entropy ในการวัดผล	ในการทดสอบแสดงให้เห็นว่า ผลลัพธ์ของ Jaccard Coefficient Pearson Correlation Coefficient และ Cosine Similarity มีประสิทธิภาพที่ใกล้เคียงกันทั้งในการทดสอบ Purity และ Entropy

ตารางที่ 4.7 ตารางเปรียบเทียบงานวิจัยที่ประยุกต์ใช้ขั้นตอนวิธีคล้ายคลึงกัน (ต่อ)

ผู้วิจัย	จุดประสงค์	ขั้นตอนวิธีที่ใช้	ผลลัพธ์
Hadi Mohammadzadeh et al. (2012)	ค้นหาหัวข้อข่าวที่มีความใกล้เคียงกับหัวข้อข่าวและเนื้อหาข่าวที่ปรากฏในเว็บไซต์ตัวอย่าง	หาความคล้ายกันของข้อความด้วย Cosine Similarity โดยวัดผลระหว่างการแปลงข้อความให้เป็น Vector ของขั้นตอนวิธี 4 ขั้นตอนวิธี ประกอบด้วย TF, TF – IDF, OSM และ TF – IDF + OSM เพื่อค้นหาขั้นตอนวิธีที่ให้ผลที่ดีที่สุดในการแนะนำหัวข้อข่าว	ขั้นตอนในการแปลงข้อความเป็น Vector ทั้ง 4 วิธี ให้คะแนนใกล้เคียงกัน โดยให้คะแนนความแม่นยำและถูกต้องอยู่ที่ประมาณ 0.96 คะแนน แต่จะมีความแตกต่างกันในบางกลุ่มข่าว ผู้วิจัยได้กล่าวไว้ว่า หากมีการวิจัยต่อในเรื่องความแตกต่างกันระหว่างแต่ละกลุ่มข่าว ผลลัพธ์อาจมีความแม่นยำขึ้น
งานวิจัยนี้	แนะนำชื่อรายงานที่เหมาะสมกับโครงสร้างข้อมูล	ผสมคำเพื่อสร้างเป็นชื่อรายงานจากคำสำคัญที่ผ่านการตัดคำและตัดคำหยุด โดยชื่อรายงานนั้นประกอบจากคำสำคัญ 4 ตำแหน่ง	แบบจำลองให้ผลลัพธ์การสร้างชื่อใกล้เคียงกับชื่อของรายงานที่ใช้งานในหน่วยงานภาครัฐในประเทศไทย โดยให้คะแนน Cosine Similarity อยู่ที่ประมาณ 0.6

## 4.2 การพัฒนาระบบ

ในส่วนนี้ผู้วิจัยจะนำเสนอเว็บแอปพลิเคชันที่พัฒนาตามแบบจำลองที่ได้ออกแบบไว้ โดยแบ่งโครงสร้างออกเป็น 4 ส่วน ดังนี้

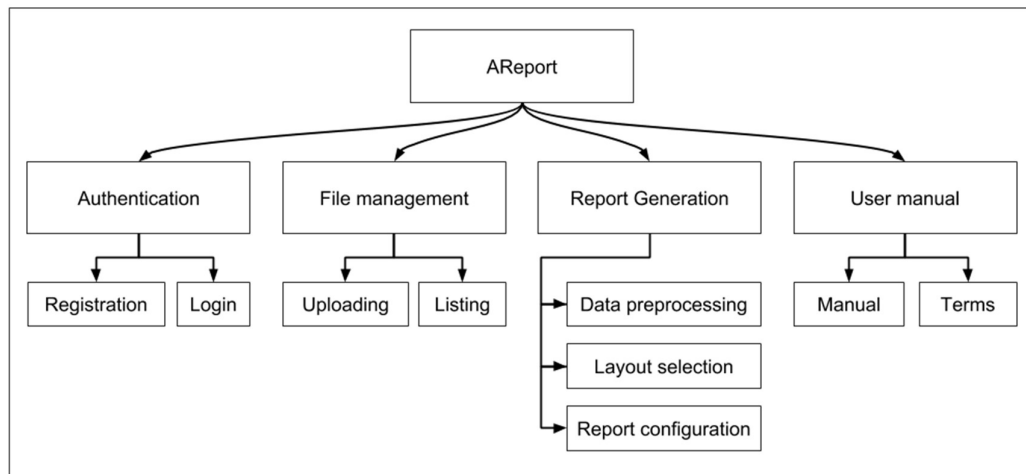
4.2.1 การยืนยันตัวตนผู้ใช้งาน (Authentication)

4.2.2 การจัดการไฟล์ข้อมูล (File management)

4.2.3 การสร้างรายงาน (Report generation)

4.2.4 คู่มือการใช้งาน (Manual)

ซึ่งรูปที่ 4.4 แสดงให้เห็นโครงสร้างของเว็บแอปพลิเคชันที่พัฒนาขึ้น โดยแสดงส่วนประกอบทั้ง 4 ส่วน



รูปที่ 4.4 โครงสร้างเว็บแอปพลิเคชัน

4.2.1 การยืนยันตัวตนผู้ใช้งาน (Authentication) เป็นส่วนสำหรับให้ผู้สร้างรายงานสร้างบัญชีผู้ใช้งานระบบ และเข้าใช้งานระบบ โดยแบ่งออกเป็น 2 ส่วน คือ

4.2.1.1 ส่วนลงทะเบียนบัญชีผู้ใช้งานระบบ (Registration) เป็นส่วนสำหรับลงทะเบียนบัญชีผู้ใช้งานระบบเพื่อเข้าใช้งานระบบ โดยผู้สร้างรายงานจะต้องกรอกข้อมูลดังนี้

- 1) ชื่อ เป็นภาษาไทยหรืออังกฤษ
- 2) นามสกุล เป็นภาษาไทยหรืออังกฤษ

- 3) ชื่อผู้สร้างรายงาน เป็นภาษาอังกฤษ
- 4) อีเมล เป็นข้อมูลสำหรับติดต่อ
- 5) รหัสผ่าน ประกอบด้วย ภาษาอังกฤษ หรือตัวเลข
- 6) ยืนยันรหัสผ่าน

ซึ่งรูปที่ 4.5 แสดงให้เห็นหน้าจอลงทะเบียนบัญชีผู้ใช้งานระบบ โดยผู้ใช้งานต้องกรอกข้อมูลให้ครบถ้วนเพื่อสร้างบัญชีผู้ใช้งานระบบ

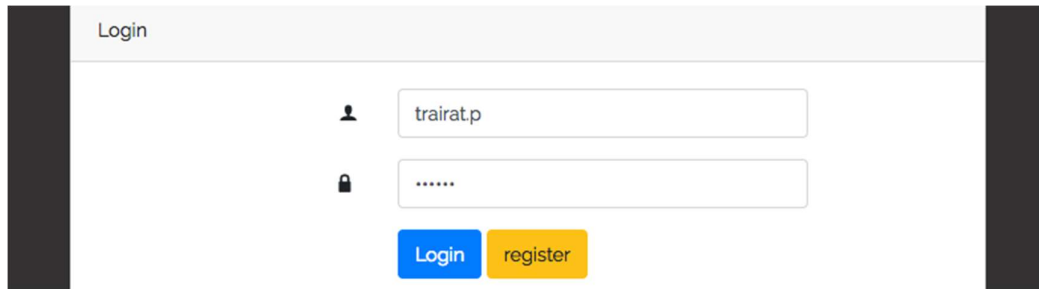
The image shows a web form titled "Register". It contains six input fields stacked vertically: "First name", "Last name", "Username", "E-Mail Address", "Password", and "Confirm Password". Below the fields are two buttons: a blue "Register" button and a yellow "Cancel" button.

รูปที่ 4.5 ภาพตัวอย่างหน้าจอลงทะเบียนผู้ใช้งานระบบ

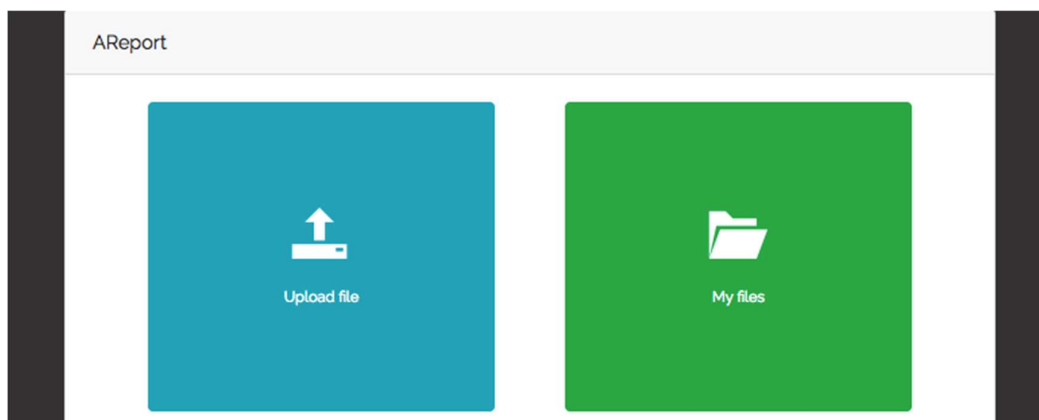
#### 4.2.1.2 ส่วนเข้าสู่ระบบ (Login)

เมื่อผู้สร้างรายงานสร้างบัญชีผู้ใช้งานระบบแล้ว สามารถเข้าสู่ระบบได้โดยกรอกชื่อบัญชีผู้ใช้งานระบบและรหัสผ่านลงในหน้าจอระบบดังรูปที่ 4.6 และเมื่อเข้าสู่ระบบแล้วระบบจะแสดงหน้าจอต้อนรับดังรูปที่ 4.7 ซึ่งประกอบด้วย 2 ส่วน คือ ส่วนสำหรับเข้าสู่หน้าจัดการไฟล์ข้อมูล และส่วนสำหรับเข้าสู่หน้าอัปโหลดไฟล์เข้าสู่ระบบ





รูปที่ 4.6 ภาพตัวอย่างหน้าจอเข้าสู่ระบบ



รูปที่ 4.7 ภาพตัวอย่างหน้าจอต้อนรับ

#### 4.2.2 การจัดการไฟล์ข้อมูล (File management)

เป็นส่วนสำหรับการจัดการไฟล์ข้อมูล โดยแบ่งออกเป็น 2 ส่วน คือ

4.2.2.1 ส่วนอัปโหลดไฟล์ข้อมูลเข้าสู่ระบบ (Upload) ใช้สำหรับอัปโหลดไฟล์ข้อมูลเข้าสู่ระบบ โดยผู้สร้างรายงานจะต้องเลือกไฟล์ ชนิด Comma Separated Values (CSV) และตัวคั่นข้อมูลต้องเป็นสัญลักษณ์ | เท่านั้น โดยส่วนนี้มีลักษณะหน้าจอดังรูปที่ 4.8



รูปที่ 4.8 ภาพตัวอย่างหน้าจออัปโหลดไฟล์

#### ขั้นตอนวิธีที่ 4.1 อัปโหลดไฟล์ข้อมูลเข้าสู่ระบบ (File upload)

```

Function: File upload
  userId = getUserIdFromSession()
  destinationPath = `uploads/` + userId + ` ` + date('Y-m-d')
  modifiedFileName = date('H-i-s') + ` ` + getFileName()
  Upload file To destinationPath With modifiedFileName
  If Upload success Then
    Save destinationPath And modifiedFileName To database
    Return redirect To file list page
  Else
    Return redirect To file upload page
  End

```

ขั้นตอนวิธีที่ 4.1 แสดงให้เห็นขั้นตอนในการอัปโหลดไฟล์ โดยมีขั้นตอนดังนี้

- 1) บันทึกรหัสผู้สร้างรายงานไว้ใน userId เพื่อใช้ในการระบุตัวตนผู้อัปโหลดไฟล์
- 2) สร้างตัวแปร destinationPath ซึ่งเก็บค่าโฟลเดอร์ปลายทางของการอัปโหลดโดยปลายทางนั้นจะมีค่าเป็น “uploads > รหัสผู้สร้างรายงาน > ปี-เดือน-วันของวันขณะที่ทำการอัปโหลดไฟล์ข้อมูล”
- 3) สร้างตัวแปร modifiedFileName ซึ่งเก็บค่าชื่อไฟล์ที่ถูกบันทึกลงในฐานข้อมูล โดยชื่อนั้นจะถูกดัดแปลงเป็น “ชั่วโมง-นาที่-วินาที-ตามด้วยชื่อดั้งเดิมของไฟล์ข้อมูล” เพื่อหลีกเลี่ยงชื่อที่อาจซ้ำกัน
- 4) อัปโหลดไฟล์ไปยังตำแหน่งไฟล์ปลายทาง
- 5) หากอัปโหลดสำเร็จจะบันทึกไฟล์และตำแหน่งไฟล์เข้าสู่ฐานข้อมูล
- 6) นำผู้สร้างรายงานไปยังหน้าแสดงไฟล์ในระบบต่อไป
- 7) หากอัปโหลดไม่สำเร็จ จะนำผู้สร้างรายงานกลับไปยังหน้าอัปโหลดไฟล์ เพื่อทำการอัปโหลดใหม่ โดยหน้าจอส่วนอัปโหลดไฟล์มีลักษณะดังรูปที่ 4.8

#### 4.2.2.2 แสดงรายชื่อไฟล์ข้อมูลในระบบ (List)

ส่วนนี้เป็นส่วนสำหรับแสดงรายชื่อไฟล์ทั้งหมดที่ผู้สร้างรายงานอัปโหลดเรียงลำดับตามวันที่ปัจจุบันไปถึงอดีต โดยส่วนนี้จะอนุญาตให้ผู้สร้างรายงานลบไฟล์ของตัวเองได้และเลือกไฟล์เพื่อใช้ในการสร้างรายงานได้ ดังรูปที่ 4.9



รูปที่ 4.9 ภาพตัวอย่างหน้าจอรายชื่อไฟล์

#### 4.2.3 การสร้างรายงาน (Report generation)

ส่วนนี้เป็นส่วนสำหรับสร้างรายงาน โดยแบ่งเป็น 3 ขั้นตอน คือ

##### 4.2.3.1 การเตรียมข้อมูลเพื่อใช้ในการคำนวณข้อมูล (Data preprocessing)

ในขั้นตอนนี้จะเป็นการนำขั้นตอนวิธีการตรวจสอบประเภทข้อมูลโดยใช้หัวตารางมาพัฒนาเป็นการทำงาน โดยระบบจะทำการตรวจสอบประเภทข้อมูลที่ตรงกับชื่อหัวตารางแล้วแสดงผลลัพธ์ในรูปแบบรายการเลือกแบบดิ่งลง โดยรายการเลือกแบบดิ่งลงจะเลือกประเภทข้อมูลที่ตรงกับหัวตารางให้เป็นตัวเลือกตั้งต้น ผู้สร้างรายงานสามารถแก้ไขประเภทข้อมูลได้ในกรณีที่ประเภทข้อมูลไม่ถูกต้อง

โดยรูปที่ 4.10 แสดงให้เห็นหน้าจอการจัดเตรียมข้อมูลสำหรับสร้างรายงาน ซึ่งในมุมมองของแต่ละแถวจะมีปุ่มสำหรับเลือกแถวนั้นในการสร้างรายงาน ในเวอร์ชันปัจจุบันจะรองรับการเลือกข้อมูลครั้งละไม่เกิน 2 แถว

รายละเอียดของไฟล์ที่ท่านกำลังจะสร้างรายงาน

ไฟล์ uploads/1/2018-03-20/06-58-16-14-20-50-file\_for\_pie\_chart.txt ของท่านประกอบด้วยข้อมูลดังนี้

ในปัจจุบันระบบรองรับการสร้างรายงานจากข้อมูลสูงสุด 2 คอลัมน์เท่านั้น

ชื่อคอลัมน์	ประเภทข้อมูล	เลือก
วัน	DATE	<input checked="" type="checkbox"/>
เดือน	DATE	<input checked="" type="checkbox"/>
ชื่อ	TEXT	<input type="checkbox"/>

[ตัดไป](#) [ย้อนกลับ](#)

รูปที่ 4.10 ภาพตัวอย่างหน้าจอสำหรับเลือกแถวข้อมูลในการสร้างรายงาน

4.2.3.2 การเลือกรูปแบบแผนภูมิที่ใช้แสดงผล (Layout selection) เป็นขั้นตอนที่ให้ผู้สร้างรายงานเลือกประเภทแผนภูมิที่ต้องการ โดยมีลักษณะหน้าจอ ดังรูปที่ 4.11

รูปแบบรายงานที่เกี่ยวข้องกับชื่อรายงานท่าน

กรุณาเลือกรูปแบบรายงาน

- PIECHART
- LINECHART
- BARCHART

[เลือกรูปแบบแผนภูมิ](#) [ย้อนกลับ](#)

รูปที่ 4.11 ภาพตัวอย่างหน้าจอการเลือกรูปแบบแผนภูมิ

#### ขั้นตอนวิธีที่ 4.2 การเลือกรูปแบบแผนภูมิที่ใช้แสดงผล (Layout selection)

```

Function: layoutSelection
layoutType = getSelectedLayout()
selectedFields = getSelectedFields()
selectedFile = getSelectedFile()
content[] = Read selectedFile
result[] = Count values Of selectedFields \From content
If layoutType Is `PIE`
    Compute percentage Of result[]
Else
    Sort result
    Render result As chart
return chart
End

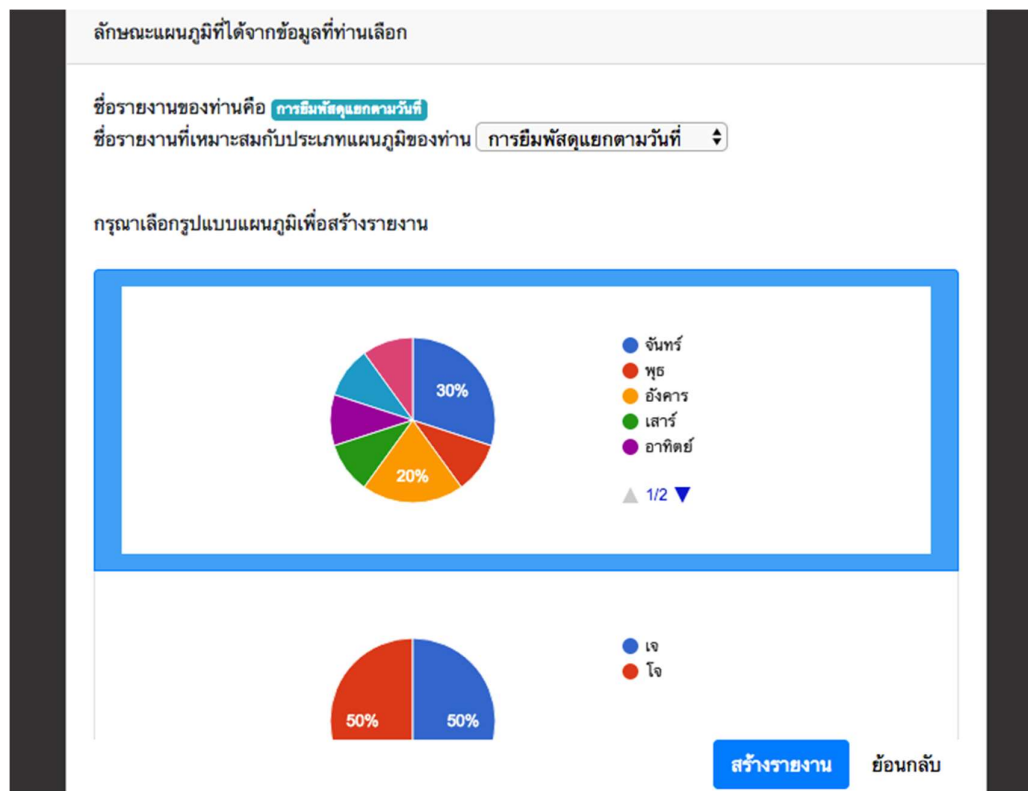
```

ขั้นตอนวิธีที่ 4.2 แสดงให้เห็นขั้นตอนการทำงานในขณะที่ผู้สร้างรายงานเลือกประเภทแผนภูมิโดยเริ่มจาก

- 1) อ่านและบันทึกข้อมูลประเภทแผนภูมิที่ผู้สร้างรายงานเลือกไว้ในตัวแปร layoutType ซึ่งตัวแปรนี้มีส่วนสำคัญในการกำหนดวิธีการคำนวณข้อมูล
- 2) อ่านและบันทึกแถวที่ผู้สร้างรายงานเลือกไว้ในตัวแปร selectedFields ซึ่งตัวแปรนี้ใช้สำหรับกำหนดว่าจะคำนวณข้อมูลจากแถวข้อมูลใด
- 3) อ่านและบันทึกที่อยู่ไฟล์ที่ผู้สร้างรายงานเลือกไว้ในตัวแปร selectedFile เพื่อกำหนดว่า ตัวอ่านข้อมูลจะอ่านข้อมูลจากไฟล์ใด
- 4) อ่านข้อมูลทั้งหมดจากไฟล์ที่สร้างรายงานเลือกตามตำแหน่งไฟล์ที่บันทึกอยู่ใน selectedFile โดยเก็บค่าจากแถวที่กำหนดจากตัวแปร selectedFields ไว้ในตัวแปร content ในลักษณะอาร์เรย์
- 5) นับความถี่ของข้อมูลแต่ละแถวในที่เก็บในตัวแปร content โดยบันทึกผลไว้ในตัวแปร result
- 6) กรณีที่ตัวแปร layoutType มีค่าเป็น PIE จะคำนวณร้อยละของข้อมูลแต่ละตัวใน result
- 7) เรียงลำดับข้อมูลที่บันทึกในตัวแปร result จากมากไปน้อย

8) สร้างแผนภูมิจากข้อมูล

9) แสดงผลลัพธ์แผนภูมิบนหน้าจอตั้งรูปที่ 4.12



รูปที่ 4.12 ภาพตัวอย่างหน้าจอการเลือกผลลัพธ์

#### 4.2.3.3 การปรับตั้งค่าสำหรับการสร้างรายงาน (Report configuration)

เป็นขั้นตอนที่จะให้ผู้สร้างรายงานตั้งค่าผลลัพธ์ที่จะปรากฏในรายงาน ประกอบด้วย

1) ชื่อรายงาน

2) แผนภูมิสรุปข้อมูล

เมื่อผู้สร้างรายงานตั้งค่าเรียบร้อยแล้ว สามารถกดปุ่ม “สร้างรายงาน” เพื่อสร้างรายงานในรูปแบบ PDF

#### 4.2.4 คู่มือการใช้งาน (User manual)

คู่มือการใช้งาน แบ่งเป็น 2 ส่วน คือ

##### 4.2.4.1 คู่มือการใช้งาน

ส่วนนี้จะแสดงวิธีการใช้งานเป็นลำดับให้กับผู้สร้างรายงาน โดยแสดงภาพและคำบรรยาย ดังรูปที่ 4.13

### คู่มือการใช้งาน

ส่วนที่	รายละเอียด													
1. การสร้างรายงาน	<p>1. ให้ทำการกดที่เมนู "My files" เพื่อเข้าสู่หน้าภาพรวมไฟล์</p> <p>2. ในหน้าภาพรวมไฟล์ ระบบจะแสดงไฟล์ทั้งหมดที่มีอยู่ในระบบ ทางด้านขวาจะมีปุ่มสำหรับการจัดการไฟล์ 2 ปุ่มคือ "Select" และ "Delete" ให้ทำการกด "Select"</p> <div style="border: 1px solid #ccc; padding: 5px; margin: 5px 0;"> <div style="display: flex; justify-content: space-between;"> <span>ไฟล์ไฟล์</span> <span>การจัดการไฟล์</span> </div> <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="border: 1px solid #ccc; padding: 2px;">file_for_pie_chart.txt</td> <td style="text-align: right; padding: 2px;"> <span style="color: green; font-weight: bold;">Select</span> <span style="color: red; font-weight: bold;">Delete</span> </td> </tr> <tr> <td style="border: 1px solid #ccc; padding: 2px;">file_for_bar_chart.txt</td> <td style="text-align: right; padding: 2px;"> <span style="color: green; font-weight: bold;">Select</span> <span style="color: red; font-weight: bold;">Delete</span> </td> </tr> </table> </div> <p>3. ในหน้ารายละเอียดของไฟล์ ระบบจะแสดงข้อมูลของไฟล์ที่ท่านเลือก โดยจะแบ่งเป็นสามส่วนด้วยกันคือ</p> <ul style="list-style-type: none"> <li>○ "ชื่อคอลัมน์" ส่วนแสดงชื่อของคอลัมน์จากไฟล์ที่ท่านเลือก</li> <li>○ "ประเภทข้อมูล" ส่วนแสดงประเภทข้อมูลที่ได้รับการตรวจสอบประเภทข้อมูลเบื้องต้นจากระบบ โดยท่านสามารถแก้ไขประเภทข้อมูลได้ตามความเหมาะสม</li> <li>○ "เลือก" ส่วนสำหรับการเลือกเพื่อใช้ในการสร้างรายงาน</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin: 5px 0;"> <table style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="width: 30%;">ชื่อคอลัมน์</th> <th style="width: 40%;">ประเภทข้อมูล</th> <th style="width: 30%;">เลือก</th> </tr> </thead> <tbody> <tr> <td>วันที่</td> <td>DATE</td> <td style="text-align: center;">☑</td> </tr> <tr> <td>ชื่อ</td> <td>TEXT</td> <td style="text-align: center;">☑</td> </tr> </tbody> </table> </div>	file_for_pie_chart.txt	<span style="color: green; font-weight: bold;">Select</span> <span style="color: red; font-weight: bold;">Delete</span>	file_for_bar_chart.txt	<span style="color: green; font-weight: bold;">Select</span> <span style="color: red; font-weight: bold;">Delete</span>	ชื่อคอลัมน์	ประเภทข้อมูล	เลือก	วันที่	DATE	☑	ชื่อ	TEXT	☑
file_for_pie_chart.txt	<span style="color: green; font-weight: bold;">Select</span> <span style="color: red; font-weight: bold;">Delete</span>													
file_for_bar_chart.txt	<span style="color: green; font-weight: bold;">Select</span> <span style="color: red; font-weight: bold;">Delete</span>													
ชื่อคอลัมน์	ประเภทข้อมูล	เลือก												
วันที่	DATE	☑												
ชื่อ	TEXT	☑												

รูปที่ 4.13 ภาพตัวอย่างหน้าจอแสดงวิธีการใช้งาน

4.2.4.2 ข้อตกลงในการใช้งาน เป็นส่วนที่แสดงข้อตกลงการใช้งาน เช่น ลักษณะของไฟล์ที่ใช้ในการสร้างรายงาน ข้อจำกัดในการสร้างรายงาน ดังรูปที่ 4.14

ข้อตกลงการใช้งาน	
ข้อที่	รายละเอียด
1. ไฟล์ข้อมูล	<ol style="list-style-type: none"> <li>ไฟล์ข้อมูลจะต้องเป็นไฟล์สกุล csv เท่านั้น</li> <li>ไฟล์ข้อมูลจะต้องขึ้นด้วยอักษร ' ' เท่านั้น  <pre> วันที่ส่งค่า 2017/01/01 Type A 12000 2017/01/03 Type A 8000 2017/01/15 Type B 3000 2017/02/01 Type A 9000 2017/02/14 Type B 6000 2017/02/17 Type A 8000 2017/02/19 Type A 30000 </pre> </li> <li>ไฟล์ข้อมูลจะต้องเข้ารหัสแบบ UTF-8 เท่านั้น</li> <li>ไฟล์จะต้องมีหัวตาราง</li> </ol>
2. ประเภทข้อมูล	<ol style="list-style-type: none"> <li>ประเภทข้อมูลมีทั้งหมด 8 ประเภท โดยรวบรวมข้อมูลจากรายงานในเว็บไซต์ <a href="http://www.data.go.th">www.data.go.th</a></li> </ol>
3. แผนภูมิ	<ol style="list-style-type: none"> <li>แผนภูมิมี่ทั้งหมด 3 ประเภท โดยการสร้างแผนภูมิจะใช้การนับความถี่ข้อมูล การหาค่าร้อยละ</li> </ol>

รูปที่ 4.14 ภาพตัวอย่างหน้าจอแสดงข้อตกลงการใช้งาน



## บทที่ 5

### สรุปผลการทดลอง

#### 5.1 สรุปและข้อเสนอแนะ

จากการออกแบบและทดลองแบบทดสอบจนถึงการพัฒนาเป็นระบบ แต่ละขั้นตอนได้ให้ผลลัพธ์ที่ตีรวมถึงกรณีศึกษาที่เกิดจากความผิดพลาด ในส่วนนี้ผู้วิจัยจะได้อธิบายปัญหาที่ได้พบ การแก้ไขและแนวทางการนำไปใช้ประโยชน์

#### 5.2 สรุปผลการทดลอง

งานวิจัยนี้มีจุดประสงค์ในการเสนอขั้นตอนวิธีในการสร้างรายงานให้กับผู้สร้างรายงานด้วยวิธีการเสนอชื่อรายงานที่เหมาะสมกับประเภทข้อมูลของข้อมูลต้นฉบับ โดยชื่อของรายงานเกิดจากการผสมคำสำคัญที่รวบรวมจากรายงานที่ใช้ในภาครัฐภายในประเทศไทย จากการออกแบบแบบจำลองที่ใช้ในการผสมคำสำคัญผู้วิจัยได้ออกแบบการทดลองโดยแบ่งออกเป็น 4 ส่วน เพื่อแสดงให้เห็นผลลัพธ์ของการทำงาน

การทดลองที่ 1 เป็นการเปรียบเทียบความใกล้เคียงของชื่อรายงานโดยใช้การคำนวณค่าระยะทาง Cosine Similarity ระหว่างชื่อรายงานที่ผ่านการผสมคำสำคัญและชื่อรายงานที่ใช้ในหน่วยงานภาครัฐด้านสังคมและสวัสดิการจำนวน 64 ฉบับ โดยสมมติฐานการทดลองที่ 1 คือ ค่าระยะทาง Cosine Similarity มีค่าเฉลี่ยที่ 0.5 ซึ่งหมายความว่า ชื่อรายงานที่ผ่านการผสมคำมีความคล้ายกันเกินร้อยละ 50 จากการทดลอง ผลลัพธ์ที่ได้แสดงให้เห็นว่าผลลัพธ์เป็นไปตามสมมติฐาน ซึ่งหมายความว่า แบบจำลองสามารถผสมคำออกมาได้ใกล้เคียงกับชื่อรายงานจริง

การทดลองที่ 2 เป็นการทดสอบกระบวนการตรวจสอบความถูกต้องของคำสำคัญโดยสมมติฐานการทดลอง คือ คำสำคัญที่ผ่านการตรวจสอบเมื่อนำมาผสมเป็นชื่อรายงานจะให้ค่าระยะทาง Cosine Similarity เฉลี่ยที่ดีกว่า จากการทดลองผลลัพธ์ที่ได้แสดงให้เห็นว่า การตรวจสอบความถูกต้องของคำสำคัญนั้นให้ผลลัพธ์ที่ดียิ่งขึ้น จากการสำรวจชื่อรายงานพบว่า ในกรณีที่คำสำคัญถูกตรวจสอบความถูกต้อง โดยปรับปรุงให้มีความสละสลวยมากยิ่งขึ้น ส่งผลให้ชื่อของรายงานอ่านได้ง่ายมากยิ่งขึ้น

การทดลองที่ 3 เป็นการนำสถิติของคำสำคัญที่ใช้ในชื่อของรายงานต้นฉบับมาปรับปรุงประเภทคำสำคัญที่ใช้ในการสร้างเป็นชื่อรายงาน สำหรับสมมติฐานการทดลอง ผู้วิจัยคาดว่าหากเราใช้คำสำคัญที่นิยมใช้ในรายงานจริง ๆ มาใช้ในการสร้างชื่อรายงานจะให้ผลลัพธ์ที่ดี ซึ่งจากผลการทดลอง ผลลัพธ์ที่ได้แสดงให้เห็นว่า การนำคำสำคัญตามสถิติที่นิยมใช้มาใช้ในการสร้างชื่อรายงานให้ผลลัพธ์ที่ดี

การทดลองที่ 4 เป็นการแบ่งกลุ่มคำสำคัญตามประเภทข้อมูลที่เกี่ยวข้องแล้วทำการทดสอบว่า หากทำการแยกกลุ่มคำสำคัญออกตามประเภทข้อมูลแล้ว ผลลัพธ์ของแต่ละกลุ่มจะให้ค่าเฉลี่ยคะแนน Cosine Similarity ที่เท่ากันหรือไม่ จากการทดลองแสดงให้เห็นว่าการแบ่งกลุ่มคำสำคัญให้ผลลัพธ์ที่แตกต่างกัน โดยมีเหตุผลมาจากจำนวนคำสำคัญในแต่ละกลุ่มแตกต่างกันจึงทำให้ผลลัพธ์แตกต่างกันเล็กน้อย

## 5.3 ปัญหาและอุปสรรค

5.3.1 เนื่องจากคำสำคัญที่ใช้ในหมวดหมู่แต่ละหมวดหมู่มีความแตกต่างกัน การนำไปใช้งานจริงจำเป็นต้องเก็บข้อมูลคำสำคัญก่อน

5.3.2 ชื่อรายงานที่ผ่านการผสมคำสำคัญไม่สละสลวยเหมือนชื่อรายงานจริง เนื่องจากข้อจำกัดในการเลือกคำสันธานมาใช้ในการผสมคำสำคัญ หากต้องการสร้างชื่อรายงานให้สละสลวย จำเป็นต้องเพิ่มขนาดของข้อความ โดยหากขนาดข้อความยาว ความน่าจะเป็นของชื่อรายงานจะมีค่าเพิ่มมากขึ้น ส่งผลให้การประมวลผลคำใช้เวลานานมากยิ่งขึ้น โดยผู้วิจัยได้ทำการทดสอบเพิ่มขนาดข้อความเป็น 6 หลัก การประมวลผลคำใช้เวลามากกว่า 5 นาที โดยขนาดข้อมูลสูงถึง 831 MB

5.3.3 คำหยุดบางคำที่ถูกตัดออกจากกระบวนการตัดคำหยุดสามารถเพิ่มความสละสลวยของชื่อรายงานได้ แต่หากใช้แบบจำลองนี้อย่างอัตโนมัติจะส่งผลให้คำเหล่านั้นถูกตัดออก ดังนั้นเมื่อนำแบบจำลองไปพัฒนาเป็นระบบควรมีการเพิ่มเติมคำหยุดบางคำตามความเหมาะสม โดยต้องอาศัยผู้ดูแลระบบในการคัดเลือกความเหมาะสมของคำ

## 5.4 ข้อเสนอแนะ

5.4.1 ออกแบบแบบจำลองเพื่อเพิ่มขนาดของข้อความมากขึ้น เพื่อความสละสลวยของข้อมูล โดยใช้ระยะเวลาในการประมวลผลน้อยลง

5.4.2 ออกแบบรูปแบบรายงานที่หลากหลายมากยิ่งขึ้น เนื่องจากงานวิจัยนี้ผู้วิจัยทำการศึกษาเฉพาะรายงานแบบตาราง

5.4.3 เทคนิคการผสมคำสามารถนำไปต่อยอดเพื่อใช้ทำเครื่องมือในการสร้างประโยคได้

5.4.4 ในบางกรณีคำหุุดอาจมีผลต่อความสละสลวยของชื่อรายงาน และอาจส่งผลกระทบต่อวิธีการในการคำนวณข้อมูล

## 5.5 ลักษณะการนำไปใช้ประโยชน์

งานวิจัยนี้แบ่งประโยชน์ในการนำไปใช้เป็น 2 ส่วน คือ

5.5.1 นำไปใช้ประโยชน์ในฐานะเครื่องมือในการสร้างรายงาน โดยแบบจำลองเมื่อพัฒนาเป็นระบบแล้ว สามารถใช้สร้างรายงานอย่างง่ายได้ ทำให้ผู้สร้างรายงานสามารถเห็นภาพรวมของข้อมูลได้โดยไม่ต้องออกแบบรายงานเพียงแต่เตรียมข้อมูลเบื้องต้นให้เหมาะสม

5.5.2 นำไปต่อยอดเพื่อพัฒนาเป็นแบบจำลองที่ใช้ผสมคำเป็นชื่อ หรือประโยค หากนำแบบจำลองนี้ไปพัฒนาให้มีความสามารถในการเรียงคำให้เหมาะสม

### เอกสารอ้างอิง

- กำชัย ทองหล่อ. (2550). *หลักภาษาไทย*. กรุงเทพฯ: รวมสาส์น.
- ตลาดหลักทรัพย์แห่งประเทศไทย. (2557). “การจัดกลุ่มอุตสาหกรรมและหมวดธุรกิจ.” (ออนไลน์) เข้าถึงได้ที่ [https://www.set.or.th/th/regulations/simplified\\_regulations/industry\\_sector\\_p1.html](https://www.set.or.th/th/regulations/simplified_regulations/industry_sector_p1.html)(วันที่ 12 ธันวาคม 2560).
- ประพนธ์ เรืองณรงค์. (2545). *ภาษาไทย*. กรุงเทพฯ: ประสานมิตร.
- ปริดาภรณ์ กาญจนสำราญวงศ์. (2560). *หลักสถิติเบื้องต้น*. นนทบุรี: ไอดีซีฯ.
- รำเนา พลอยขาว. (2549). “ชนิดของคำ.” (ออนไลน์) เข้าถึงได้ที่ <http://www.prakan.ac.th/Link-Data/e-learning-49/e-learning/page5.htm>(วันที่ 7 ธันวาคม 2560).
- วิรัช ศรีเลิศล้ำวาณิช. (2557). “*Word segmentation for Thai in machine translation system การตัดคำในระบบแปลภาษา*.” (ออนไลน์) เข้าถึงได้ที่ [https://www.researchgate.net/publication/243659316\\_Word\\_segmentation\\_for\\_Thai\\_in\\_machine\\_translation\\_system](https://www.researchgate.net/publication/243659316_Word_segmentation_for_Thai_in_machine_translation_system)(วันที่ 28 มิถุนายน 2561).
- ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ. (2560). “โปรแกรมตัดคำภาษาไทย LexTo Plus.” (ออนไลน์) เข้าถึงได้ที่ <https://www.nectec.or.th/innovation/innovationsoftware/lextoplus.html>(วันที่ 11 ตุลาคม 2560).
- ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ. (2560). “ระบบคลังสื่อประสมและข้อความกำกับ.” (ออนไลน์) เข้าถึงได้ที่ <https://www.nectec.or.th/corpus/index.php?league=pm>(วันที่ 11 ตุลาคม 2560).
- สมพงษ์ พันธุ์รัตน์. (2561). “การทดสอบสมมติฐาน.” (ออนไลน์) เข้าถึงได้ที่ [https://home.kku.ac.th/sompo\\_pu/spweb/](https://home.kku.ac.th/sompo_pu/spweb/)(วันที่ 28 มิถุนายน 2561).
- สายชล สีนสมบูรณ์ทอง. (2558). “การทำเหมืองข้อมูล” กรุงเทพฯ: จามจุรีโปรดักท์.

- สำนักงานพัฒนารัฐบาลดิจิทัล. (2559). “Government Open Data of Thailand.” (ออนไลน์) เข้าถึงได้ที่ <https://data.go.th/Datasets.aspx>(วันที่ 28 มิถุนายน 2561).
- สุรางค์ จันทราภรณ์. (2547). “ความรู้ เรื่อง คำวิเศษณ์.” (ออนไลน์) เข้าถึงได้ที่ <http://www.kr.ac.th/ebook2/surang/04.html>(วันที่ 28 มิถุนายน 2561).
- สุรางค์ จันทราภรณ์. (2547). “ความรู้ เรื่อง คำสรรพนาม.” (ออนไลน์) เข้าถึงได้ที่ <http://www.kr.ac.th/ebook2/surang/02.html>(วันที่ 28 มิถุนายน 2561).
- สุรางค์ จันทราภรณ์. (2547). “คำนาม.” (ออนไลน์) เข้าถึงได้ที่ <http://www.kr.ac.th/ebook2/surang/01.html>(วันที่ 28 มิถุนายน 2561).
- Carlos V. (2001). “Java and reports, some solutions for the past, present and future.” *Proceedings of the 29th annual ACM SIGUCCS conference on User services (SIGUCCS '01)*. New York, NY, USA: ACM. 275-278.
- DataCamp. (2017). “Line Charts.” (Online) Available on <https://www.statmethods.net/graphs/line.html>(7 December 2017).
- DataCamp. (2017). “Pie Charts.” (Online) Available on <https://www.statmethods.net/graphs/pie.html>(7 December 2017).
- DataCamp. (2017). “Quick R Graph.” (Online) Available on <https://www.statmethods.net/graphs/index.html>(7 June 2017).
- Ewan K., Edward L. Steven B. (2009). “Natural Language Processing with Python.” O'Reilly Media.
- Flavius F., Marnix M., and Frederik H. (2012). “Semantics-based news recommendation.” *Proceedings of the 2nd International Conference on Web Intelligence, Mining and Semantics (WIMS '12)*. New York, NY, USA: ACM.
- Gideon D., Yoelle M., and Idan S Anna S. (2015). “Learning from the past: answering new questions with past answers.” *Proceedings of the 21st international conference on World Wide Web (WWW '12)*. New York, NY, USA: ACM. 759-768.

- Ho C. and Luk, Robert W. and Wong, Kam F. and Kwok, Kui L. (2008). "Interpreting TF-IDF Term Weights As Making Relevance Decisions." *Transactions on Information Systems (TOIS)*. New York, NY, USA: ACM. 1-37.
- IBM Knowledge Center. (2018). "Chart types." (Online) Available on [https://www.ibm.com/support/knowledgecenter/th/SSEP7J\\_10.2.1/com.ibm.swg.ba.cognos.dsk\\_ug.10.2.0.1.doc/c\\_dsk\\_charts\\_types.html](https://www.ibm.com/support/knowledgecenter/th/SSEP7J_10.2.1/com.ibm.swg.ba.cognos.dsk_ug.10.2.0.1.doc/c_dsk_charts_types.html)(7 June 2018).
- Jaspersoft Corporation. (2017). "*JasperReports® Library*." (Online) Available on <https://community.jaspersoft.com/project/jasperreports-library>(7 December 2017).
- Konrad V., and Daniel S Veronika T. (2013). "Query matching for report recommendation." *Proceedings of the 22nd ACM international conference on Conference on information & knowledge management (CIKM '13)*. New York, NY, USA: ACM. 1391-1400.
- Matthias N. and Tim K. (2009). "Generating SQL/XML query and update statements." *Proceedings of the 18th ACM conference on Information and knowledge management (CIKM '09)*. New York, NY, USA: ACM. 1187-1196.
- ORACLE. (2018). "Data Types." (Online) Available on <https://dev.mysql.com/doc/refman/8.0/en/data-types.html>(28 June 2018).
- ReportLab. (2017). "*ReportLab Chart Gallery*." (Online) Available on <https://www.reportlab.com/chartgallery/>(28 June 2018).
- Valentin J. and Maarten de R. (2005). "Retrieving answers from frequently asked questions pages on the web." *Proceedings of the 14th ACM international conference on Information and knowledge management (CIKM '05)*. New York, NY, USA: ACM. 76-83.

Victoria state government. (2017). "*Dot Plots and Stem-and-Leaf Plots - More About.*" (Online) Available on <https://www.education.vic.gov.au/school/teachers/teachingresources/discipline/maths/continuum/Pages/dotplots.aspx>(27 June 2018).



## ภาคผนวก ก

## คำจำกัดความคำสำคัญในงานวิจัย

ตารางที่ ก.1 ตารางคำจำกัดความของ “การผสมคำเพื่อสร้างชื่อ”

คำภาษาไทย	การผสมคำเพื่อสร้างชื่อ
คำภาษาอังกฤษ	Report title combination
ความหมาย	การนำคำสำคัญมาเรียงกัน โดยใส่คำสันธานเชื่อมระหว่างคำสำคัญที่ดึงมาจากฐานข้อมูลคำสำคัญ
ตัวอย่าง	-

ตารางที่ ก.2 ตารางคำจำกัดความของ “การสร้างรายงานแบบอัตโนมัติ”

คำภาษาไทย	การสร้างรายงานแบบอัตโนมัติ
คำภาษาอังกฤษ	Automatic report generation
ความหมาย	การสร้างรายงาน โดยแสดงชื่อที่เป็นไปได้ทั้งหมดให้กับผู้สร้างรายงาน และสร้างรายงานตามที่คุณสร้างรายงานเลือก
ตัวอย่าง	-

ตารางที่ ก.3 ตารางคำจำกัดความของ “การเปรียบเทียบชื่อ”

คำภาษาไทย	การเปรียบเทียบชื่อ
คำภาษาอังกฤษ	Name comparing
ความหมาย	การเปรียบเทียบคำจากข้อมูลต้นฉบับกับข้อมูลในฐานข้อมูล โดยการใช้ Regular expression
ตัวอย่าง	-

ตารางที่ ก.4 ตารางคำจำกัดความของ “ข้อมูลต้นฉบับ”

คำภาษาไทย	ข้อมูลต้นฉบับ
คำภาษาอังกฤษ	Raw data
ความหมาย	ข้อมูลที่มีลักษณะเป็นตารางพร้อมกับชื่อแถว
ตัวอย่าง	-

ตารางที่ ก.5 ตารางคำจำกัดความของ “ข้อมูลที่ผ่านกระบวนการเตรียมข้อมูล”

คำภาษาไทย	ข้อมูลที่ผ่านกระบวนการเตรียมข้อมูล
คำภาษาอังกฤษ	Pre-processed data
ความหมาย	ข้อมูลผ่านการเตรียมเพื่อใช้ในการคำนวณ
ตัวอย่าง	2018/01/01 จะถูกแปลงเป็น 2018, 2018/01 และ 2018/01/01 เพื่อใช้ในการคำนวณเวลาในสามรูปแบบ

ตารางที่ ก.6 ตารางคำจำกัดความของ “ความถี่ข้อมูล”

คำภาษาไทย	ความถี่ข้อมูล
คำภาษาอังกฤษ	Counted data
ความหมาย	ข้อมูลที่ถูกนับ โดยบันทึกคีย์เป็นค่าของข้อมูล และ บันทึกค่าเป็นความถี่ของข้อมูลที่เกิดขึ้น
ตัวอย่าง	[[ Word: 'เฉลี่ย', Count: '5' }], { Word: 'มากกว่า', Count: '3' }]

ตารางที่ ก.7 ตารางคำจำกัดความของ “คำที่ถูกตัด”

คำภาษาไทย	คำที่ถูกตัด
คำภาษาอังกฤษ	Segmented words
ความหมาย	ประโยคที่ถูกตัดคำ โดยแต่ละคำจะคั่นด้วยสัญลักษณ์
ตัวอย่าง	รายได้เฉลี่ยต่อปีของคนและครัวเรือน  (รายได้จังหวัด)  จากข้อมูล  จปฐ  ปี  2558  จำนวนครัวเรือนที่คนในครัวเรือนมีรายได้เฉลี่ยน้อยกว่า  30,000  บาทต่อปี  จากข้อมูล  จปฐ  ปี  2555 - 2558  รายจ่ายเฉลี่ยต่อปีของคนและครัวเรือน  (รายได้จังหวัด)  จากข้อมูล  จปฐ  ปี  2558  อัตราการว่างงาน  จำนวนตามเพศ  พ.ศ.  2549  -   2558  จำนวนประชากรอายุ  15  ปีขึ้นไป  จำนวนตามสถานภาพแรงงาน  พ.ศ.  ทั่วราชอาณาจักร  พ.ศ.  2549  -   2558  ตำแหน่งว่าง  จำนวนตามอาชีพ  และวุฒิการศึกษา  ทั่วราชอาณาจักร  ปี  2558  ผู้ลงทะเบียนสมัครงาน  จำนวนตามอาชีพ  และวุฒิการศึกษา  ทั่วราชอาณาจักร  ปี  2558  การบรรจงาน  จำนวนตามอาชีพ  และวุฒิการศึกษา  ทั่วราชอาณาจักร  ปี  2558  ข้อมูลการจัดหางานในประเทศ  จำนวนตามจังหวัด  ทั่วราชอาณาจักร  ปี  2558

ตารางที่ ก.8 ตารางคำจำกัดความของ “คำสำคัญ”

คำภาษาไทย	คำสำคัญ
คำภาษาอังกฤษ	Important words
ความหมาย	คำนามที่เป็นส่วนประกอบในการสร้างเป็นชื่อรายงาน เกิดจากการนำชื่อของรายงานมาผ่านกระบวนการคัดกรองคำสำคัญ
ตัวอย่าง	ประชากร รายได้ ค่าใช้จ่าย

ตารางที่ ก.9 ตารางคำจำกัดความของ “คำสำคัญที่ค้นหาประเภทข้อมูลและได้รับการยืนยันแล้ว”

คำภาษาไทย	คำสำคัญที่ค้นหาประเภทข้อมูลและได้รับการยืนยันแล้ว
คำภาษาอังกฤษ	Verified Important words
ความหมาย	คำสำคัญที่ได้รับการตรวจสอบความถูกต้องจากผู้ดูแลระบบแล้ว
ตัวอย่าง	[[ Word: ‘เฉลี่ย’, Type: ‘Number’, Active: ‘True’ ]]

ตารางที่ ก.10 ตารางคำจำกัดความของ “คำสำคัญที่ค้นหาประเภทข้อมูลแล้ว”

คำภาษาไทย	คำสำคัญที่ค้นหาประเภทข้อมูลแล้ว
คำภาษาอังกฤษ	Important words with tagged data type
ความหมาย	อาร์เรย์ของชื่อคำสำคัญและประเภทข้อมูลของคำสำคัญนั้น
ตัวอย่าง	[[ Word: ‘จำนวน’, Type: ‘Number’ }, { Word: ‘ชื่อ’, Type: ‘Text’ }]

ตารางที่ ก.11 ตารางคำจำกัดความของ “คำหยุด”

คำภาษาไทย	คำหยุด
คำภาษาอังกฤษ	Stop words
ความหมาย	คำที่รวบรวมมาจากรายงานต้นฉบับ แบ่งได้เป็นดังนี้ คำสามัญนาม คำวิสามัญนาม คำลักษณะนาม คำสรรพนาม คำสันธาน คำวิเศษณ์ คำกริยา สัญลักษณ์พิเศษ และ ตัวเลข
ตัวอย่าง	และ 2018

ตารางที่ ก.12 ตารางคำจำกัดความของ “ชื่อรายงาน”

คำภาษาไทย	ชื่อรายงาน
คำภาษาอังกฤษ	Report title
ความหมาย	ชื่อรายงานเกิดจากการรวมกลุ่มคำสำคัญให้เป็นประโยคที่สื่อถึงข้อมูลภายใน
ตัวอย่าง	รายงานแสดงค่าใช้จ่ายรายเดือนของประชาชนแต่ละจังหวัด รายงานจำนวนประชากรที่มีงานทำแยกตามจังหวัด รายงานอุบัติเหตุในช่วงสงกรานต์

ตารางที่ ก.13 ตารางคำจำกัดความของ “ชื่อแถว”

คำภาษาไทย	ชื่อแถว
คำภาษาอังกฤษ	Column name
ความหมาย	ชื่อของแต่ละแถวข้อมูล
ตัวอย่าง	-

ตารางที่ ก.14 ตารางคำจำกัดความของ “ชื่อแถวที่ค้นหาประเภทข้อมูลและได้รับการยืนยันแล้ว”

คำภาษาไทย	ชื่อแถวที่ค้นหาประเภทข้อมูลและได้รับการยืนยันแล้ว
คำภาษาอังกฤษ	Verified column name
ความหมาย	ชื่อแถวที่ได้รับการตรวจสอบความถูกต้องจากผู้ดูแลระบบแล้ว
ตัวอย่าง	[[ Word: ‘เดือน’, Type: ‘Date’, Active: ‘True’ ]]

ตารางที่ ก.15 ตารางคำจำกัดความของ “ชื่อแถวที่ค้นหาประเภทข้อมูลแล้ว”

คำภาษาไทย	ชื่อแถวที่ค้นหาประเภทข้อมูลแล้ว
คำภาษาอังกฤษ	Column name with tagged data type
ความหมาย	อาร์เรย์ของชื่อแถวและประเภทข้อมูลของชื่อนั้น
ตัวอย่าง	[[ Word: ‘เดือน’, Type: ‘Date’ ]]

ตารางที่ ก.16 ตารางคำจำกัดความของ “ตารางบันทึกคำสำคัญ”

คำภาษาไทย	ตารางบันทึกคำสำคัญ
คำภาษาอังกฤษ	Important words table
ความหมาย	ตารางบันทึกคำสำคัญพร้อมกับชนิดข้อมูลของคำสำคัญ ประกอบด้วยข้อมูล 4 แถว คือ รหัสประจำคำสำคัญ ชนิดข้อมูลของคำสำคัญ คำสำคัญ และ ตำแหน่งของคำสำคัญเมื่อถูกใช้ในการสร้างชื่อรายงาน
ตัวอย่าง	<pre> [   ID: 1,   Datatype: 'TEXT',   Word: 'รายชื่อ',   Position: 'PRE' ] </pre>

ตารางที่ ก.17 ตารางคำจำกัดความของ “ตารางบันทึกคำหยุด”

คำภาษาไทย	ตารางบันทึกคำหยุด
คำภาษาอังกฤษ	Stop words table
ความหมาย	ตารางบันทึกชื่อคำหยุด ประกอบด้วยข้อมูล 2 แถว คือ รหัสคำหยุด และ คำหยุด
ตัวอย่าง	<pre> [   ID: 1,   Word: 'ต่อ' }, {   ID: 2,   Word: 'และ' } ] </pre>

ตารางที่ ก.18 ตารางคำจำกัดความของ “ตารางบันทึกชื่อแถว”

คำภาษาไทย	ตารางบันทึกชื่อแถว
คำภาษาอังกฤษ	Column name table
ความหมาย	ตารางบันทึกชื่อของแถวในรายงานต้นฉบับ ประกอบด้วยข้อมูล 3 แถว คือ รหัสชื่อแถว ชื่อแถว และ ประเภทข้อมูลของชื่อแถว
ตัวอย่าง	<pre> [   ID: 1,   Datatype: 'DATE',   Word: 'วัน' }, {   ID: 2,   Datatype: 'AMOUNT',   Word: 'หน่วย' }] </pre>

ตารางที่ ก.19 ตารางคำจำกัดความของ “ประเภทข้อมูล”

คำภาษาไทย	ประเภทข้อมูล
คำภาษาอังกฤษ	Data type
ความหมาย	ประเภทข้อมูลที่ปรากฏแต่ละแถวข้อมูล
ตัวอย่าง	Date Amount Text

ตารางที่ ก.20 ตารางคำจำกัดความของ “ประเภทข้อมูลข้อความ”

คำภาษาไทย	ประเภทข้อมูลข้อความ
คำภาษาอังกฤษ	Text
ความหมาย	ข้อมูลที่มีลักษณะเป็นคำมีขนาดไม่เกิน 30 ตัวอักษร สามารถนำมาจัดกลุ่มได้ในกรณีที่มีการเกิดซ้ำกัน อาจเป็นข้อมูลที่มีลักษณะเป็นชื่อสถานที่ ชื่อตำบล ชื่ออำเภอ โดยข้อมูลประเภทนี้จะเก็บค่าอย่างใดอย่างหนึ่ง ไม่ได้เก็บในลักษณะที่รวมข้อมูลเป็นข้อมูลเดียว
ตัวอย่าง	ประเภทของสิ่งต่าง ๆ เช่น ประเภทที่หนึ่ง ประเภทที่สอง ชื่อสถานที่ เช่น ภูเก็ต กระบี่

ตารางที่ ก.21 ตารางคำจำกัดความของ “ประเภทข้อมูลจำนวน”

คำภาษาไทย	ประเภทข้อมูลจำนวน
คำภาษาอังกฤษ	Numeric
ความหมาย	ข้อมูลที่มีลักษณะเป็นตัวเลข
ตัวอย่าง	1 3 5 7 9

ตารางที่ ก.22 ตารางคำจำกัดความของ “ประเภทข้อมูลที่ถูกละเลือก”

คำภาษาไทย	ประเภทข้อมูลที่ถูกละเลือก
คำภาษาอังกฤษ	Selected data type
ความหมาย	ประเภทข้อมูลของแต่ละแถวข้อมูลที่ถูกละเลือก
ตัวอย่าง	-

ตารางที่ ก.23 ตารางคำจำกัดความของ “ประเภทข้อมูลวันเวลา”

คำภาษาไทย	ประเภทข้อมูลวันเวลา
คำภาษาอังกฤษ	Date
ความหมาย	ข้อมูลที่มีลักษณะเป็น วัน เดือน ปี เวลา ชั่วโมง ชื่อวัน ชื่อเดือน
ตัวอย่าง	2018/01/01 11:30 2018/01/01 จันทร์



ตารางที่ ก.24 ตารางคำจำกัดความของ “ประเภทแผนภูมิที่ถูกเลือก”

คำภาษาไทย	ประเภทแผนภูมิที่ถูกเลือก
คำภาษาอังกฤษ	Selected chart layout
ความหมาย	ประเภทแผนภูมิที่ถูกเลือกโดยผู้สร้างรายงาน
ตัวอย่าง	Bar chart

ตารางที่ ก.25 ตารางคำจำกัดความของ “ผู้สร้างรายงาน”

คำภาษาไทย	ผู้สร้างรายงาน
คำภาษาอังกฤษ	User
ความหมาย	ผู้ใช้งานระบบในการสร้างรายงาน
ตัวอย่าง	-

ตารางที่ ก.26 ตารางคำจำกัดความของ “รายงาน”

คำภาษาไทย	รายงาน
คำภาษาอังกฤษ	Report
ความหมาย	งานวิจัยนี้ รายงาน คือชุดเอกสารที่แสดงภาพสรุปข้อมูลในรูปแบบแผนภูมิ และแสดงข้อมูลในรูปแบบตาราง
ตัวอย่าง	-

ตารางที่ ก.27 ตารางคำจำกัดความของ “รายงานที่ได้ดาวน์โหลดมา”

คำภาษาไทย	รายงานที่ได้ดาวน์โหลดมา
คำภาษาอังกฤษ	Downloaded reports
ความหมาย	รายงานที่ได้ดาวน์โหลดมาจากเว็บไซต์ <a href="http://www.data.go.th">http://www.data.go.th</a>
ตัวอย่าง	-

ตารางที่ ก.28 ตารางคำจำกัดความของ “วิธีการคำนวณข้อมูล”

คำภาษาไทย	วิธีการคำนวณข้อมูล
คำภาษาอังกฤษ	Computation method
ความหมาย	วิธีการในการคำนวณข้อมูล โดยจะสอดคล้องกับแถวข้อมูลที่ถูกเลือก
ตัวอย่าง	-

ตารางที่ ก.29 ตารางคำจำกัดความของ “อาร์เรย์”

คำภาษาไทย	อาร์เรย์
คำภาษาอังกฤษ	Array
ความหมาย	ชุดข้อมูลประกอบด้วยคีย์และค่า
ตัวอย่าง	[[ ID: 1, CONTENT: 'Hello' }], [ ID: 2, CONTENT: 'World' ]]

ตารางที่ ก.30 ตารางคำจำกัดความของ “แถวข้อมูลที่ถูกเลือก”

คำภาษาไทย	แถวข้อมูลที่ถูกเลือก
คำภาษาอังกฤษ	Selected columns
ความหมาย	แถวที่ถูกเลือกในการสร้างรายงานโดยผู้สร้างรายงาน
ตัวอย่าง	-

ตารางที่ ก.31 ตารางคำจำกัดความของ “โครงสร้างข้อมูล”

คำภาษาไทย	โครงสร้างข้อมูล
คำภาษาอังกฤษ	Data schema
ความหมาย	ประเภทข้อมูลของแต่ละแถวที่ประกอบกันเป็นรายงาน
ตัวอย่าง	<pre> [[   Column_id: 1,   Column_name: 'ชื่อ',   Column_type: 'TEXT' }, {   Column_id: 2,   Column_name: 'เงินเดือน',   Column_type: 'AMOUNT' }] </pre>

ตารางที่ ก.32 ตารางคำจำกัดความของ “ไม่ทราบประเภทข้อมูล”

คำภาษาไทย	ไม่ทราบประเภทข้อมูล
คำภาษาอังกฤษ	Unknown
ความหมาย	ข้อมูลที่ไม่สามารถค้นหาประเภทข้อมูลในฐานข้อมูลชื่อแถวได้
ตัวอย่าง	ค่าที่เขียนผิด

## ภาคผนวก ข

## รายชื่อรายงานต้นฉบับ

ตารางที่ ข.1 รายงานจากหมวดหมู่สังคมและสวัสดิการ เรียงลำดับตามความนิยมประจำเดือน ตุลาคม 2560

ลำดับ	ชื่อรายงาน
1	รายได้เฉลี่ยต่อเดือนต่อครัวเรือน จำแนกตามภาค และจังหวัด พ.ศ. 2541 - 2558
2	รายได้เฉลี่ยต่อปีของคนและครัวเรือน (รายจังหวัด) จากข้อมูล จปฐ ปี 2558
3	จำนวนครัวเรือนที่คนในครัวเรือนมีรายได้เฉลี่ยน้อยกว่า 30,000 บาทต่อปี จากข้อมูล จปฐ ปี 2555 - 2558
4	รายจ่ายเฉลี่ยต่อปีของคนและครัวเรือน (รายจังหวัด) จากข้อมูล จปฐ ปี 2558
5	อัตราการว่างงาน จำแนกตามเพศ พ.ศ. 2549 - 2558
6	จำนวนประชากรอายุ 15 ปีขึ้นไป จำแนกตามสถานภาพแรงงาน เพศ ทวีราชอาณาจักร พ.ศ. 2549 - 2558
7	ตำแหน่งงานว่าง จำแนกตามอาชีพ และวุฒิการศึกษา ทวีราชอาณาจักร ปี 2558
8	ผู้ลงทะเบียนสมัครงาน จำแนกตามอาชีพ และวุฒิการศึกษา ทวีราชอาณาจักร ปี 2558
9	การบรรจุนงาน จำแนกตามอาชีพ และวุฒิการศึกษา ทวีราชอาณาจักร ปี 2558
10	ข้อมูลการจัดหางานในประเทศ จำแนกตามจังหวัด ทวีราชอาณาจักร ปี 2558
11	ผู้ลงทะเบียนสมัครงาน ตำแหน่งงานว่าง และการบรรจุนงาน จำแนกตามอายุ ทวีราชอาณาจักร ปี 2558
12	รายได้เฉลี่ยต่อเดือนของครัวเรือน จำแนกตามแหล่งที่มาของรายได้และสถานะทางเศรษฐกิจสังคมของครัวเรือน ทวีราชอาณาจักร พ.ศ. 2558

ตารางที่ ข.1 รายงานจากหมวดหมู่สังคมและสวัสดิการ เรียงลำดับตามความนิยมประจำเดือน ตุลาคม 2560 (ต่อ)

ลำดับ	ชื่อรายงาน
13	ผู้ลงทะเบียนสมัครงาน ตำแหน่งงานว่าง และการบรรจุนงาน จำแนกตามระดับการศึกษา ทั่วราชอาณาจักร ปี 2558
14	ตำแหน่งงานว่าง จำแนกตามอุตสาหกรรม และวุฒิการศึกษา ทั่วราชอาณาจักร ปี 2558
15	ผู้ลงทะเบียนสมัครงาน ตำแหน่งงานว่าง และการบรรจุนงาน จำแนกตามอาชีพ และราย อาชีพที่มีความต้องการมาก ทั่วราชอาณาจักร ปี 2558
16	ตำแหน่งงานว่าง จำแนกตามอาชีพ และอายุ ทั่วราชอาณาจักร ปี 2558
17	การบรรจุนงาน จำแนกตามอาชีพ และอายุ ทั่วราชอาณาจักร ปี 2558
18	รายได้เฉลี่ยต่อเดือนของครัวเรือน จำแนกตามแหล่งที่มาของรายได้ รายภาคและเขตการ ปกครอง ทั่วราชอาณาจักร พ.ศ. 2558
19	จำนวนคนต่างด้าวคงเหลือทำงานอยู่ในประเทศไทย (ประเภททั่วไป) จำแนกตามสัญชาติ และอาชีพ ณ ธันวาคม 2558
20	ผู้ลงทะเบียนสมัครงาน จำแนกตามอาชีพ และอายุ ทั่วราชอาณาจักร ปี 2558
21	ตำแหน่งงานว่าง จำแนกตามอุตสาหกรรม และอายุ ทั่วราชอาณาจักร ปี 2558
22	การบรรจุนงาน จำแนกตามอุตสาหกรรม และอายุ ทั่วราชอาณาจักร ปี 2558
23	จำนวน และอัตราการว่างงาน จำแนกตามภาค พ.ศ. 2552 - 2559
24	รายได้เฉลี่ยต่อเดือนของครอบครัวข้าราชการ จำแนกตามประเภทของรายได้ และ ตำแหน่งประเภทของข้าราชการ พ.ศ. 2555
25	รายได้เฉลี่ยต่อเดือนของครอบครัวข้าราชการ จำแนกตามประเภทของรายได้ และ ตำแหน่งประเภทของข้าราชการ รายภาค พ.ศ. 2555
26	ตำแหน่งงานว่าง จำแนกตามอุตสาหกรรม และอายุ ทั่วราชอาณาจักร ปี 2559
27	ตำแหน่งงานว่าง จำแนกตามอาชีพ และวุฒิการศึกษา ทั่วราชอาณาจักร ปี 2559
28	ผู้ลงทะเบียนสมัครงาน จำแนกตามอาชีพ และวุฒิการศึกษา ทั่วราชอาณาจักร ปี 2559
29	ร้อยละของครัวเรือนที่มีหนี้สินและมูลค่าหนี้สินเฉลี่ยต่อครัวเรือนทั้งสิ้น จำแนกตามภาค และจังหวัด ระหว่างปี 2550 - 2556

ตารางที่ ข.1 รายงานจากหมวดหมู่สังคมและสวัสดิการ เรียงลำดับตามความนิยมประจำเดือน ตุลาคม 2560 (ต่อ)

ลำดับ	ชื่อรายงาน
30	ส่วนแบ่งรายได้ประจำของครัวเรือน โดยการจำแนกครัวเรือนเป็น 5 กลุ่ม รายภาค ทั่วราชอาณาจักร พ.ศ. 2558
31	ผู้ลงทะเบียนสมัครงาน จำแนกตามอาชีพ และอายุ ทั่วราชอาณาจักร ปี 2559
32	จำนวนประชากรอายุ 15 ปีขึ้นไป จำแนกตามระดับการศึกษาที่สำเร็จ สถานภาพแรงงาน ทั่วราชอาณาจักร ไตรมาสที่ 1 พ.ศ. 2560
33	ค่าใช้จ่ายเฉลี่ยต่อเดือนของครอบครัวข้าราชการ จำแนกตามประเภทของค่าใช้จ่าย และตำแหน่งประเภทของข้าราชการ พ.ศ. 2555
34	จำนวน และร้อยละของประชากรอายุ 15 ปีขึ้นไป จำแนกตามสถานภาพแรงงาน พ.ศ. 2552 - 2559
35	ร้อยละของครัวเรือน จำแนกตามรายได้ทั้งสิ้นเฉลี่ยต่อเดือน และสถานะทางเศรษฐกิจสังคมของครัวเรือน ทั่วราชอาณาจักร พ.ศ. 2558
36	จำนวนผู้มีงานทำที่อยู่ในแรงงานในระบบและนอกระบบ จำแนกตามสถานภาพการทำงาน เพศ และภาค ทั่วราชอาณาจักร พ.ศ. 2559
37	ค่าใช้จ่ายทั้งสิ้นเฉลี่ยต่อเดือนต่อครัวเรือน และการเปลี่ยนแปลงต่อปี จำแนกตามภาคและจังหวัด ระหว่างปี 2549 - 2556
38	จำนวนผู้มีงานทำ จำแนกตามอาชีพ และอุตสาหกรรม ทั่วราชอาณาจักร ไตรมาสที่ 1 พ.ศ. 2560
39	ค่าใช้จ่ายเฉลี่ยต่อเดือนของครัวเรือน จำแนกตามประเภทค่าใช้จ่าย รายภาคและเขตการปกครอง ทั่วราชอาณาจักร พ.ศ. 2558
40	จำนวนประชากรอายุ 15 ปีขึ้นไป จำแนกตามสถานภาพสมรส สถานภาพแรงงาน ทั่วราชอาณาจักร ไตรมาสที่ 1 พ.ศ. 2560
41	จำนวนครัวเรือนและจำนวนประชากรอายุ 15 ปีขึ้นไปที่รับงานมาทำที่บ้าน จำแนกตามประเภทผู้รับงาน เพศ เขตการปกครอง และภาค พ.ศ. 2550

ตารางที่ ข.1 รายงานจากหมวดหมู่สังคมและสวัสดิการ เรียงลำดับตามความนิยมประจำเดือน ตุลาคม 2560 (ต่อ)

ลำดับ	ชื่อรายงาน
42	จำนวนครัวเรือนที่เป็นหนี้ จำแนกตามวัตถุประสงค์ของการกู้ยืม ประเภทหนี้สิน และสถานะทางเศรษฐกิจสังคมของครัวเรือน ทัวราชอาณาจักร พ.ศ. 2558
43	ค่าใช้จ่ายทั้งสิ้นเฉลี่ยต่อเดือนต่อคน และการเปลี่ยนแปลงต่อปี จำแนกตามภาคและจังหวัด ระหว่างปี 2549 - 2556
44	ร้อยละของข้าราชการ จำแนกตามอายุ อายุราชการ ระดับการศึกษา สถานภาพสมรส ลักษณะการอยู่อาศัย รายได้ของคู่สมรส และตำแหน่งประเภทของข้าราชการ รายงาน พ.ศ. 2555
45	มูลค่าทรัพย์สินของครัวเรือน ประเภทบ้าน ที่ดิน สิ่งปลูกสร้าง ยานพาหนะ และสินทรัพย์ทางการเงิน รายงาน ทัวราชอาณาจักร พ.ศ. 2558
46	จำนวนลูกจ้าง จำแนกตามอุตสาหกรรม ระดับของรายได้ ทัวราชอาณาจักร ไตรมาสที่ 1 พ.ศ. 2560
47	จำนวนผู้จ้างงานทำที่อยู่ในแรงงานในระบบและนอกระบบ จำแนกตามอุตสาหกรรม เพศ และภาค ทัวราชอาณาจักร พ.ศ. 2559
48	จำนวนลูกจ้างเอกชน จำแนกตามค่าจ้างเฉลี่ย ผลประโยชน์อื่นที่ได้รับ อุตสาหกรรม ทัวราชอาณาจักร ไตรมาสที่ 1 พ.ศ. 2560
49	ร้อยละของข้าราชการ จำแนกตามภาวะทางการเงินและวิธีการแก้ปัญหาเมื่อรายได้ไม่พอกับค่าใช้จ่าย และตำแหน่งประเภทของข้าราชการ พ.ศ. 2555
50	จำนวนหนี้สินเฉลี่ยต่อครัวเรือนทั้งสิ้น จำแนกตามวัตถุประสงค์ของการกู้ยืม รายงานและเขตการปกครอง ทัวราชอาณาจักร พ.ศ. 2558
51	ร้อยละของครัวเรือนที่มีหนี้สินและมูลค่าหนี้สินเฉลี่ยต่อครัวเรือนที่เป็นหนี้ จำแนกตามภาคและจังหวัด ระหว่างปี 2550 - 2556
52	ค่าใช้จ่ายเฉลี่ยต่อเดือนของครัวเรือน จำแนกตามประเภทค่าใช้จ่าย และสถานะทางเศรษฐกิจสังคมของครัวเรือน ทัวราชอาณาจักร พ.ศ. 2558

ตารางที่ ข.1 รายงานจากหมวดหมู่สังคมและสวัสดิการ เรียงลำดับตามความนิยมประจำเดือน  
ตุลาคม 2560 (ต่อ)

ลำดับ	ชื่อรายงาน
53	จำนวนผู้มีงานทำ จำแนกตามระดับการศึกษาที่สำเร็จ อุตสาหกรรม ทวีราชอาณาจักร ไตรมาสที่ 1 พ.ศ. 2560
54	จำนวน และร้อยละของผู้มีงานทำ จำแนกตามชั่วโมงการทำงานต่อสัปดาห์ พ.ศ. 2552 - 2559
55	จำนวนผู้ว่างงาน ที่หางานทำ จำแนกตามระยะเวลาที่หางานทำ ระดับการศึกษาที่สำเร็จ ทวีราชอาณาจักร ไตรมาสที่ 1 พ.ศ. 2560
56	จำนวนผู้มีงานทำ จำแนกตามสถานภาพทำงาน อาชีพ ทวีราชอาณาจักร ไตรมาสที่ 1 พ.ศ. 2560
57	จำนวนผู้มีงานทำ จำแนกตามสถานภาพการทำงาน จำนวนชั่วโมงทำงานต่อสัปดาห์ การมีเวลาเหลือพอและต้องการทำงานเพิ่มขึ้น ทวีราชอาณาจักร ไตรมาสที่ 1 พ.ศ. 2560
58	จำนวนลูกจ้าง จำแนกตามค่าจ้างเฉลี่ย ผลประโยชน์อื่นๆ ที่ได้รับ อุตสาหกรรม ทวีราชอาณาจักร ไตรมาสที่ 1 พ.ศ. 2560
59	จำนวน และร้อยละของผู้มีงานทำ จำแนกตามอุตสาหกรรม พ.ศ. 2552 - 2559
60	จำนวนผู้มีงานทำ จำแนกตามอาชีพ จำนวนชั่วโมงทำงานต่อสัปดาห์ การมีเวลาเหลือพอและต้องการทำงานเพิ่มขึ้น ทวีราชอาณาจักร ไตรมาสที่ 1 พ.ศ. 2560
61	จำนวนผู้มีงานทำ จำแนกตามสถานภาพทำงาน อุตสาหกรรม ทวีราชอาณาจักร ไตรมาสที่ 1 พ.ศ. 2560
62	จำนวนผู้มีงานทำ จำแนกตามระดับการศึกษาที่สำเร็จ อาชีพ ทวีราชอาณาจักร ไตรมาสที่ 1 พ.ศ. 2560
63	จำนวนผู้ว่างงาน ที่เคยทำงานมาก่อน จำแนกตามอาชีพที่ทำครั้งสุดท้าย อุตสาหกรรมที่ ทำครั้งสุดท้าย ทวีราชอาณาจักร ไตรมาสที่ 1 พ.ศ. 2560
64	จำนวนลูกจ้างรัฐบาล จำแนกตามค่าจ้างเฉลี่ย ผลประโยชน์อื่นๆ ที่ได้รับ อุตสาหกรรม ทวี ราชอาณาจักร ไตรมาสที่ 1 พ.ศ. 2560



ตารางที่ ข.2 รายงานจากหมวดหมู่คมนาคมและโลจิสติกส์ เรียงลำดับตามความนิยมประจำเดือน  
ตุลาคม 2560

ลำดับ	ชื่อรายงาน
1	สถิติขนส่งทางอากาศปริมาณการจราจรทางอากาศของสายการบินต้นทุนต่ำปี 2553
2	จำนวนรถจดทะเบียน (สะสม) ณ วันที่ 31 มกราคม 2560
3	สถิติการขนส่งทางอากาศภายใต้ของท่าอากาศยานสังกัดกรมการบินพลเรือน ประจำปี 2014
4	สถิติขนส่งทางอากาศปริมาณการจราจรทางอากาศของสายการบินต้นทุนต่ำปี 2557
5	ข้อมูลด้านโครงสร้างพื้นฐานและสาธารณูปโภค องค์กรปกครองส่วนท้องถิ่น
6	ตารางแสดงปริมาณการขนส่งสินค้า การขนส่งสินค้าภายในประเทศ และต้นทุนการขนส่งสินค้า
7	สถิติการขนส่งทางอากาศของท่าอากาศยานสังกัดเอกชน ประจำปี 2014
8	ตารางจำนวนผู้โดยสารและเที่ยวเรือโดยสาร เรือด่วนเลียบริมฝั่งแม่น้ำเจ้าพระยา ปี 2548 - 2557
9	จำนวนใบอนุญาตขับรถ และใบอนุญาตผู้ประจำรถสะสม 31 ธันวาคม 2558
10	ตารางปริมาณสินค้าที่ขนส่งทางน้ำเฉพาะแห่ง บริเวณแม่น้ำเจ้าพระยาและป่าสัก ปี 2547-2556 (ปฏิทิน)
11	จำนวนใบอนุญาตขับรถ และใบอนุญาตผู้ประจำรถสะสม 31 มีนาคม 2559
12	ตารางปริมาณการขนส่งสินค้าในแม่น้ำ ปี 2557 จำแนกตามประเภทสินค้า และเรือชาล่อง ขาขึ้น
13	จำนวนผู้โดยสารเฉลี่ยต่อวันของเรือยนต์ข้ามฟาก จำแนกตามเส้นทางเดินเรือ (เรือเข้า - เรือออก) ปี 2548 - 2557
14	สรุปอัตราความเร็วเฉลี่ยของรถยนต์ส่วนบุคคล พ.ศ. 2557 (ช่วงเร่งด่วนเย็น 16.00 - 19.00)
15	ปริมาณเที่ยวบินภายใน Bangkok FIR ปีงบประมาณ 2558

ตารางที่ ข.3 รายงานจากหมวดหมู่เศรษฐกิจ การเงิน และอุตสาหกรรม เรียงลำดับตามความนิยม ประจำเดือน ตุลาคม 2560

ลำดับ	ชื่อรายงาน
1	ข้อมูลการจดทะเบียนนิติบุคคล พ.ศ. 2558 รายชื่อนิติบุคคลจัดตั้งใหม่ เดือนมกราคม
2	ข้อมูลการจดทะเบียนนิติบุคคล พ.ศ.2559
3	ค่าใช้จ่ายเฉลี่ยต่อเดือนต่อครัวเรือน จำแนกตามภาค และจังหวัด พ.ศ. 2549 - 2558
4	รายได้ประชาชาติของประเทศไทย แบบปริมาณลูกโซ่ พ.ศ. 2557
5	จำนวนหนี้สินโดยเฉลี่ยต่อครัวเรือน จำแนกตามภาค และจังหวัด พ.ศ. 2541 - 2558
6	ปริมาณการชำระเงินผ่านระบบการชำระเงินและช่องทางต่าง ๆ
7	มูลค่าการชำระเงินทางอิเล็กทรอนิกส์ผ่านอุปกรณ์อย่างหนึ่งอย่างใด หรือผ่านทาง เครือข่าย (e-Payment)
8	สรุปผลการดำเนินการจัดซื้อจัดจ้างรายเดือน ปีงบประมาณ พ.ศ.2558
9	สถิติร้องเรียนจากการประกอบกิจการโรงงานอุตสาหกรรม ประจำปีงบประมาณ 2556
10	สรุปผลการดำเนินการ 6 เดือน ประจำปีงบประมาณ พ.ศ. 2558
11	สีการใช้ประโยชน์อาคารในแผนที่การใช้อาคารมาตราส่วน 1:4000
12	จำนวนเหรียญกษาปณ์หมุนเวียนในระบบเศรษฐกิจ ปีงบประมาณ 2558
13	ข้อมูลโรงงานประเภท 101, 105 และ 106 เดือนมกราคม พ.ศ.2559
14	ข้อมูลโรงงาน จำพวกที่ 3 ในเขต กทม. เดือนมกราคม พ.ศ.2559
15	ข้อมูลโรงงานอุตสาหกรรมแจ้งประกอบ จำพวกที่ 3 ในเขต กรุงเทพมหานคร

ตารางที่ ข.4 รายงานจากหมวดหมู่สาธารณสุข เรียงลำดับตามความนิยม ประจำเดือน ตุลาคม 2560

ลำดับ	ชื่อรายงาน
1	ผู้บาดเจ็บและเสียชีวิต ในช่วงเทศกาล ปี 51-58_CutName_NewYear_Edit
2	ราคากลางยา แผนปัจจุบัน
3	ข้อมูลหน่วยงานบริการสุขภาพในกรุงเทพมหานคร

ตารางที่ ข.4 รายงานจากหมวดหมู่สาธารณสุข เรียงลำดับตามความนิยมประจำเดือน ตุลาคม 2560 (ต่อ)

ลำดับ	ชื่อรายงาน
4	ข้อมูลด้านการป้องกันและบรรเทาสาธารณภัย องค์กรปกครองส่วนท้องถิ่น
5	ข้อมูลหน่วยงานบริการสุขภาพจังหวัดภูเก็ต
6	จำนวนพาหนะปฏิบัติการฉุกเฉิน ที่ขึ้นทะเบียนในระบบการแพทย์ฉุกเฉิน (ITEMS)
7	จำนวนหน่วยปฏิบัติการฉุกเฉิน ที่ขึ้นทะเบียนในระบบการแพทย์ฉุกเฉิน (ITEMS)
8	จำนวนผู้ปฏิบัติการฉุกเฉิน ที่ขึ้นทะเบียนในระบบการแพทย์ฉุกเฉิน (ITEMS)
9	จำนวนชุดปฏิบัติการฉุกเฉิน ที่ขึ้นทะเบียนในระบบการแพทย์ฉุกเฉิน (ITEMS)
10	จำนวนผู้ปฏิบัติการฉุกเฉิน ที่ขึ้นทะเบียนในระบบการแพทย์ฉุกเฉิน (ITEMS)
11	จำนวนชุดปฏิบัติการฉุกเฉิน ที่ขึ้นทะเบียนในระบบการแพทย์ฉุกเฉิน (ITEMS)
12	รายงานที่ตั้งศูนย์รับแจ้งเหตุและสั่งการจังหวัด
13	รายงานที่ตั้งงานการแพทย์ฉุกเฉินจังหวัด
14	จำนวนหน่วยปฏิบัติการฉุกเฉิน ที่ขึ้นทะเบียนในระบบการแพทย์ฉุกเฉิน (ITEMS)

ตารางที่ ข.5 รายงานจากหมวดหมู่งบประมาณและการใช้จ่ายของภาครัฐ เรียงลำดับตามความนิยม ประจำเดือน ตุลาคม 2560

ลำดับ	ชื่อรายงาน
1	ประมาณการรายได้ และผลการจัดเก็บรายได้ ปีงบประมาณ 2558
2	รายงานสรุปผลการดำเนินงานจัดซื้อจัดจ้างของสำนักงานปลัดกระทรวงยุติธรรม ประจำปีงบประมาณ 2558 - 2560
3	ผลการเบิกจ่ายงบประมาณของสำนักงาน ป.ป.ท.
4	รายงานการจัดซื้อจัดจ้างของหน่วยงานภาครัฐ เดือนมีนาคม 2558
5	รายงานรายละเอียดผลการจัดซื้อจัดจ้าง เดือน มิถุนายน 2558
6	รายงานด้านเศรษฐกิจ ปี 2554 - 2557

ตารางที่ ข.5 รายงานจากหมวดหมู่งบประมาณและการใช้จ่ายของภาครัฐ เรียงลำดับตามความนิยม ประจำเดือน ตุลาคม 2560 (ต่อ)

ลำดับ	ชื่อรายงาน
7	สถิติการดำเนินงานและฐานะการเงินสหกรณ์ในภาคการเกษตร ในรอบ 10 ปี (2548 - 2557)
8	รายงานรายละเอียดผลการจัดซื้อจัดจ้าง เดือน กันยายน 2558 ถึง มีนาคม 2559
9	รายงานการจัดซื้อจัดจ้างของหน่วยงานภาครัฐ เดือนสิงหาคม 2559 (ครึ่งเดือนแรก)

ตารางที่ ข.6 รายงานจากหมวดหมู่พลังงาน ทรัพยากรธรรมชาติ และสิ่งแวดล้อม เรียงลำดับตามความนิยมประจำเดือน ตุลาคม 2560

ลำดับ	ชื่อรายงาน
1	ปริมาณการจำหน่ายรายจังหวัด น้ำมันดีเซลหมุนเร็วธรรมดา สาขาสถานีบริการ
2	ข้อมูลคุณภาพอากาศประเทศไทย (PM10) รายชั่วโมง พ.ศ.2557 กรมควบคุมมลพิษ
3	พื้นที่ป่าไม้ แยกรายจังหวัด ปี 2558
4	รายงานพลังงานทดแทนของประเทศไทย 2557 (Thailand Alternative Energy Situation 2014)
5	ปริมาณการจำหน่ายรายจังหวัด แก๊สโซฮอล์ 91 สาขาสถานีบริการ
6	ปริมาณการจำหน่ายรายจังหวัด แก๊สโซฮอล์ อี 20 สาขาสถานีบริการ
7	ปริมาณการจำหน่ายรายจังหวัด แก๊สโซฮอล์ 95 สาขาสถานีบริการ
8	ข้อมูลด้านสิ่งแวดล้อม องค์กรปกครองส่วนท้องถิ่น
9	แหล่งสิ่งแวดล้อมศิลปกรรมของประเทศไทย
10	ข้อมูลผลการตรวจวัดคุณภาพน้ำจากสถานีตรวจวัดคุณภาพน้ำอัตโนมัติแบบ real-time ของกรมควบคุมมลพิษ ในลุ่มน้ำเจ้าพระยา
11	ปริมาณการจำหน่ายรายจังหวัด ก๊าซแอลพีจี สาขาสถานีบริการ
12	การผลิต การใช้ และการนำเข้าสุทธิ พลังงานเชิงพาณิชย์ขั้นต้น จำแนกตามเชื้อเพลิง
13	การจัดสรรค่าภาคหลวงให้ อบต. อบจ. และเทศบาลทั่วประเทศ ปี 2553 - 2558

ตารางที่ ข.6 รายงานจากหมวดหมู่พลังงาน ทรัพยากรธรรมชาติ และสิ่งแวดล้อม เรียงลำดับตามความนิยมประจำเดือน ตุลาคม 2560 (ต่อ)

ลำดับ	ชื่อรายงาน
14	รายงานดุลยภาพพลังงานของประเทศไทย ปี 2558 (Energy Balance of Thailand 2015)
15	ปริมาณการจำหน่ายรายจังหวัด แก๊สโซฮอล์ อี 85 สาขาสถานีบริการ
16	ข้อมูลผลการตรวจวัดคุณภาพน้ำจากสถานีตรวจวัดคุณภาพน้ำอัตโนมัติแบบ real-time ของกรมควบคุมมลพิษ ในลุ่มน้ำท่าจีน
17	ปริมาณการจำหน่ายรายจังหวัด แก๊สโซฮอล์ 91 สาขาอุตสาหกรรม
18	Mineral Production of Thailand (2011 - 2015) for the period of January - March 2015
19	ปริมาณการจำหน่ายรายจังหวัด ก๊าซแอลพีจี สาขาอุตสาหกรรม
20	ผลการดำเนินงานด้านพลังงานทดแทน เดือน ม.ค. - มิ.ย. 2558 (ปีปฏิทิน) และ เดือน ต.ค. 2557 - มิ.ย. 2558 (ปีงบประมาณ)
21	ปริมาณการจำหน่ายรายจังหวัด ดีเซลพื้นฐาน สาขาอุตสาหกรรม
22	ปริมาณการจำหน่ายรายจังหวัด แก๊สโซฮอล์ อี 20 สาขาการขนส่ง

ตารางที่ ข.7 รายงานจากหมวดหมู่กฎหมาย ศาล และอาชญากรรม เรียงลำดับตามความนิยมประจำเดือน ตุลาคม 2560

ลำดับ	ชื่อรายงาน
1	ข้อมูลสถิติการจับกุมคดียาเสพติด ประจำปีงบประมาณ 2558
2	สถิติการรับเรื่องร้องเรียนการทุจริตในภาครัฐของสำนักงาน ป.ป.ท.
3	สรุปผลการดำเนินการมาตรการริบทรัพย์สินค้าจากคดียาเสพติด เดือน ม.ค. - ส.ค. ปี 2558
4	ข้อมูลสถิติการคุมประพฤติดีมาแล้วชั่วช่วงเทศกาลปีใหม่ตั้งแต่ พ.ศ. 2549 - พ.ศ 2559
5	ข้อมูลสถิติคดีสืบเสาะและพินิจของสำนักงานคุมประพฤติทั่วประเทศ ปีงบประมาณ 2546 - 2559 กรมคุมประพฤติ

ตารางที่ ข.7 รายงานจากหมวดหมู่กฎหมาย ศาล และอาชญากรรม เรียงลำดับตามความนิยม ประจำเดือน ตุลาคม 2560 (ต่อ)

ลำดับ	ชื่อรายงาน
6	สถิติแสดงผลการดำเนินงานด้านปราบปรามการทุจริตของคณะกรรมการ ป.ป.ช. (ตั้งแต่ปีงบประมาณ พ.ศ. 2550 - 2556)
7	ข้อมูลสถิติคดีฟืนฟูฯ จำแนกชนิดยาเสพติดของสำนักงานคุมประพฤติทั่วประเทศ ปีงบประมาณ 2550 - 2559
8	ข้อมูลสถิติการคุมความประพฤติคดีเมาแล้วขับ ช่วงเทศกาลสงกรานต์ ตั้งแต่ พ.ศ. 2549 - พ.ศ. 2558
9	ข้อมูลสถิติงานควบคุมและสอดส่อง จำแนกภารกิจ เพศ อายุของผู้ถูกคุมประพฤติ ของสำนักงานคุมประพฤติทั่วประเทศ ปีงบประมาณ 2550 - 2559
10	ข้อมูลสถิติคดีตรวจพิสูจน์ จำแนกชนิดยาเสพติดของสำนักงานคุมประพฤติทั่วประเทศ ปีงบประมาณ 2550 - 2559
11	ข้อมูลสถิติคดีควบคุมและสอดส่องของสำนักงานคุมประพฤติ กรมคุมประพฤติ ปีงบประมาณ 2546 - 2559
12	ข้อมูลสถิติรายงานอายุและเพศของจำเลย งานสืบเสาะและพินิจ ของสำนักงานคุมประพฤติทั่วประเทศ ประจำปีงบประมาณ 2550 - 2559 กรมคุมประพฤติ
13	สถิติผลการดำเนินงานด้านการตรวจสอบทรัพย์สินและหนี้สิน (ตั้งแต่ปีงบประมาณ พ.ศ. 2550 - 2556)
14	ข้อมูลสถิติการคุมประพฤติทั้งผู้ใหญ่และเด็กตั้งแต่ปี พ.ศ. 2548 - 31 มกราคม 2559
15	ข้อมูลสถิติคดีควบคุมและสอดส่องจำแนกตามฐานความผิดของสำนักงานคุมประพฤติทั่วประเทศ ปีงบประมาณ 2553 - 2559
16	ข้อมูลสถิติคดีฟืนฟูฯ จำแนกตามฐานความผิดของสำนักงานคุมประพฤติทั่วประเทศ ปีงบประมาณ 2550 - 2559

ตารางที่ ข.8 รายงานจากหมวดหมู่เกษตรกรรมและชลประทาน เรียงลำดับตามความนิยม  
ประจำเดือน ตุลาคม 2560

ลำดับ	ชื่อรายงาน
1	ทะเบียนรายชื่อผู้ได้รับการรับรองมาตรฐานฟาร์มเพาะเลี้ยงสัตว์น้ำ กรมประมง (พ.ศ. 2553 - 2558)
2	ข้อมูลงานวิจัยด้านการประมงชายฝั่งของกรมประมง (พ.ศ. 2541 - 2558)
3	ข้อมูลการพัฒนาแหล่งน้ำขนาดเล็กกรมพัฒนาที่ดิน ปี 2556-2557
4	สถิติการดำเนินงานและฐานะการเงินสหกรณ์ในภาคการเกษตร ในรอบ 10 ปี (2549 - 2558)

## ภาคผนวก ค

## คำหยุดที่ปรากฏในชื่อของรายงาน

ส่วนนี้ผู้วิจัยจะแสดงคำหยุดที่ปรากฏในชื่อของรายงาน โดยแสดงชนิดของคำอ้างอิงตามชนิดของคำของคลังคำ ORCHID (ศูนย์เทคโนโลยีอิเล็กทรอนิกส์ และคอมพิวเตอร์แห่งชาติ, 2561)

ตารางที่ ค.1 ตารางแสดงคำหยุดและชนิดของคำเหล่านั้นที่พบในชื่อของรายงาน

ลำดับ	คำหยุด	ชนิดของคำ
1	มาก	ADVN (ADVERB WITH NORMAL FORM)
2	บาท	CMTR (MEASUREMENT CLASSIFIER)
3	ราย	CNIT (UNIT CLASSIFIER)
4	แห่ง	CNIT (UNIT CLASSIFIER)
5	ทั้งสิ้น	DDAN (DEFINITE DETERMINER, AFTER NOUN)
6	อื่นๆ	DIAC (INDEFINITE DETERMINER, FOLLOWING)
7	การ	FIXN (NOMINAL PREFIX)
8	และ	JCRG (COORDINATING CONJUNCTION)
9	หรือ	JCRG (COORDINATING CONJUNCTION)
10	โดย	JSBR (SUBORDINATING CONJUNCTION)
11	พ.ย.	NCMN (COMMON NOUN)
12	ก.ย.	NCMN (COMMON NOUN)
13	ปี	NCMN (COMMON NOUN)
14	มิ.ย.	NCMN (COMMON NOUN)
15	มีนาคม	NCMN (COMMON NOUN)



ตารางที่ ค.1 ตารางแสดงคำหยุดและชนิดของคำเหล่านั้นที่พบในชื่อของรายงาน (ต่อ)

ลำดับ	คำหยุด	ชนิดของคำ
16	ต.ค.	NCMN (COMMON NOUN)
17	สิงหาคม	NCMN (COMMON NOUN)
18	มิถุนายน	NCMN (COMMON NOUN)
19	ม.ค.	NCMN (COMMON NOUN)
20	พ.	NCMN (COMMON NOUN)
21	ตุลาคม	NCMN (COMMON NOUN)
22	อังคาร	NCMN (COMMON NOUN)
23	ราชอาณาจักร	NCMN (COMMON NOUN)
24	พ.ค.	NCMN (COMMON NOUN)
25	ช่วง	NCMN (COMMON NOUN)
26	พ.ศ.	NCMN (COMMON NOUN)
27	พฤษภาคม	NCMN (COMMON NOUN)
28	จำนวน	NCMN (COMMON NOUN)
29	พฤ.	NCMN (COMMON NOUN)
30	มี.ค.	NCMN (COMMON NOUN)
31	ก.ค.	NCMN (COMMON NOUN)
32	เสาร์	NCMN (COMMON NOUN)
33	เดือน	NCMN (COMMON NOUN)
34	ก.พ.	NCMN (COMMON NOUN)
35	ข้อมูล	NCMN (COMMON NOUN)
36	ศุกร์	NCMN (COMMON NOUN)
37	แบบ	NCMN (COMMON NOUN)
38	จันทร์	NCMN (COMMON NOUN)
39	ธันวาคม	NCMN (COMMON NOUN)

ตารางที่ ค.1 ตารางแสดงคำหยุดและชนิดของคำเหล่านั้นที่พบในชื่อของรายงาน (ต่อ)

ลำดับ	คำหยุด	ชนิดของคำ
40	ส.ค.	NCMN (COMMON NOUN)
41	ส.	NCMN (COMMON NOUN)
42	ครัวเรือน	NCMN (COMMON NOUN)
43	พฤศจิกายน	NCMN (COMMON NOUN)
44	ศ.	NCMN (COMMON NOUN)
45	จ.	NCMN (COMMON NOUN)
46	ป.	NCMN (COMMON NOUN)
47	ธ.ค.	NCMN (COMMON NOUN)
48	อา.	NCMN (COMMON NOUN)
49	ป.ป.ช.	NCMN (COMMON NOUN)
50	พุธ	NCMN (COMMON NOUN)
51	พฤหัสบดี	NCMN (COMMON NOUN)
52	กรกฎาคม	NCMN (COMMON NOUN)
53	กันยายน	NCMN (COMMON NOUN)
54	เมษายน	NCMN (COMMON NOUN)
55	เม.ย.	NCMN (COMMON NOUN)
56	.	NCNM (CARDINAL NUMBER)
57	กทม.	NPRP (PROPER NOUN)
58	กุมภาพันธ์	NPRP (PROPER NOUN)
59	มกราคม	NPRP (PROPER NOUN)
60	จปฐ	NPRP (PROPER NOUN)
61	อาทิตย์	NPRP (PROPER NOUN)
62	ศ.	NTTL (TITLE NOUN)
63	อ.	NTTL (TITLE NOUN)

ตารางที่ ค.1 ตารางแสดงคำหยุดและชนิดของคำเหล่านั้นที่พบในชื่อของรายงาน (ต่อ)

ลำดับ	คำหยุด	ชนิดของคำ
64	ที่	PREL (RELATIVE PRONOUN)
65	ฯ	PUNC (PUNCTUATION)
66	ต่อ	RPRE (PREPOSITION)
67	รอบ	RPRE (PREPOSITION)
68	จาก	RPRE (PREPOSITION)
69	เฉพาะ	RPRE (PREPOSITION)
70	ตาม	RPRE (PREPOSITION)
71	ถึง	RPRE (PREPOSITION)
72	ของ	RPRE (PREPOSITION)
73	ทั่ว	RPRE (PREPOSITION)
74	ณ	RPRE (PREPOSITION)
75	ใน	RPRE (PREPOSITION)
76	จำแนก	VACT (ACTIVE VERB)
77	น้อย	VATT (ATTRIBUTIVE VERB)
78	มี	VSTA (STATIVE VERB)
79	สำเร็จ	VSTA (STATIVE VERB)
80	สะสม	VSTA (STATIVE VERB)
81	ได้รับ	VSTA (STATIVE VERB)
82	สังกัด	VSTA (STATIVE VERB)
83	เป็น	VSTA (STATIVE VERB)
84	อยู่	XVAE (POST-VERB AUXILIARY)

## ภาคผนวก ง

### สถิติคำและชนิดของคำที่ปรากฏในรายงาน

ส่วนนี้ผู้วิจัยจะแสดงอันดับคำ 10 อันดับแรกที่ปรากฏในชื่อของรายงานแยกตามตำแหน่งของคำที่ผ่านการตัดคำด้วยเครื่องมือ LexTo โดยแสดงชนิดของคำอ้างอิงตามชนิดของคำของคลังคำ ORCHID (ศูนย์เทคโนโลยีอิเล็กทรอนิกส์ และคอมพิวเตอร์แห่งชาติ, 2561) และแสดงความถี่ของคำที่ปรากฏ สำหรับบางตำแหน่งอาจมีจำนวนคำที่ปรากฏไม่ถึง 10 คำ จะแสดงตามจำนวนสูงสุดที่ปรากฏ ซึ่งทั้งหมดถูกจัดตามส่วนของชื่อรายงานโดยแสดงอัตราส่วนดังรูปที่ ง.1 รูปที่ ง.2 และรูปที่ ง.3

ตารางที่ ง.1 คำที่ปรากฏในชื่อรายงานในตำแหน่งคำที่ 1 เรียงตามลำดับความถี่

ลำดับ	คำ	ชนิดของคำ	ความถี่ของคำที่ปรากฏ
1	จำนวน	NCMN (COMMON NOUN)	27
2	ผู้	PPRS (PERSONAL PRONOUN)	7
3	รายได้	NCMN (COMMON NOUN)	6
4	ตำแหน่งงาน	NCMN (COMMON NOUN)	6
5	ร้อยละ	NCMN (COMMON NOUN)	5
6	ค่าใช้จ่าย	NCMN (COMMON NOUN)	5
7	การบรรจุ	NCMN (COMMON NOUN)	3
8	รายจ่าย	NCMN (COMMON NOUN)	1
9	อัตรา	NCMN (COMMON NOUN)	1
10	ข้อมูล	NCMN (COMMON NOUN)	1

ตารางที่ ง.2 คำที่ปรากฏในชื่อรายงานในตำแหน่งคำที่ 2 เรียงตามลำดับความถี่

ลำดับ	คำ	ชนิดของคำ	ความถี่ของคำที่ปรากฏ
1	เฉลี่ย	VACT (ACTIVE VERB)	10
2	ผู้	PPRS (PERSONAL PRONOUN)	9
3	ลงทะเบียน	VACT (ACTIVE VERB)	7
4	ว่าง	VSTA (STATIVE VERB)	6
5	ของ	RPRE (PREPOSITION)	5
6	และ	JCRG (COORDINATING CONJUNCTION)	4
7	ครัวเรือน	NCMN (COMMON NOUN)	3
8	ประชากร	NCMN (COMMON NOUN)	3
9	งาน	NCMN (COMMON NOUN)	3
10	ลูกจ้าง	NCMN (COMMON NOUN)	3

ตารางที่ ง.3 คำที่ปรากฏในชื่อรายงานในตำแหน่งคำที่ 3 เรียงตามลำดับความถี่

ลำดับ	คำ	ชนิดของคำ	ความถี่ของคำที่ปรากฏ
1	จำแนก	VACT (ACTIVE VERB)	12
2	ต่อ	RPRE (PREPOSITION)	10
3	มี	VSTA (STATIVE VERB)	9
4	สมัครงาน	NCMN (COMMON NOUN)	7
5	อายุ	NCMN (COMMON NOUN)	3
6	ครัวเรือน	NCMN (COMMON NOUN)	3
7	ร้อยละ	NCMN (COMMON NOUN)	3
8	เฉลี่ย	VACT (ACTIVE VERB)	3
9	ที่	PREL (RELATIVE PRONOUN)	2
10	ข้าราชการ	NCMN (COMMON NOUN)	2

ตารางที่ ง.4 คำที่ปรากฏในชื่อรายงานในตำแหน่งคำที่ 4 เรียงตามลำดับความถี่

ลำดับ	คำ	ชนิดของคำ	ความถี่ของคำที่ปรากฏ
1	ตาม	RPRE (PREPOSITION)	12
2	งาน	NCMN (COMMON NOUN)	10
3	จำแนก	VACT (ACTIVE VERB)	9
4	เดือน	NCMN (COMMON NOUN)	8
5	ของ	RPRE (PREPOSITION)	4
6	ตำแหน่งงาน	NCMN (COMMON NOUN)	3
7	ต่อ	RPRE (PREPOSITION)	3
8	ปี	NCMN (COMMON NOUN)	2
9	ที่	PREL (RELATIVE PRONOUN)	2
10	คนใน	NCMN (COMMON NOUN)	1

ตารางที่ ง.5 คำที่ปรากฏในชื่อรายงานในตำแหน่งคำที่ 5 เรียงตามลำดับความถี่

ลำดับ	คำ	ชนิดของคำ	ความถี่ของคำที่ปรากฏ
1	ของ	RPRE (PREPOSITION)	9
2	ตาม	RPRE (PREPOSITION)	9
3	ทำ	VACT (ACTIVE VERB)	9
4	อาชีพ	NCMN (COMMON NOUN)	5
5	อุตสาหกรรม	NCMN (COMMON NOUN)	5
6	ครัวเรือน	NCMN (COMMON NOUN)	3
7	ปี	NCMN (COMMON NOUN)	3
8	ว่าง	VSTA (STATIVE VERB)	3
9	มี	VSTA (STATIVE VERB)	2
10	ประชากร	NCMN (COMMON NOUN)	2

ตารางที่ ง.6 คำที่ปรากฏในชื่อรายงานในตำแหน่งคำที่ 6 เรียงตามลำดับความถี่

ลำดับ	คำ	ชนิดของคำ	ความถี่ของคำที่ปรากฏ
1	และ	JCRG (COORDINATING CONJUNCTION)	12
2	จำแนก	VACT (ACTIVE VERB)	9
3	ครัวเรือน	NCMN (COMMON NOUN)	5
4	อาชีพ	NCMN (COMMON NOUN)	4
5	มี	VSTA (STATIVE VERB)	3
6	ขึ้นไป	DIAQ (INDEFINITE DETERMINER)	3
7	ครอบครัว	NCMN (COMMON NOUN)	3
8	อายุ	NCMN (COMMON NOUN)	3
9	คน	NCMN (COMMON NOUN)	2
10	หนังสือ	NCMN (COMMON NOUN)	2

ตารางที่ ง.7 คำที่ปรากฏในชื่อรายงานในตำแหน่งคำที่ 7 เรียงตามลำดับความถี่

ลำดับ	คำ	ชนิดของคำ	ความถี่ของคำที่ปรากฏ
1	จำแนก	VACT (ACTIVE VERB)	9
2	ตาม	RPRE (PREPOSITION)	9
3	และ	JCRG (COORDINATING CONJUNCTION)	8
4	อายุ	NCMN (COMMON NOUN)	6
5	วุฒิการศึกษา	NCMN (COMMON NOUN)	4
6	การบรรจุ	NCMN (COMMON NOUN)	3
7	ข้าราชการ	NCMN (COMMON NOUN)	3
8	งาน	NCMN (COMMON NOUN)	3
9	ใน	RPRE (PREPOSITION)	2
10	ผลประโยชน์	NCMN (COMMON NOUN)	2

ตารางที่ ง.8 คำที่ปรากฏในชื่อรายงานในตำแหน่งคำที่ 8 เรียงตามลำดับความถี่

ลำดับ	คำ	ชนิดของคำ	ความถี่ของคำที่ปรากฏ
1	ตาม	RPRE (PREPOSITION)	9
2	ทั่ว	RPRE (PREPOSITION)	9
3	ครัวเรือน	NCMN (COMMON NOUN)	3
4	งาน	NCMN (COMMON NOUN)	3
5	จำแนก	VACT (ACTIVE VERB)	3
6	และ	JCRG (COORDINATING CONJUNCTION)	3
7	ทำ	VACT (ACTIVE VERB)	3
8	เฉลี่ย	VACT (ACTIVE VERB)	2
9	วุฒิการศึกษา	NCMN (COMMON NOUN)	2
10	อายุ	NCMN (COMMON NOUN)	2

ตารางที่ ง.9 คำที่ปรากฏในชื่อรายงานในตำแหน่งคำที่ 9 เรียงตามลำดับความถี่

ลำดับ	คำ	ชนิดของคำ	ความถี่ของคำที่ปรากฏ
1	ราชอาณาจักร	NCMN (COMMON NOUN)	9
2	ทั่ว	RPRE (PREPOSITION)	6
3	จำแนก	VACT (ACTIVE VERB)	6
4	ที่	PREL (RELATIVE PRONOUN)	4
5	ตาม	RPRE (PREPOSITION)	3
6	แหล่งที่มา	NCMN (COMMON NOUN)	2
7	หนังสือ	NCMN (COMMON NOUN)	2
8	ระดับการศึกษา	NCMN (COMMON NOUN)	2
9	ขึ้นไป	DIAQ (INDEFINITE DETERMINER)	2
10	ใน	RPRE (PREPOSITION)	2



ตารางที่ ง.10 คำที่ปรากฏในชื่อรายงานในตำแหน่งคำที่ 10 เรียงตามลำดับความถี่

ลำดับ	คำ	ชนิดของคำ	ความถี่ของคำที่ปรากฏ
1	ปี	NCMN (COMMON NOUN)	9
2	ราชอาณาจักร	NCMN (COMMON NOUN)	6
3	ตาม	RPRE (PREPOSITION)	6
4	ของ	RPRE (PREPOSITION)	3
5	ประเภท	NCMN (COMMON NOUN)	3
6	ที่	PREL (RELATIVE PRONOUN)	3
7	และ	JCRG (COORDINATING CONJUNCTION)	2
8	ราย	CNIT (UNIT CLASSIFIER)	2
9	เฉลี่ย	VACT (ACTIVE VERB)	2
10	จำแนก	VACT (ACTIVE VERB)	2

ตารางที่ ง.11 คำที่ปรากฏในชื่อรายงานในตำแหน่งคำที่ 11 เรียงตามลำดับความถี่

ลำดับ	คำ	ชนิดของคำ	ความถี่ของคำที่ปรากฏ
1	ปี	NCMN (COMMON NOUN)	7
2	และ	JCRG (COORDINATING CONJUNCTION)	4
3	อุตสาหกรรม	NCMN (COMMON NOUN)	4
4	จังหวัด	NCMN (COMMON NOUN)	3
5	ของ	RPRE (PREPOSITION)	3
6	ทั่ว	RPRE (PREPOSITION)	3
7	รายได้	NCMN (COMMON NOUN)	2
8	อาชีพ	NCMN (COMMON NOUN)	2
9	ต่อ	RPRE (PREPOSITION)	2
10	ตาม	RPRE (PREPOSITION)	2

ตารางที่ ง.12 คำที่ปรากฏในชื่อรายงานในตำแหน่งคำที่ 12 เรียงตามลำดับความถี่

ลำดับ	คำ	ชนิดของคำ	ความถี่ของคำที่ปรากฏ
1	ทั่ว	RPRE (PREPOSITION)	8
2	จำแนก	VACT (ACTIVE VERB)	3
3	รายได้	NCMN (COMMON NOUN)	3
4	ราชอาณาจักร	NCMN (COMMON NOUN)	3
5	พ.ศ.	NCMN (COMMON NOUN)	2
6	และ	JCRG (COORDINATING CONJUNCTION)	2
7	ราย	CNIT (UNIT CLASSIFIER)	2
8	ครัวเรือน	NCMN (COMMON NOUN)	2
9	สถานภาพ	NCMN (COMMON NOUN)	2
10	สถานะ	NCMN (COMMON NOUN)	2

ตารางที่ ง.13 คำที่ปรากฏในชื่อรายงานในตำแหน่งคำที่ 13 เรียงตามลำดับความถี่

ลำดับ	คำ	ชนิดของคำ	ความถี่ของคำที่ปรากฏ
1	ราชอาณาจักร	NCMN (COMMON NOUN)	8
2	และ	JCRG (COORDINATING CONJUNCTION)	5
3	ราย	CNIT (UNIT CLASSIFIER)	3
4	ตาม	RPRE (PREPOSITION)	3
5	ไตรมาส	NCMN (COMMON NOUN)	3
6	จาก	RPRE (PREPOSITION)	2
7	ภาค	NCMN (COMMON NOUN)	2
8	แรงงาน	NCMN (COMMON NOUN)	2
9	ทาง	NCMN (COMMON NOUN)	2
10	ระบบ	NCMN (COMMON NOUN)	2

ตารางที่ ง.14 คำที่ปรากฏในชื่อรายงานในตำแหน่งคำที่ 14 เรียงตามลำดับความถี่

ลำดับ	คำ	ชนิดของคำ	ความถี่ของคำที่ปรากฏ
1	ไตรมาส	NCMN (COMMON NOUN)	5
2	ภาค	NCMN (COMMON NOUN)	4
3	พ.ศ.	NCMN (COMMON NOUN)	3
4	ตำแหน่ง	NCMN (COMMON NOUN)	3
5	จำแนก	VACT (ACTIVE VERB)	3
6	ที่	PREL (RELATIVE PRONOUN)	3
7	ข้อมูล	NCMN (COMMON NOUN)	2
8	ปี	NCMN (COMMON NOUN)	2
9	ทั่ว	RPRE (PREPOSITION)	2
10	ทำ	VACT (ACTIVE VERB)	2

ตารางที่ ง.15 คำที่ปรากฏในชื่อรายงานในตำแหน่งคำที่ 15 เรียงตามลำดับความถี่

ลำดับ	คำ	ชนิดของคำ	ความถี่ของคำที่ปรากฏ
1	ที่	PREL (RELATIVE PRONOUN)	7
2	และ	JCRG (COORDINATING CONJUNCTION)	4
3	ประเภท	NCMN (COMMON NOUN)	3
4	ตาม	RPRE (PREPOSITION)	3
5	จปฐ	NCMN (COMMON NOUN)	2
6	ราชอาณาจักร	NCMN (COMMON NOUN)	2
7	ศร	NCMN (COMMON NOUN)	2
8	ทั่ว	RPRE (PREPOSITION)	2
9	เวลา	NCMN (COMMON NOUN)	2
10	ข้อมูล	NCMN (COMMON NOUN)	1

ตารางที่ ง.16 คำที่ปรากฏในชื่อรายงานในตำแหน่งคำที่ 16 เรียงตามลำดับความถี่

ลำดับ	คำ	ชนิดของคำ	ความถี่ของคำที่ปรากฏ
1	พ.ศ.	NCMN (COMMON NOUN)	4
2	ของ	RPRE (PREPOSITION)	3
3	ปี	NCMN (COMMON NOUN)	2
4	จังหวัด	NCMN (COMMON NOUN)	2
5	ราชอาณาจักร	NCMN (COMMON NOUN)	2
6	และ	JCRG (COORDINATING CONJUNCTION)	2
7	อุตสาหกรรม	NCMN (COMMON NOUN)	2
8	เหลือ	VSTA (STATIVE VERB)	2
9	มี	VSTA (STATIVE VERB)	1
10	ทั่ว	RPRE (PREPOSITION)	1

ตารางที่ ง.17 คำที่ปรากฏในชื่อรายงานในตำแหน่งคำที่ 17 เรียงตามลำดับความถี่

ลำดับ	คำ	ชนิดของคำ	ความถี่ของคำที่ปรากฏ
1	พ.ศ.	NCMN (COMMON NOUN)	7
2	ข้าราชการ	NCMN (COMMON NOUN)	3
3	ที่	PREL (RELATIVE PRONOUN)	3
4	สังคม	NCMN (COMMON NOUN)	2
5	เพศ	NCMN (COMMON NOUN)	2
6	ระหว่าง	RPRE (PREPOSITION)	2
7	ตำแหน่ง	NCMN (COMMON NOUN)	2
8	ปี	NCMN (COMMON NOUN)	1
9	ความต้องการ	NCMN (COMMON NOUN)	1
10	ราชอาณาจักร	NCMN (COMMON NOUN)	1

ตารางที่ ง.18 คำที่ปรากฏในชื่อรายงานในตำแหน่งคำที่ 18 เรียงตามลำดับความถี่

ลำดับ	คำ	ชนิดของคำ	ความถี่ของคำที่ปรากฏ
1	และ	JCRG (COORDINATING CONJUNCTION)	5
2	พ.ศ.	NCMN (COMMON NOUN)	4
3	ของ	RPRE (PREPOSITION)	2
4	ปี	NCMN (COMMON NOUN)	2
5	ประเภท	NCMN (COMMON NOUN)	2
6	สังคม	NCMN (COMMON NOUN)	1
7	มาก	VATT (ATTRIBUTIVE VERB)	1
8	ธันวาคม	NCMN (COMMON NOUN)	1
9	ราย	CNIT (UNIT CLASSIFIER)	1
10	จังหวัด	NCMN (COMMON NOUN)	1

ตารางที่ ง.19 คำที่ปรากฏในชื่อรายงานในตำแหน่งคำที่ 19 เรียงตามลำดับความถี่

ลำดับ	คำ	ชนิดของคำ	ความถี่ของคำที่ปรากฏ
1	ของ	RPRE (PREPOSITION)	3
2	ภาค	NCMN (COMMON NOUN)	3
3	ทั่ว	RPRE (PREPOSITION)	2
4	พ.ศ.	NCMN (COMMON NOUN)	2
5	ครัวเรือน	NCMN (COMMON NOUN)	2
6	ต้องการ	VACT (ACTIVE VERB)	2
7	ระหว่าง	RPRE (PREPOSITION)	1
8	ประเภท	CLTV (COLLECTIVE CLASSIFIER)	1
9	สังคม	NCMN (COMMON NOUN)	1
10	จังหวัด	NCMN (COMMON NOUN)	1

ตารางที่ ง.20 คำที่ปรากฏในชื่อรายงานในตำแหน่งคำที่ 20 เรียงตามลำดับความถี่

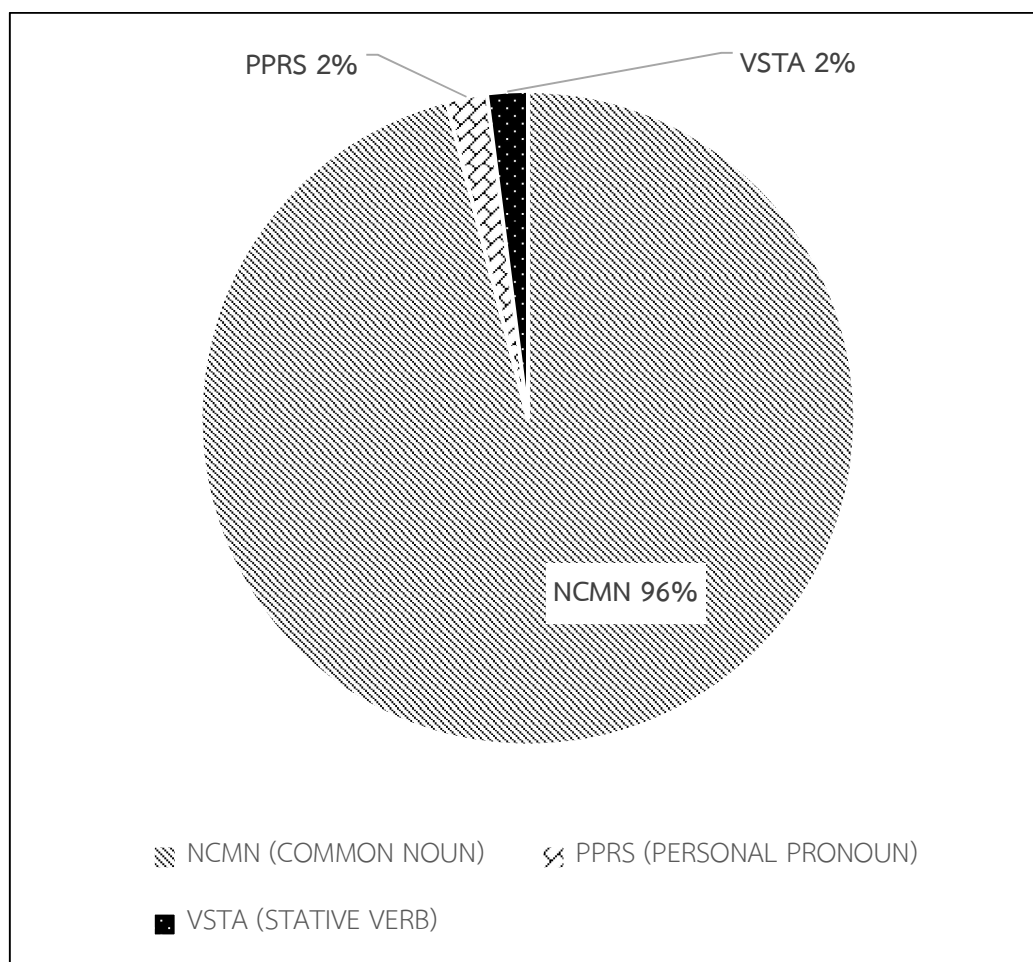
ลำดับ	คำ	ชนิดของคำ	ความถี่ของคำที่ปรากฏ
1	ทั่ว	RPRE (PREPOSITION)	5
2	ราชอาณาจักร	NCMN (COMMON NOUN)	2
3	ข้าราชการ	NCMN (COMMON NOUN)	2
4	ทำงาน	VACT (ACTIVE VERB)	2
5	ครัวเรือน	NCMN (COMMON NOUN)	1
6	พ.ศ.	NCMN (COMMON NOUN)	1
7	ปี	NCMN (COMMON NOUN)	1
8	ผู้รับ	NCMN (COMMON NOUN)	1
9	ของ	RPRE (PREPOSITION)	1
10	ระหว่าง	RPRE (PREPOSITION)	1

ตารางที่ ง.21 คำที่ปรากฏในชื่อรายงานในตำแหน่งคำที่ 21 เรียงตามลำดับความถี่

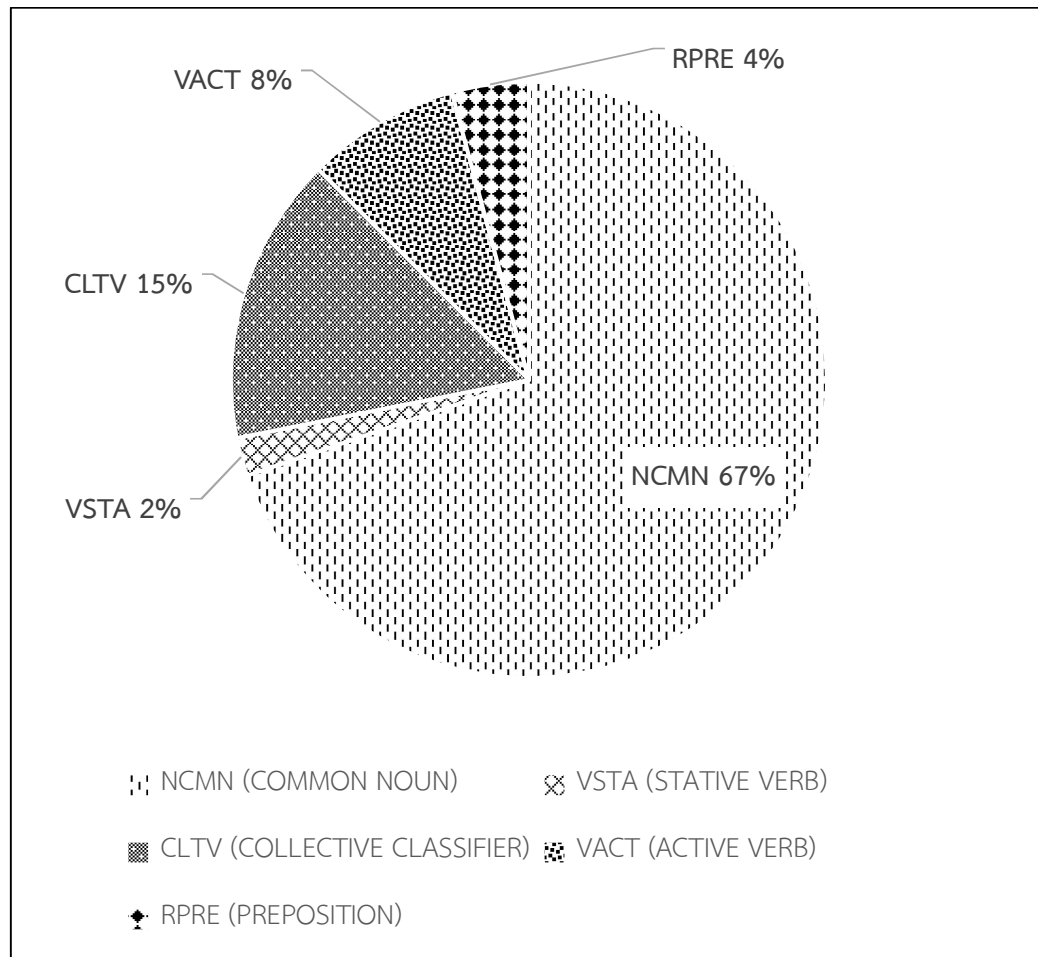
ลำดับ	คำ	ชนิดของคำ	ความถี่ของคำที่ปรากฏ
1	ราชอาณาจักร	NCMN (COMMON NOUN)	5
2	ปี	NCMN (COMMON NOUN)	2
3	เพิ่มขึ้น	VSTA (STATIVE VERB)	2
4	ทั่ว	RPRE (PREPOSITION)	1
5	งาน	NCMN (COMMON NOUN)	1
6	ครัวเรือน	NCMN (COMMON NOUN)	1
7	ราย	CNIT (UNIT CLASSIFIER)	1
8	พ.ศ.	NCMN (COMMON NOUN)	1
9	ไตรมาส	NCMN (COMMON NOUN)	1

ตารางที่ ง.22 คำที่ปรากฏในชื่อรายงานในตำแหน่งคำที่ 22 เรียงตามลำดับความถี่

ลำดับ	คำ	ชนิดของคำ	ความถี่ของคำที่ปรากฏ
1	พ.ศ.	NCMN (COMMON NOUN)	4
2	ทั่ว	RPRE (PREPOSITION)	3
3	ราชอาณาจักร	NCMN (COMMON NOUN)	1
4	เพศ	NCMN (COMMON NOUN)	1
5	ภาค	NCMN (COMMON NOUN)	1
6	ที่	PREL (RELATIVE PRONOUN)	1
7	ไตรมาส	NCMN (COMMON NOUN)	1

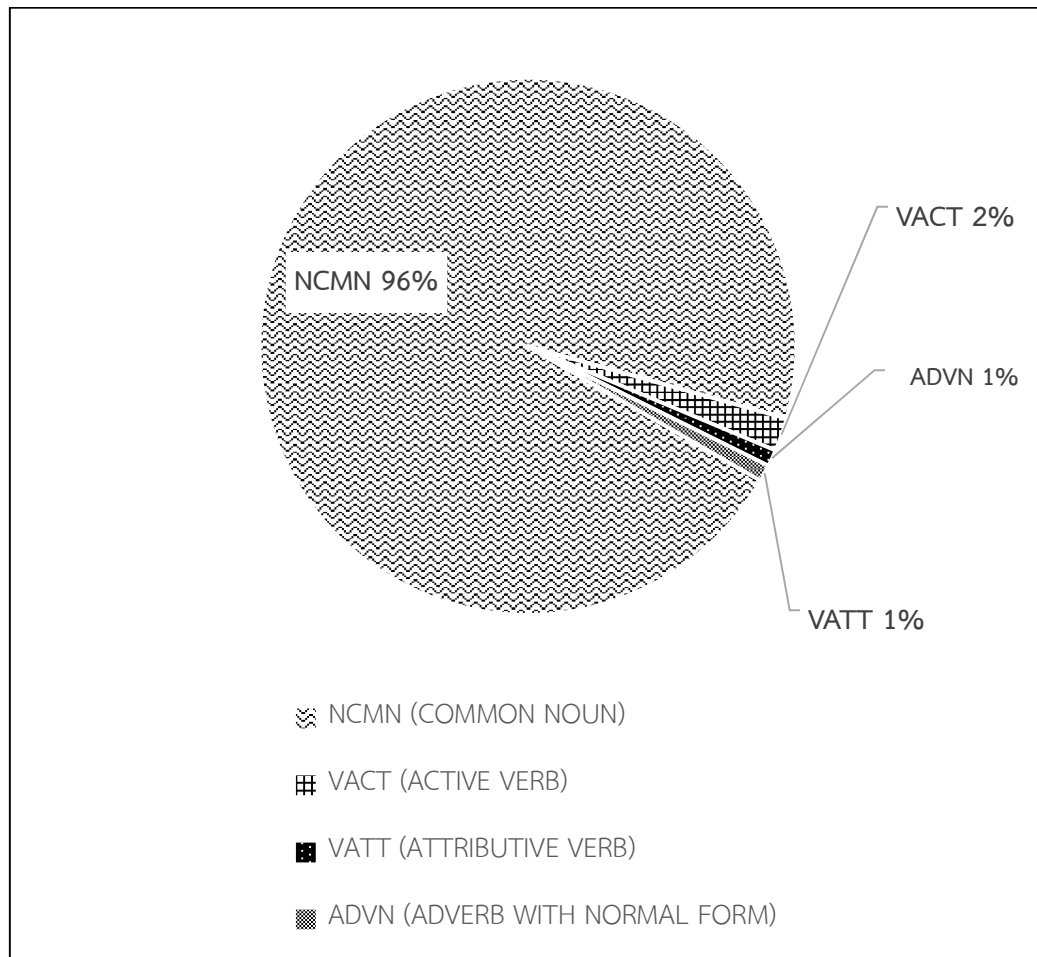


รูปที่ ง.1 อัตราส่วนชนิดของคำที่ปรากฏในตำแหน่งส่วนเริ่มต้นชื่อรายงาน (Beginning Phrase)



รูปที่ ง.2 อัตราส่วนชนิดของคำที่ปรากฏในตำแหน่งส่วนกลางของชื่อรายงาน (Middle Phrase)





รูปที่ ง.3 อัตราส่วนชนิดของคำที่ปรากฏในตำแหน่งส่วนท้ายชื่อรายงาน (Ending Phrase)

## ภาคผนวก จ

### ตัวอย่างชื่อรายงานที่แบบจำลองสร้างขึ้นจากชื่อรายงาน

ส่วนนี้ผู้วิจัยจะแสดงตัวอย่างของชื่อรายงานที่ผ่านการตัดคำและการตัดคำหยุด เพื่อให้ได้คำสำคัญที่ใช้ในการสร้างเป็นชื่อรายงาน โดยแสดงตัวอย่าง 10 อันดับของรายงานที่ได้รับ ความนิยมในเดือนตุลาคม 2560 จากหมวดหมู่สังคมและสวัสดิการ โดยตัวอย่างแบ่งออกเป็น 6 ส่วน

- 4) ชื่อรายงาน แสดงชื่อของรายงาน
- 5) ชื่อรายงานที่ผ่านการตัดคำ แสดงคำที่ผ่านการตัดคำด้วยเครื่องมือ LexTo
- 6) ชื่อรายงานที่ผ่านการตัดคำหยุด แสดงคำที่ผ่านการตัดคำหยุดโดยแบบจำลอง
- 7) คำสำคัญ แสดงคำสำคัญที่ได้จากชื่อรายงาน โดยแบ่งเป็น ส่วนเริ่มต้นของชื่อ รายงาน ส่วนกลางของชื่อรายงาน และส่วนท้ายของชื่อรายงาน
- 8) คำสำคัญที่ได้ตรวจสอบ แสดงคำสำคัญที่ผ่านการตรวจสอบความถูกต้อง และปรับปรุงให้มีความสละสลวยแล้ว
- 9) ชื่อรายงานที่เกิดขึ้น แสดงชื่อรายงานที่เกิดขึ้น โดยเป็นชื่อที่เกิดจากการผสมคำ สำคัญที่ผ่านการตรวจสอบความถูกต้อง และปรับปรุงให้มีความสละสลวยแล้ว

ตารางที่ จ.1 ตัวอย่างชื่อรายงานที่แบบจำลองสร้างขึ้นจากรายงานที่ 1

ชื่อรายงาน	รายได้เฉลี่ยต่อเดือนต่อครัวเรือน จำแนกตามภาค และจังหวัด พ.ศ. 2541 - 2558		
ชื่อรายงานที่ผ่านการตัดคำ	รายได้ เฉลี่ย ต่อ เดือน ต่อ ครัวเรือน   จำแนก ตาม ภาค   และ จังหวัด   พ.ศ.   2541     2558		
ชื่อรายงานที่ผ่านการตัดคำหยุด	รายได้ เฉลี่ย ภาค จังหวัด		
ตำแหน่งคำสำคัญ	ส่วนเริ่มต้น	ส่วนกลาง	ส่วนท้าย
คำสำคัญ	รายได้	เฉลี่ย ภาค	จังหวัด
คำสำคัญที่ได้ตรวจสอบ	รายได้	เฉลี่ยของแต่ละภาค	จำแนกตามจังหวัด
ชื่อรายงานที่เกิดขึ้น	รายได้เฉลี่ยของแต่ละภาคจำแนกตามจังหวัด		

ตารางที่ จ.2 ตัวอย่างชื่อรายงานที่แบบจำลองสร้างขึ้นจากรายงานที่ 2

ชื่อรายงาน	รายได้เฉลี่ยต่อปีของคนและครัวเรือน (รายจังหวัด) จากข้อมูล จปฐ ปี 2558		
ชื่อรายงานที่ผ่านการตัดคำ	รายได้ เฉลี่ย ต่อ ปี ของ คน และ ครัวเรือน   ( ราย จังหวัด )   จาก ข้อมูล   จปฐ   ปี   2558		
ชื่อรายงานที่ผ่านการตัดคำหยุด	รายได้ เฉลี่ย คน จังหวัด		
ตำแหน่งคำสำคัญ	ส่วนเริ่มต้น	ส่วนกลาง	ส่วนท้าย
คำสำคัญ	รายได้	เฉลี่ย คน	จังหวัด
คำสำคัญที่ได้ตรวจสอบ	รายได้	เฉลี่ยต่อคน	จำแนกตามจังหวัด
ชื่อรายงานที่เกิดขึ้น	รายได้เฉลี่ยต่อคนจำแนกตามจังหวัด		

ตารางที่ จ.3 ตัวอย่างชื่อรายงานที่แบบจำลองสร้างขึ้นจากชื่อรายงานที่ 3

ชื่อรายงาน	จำนวนครัวเรือนที่คนในครัวเรือนมีรายได้เฉลี่ยน้อยกว่า 30,000 บาทต่อปี จากข้อมูล จปฐ ปี 2555-2558		
ชื่อรายงานที่ผ่านการตัดคำ	จำนวน ครัวเรือนที่ คนใน ครัวเรือน มี รายได้ เฉลี่ย น้อยกว่า   30,000   บาท ต่อปี   จาก ข้อมูล   จปฐ   ปี   2555 - 2558		
ชื่อรายงานที่ผ่านการตัดคำหยุด	คนใน รายได้ เฉลี่ย น้อยกว่า		
ตำแหน่งคำสำคัญ	ส่วนเริ่มต้น	ส่วนกลาง	ส่วนท้าย
คำสำคัญ	คนใน	รายได้ เฉลี่ย	น้อยกว่า
คำสำคัญที่ได้ตรวจสอบ	จำนวนคน	มีรายได้เฉลี่ย	น้อยกว่า
ชื่อรายงานที่เกิดขึ้น	จำนวนคนมีรายได้เฉลี่ยน้อยกว่า		

ตารางที่ จ.4 ตัวอย่างชื่อรายงานที่แบบจำลองสร้างขึ้นจากชื่อรายงานที่ 4

ชื่อรายงาน	รายจ่ายเฉลี่ยต่อปีของคนและครัวเรือน (รายจังหวัด) จากข้อมูล จปฐ ปี 2558		
ชื่อรายงานที่ผ่านการตัดคำ	รายจ่าย เฉลี่ย ต่อปี ของ คน และ ครัวเรือน   (ราย จังหวัด)   จาก ข้อมูล   จปฐ   ปี   2558		
ชื่อรายงานที่ผ่านการตัดคำหยุด	รายจ่าย เฉลี่ย คน จังหวัด		
ตำแหน่งคำสำคัญ	ส่วนเริ่มต้น	ส่วนกลาง	ส่วนท้าย
คำสำคัญ	รายจ่าย	เฉลี่ย คน	จังหวัด
คำสำคัญที่ได้ตรวจสอบ	รายจ่าย	เฉลี่ยต่อคน	จำแนกตามจังหวัด
ชื่อรายงานที่เกิดขึ้น	รายได้เฉลี่ยต่อคนจำแนกตามจังหวัด		

ตารางที่ จ.5 ตัวอย่างชื่อรายงานที่แบบจำลองสร้างขึ้นจากชื่อรายงานที่ 5

ชื่อรายงาน	อัตรการว่างงาน จำแนกตามเพศ พ.ศ. 2549 - 2558		
ชื่อรายงานที่ผ่านการตัดคำ	อัตรา การ ว่างงาน   จำแนก ตาม เพศ   พ.ศ.   2549   -   2558		
ชื่อรายงานที่ผ่านการตัดคำหยุด	อัตรา ว่างงาน เพศ		
ตำแหน่งคำสำคัญ	ส่วนเริ่มต้น	ส่วนกลาง	ส่วนท้าย
คำสำคัญ	อัตรา	ว่างงาน	เพศ
คำสำคัญที่ได้ตรวจสอบ	อัตรการ	ว่างงาน	จำแนกตามเพศ
ชื่อรายงานที่เกิดขึ้น	อัตรการว่างงานจำแนกตามเพศ		

ตารางที่ จ.6 ตัวอย่างชื่อรายงานที่แบบจำลองสร้างขึ้นจากชื่อรายงานที่ 6

ชื่อรายงาน	จำนวนประชากรอายุ 15 ปีขึ้นไป จำแนกตามสถานภาพแรงงาน เพศ ทั่วราชอาณาจักร พ.ศ. 2549 - 2558		
ชื่อรายงานที่ผ่านการตัดคำ	จำนวน ประชากร อายุ   15   ปี ขึ้นไป   จำแนก ตาม สถานภาพ แรงงาน   เพศ   ทั่ว ราชอาณาจักร   พ.ศ.   2549   -   2558		
ชื่อรายงานที่ผ่านการตัดคำหยุด	ประชากร อายุ ขึ้นไป สถานภาพ แรงงาน เพศ		
ตำแหน่งคำสำคัญ	ส่วนเริ่มต้น	ส่วนกลาง	ส่วนท้าย
คำสำคัญ	ประชากร	อายุขึ้นไป  สถานภาพ แรงงาน	เพศ
คำสำคัญที่ได้ตรวจสอบ	จำนวนประชากร	อายุ ขึ้นไป	จำแนกตามสถานภาพ แรงงาน จำแนกตาม เพศ
ชื่อรายงานที่เกิดขึ้น	จำนวนประชากรอายุ ขึ้นไปจำแนกตามสถานภาพแรงงาน จำนวนประชากรอายุ ขึ้นไปจำแนกตามเพศ		

ตารางที่ จ.7 ตัวอย่างชื่อรายงานที่แบบจำลองสร้างขึ้นจากชื่อรายงานที่ 7

ชื่อรายงาน	ตำแหน่งงานว่าง จำแนกตามอาชีพ และวุฒิการศึกษา ทั่วราชอาณาจักร ปี 2558		
ชื่อรายงานที่ผ่านการตัดคำ	ตำแหน่งงานว่าง จำแนกตามอาชีพ และวุฒิการศึกษา ทั่วราชอาณาจักร ปี 2558		
ชื่อรายงานที่ผ่านการตัดคำหยุด	ตำแหน่งงานว่างอาชีพวุฒิการศึกษา		
ตำแหน่งคำสำคัญ	ส่วนเริ่มต้น	ส่วนกลาง	ส่วนท้าย
คำสำคัญ	ตำแหน่งงาน	ว่างอาชีพ	วุฒิการศึกษา
คำสำคัญที่ได้ตรวจสอบ	ตำแหน่งงานว่าง	-	จำแนกตามอาชีพ จำแนกตามวุฒิการศึกษา
ชื่อรายงานที่เกิดขึ้น	ตำแหน่งงานว่างจำแนกตามอาชีพ ตำแหน่งงานว่างจำแนกตามวุฒิการศึกษา		

ตารางที่ จ.8 ตัวอย่างชื่อรายงานที่แบบจำลองสร้างขึ้นจากชื่อรายงานที่ 8

ชื่อรายงาน	ผู้ลงทะเบียนสมัครงาน จำแนกตามอาชีพ และวุฒิการศึกษา ทั่วราชอาณาจักร ปี 2558		
ชื่อรายงานที่ผ่านการตัดคำ	ผู้ลงทะเบียนสมัครงาน จำแนกตามอาชีพ และวุฒิการศึกษา ทั่วราชอาณาจักร ปี 2558		
ชื่อรายงานที่ผ่านการตัดคำหยุด	ผู้ลงทะเบียนสมัครงานอาชีพวุฒิการศึกษา		
ตำแหน่งคำสำคัญ	ส่วนเริ่มต้น	ส่วนกลาง	ส่วนท้าย
คำสำคัญ	ผู้	ลงทะเบียนสมัครงานอาชีพ	วุฒิการศึกษา
คำสำคัญที่ได้ตรวจสอบ	ผู้ลงทะเบียน	สมัครงาน	จำแนกตามอาชีพ จำแนกตามวุฒิการศึกษา
ชื่อรายงานที่เกิดขึ้น	ผู้ลงทะเบียนสมัครงานจำแนกตามอาชีพ ผู้ลงทะเบียนสมัครงานจำแนกตามวุฒิการศึกษา		

ตารางที่ จ.9 ตัวอย่างชื่อรายงานที่แบบจำลองสร้างขึ้นจากชื่อรายงานที่ 9

ชื่อรายงาน	การบรรจุงาน จำแนกตามอาชีพ และวุฒิการศึกษา ทั่วราชอาณาจักร ปี 2558		
ชื่อรายงานที่ผ่านการตัดคำ	การบรรจุงาน จำแนกตาม อาชีพ และ วุฒิการศึกษา ทั่ว ราชอาณาจักร ปี 2558		
ชื่อรายงานที่ผ่านการตัดคำหยุด	การบรรจุงาน อาชีพ วุฒิการศึกษา		
ตำแหน่งคำสำคัญ	ส่วนเริ่มต้น	ส่วนกลาง	ส่วนท้าย
คำสำคัญ	การบรรจุ	งาน อาชีพ	วุฒิการศึกษา
คำสำคัญที่ได้ตรวจสอบ	การบรรจุงาน	-	จำแนกตามอาชีพ  จำแนกตามวุฒิ การศึกษา
ชื่อรายงานที่เกิดขึ้น	การบรรจุงานจำแนกตามอาชีพ การบรรจุงานจำแนกตามวุฒิการศึกษา		

ตารางที่ จ.10 ตัวอย่างชื่อรายงานที่แบบจำลองสร้างขึ้นจากชื่อรายงานที่ 10

ชื่อรายงาน	ข้อมูลการจัดหางานในประเทศ จำแนกตามจังหวัด ทั่วราชอาณาจักร ปี 2558		
ชื่อรายงานที่ผ่านการตัดคำ	ข้อมูล การจัดหางาน ในประเทศ จำแนกตาม จังหวัด ทั่ว ราชอาณาจักร ปี 2558		
ชื่อรายงานที่ผ่านการตัดคำหยุด	จัดหางาน ในประเทศ จังหวัด		
ตำแหน่งคำสำคัญ	ส่วนเริ่มต้น	ส่วนกลาง	ส่วนท้าย
คำสำคัญ	จัดหา	งาน ในประเทศ	จังหวัด
คำสำคัญที่ได้ตรวจสอบ	จำนวนการจัดหางาน	ในประเทศ	จำแนกตามจังหวัด
ชื่อรายงานที่เกิดขึ้น	จำนวนการจัดหางานในประเทศจำแนกตามจังหวัด		

## ภาคผนวก ฉ

## คำสำคัญที่ปรากฏในชื่อรายงาน

ส่วนนี้ผู้วิจัยจะแสดงคำสำคัญที่ปรากฏในชื่อรายงาน โดยแสดงเป็น 3 ส่วน คือ ส่วนเริ่มต้นชื่อรายงาน ส่วนกลางของชื่อรายงาน และส่วนท้ายชื่อรายงาน ซึ่งคำสำคัญทั้งหมดผ่านการตรวจสอบความถูกต้อง และปรับแก้ให้มีความสละสลวยแล้ว

ตารางที่ ฉ.1 คำสำคัญที่ปรากฏในชื่อรายงาน

ลำดับ	คำสำคัญ	ตำแหน่งที่ปรากฏ
1	รายได้	ส่วนเริ่มต้นชื่อรายงาน
2	มูลค่าสินทรัพย์	ส่วนเริ่มต้นชื่อรายงาน
3	มูลค่าทรัพย์สิน	ส่วนเริ่มต้นชื่อรายงาน
4	การบรรจุงาน	ส่วนเริ่มต้นชื่อรายงาน
5	ตำแหน่งงาน	ส่วนเริ่มต้นชื่อรายงาน
6	อัตราการว่างงาน	ส่วนเริ่มต้นชื่อรายงาน
7	จำนวนลูกจ้างรัฐบาล	ส่วนเริ่มต้นชื่อรายงาน
8	ตำแหน่งงานว่าง	ส่วนเริ่มต้นชื่อรายงาน
9	ข้อมูลการจัดหางาน	ส่วนเริ่มต้นชื่อรายงาน
10	แสดงหนี้สิน	ส่วนเริ่มต้นชื่อรายงาน
11	แสดงทรัพย์สิน	ส่วนเริ่มต้นชื่อรายงาน
12	จำนวนคนต่างด้าว	ส่วนเริ่มต้นชื่อรายงาน
13	จำนวนคนใน	ส่วนเริ่มต้นชื่อรายงาน
14	จำนวนประชากร	ส่วนเริ่มต้นชื่อรายงาน
15	ค่าใช้จ่าย	ส่วนเริ่มต้นชื่อรายงาน
16	จำนวนลูกจ้าง	ส่วนเริ่มต้นชื่อรายงาน



ตารางที่ ๑.1 คำสำคัญที่ปรากฏในชื่อรายงาน (ต่อ)

ลำดับ	คำสำคัญ	ตำแหน่งที่ปรากฏ
17	รายจ่าย	ส่วนเริ่มต้นชื่อรายงาน
18	จำนวนผู้ลงทะเบียน	ส่วนเริ่มต้นชื่อรายงาน
19	การอยู่อาศัย	ส่วนเริ่มต้นชื่อรายงาน
20	จำนวนสิ่งปลูกสร้าง	ส่วนเริ่มต้นชื่อรายงาน
21	สถานภาพการทำงาน	ส่วนเริ่มต้นชื่อรายงาน
22	ภาวะการมีงานทำ	ส่วนเริ่มต้นชื่อรายงาน
23	จำนวนยานพาหนะ	ส่วนเริ่มต้นชื่อรายงาน
24	ค่าจ้าง	ส่วนเริ่มต้นชื่อรายงาน
25	การยืม	ส่วนเริ่มต้นชื่อรายงาน
26	การเป็นหนี้	ส่วนเริ่มต้นชื่อรายงาน
27	จำนวนคู่สมรส	ส่วนเริ่มต้นชื่อรายงาน
28	ลักษณะ	ส่วนเริ่มต้นชื่อรายงาน
29	การบรรจุงาน	ส่วนเริ่มต้นชื่อรายงาน
30	จำนวนคน	ส่วนเริ่มต้นชื่อรายงาน
31	จำนวนแรงงาน	ส่วนเริ่มต้นชื่อรายงาน
32	ที่มีการเปลี่ยนแปลง	ส่วนกลางของชื่อรายงาน
33	ตามรูปแบบการทำงาน	ส่วนกลางของชื่อรายงาน
34	เฉลี่ย	ส่วนกลางของชื่อรายงาน
35	รายชั่วโมง	ส่วนกลางของชื่อรายงาน
36	ภายในระยะเวลา	ส่วนกลางของชื่อรายงาน
37	นอกประเทศ	ส่วนกลางของชื่อรายงาน
38	ตามวัตถุประสงค์	ส่วนกลางของชื่อรายงาน
39	ในแต่ละครอบครัว	ส่วนกลางของชื่อรายงาน
40	ในประเทศ	ส่วนกลางของชื่อรายงาน
41	ระหว่างประเทศ	ส่วนกลางของชื่อรายงาน
42	จำแนกตามระดับการศึกษา	ส่วนท้ายชื่อรายงาน
43	จำแนกตามกลุ่มแรงงาน	ส่วนท้ายชื่อรายงาน
44	จำแนกตามสัปดาห์	ส่วนท้ายชื่อรายงาน

ตารางที่ ๑.1 คำสำคัญที่ปรากฏในชื่อรายงาน (ต่อ)

ลำดับ	คำสำคัญ	ตำแหน่งที่ปรากฏ
45	จำแนกตามวุฒิการศึกษา	ส่วนท้ายชื่อรายงาน
46	ในสังคม	ส่วนท้ายชื่อรายงาน
47	ของข้าราชการ	ส่วนท้ายชื่อรายงาน
48	จำแนกตามไตรมาส	ส่วนท้ายชื่อรายงาน
49	จำแนกตามเขตการปกครอง	ส่วนท้ายชื่อรายงาน
50	จำแนกตามภาค	ส่วนท้ายชื่อรายงาน
51	จำแนกตามอาชีพ	ส่วนท้ายชื่อรายงาน
52	จำแนกตามอายุ	ส่วนท้ายชื่อรายงาน
53	จำแนกตามกลุ่มอุตสาหกรรม	ส่วนท้ายชื่อรายงาน
54	จำแนกตามจังหวัด	ส่วนท้ายชื่อรายงาน
55	จำแนกตามเพศ	ส่วนท้ายชื่อรายงาน
56	น้อยกว่าสัญชาติ	ส่วนท้ายชื่อรายงาน
57	น้อยกว่าประเภท	ส่วนท้ายชื่อรายงาน
58	น้อยกว่าแหล่งที่มา	ส่วนท้ายชื่อรายงาน
59	น้อยกว่าตำแหน่ง	ส่วนท้ายชื่อรายงาน
60	น้อยกว่าสถานภาพสมรส	ส่วนท้ายชื่อรายงาน

## ประวัติผู้เขียน

ชื่อ สกุล	นายไตรรัตน์ พัฒน์เอี่ยม	
รหัสประจำตัว	5830223010	
นักศึกษา		
วุฒิการศึกษา		
วุฒิ	ชื่อสถาบัน	ปีที่สำเร็จการศึกษา
วิทยาศาสตร์บัณฑิต (วิศวกรรมซอฟต์แวร์)	มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตภูเก็ต	2555

## การตีพิมพ์เผยแพร่ผลงาน

ไตรรัตน์ พัฒน์เอี่ยม และ ณิชฐพงศ์ ทองเทพ. (2560). “การสร้างรายงานอัตโนมัติด้วยการสกัดชื่อรายงานและโครงสร้างข้อมูล.”, เอกสารการประชุมวิชาการ ระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 13 ประจำปี 2560, โรงแรมอโนมาแกรนด์ กรุงเทพฯ: 6-7 กรกฎาคม 2560.