


การวิเคราะห์ปัจจัยที่มีผลต่อการสำเร็จการศึกษาของนักศึกษาภาควิชาวิทยาการคอมพิวเตอร์  
มหาวิทยาลัยสยาม

Analysis of Factors Affecting Graduation of Computer Science Students Siam University

The logo of Siam University is a large, circular emblem in the background. It features a central shield with a crown on top, surrounded by a wreath. The shield is flanked by two stars. The outer ring of the logo contains the text 'SIAM UNIVERSITY' in English and Thai. The Thai text 'มหาวิทยาลัยสยาม' is written in a stylized font along the top inner edge of the ring.

นายณฤทธิ	จตุรัสพันแสง	6004800055
นายนิธิศ	ชวรัตน์เจริญกิจ	6004800062

ปริญญาานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรบัณฑิต  
ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์  
มหาวิทยาลัยสยาม  
ปีการศึกษา 2563

หัวข้อปริญญานิพนธ์

การวิเคราะห์ปัจจัยที่มีผลต่อการสำเร็จการศึกษาของนักศึกษา  
ภาควิชาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยสยาม  
Analysis of Factors Affecting Graduation of Computer Science  
Students Siam University

หน่วยกิตของปริญญานิพนธ์

3 หน่วยกิต

รายชื่อคณะผู้จัดทำ

นายณฤทธิ์ จัตุรัสพันแสง 6004800055

นายนิธิศ ชวรัตน์เจริญกิจ 6004800062

อาจารย์ที่ปรึกษา

อาจารย์จรรยา แหยมเจริญ

ระดับการศึกษา

วิทยาศาสตรบัณฑิต

ภาควิชา

วิทยาการคอมพิวเตอร์

ปีการศึกษา

2563

อนุมัติให้ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรบัณฑิต  
สาขาวิชาวิทยาการคอมพิวเตอร์

คณะกรรมการสอบปริญญานิพนธ์

  
.....ประธานกรรมการ  
( ผู้ช่วยศาสตราจารย์ ดร.นุชงค์ อูทโยภาส )

  
.....กรรมการ  
( อาจารย์เอก บำรุงศรี )

  
.....อาจารย์ที่ปรึกษา  
( อาจารย์จรรยา แหยมเจริญ )

หัวข้อปริญญานิพนธ์	การวิเคราะห์ปัจจัยที่มีผลต่อการสำเร็จการศึกษาของนักศึกษา ภาควิชาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยสยาม		
หน่วยกิตของปริญญานิพนธ์	3 หน่วยกิต		
รายชื่อคณะผู้จัดทำ	นายณฤทธิ	จตุรัสพันธ์แสง	6004800055
	นายนิริศ	ชวรัตน์เจริญกิจ	6004800062
อาจารย์ที่ปรึกษา	อาจารย์จรรยา แหยมเจริญ		
ระดับการศึกษา	วิทยาศาสตรบัณฑิต		
ภาควิชา	วิทยาการคอมพิวเตอร์		
ปีการศึกษา	2563		

### บทคัดย่อ

วัตถุประสงค์ในการจัดทำปริญญานิพนธ์นี้เพื่อวิเคราะห์ปัจจัยที่ส่งผลต่อการสำเร็จการศึกษาของนักศึกษาภาควิชาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยสยาม เนื่องจากมีจำนวนนักศึกษาที่เข้าศึกษาต่อในหลักสูตรวิทยาการคอมพิวเตอร์ มีอัตราต้อออกและย้ายสาขาที่สูงเกินกว่าร้อยละ 50 จากจำนวนนักศึกษาที่เข้าศึกษาต่อในปีการศึกษานั้นๆ และมีอัตราการสำเร็จการศึกษาตามเกณฑ์การศึกษา 4 ปี ที่ต่ำ ทางภาควิชาจึงต้องการทราบสาเหตุที่แท้จริง จึงได้มอบหมายให้คณะผู้จัดทำทำการวิเคราะห์หาสาเหตุดังกล่าว โดยได้กำหนดปัจจัยที่มีผลต่อการสำเร็จการศึกษาของนักศึกษาในชั้นต้นไว้ 5 ปัจจัย ได้แก่ 1) วุฒิการศึกษาที่จบก่อนเข้าศึกษาต่อในหลักสูตรวิทยาการคอมพิวเตอร์ 2) เกรดเฉลี่ยสะสม (GPA) 3) รายวิชาบังคับที่ต้องเรียนก่อน (Pre-Requisite) 4) ผลการศึกษาตามรายวิชา และ 5) จำนวนหน่วยกิตสะสมที่ลงทะเบียนเรียนและผ่าน ในการวิเคราะห์ข้อมูลได้ใช้ข้อมูลการลงทะเบียนเรียนและผลการเรียนของนักศึกษาที่เข้าศึกษาต่อในหลักสูตรวิทยาการคอมพิวเตอร์ตั้งแต่ปีการศึกษา 2555 ถึง 2560 ขั้นตอนในการดำเนินการวิเคราะห์ประกอบด้วย 1) กำหนดวัตถุประสงค์ในการวิเคราะห์ข้อมูล 2) รวบรวมข้อมูล 3) เตรียมข้อมูล 4) จัดกลุ่มข้อมูล และ 5) นำเสนอแผนภาพข้อมูล เครื่องมือในการวิเคราะห์ข้อมูล ได้แก่ โปรแกรม Tableau ผลลัพธ์ที่ได้จากการวิเคราะห์ ปัจจัยที่มีผลต่อการสำเร็จการศึกษาของนักศึกษาภาควิชาวิทยาการคอมพิวเตอร์ ประกอบด้วย เกรดเฉลี่ยสะสม ผลการศึกษาตามรายวิชา และจำนวนหน่วยกิตสะสมที่ลงทะเบียนเรียนและผ่าน โดยปัจจัยที่ไม่มีผลต่อการสำเร็จการศึกษาของนักศึกษาภาควิชาวิทยาการคอมพิวเตอร์ ประกอบด้วย วุฒิการศึกษาที่จบก่อนเข้าศึกษาต่อ และรายวิชาบังคับที่ต้องเรียนก่อน ซึ่งอาจารย์ที่ปรึกษาชั้นปีและอาจารย์ผู้สอนจะนำผลการวิเคราะห์ไปใช้เพื่อกำกับดูแลนักศึกษาและปรับปรุงวิธีการเรียนการสอนให้มีประสิทธิภาพยิ่งขึ้น

**คำสำคัญ:** ปัจจัยที่ส่งผลต่อการสำเร็จการศึกษา, การวิเคราะห์ข้อมูล, การจัดกลุ่มข้อมูล

**Title** Analysis of Factors Affecting Graduation of Computer Science Students  
Siam University

**Credit** 3 Units

**By** Mr. Narit Chatturatphansang 6004800055  
Mr. Nithit Chavaratjaroenkit 6004800062

**Advisor** Miss Janya Yamchareon

**Program** Bachelor of Science

**Major** Computer Science

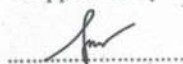
**Academic year** 2021

### Abstract

The objective of this project was to analyze the factors affecting the graduation of computer science students. According to the high number of students entered in computer science program, there is a dropout rate that was higher than 50 percent, and has a low graduation rate. The department wants to know the real reason and assigned to our team to analyze to find the real factor. The term initially identified five factors affecting students' graduation: 1) completed educational qualifications before entering into the computer science programs; 2) grade point average (GPA); 3) courses required study first (prerequisite); 4) study results in each course; and 5) the number of cumulative credits registered and passed. In the data analysis, enrollment and academic performance data of students admitted to the computer science programs from 2012 to 2017 were used. The analysis process consisted of: 1) research understanding; 2) data understanding; 3) data preparation; 4) data clustering; and 5) data visualization. Data analysis tool was Tableau. The study found, the factors affecting the graduation of computer science students consisted of grade point average, study results in each course, and the number of cumulative credits registered and passed. The factors that did not affect the graduation of students consisted of the completed educational qualifications before entering in the computer science program and courses required to study first. The advisors and teachers will use the results of the analysis to supervise students and improve teaching methods to be more effective.

**Keywords:** graduation, data analytics, data clustering

Approved by



Approved by



## กิตติกรรมประกาศ

### (Acknowledgment)

การจัดทำปฏิญานិพนธ์ฉบับนี้สำเร็จได้นั้น คณะผู้จัดทำได้รับความกรุณาจากอาจารย์ผู้สอนทุกท่านที่ให้ข้อมูลต่างๆ ส่งผลให้คณะผู้จัดทำได้รับความรู้และประสบการณ์ต่างๆ ที่มีค่ามากมายสำหรับปฏิญานิพนธ์ฉบับนี้สำเร็จลงได้ด้วยดีจากความร่วมมือและสนับสนุนจากหลายฝ่ายดังนี้

- |                 |                 |                        |
|-----------------|-----------------|------------------------|
| 1. อาจารย์จรรยา | แหมมเจริญ       | อาจารย์ที่ปรึกษา       |
| 2. นายณฤทธิ์    | จตุรัสพันแสง    | ตำแหน่ง : Data Analyst |
| 3. นายนิธิศ     | ชวรัตน์เจริญกิจ | ตำแหน่ง : Data Analyst |

คณะผู้จัดทำใคร่ขอขอบพระคุณคณะกรรมการสอบปฏิญานิพนธ์ ที่ได้ให้คำแนะนำสำคัญในการสอบปฏิญานิพนธ์ฉบับนี้ และผู้มีส่วนร่วมทุกท่าน รวมทั้งผู้ที่ไม่ได้กล่าวนามที่มีส่วนร่วมในการให้ข้อมูลให้ความช่วยเหลือ และเป็นທີ່ปรึกษาให้คำแนะนำต่างๆ จนทำให้งานทุกอย่างประสบความสำเร็จไปด้วยดี และจัดทำรายงานฉบับนี้จนเสร็จสมบูรณ์ ซึ่งคณะผู้จัดทำขอขอบพระคุณเป็นอย่างสูงไว้ ณ ที่นี้ด้วย

คณะผู้จัดทำ

นายณฤทธิ์ จตุรัสพันแสง

นายนิธิศ ชวรัตน์เจริญกิจ

# สารบัญ

หน้า

บทคัดย่อ.....	ก
Abstract .....	ข
กิตติกรรมประกาศ (Acknowledgment).....	ค
บทที่ 1 บทนำ .....	1
1.1 ที่มาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์ของปริญญาานิพนธ์ .....	1
1.3 ขอบเขตของปริญญาานิพนธ์ .....	2
1.4 ประโยชน์ที่คาดว่าจะได้รับ .....	2
1.5 ขั้นตอนและวิธีการดำเนินงานปริญญาานิพนธ์ .....	2
1.6 ระยะเวลาดำเนินงานปริญญาานิพนธ์ .....	4
1.7 อุปกรณ์และเครื่องมือที่ใช้ในการวิเคราะห์ข้อมูล .....	4
บทที่ 2 การทบทวนวรรณกรรมที่เกี่ยวข้อง .....	6
2.1 Data Analytics.....	6
2.2 ประเภทของข้อมูลสำหรับการวิเคราะห์ข้อมูล .....	8
2.3 Clustering.....	9
2.4 Data Preparation .....	10
2.5 Data Visualization .....	11
2.6 Tableau .....	14
บทที่ 3 การวิเคราะห์ข้อมูล .....	15
3.1 รายละเอียดของปริญญาานิพนธ์ .....	15
3.2 ขั้นตอนในการวิเคราะห์ข้อมูล.....	16
3.2.1 กำหนดวัตถุประสงค์ในการวิเคราะห์ข้อมูล (Research Understanding).....	16
3.2.2 รวบรวมข้อมูล (Data Understanding) .....	17
3.2.3 เตรียมข้อมูล (Data Preparation).....	17
3.2.4 จัดกลุ่มข้อมูล (Data Clustering).....	21

## สารบัญ (ต่อ)

	หน้า
บทที่ 4 การออกแบบทางกายภาพ.....	40
4.1 การวิเคราะห์ข้อมูลแบบจัดกลุ่ม GPA ของนักศึกษาที่กำลังศึกษาอยู่.....	40
4.2 การวิเคราะห์ภาพรวมผลการศึกษาของแต่ละรายวิชา.....	41
4.3 การวิเคราะห์ภาพรวมผลการศึกษาแต่ละรายวิชาตามปีการศึกษา.....	42
4.4 การวิเคราะห์เกรดตามรายวิชาของนักศึกษาตามปีการศึกษา.....	43
4.5 การวิเคราะห์เกรดเฉลี่ยสะสมและจำนวนรายวิชาที่ลงทะเบียนของนักศึกษาที่กำลังศึกษา.....	44
4.6 การวิเคราะห์ข้อมูลแบบจัดกลุ่ม GPA ของนักศึกษาที่จบศึกษาแล้ว.....	47
4.7 การวิเคราะห์ภาพรวมปีการศึกษาที่นักศึกษาเข้าศึกษาต่อในหลักสูตรและปีการศึกษาที่จบการศึกษา.....	48
4.8 การวิเคราะห์เกรดของแต่ละรายวิชาตามปีการศึกษาของนักศึกษาที่จบการศึกษาแล้ว.....	49
4.9 การวิเคราะห์ภาพรวมเกรดเฉลี่ยสะสมและจำนวนรายวิชาที่ลงทะเบียนเรียนของนักศึกษาที่จบการศึกษาแล้ว.....	50
4.10 การวิเคราะห์ภาพรวมผลการศึกษาของแต่ละรายวิชาของนักศึกษาที่ตกออกหรือย้ายออก.....	50
4.11 การวิเคราะห์ภาพรวมจำนวนเกรดของแต่ละรายวิชาและเกรดในแต่ละรายวิชาที่นักศึกษาลงทะเบียนเรียน.....	51
4.12 การวิเคราะห์ภาพรวมของเกรดเฉลี่ยสะสมและจำนวนรายวิชาที่ลงทะเบียนเรียนของนักศึกษาที่ตกออกหรือย้ายออก.....	52
4.13 การวิเคราะห์ภาพรวมจำนวนนักศึกษานักศึกษาตามสถานภาพของนักศึกษา ที่มีสถานภาพนักศึกษาอยู่ระหว่าง 50 และ 72.....	54
4.14 การวิเคราะห์ภาพรวมสถานภาพต่าง ๆ ของจำนวนคนที่เข้าศึกษาต่อในแต่ละปีการศึกษา และจำนวนวุฒิการศึกษาที่จบก่อนเข้าศึกษาต่อในหลักสูตร.....	55
บทที่ 5 สรุปผลและข้อเสนอแนะ.....	58
5.1 สรุปผลปริญญานิพนธ์.....	58
5.2 ข้อเสนอแนะ.....	59
บรรณานุกรม.....	60

## สารบัญรูปภาพ

หน้า

รูปที่ 2.1 กระบวนการเตรียมข้อมูล (Data Preparation) .....	10
รูปที่ 2.2 เวิร์กโฟลว์ของ Tableau .....	14
รูปที่ 3.1 ขั้นตอนการวิเคราะห์ปัจจัยที่มีผลต่อการสำเร็จการศึกษาของนักศึกษาภาควิชาวิทยาการ คอมพิวเตอร์ .....	16
รูปที่ 3.2 ตัวอย่างข้อมูลที่ได้จากระบบของสำนักทะเบียนและวัดผล .....	18
รูปที่ 3.3 ข้อมูลที่ผ่านการ Cleaning แล้ว .....	18
รูปที่ 3.4 ตัวอย่างข้อมูลนักศึกษาสำเร็จการศึกษาแล้ว .....	19
รูปที่ 3.5 ตัวอย่างข้อมูลนักศึกษาที่กำลังศึกษาอยู่ในปัจจุบัน .....	20
รูปที่ 3.6 ตัวอย่างข้อมูลนักศึกษาที่ไม่สำเร็จการศึกษา .....	20
รูปที่ 3.7 WorkBook Name: Status40_Number of people finished .....	22
รูปที่ 3.8 WorkBook Name: Status40_People Study and Course grades included .....	23
รูปที่ 3.9 แสดงข้อมูลการศึกษาของนักศึกษาเป็นรายบุคคล .....	24
รูปที่ 3.10 WorkBook Name: Status40_Total grades and Cumulative GPA and Register Subject .....	24
รูปที่ 3.11 แสดงข้อมูลการศึกษาของนักศึกษาเป็นรายบุคคล .....	25
รูปที่ 3.12 Work Book Name: Status40_Total grades and Cumulative GPA and Register Subject .....	26
รูปที่ 3.13 Workbooks Name: Status10_Cluster GPA .....	27
รูปที่ 3.14 Workbooks Name: Status10_People Study and Course grades included .....	28
รูปที่ 3.15 แสดงข้อมูลการศึกษาของนักศึกษาเป็นรายบุคคล .....	29
รูปที่ 3.16 Workbooks Name: Status10_Total grades and Cumulative GPA and Register Subject .....	30
รูปที่ 3.17 Workbooks Name: Status10_Process Course grades and Pie Grade .....	31
รูปที่ 3.18 Workbooks Name: Status50-72_People Study and Course grades included .....	32
รูปที่ 3.19 แสดงข้อมูลรายละเอียดเกรดในแต่ละวิชา.....	33
รูปที่ 3.20 Workbooks Name: Status50-72_Process Course grades and Pie Grade .....	34



## สารบัญรูปภาพ (ต่อ)

หน้า

รูปที่ 3.21 Workbooks Name: Status50-72_Total grades and Cumulative GPA and Register Subject.....	35
รูปที่ 3.22 Workbooks Name: Status50-72_Total number of status and people .....	36
รูปที่ 3.23 กราฟแท่งแสดงจำนวนข้อมูล โรงเรียนของนักศึกษา ก่อนเข้ามหาวิทยาลัยในแต่ละปี ..	37
รูปที่ 3.24 Dashboard 1 .....	38
รูปที่ 3.25 Dashboard 2.....	39
รูปที่ 4.1 แสดงผลการวิเคราะห์ข้อมูลแบบจัดกลุ่ม GPA ของนักศึกษาที่กำลังศึกษา.....	40
รูปที่ 4.2 แสดงภาพรวมผลการลงทะเบียนของนักศึกษาแต่ละคน.....	41
รูปที่ 4.3 แสดงภาพรวมผลการศึกษาของแต่ละรายวิชา.....	42
รูปที่ 4.4 ผลลัพธ์การศึกษาแต่ละรายวิชาตามปีการศึกษา .....	42
รูปที่ 4.5 ผลลัพธ์การศึกษาและจำนวนวิชาที่ลงทะเบียนตามภาคการศึกษา .....	43
รูปที่ 4.6 แสดงผลการวิเคราะห์ผลการศึกษาโดยรวม .....	44
รูปที่ 4.7 แสดงภาพรวมเกรดเฉลี่ยสะสม (GPA) .....	44
รูปที่ 4.8 ผลลัพธ์รายวิชาที่ลงในภาคการศึกษาในแต่ละปี .....	45
รูปที่ 4.9 ผลลัพธ์การนับจำนวนเกรดที่ลงทะเบียนเรียน .....	46
รูปที่ 4.10 แสดงภาพรวมเกรดเฉลี่ยสะสมและจำนวนรายวิชาที่ลงทะเบียนเรียนของนักศึกษาที่กำลัง ศึกษาอยู่.....	46
รูปที่ 4.11 แสดงการจัดกลุ่ม GPA ของนักศึกษาที่จบการศึกษาแล้ว.....	47
รูปที่ 4.12 แสดงจำนวนนักศึกษาที่จบการศึกษาตามเกณฑ์และจบการศึกษาเกินเกณฑ์ .....	48
รูปที่ 4.13 แสดงภาพรวมของจำนวนนักศึกษาที่จบการศึกษา .....	49
รูปที่ 4.14 แสดงผลเกรดของแต่ละรายวิชาตามปีการศึกษาของนักศึกษาที่จบการศึกษาแล้ว .....	49
รูปที่ 4.15 แสดงภาพรวมเกรดเฉลี่ยสะสมและจำนวนรายวิชาที่ลงทะเบียนเรียนของนักศึกษาที่จบ การศึกษาแล้ว .....	50
รูปที่ 4.16 แสดงภาพรวมผลการศึกษาของแต่ละรายวิชาของนักศึกษาที่ตกออกหรือย้ายออก....	51
รูปที่ 4.17 แสดงภาพรวมจำนวนเกรดของแต่ละรายวิชาและเกรดในแต่ละรายวิชาที่นักศึกษา ลงทะเบียนเรียน .....	51
รูปที่ 4.18 แสดงผลลัพธ์เกรดเฉลี่ยสะสมต่อภาคการศึกษา.....	52

## สารบัญรูปภาพ (ต่อ)

หน้า

รูปที่ 4.19 แสดงผลลัพธ์ผลการเรียนเกรดโดยรวมของนักศึกษาแต่ละคน .....	52
รูปที่ 4.20 แสดงผลลัพธ์จำนวนรายวิชาที่ลงทะเบียนในแต่ละภาคการศึกษา .....	53
รูปที่ 4.21 แสดงภาพรวมของเกรดเฉลี่ยสะสมและจำนวนรายวิชาที่ลงทะเบียนเรียนของนักศึกษาที่ ตกออกหรือย้ายออก .....	53
รูปที่ 4.22 แสดงจำนวนนักศึกษานักศึกษาตามสถานภาพของนักศึกษา ที่มีสถานภาพนักศึกษาอยู่ ระหว่าง 50 และ 72 .....	54
รูปที่ 4.23 แสดงผลสถานภาพนักศึกษาเป็นรายบุคคล .....	54
รูปที่ 4.24 แสดงจำนวนนักศึกษาที่เข้าศึกษาต่อในแต่ละปีการศึกษา .....	55
รูปที่ 4.25 แสดงจำนวนนักศึกษาแยกตามสถานภาพ .....	56
รูปที่ 4.26 แสดงจำนวนนักศึกษาแยกตามวุฒิการศึกษาที่จบก่อนเข้าศึกษาต่อในหลักสูตร .....	56
รูปที่ 4.27 แสดงภาพรวมสถานภาพต่าง ๆ ของจำนวนคนที่เข้าศึกษาต่อในแต่ละปีการศึกษา และ จำนวนวุฒิการศึกษาที่จบก่อนเข้าศึกษาต่อในหลักสูตร .....	57

# บทที่ 1

## บทนำ

### 1.1 ที่มาและความสำคัญของปัญหา

ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยสยาม เปิดสอนหลักสูตร วิทยาการคอมพิวเตอร์ ระดับปริญญาตรี ในแต่ละปีการศึกษามีนักเรียนที่เข้าศึกษาต่อในหลักสูตร และย้ายออกหรือเรียนไม่สำเร็จการศึกษาตามเกณฑ์ 4 ปีการศึกษา เป็นจำนวนมาก โดยสาเหตุที่ได้จากการสอบถามนักศึกษาที่ขอย้ายคณะ มีหลากหลายสาเหตุ ได้แก่ เรียนไม่ไหวเนื้อหายากเกินไป ยังไม่เข้าใจสิ่งที่ชอบ ต้องการไปเรียนกับเพื่อน เป็นต้น นักศึกษาไม่สำเร็จการศึกษาหรือต้อออก มีทั้ง พันสภาพนักศึกษานี้เนื่องจากผลการเรียน ลาออก ขาดการติดต่อกับทางมหาวิทยาลัยเป็นเวลามากกว่า 1 ปีการศึกษา และจบการศึกษาล่าช้า ใช้เวลาในการศึกษามากกว่า 4 ปีการศึกษา คณะผู้จัดทำจึงได้ ร่วมกับอาจารย์ในภาควิชาฯ ทำการศึกษาปัจจัยที่มีผลต่อการสำเร็จการศึกษาของนักศึกษาภาควิชา วิทยาการคอมพิวเตอร์ จากข้อมูลก่อนเข้าศึกษาต่อและข้อมูลผลการเรียนในระหว่างที่ศึกษาอยู่ใน หลักสูตรฯ ของนักศึกษาที่เข้าศึกษาหลักสูตรวิทยาการคอมพิวเตอร์ตั้งแต่ปีการศึกษา 2555 ถึงปี การศึกษา 2560 เนื่องจากใช้โครงสร้างหลักสูตรเดียวกัน โดยแบ่งเป็น 2 ส่วน ดังนี้

#### 1.4.4 การศึกษาปัจจัยที่มีผลต่อการสำเร็จการศึกษาตามหลักสูตร โดยวิเคราะห์จาก

- สาขาวิชาที่จบก่อนเข้าศึกษาต่อในหลักสูตรวิทยาการคอมพิวเตอร์
- ข้อมูลการลงทะเบียนเรียน
- ข้อมูลผลการศึกษาแต่ละรายวิชา

#### 1.4.5 การตรวจสอบความเสี่ยงที่จะไม่สำเร็จการศึกษาตามเกณฑ์

ในการวิเคราะห์ข้อมูลได้นำข้อมูล (Raw Data) ที่ได้จากการรวบรวมมาจากแหล่งข้อมูล ต่างๆ มาทำการแปลงให้อยู่ในรูปแบบที่สามารถนำไปวิเคราะห์ได้ (Data Preparation) เมื่อได้ชุด ข้อมูล (Dataset) แล้วนำมาทำการจัดกลุ่ม (Clustering) และนำเสนอข้อมูลที่เข้าใจง่ายด้วยแผนภาพ (Data Visualization)

### 1.2 วัตถุประสงค์ของปริญญานิพนธ์

เพื่อศึกษาปัจจัยต่อไปนีว่ามีผลต่อการสำเร็จการศึกษาของนักศึกษาภาควิชาวิทยาการ คอมพิวเตอร์หรือไม่

- 1) วุฒิการศึกษาที่จบก่อนเข้าศึกษาต่อในหลักสูตรวิทยาการคอมพิวเตอร์
- 2) เกรดเฉลี่ยสะสม (GPA)
- 3) รายวิชาบังคับที่ต้องเรียนก่อน (Pre-Requisite)

- 4) ผลการศึกษาตามรายวิชา
- 5) จำนวนหน่วยกิตสะสมที่ลงทะเบียนเรียนและผ่าน

### 1.3 ขอบเขตปริญญานิพนธ์

- 1.3.1 รวบรวมข้อมูลจากแหล่งข้อมูลสำนักทะเบียนและวัดผล โดยใช้ข้อมูลของนักศึกษาที่เข้าศึกษาในหลักสูตรวิทยาการคอมพิวเตอร์ตั้งแต่ปีการศึกษา 2555 – 2560
- 1.3.2 ศึกษาโครงสร้างหลักสูตรของภาควิชาวิทยาการคอมพิวเตอร์ที่ใช้ตั้งแต่ปีการศึกษา 2555 จนถึงปัจจุบัน
- 1.3.3 เตรียมข้อมูลให้อยู่ในรูปแบบที่สามารถนำไปวิเคราะห์ได้
- 1.3.4 วิเคราะห์ข้อมูล ด้วยอัลกอริทึมการจัดกลุ่มข้อมูล (Clustering) ด้วยปัจจัยที่ต้องการศึกษา ได้แก่
  - 1.3.4.1 สาขาวิชาที่จบ ก่อนเข้าศึกษาต่อในหลักสูตรวิทยาการคอมพิวเตอร์
  - 1.3.4.2 เกรดเฉลี่ยสะสม (GPA)
  - 1.3.4.3 รายวิชาบังคับที่ต้องเรียนก่อน (Pre-Requisite)
  - 1.3.4.4 ผลการศึกษาตามรายวิชา
  - 1.3.4.5 จำนวนหน่วยกิตสะสมที่ลงทะเบียนเรียนและผ่าน
- 1.3.5 นำเสนอผลการวิเคราะห์ด้วยแผนภาพข้อมูล (Data Visualization)

### 1.4 ประโยชน์ที่คาดว่าจะได้รับ

- 1.4.1 ทำให้ทราบถึงปัจจัยที่ส่งผลต่อการสำเร็จการศึกษาของนักศึกษาภาควิชาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยสยาม
- 1.4.2 สามารถหาแนวทางปรับปรุงแก้ไขเพื่อให้นักศึกษาสามารถสำเร็จการศึกษาตามเกณฑ์
- 1.4.3 เพิ่มจำนวนนักศึกษาสำเร็จการศึกษาได้มากขึ้น และลดจำนวนการตกออกของนักศึกษาระหว่างการศึกษา

### 1.5 ขั้นตอนและวิธีการดำเนินงานปริญญานิพนธ์

- 1.5.1 การศึกษาและรวบรวมข้อมูล (Data Collection)

ทำการศึกษาจากข้อมูลที่รวบรวมได้จากการสอบถามและพูดคุยกับนักศึกษาที่ต้องการย้ายคณะหรือลาออกและอาจารย์ที่ปรึกษาชั้นปี จากเอกสารคำร้องที่นักศึกษานำมาเสนอต่ออาจารย์ที่ปรึกษา รวมถึงการประชุมร่วมกันระหว่างคณะผู้จัดทำและอาจารย์ที่ปรึกษา ทำการวิเคราะห์และคาดการณ์ถึงสาเหตุเบื้องต้นที่ทำให้เกิดปัญหาของการย้ายคณะ ลาออก และตกออกของ

นักศึกษา รวมถึงสาเหตุที่นักศึกษาไม่สามารถสำเร็จการศึกษาตามเกณฑ์การศึกษา 4 ปีได้ รวมถึงกำหนดข้อมูลที่เป็นต้องใช้เพื่อการวิเคราะห์ว่าจะต้องประกอบด้วยแอททริบิวต์อะไรบ้าง และจะดำเนินการจัดหามาได้อย่างไร

#### 1.5.2 การเตรียมข้อมูล (Data Preparation)

การนำข้อมูลที่รวบรวมมาได้ นั้น มาทำการ Data Cleaning คือการทำความสะอาดข้อมูล หรือการทำข้อมูลให้สมบูรณ์มากขึ้น เป็นกระบวนการตรวจสอบและแก้ไขรายการข้อมูลที่ไม่ถูกต้องหรือมีความผิดปกติออกไปจากชุดข้อมูล ตารางหรือฐานข้อมูล ซึ่งเป็นหลักสำคัญของฐานข้อมูล เพราะหมายถึงความไม่สมบูรณ์ ความไม่ถูกต้อง ความไม่สัมพันธ์กับข้อมูลอื่น ๆ เป็นต้น จึงต้องมีการแทนที่ การปรับปรุง หรือการลบข้อมูลที่ไม่ถูกต้องเหล่านี้ออกไป เพื่อให้ข้อมูลมีคุณภาพและปรับเปลี่ยนโครงสร้างข้อมูลให้เป็นรูปแบบที่เหมาะสมต่อการนำข้อมูลไปวิเคราะห์

#### 1.5.3 การวิเคราะห์ข้อมูล (Data Analytics)

ขั้นตอนการวิเคราะห์ข้อมูลจะนำข้อมูลที่เตรียมข้อมูลไว้มาทำการจัดกลุ่ม (Clustering) ตามปัจจัยที่ได้คาดการณ์ไว้ โดยแบ่งเป็น 3 ส่วน ดังนี้

- 1.5.3.1 ปัจจัยที่ทำให้นักศึกษาไม่สามารถจบการศึกษาได้ตามหลักสูตร 4 ปี จากข้อมูลนักศึกษาที่จบไปแล้ว
- 1.5.3.2 ปัจจัยที่มีแนวโน้มว่าจะทำให้นักศึกษาไม่สามารถจบการศึกษาได้ตามหลักสูตร 4 ปี จากข้อมูลนักศึกษาที่กำลังศึกษาอยู่ในปัจจุบัน
- 1.5.3.3 ปัจจัยที่ทำให้นักศึกษาตกออก หรือ ย้ายออก ในระหว่างที่กำลังศึกษาอยู่ จากข้อมูลนักศึกษาที่มีทั้งหมด

ทำการวิเคราะห์ข้อมูลเพื่อจัดกลุ่มข้อมูลตามเงื่อนไขที่กำหนดด้วยโปรแกรม

Tableau

#### 1.5.4 การนำเสนอแผนภาพของข้อมูล (Data Visualization)

เป็นการสรุปและแสดงข้อมูลที่วิเคราะห์ได้ในรูปแบบของแผนภาพ กราฟ แผนภูมิ ซึ่งสามารถข้อมูลในเชิงลึก โดยแสดงภาพรวมในเบื้องต้นและสามารถเลื่อนเข้าไปดูข้อมูลรายละเอียดที่ซ่อนอยู่ได้ (Drill Down) เพื่อให้เห็นความหมายหรือนัยสำคัญของข้อมูลได้ง่ายยิ่งขึ้น โดยในการนำเสนอด้วยกราฟหรือแผนภูมิจะเลือกชนิดของกราฟให้เหมาะสมกับสารสนเทศที่ต้องการนำเสนอ ประกอบด้วย Line Chart, Bar Chart, Radar Chart, Area Chart และจัดทำเป็น Dashboard เพื่อให้ผู้อ่าน เข้าใจได้ง่ายขึ้น

### 1.5.5 การจัดทำเอกสารปริญญานิพนธ์ (Documentation)

เป็นการจัดทำเอกสารเพื่อนำเสนอแนวทางในการจัดทำปริญญานิพนธ์ โดยมีวิธีการและขั้นตอนการดำเนินงาน รวมถึงผลสรุปที่ได้จากการวิเคราะห์ข้อมูลที่สามารถใช้เป็นแหล่งอ้างอิงและทำการวิเคราะห์ข้อมูลต่อไปได้

## 1.6 แผนและระยะเวลาในการดำเนินงานปริญญานิพนธ์

ตารางที่ 1.1 ระยะเวลาในการดำเนินงานปริญญานิพนธ์

กิจกรรม	2562					2563					
	ก.ย.	ต.ค.	พ.ย.	ธ.ค.	ม.ค.	ก.พ.	มี.ค.	เม.ย.	พ.ค.	มิ.ย.	ก.ค.
1. การรวบรวมข้อมูล			←	→							
2. การเตรียมข้อมูล					←	→			→		
3. การวิเคราะห์ข้อมูล					←	→			→		
4. การนำเสนอข้อมูลด้วยแผนภาพ					←	→			→		
5. จัดทำเอกสารปริญญานิพนธ์								←	→		→

## 1.7 อุปกรณ์และเครื่องมือที่ใช้ในการวิเคราะห์ข้อมูล

### 1.7.1 ฮาร์ดแวร์

#### 1.7.1.1 เครื่องคอมพิวเตอร์ โน้ตบุ๊ก Acer รุ่น Nitro 5

- Intel Core i5-10300
- NVIDIA GTX 1650 Ti
- 16 GB DDR5 2933 MHz
- 512 GB SSD PCIe M.2

#### 1.7.1.2 เครื่องคอมพิวเตอร์ โน้ตบุ๊ก Asus TUF Gaming รุ่น Ryzen 5

- Ryzen 5 3550H
- NVIDIA GeForce 1650

- 8.00 GB
- 64-bit Operating System, x64 based processor

## 1.7.2 ซอฟต์แวร์

1.7.2.1 ระบบปฏิบัติการ Microsoft Window **10**

1.7.2.2 โปรแกรม Microsoft Excel 2019

1.7.2.3 โปรแกรม Tableau



## บทที่ 2

### การทบทวนวรรณกรรมที่เกี่ยวข้อง

ในการจัดทำวิทยานิพนธ์นี้คณะผู้จัดทำได้ทำการศึกษาค้นคว้าหลักการ แนวคิด ทฤษฎี และเทคโนโลยีที่เกี่ยวข้อง เพื่อนำมาใช้เป็นแหล่งอ้างอิงและประยุกต์ใช้ในการวิเคราะห์ข้อมูลเพื่อหาปัจจัยที่มีผลต่อการสำเร็จการศึกษาของนักศึกษาภาควิชาวิทยาการคอมพิวเตอร์ และการนำเสนอแผนภาพของข้อมูลที่วิเคราะห์ได้ โดยมีประกอบด้วย

#### 2.1 Data Analytics <sup>1</sup>

Data Analytics คือ การวิเคราะห์ข้อมูลที่มีอยู่ในปัจจุบันและอดีต เพื่อทำนายอนาคต โดยมีวัตถุประสงค์เพื่อปรับปรุงการดำเนินงานขององค์กร Data Analytics เป็นเครื่องมือที่สำคัญในปัจจุบันทั้งในภาคธุรกิจ การศึกษา และภาครัฐ โดย Data Analytics แบ่งเป็น 3 แบบ คือ

##### **Descriptive Analytics**

เป็นรูปแบบการใช้ข้อมูลแบบพื้นฐานโดยเน้นการอธิบายว่ากำลังเกิดอะไรขึ้น สามารถอธิบายถึงสาเหตุการเกิดต่าง ๆ Descriptive Analytics ตัวอย่างเช่น รายงานธุรกิจ รายงานด้านการรายงานผลดำเนินงานที่ผ่านมา เป็นข้อมูลพื้นฐานที่แสดงผลในการทำกิจกรรมต่าง ๆ

##### **Predictive Analytics**

เป็นรูปแบบการใช้ข้อมูลที่มีความซับซ้อนมากขึ้น โดยจะเป็นการ “พยากรณ์” หรือ “ทำนาย” สิ่งที่กำลังจะเกิดขึ้น โดยใช้ข้อมูลในอดีตร่วมกับโมเดลทางคณิตศาสตร์ หรือร่วมกับการทำ Data Mining นอกจากนี้ Predictive Analytics ยังทำให้สามารถวิเคราะห์หาโอกาสและความเสี่ยงต่างๆ ที่จะเกิดขึ้นในอนาคตได้ ตัวอย่างเช่นการรู้แนวโน้มทางการตลาด การพยากรณ์ยอดขาย

##### **Prescriptive Analytics**

เป็นรูปแบบการวิเคราะห์ข้อมูลที่มีความซับซ้อนและยากที่สุด ไม่เพียงพยากรณ์หรือทำนายว่าจะเกิดอะไรขึ้น แต่ยังให้คำแนะนำในทางเลือกต่างๆ และผลของแต่ละทางเลือกว่าจะเป็นอย่างไร โมเดลของ Prescriptive Analytics นั้นจะสามารถปรับเปลี่ยนได้ตามข้อมูลที่เพิ่มเติมเข้ามามากขึ้น และ Prescriptive Analytics นี้ยังเป็นการใช้ข้อมูลที่มากที่สุด และเกี่ยวข้องกับ Big Data เป็นอย่างมาก

<sup>1</sup> อ้างอิง <https://work360.in.th/digital-marketing/data-analytics/what-is-data-analytics/>

<sup>1</sup> อ้างอิง <https://katalyst.kasikornbank.com/th/blog/Pages/what-is-big-data.html>

<sup>1</sup> อ้างอิง <https://medium.com/tni-university/machine-learning-คืออะไร-1ebf94a6b431>

<sup>1</sup> อ้างอิง <https://medium.com/datacubator/validation-set-สำคัญไฉน-1abf22a68b75>



ชุดข้อมูล (Dataset) ที่ใช้ในการวิเคราะห์ข้อมูลจะมีขนาดใหญ่และถูกเก็บบันทึกไว้ผ่านวิธีการต่างๆ นั้นสามารถแบ่งเป็นหมวดหมู่ตามโครงสร้างของชุดข้อมูลได้ดังนี้

- ข้อมูลที่มีโครงสร้าง (Structured Data)

หมายถึง ชุดข้อมูลที่มีการจัดเรียงโครงสร้างอย่างเป็นระเบียบ มีความชัดเจน หรือระบุได้ด้วยตัวเลข พร้อมใช้งานได้ทันที เช่น จำนวนการซื้อขายกับลูกค้า เบอร์เซ็นต์ความเคลื่อนไหวภายในตลาดหุ้น เป็นต้น

- ข้อมูลที่ไม่มีโครงสร้าง (Unstructured Data)

หมายถึง ชุดข้อมูลที่มีโครงสร้างไม่ชัดเจน หรือไม่สามรถระบุความแน่นอนของข้อมูลนั้นๆ ได้ ยังไม่สามารถประมวลผลเพื่อนำไปใช้ได้ทันที อย่างเช่น บทความหาโต้ตอบกับลูกค้าทาง Social Media

- ข้อมูลกึ่งมีโครงสร้าง (Semi-Structured Data)

หมายถึง ชุดข้อมูลที่มีโครงสร้างระดับหนึ่งแต่ยังไม่สมบูรณ์ เช่น สถานะใน Social Media เป็นข้อมูลที่ไม่มีโครงสร้าง แต่ในกรณีที่มี Hashtag (#) เข้ามาช่วยในการจัดหมวดหมู่ จะทำให้ข้อมูลมีความเป็นระเบียบขึ้นมา

ในการวิเคราะห์ข้อมูลจะมีการแบ่งข้อมูลออกเป็น 2 ชุดข้อมูล คือ

- Training Set เป็นข้อมูลที่มีใช้ในการสอนโมเดลของเราให้มีความสามารถในการแบ่งแยกและจดจำคุณลักษณะ
- Testing Set เป็นชุดข้อมูลที่ใช้ทดสอบโมเดล หรือใช้ในการตรวจสอบโมเดล

จึงสามารถแบ่งการวิเคราะห์ข้อมูลตามผู้สอน ได้ดังนี้

- การเรียนรู้แบบมีผู้สอน (Supervised Learning) ในข้อมูลที่ใช้ฝึก (Training Data) จะมีมนุษย์มาคอยแยกประเภทหรือบอกผลลัพธ์ (Label) ที่ควรจะเป็นไป จากนั้นจะนำข้อมูลที่ใช้ฝึกแล้วไปผ่านอัลกอริทึมสำหรับสร้างโมเดลที่ไว้ทำนายผลลัพธ์ ประเภทของการเรียนรู้แบบมีผู้สอนจะแยกออกเป็นสองประเภท

1. การแยกประเภท (Classification)

2. การถดถอย (Regression)

- การเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) อัลกอริทึมจะตรวจสอบเฉพาะข้อมูลที่ป้อนเข้ามาเท่านั้น โดยปราศจากการให้ผลลัพธ์ที่จะเกิดขึ้น เช่น การสำรวจข้อมูลประชากรเพื่อหาแบบแผน (Pattern) ของข้อมูลนั้น สามารถใช้เมื่อต้องการที่จะรู้ว่า Machine แบ่งประเภทได้อย่างไร และต้องการให้อัลกอริทึมค้นหาแบบแผนและแบ่งประเภทข้อมูล โดยการเรียนรู้แบบไม่มีผู้สอนแบ่งออกเป็น 2 ประเภท คือ

1. การจับกลุ่มของข้อมูล (Clustering)

2. การลดจำนวนมิติเพื่อบีบอัดข้อมูล (Dimensionality Reduction)

## 2.2 ประเภทของข้อมูลสำหรับการวิเคราะห์ข้อมูล<sup>2</sup>

### 2.2.1 แบ่งตามแหล่งที่มา

- ข้อมูลปฐมภูมิ (Primary Data) คือข้อมูลที่เก็บรวบรวมมาจากแหล่งข้อมูลขั้นต้นหรือได้มาจากแหล่งข้อมูลโดยตรง เช่น ข้อมูลนักเรียนที่ได้มาจากการตอบแบบสอบถาม การสำรวจ การสัมภาษณ์ การวัด การสังเกต การทดลอง เป็นต้น ซึ่งข้อมูลที่ได้จะมีความถูกต้อง ทันสมัย และเป็นปัจจุบัน
- ข้อมูลทุติยภูมิ (Secondary Data) คือข้อมูลที่ได้จากแหล่งที่รวบรวมข้อมูลไว้แล้ว โดยผู้หนึ่งผู้ใด หรือหน่วยงานได้ทำการเก็บรวบรวมหรือเรียบเรียงไว้ ซึ่งข้อมูลสามารถนำมาใช้อ้างอิงได้เลย เช่น ข้อมูลสำมะโนประชากร จากสำนักงานสถิติแห่งชาติ ข้อมูลปริมาณน้ำฝน จากกรมชลประทาน เป็นต้น

### 2.2.2 แบ่งตามลักษณะของข้อมูล

- ข้อมูลเชิงปริมาณ (Quantitative Data) คือข้อมูลที่วัดออกมาเป็นตัวเลข เช่น ผลสัมฤทธิ์ทางการเรียนวิชาต่าง ๆ ที่วัดออกมาเป็นคะแนน คุณลักษณะด้านจิตพิสัย เช่น ความสนใจ คุณลักษณะทางกาย เช่น ส่วนสูง ความเร็วในการวิ่ง เป็นต้น
- ข้อมูลเชิงคุณลักษณะหรือเชิงคุณภาพ (Qualitative Data) คือข้อมูลที่ไม่ได้วัดออกมาเป็นตัวเลขแต่จะแสดงถึงคุณลักษณะของสิ่งนั้น เช่น เพศ ศาสนา สถานภาพสมรส อาชีพ ข้อความที่เป็นความคิดเห็น ผลการสังเกตที่เขียนในรูปบรรยาย

### 2.2.3 แบ่งตามสภาพของข้อมูลที่เกี่ยวข้องกับกลุ่มตัวอย่าง

- ข้อมูลส่วนบุคคล (Personal Data) คือข้อมูลที่เกี่ยวข้องกับข้อเท็จจริงส่วนตัวของกลุ่มตัวอย่าง เช่น ชื่อสกุล อายุ เพศ อาชีพ ศาสนา เป็นต้น
- ข้อมูลสิ่งแวดล้อม (Environmental Data) คือข้อมูลที่เป็นข้อเท็จจริงเกี่ยวกับสิ่งแวดล้อมของกลุ่มตัวอย่าง เช่น ลักษณะท้องถิ่นที่กลุ่มตัวอย่างอาศัยอยู่ เป็นต้น
- ข้อมูลพฤติกรรม (Behavioral Data) คือข้อมูลที่เป็นคุณลักษณะที่มีอยู่ในตัวของกลุ่มตัวอย่าง เช่น คุณลักษณะด้านความสามารถของสมอง ได้แก่ ผลสัมฤทธิ์ทางวิชาการหรือการเรียน เช่น ความรู้ความเข้าใจ ความถนัด และการกระทำสิ่งต่าง ๆ

<sup>2</sup> อ้างอิง <https://riccosmartdata.com/data-type/>

#### 2.2.4 แบ่งตามการนำไปใช้กับคอมพิวเตอร์ได้

- ข้อมูลตัวเลข (Numeric Data) หมายถึง ข้อมูลที่สามารถนำไปคำนวณได้ เช่น จำนวนเงินเดือน ราคาสินค้า ซึ่งอาจอยู่ในรูปของจำนวนเต็ม ทศนิยม เศษส่วน เป็นต้น
- ข้อมูลตัวอักษร (Text Data) หมายถึง ข้อมูลที่ไม่สามารถนำไปคำนวณได้ แต่อาจนำไปเรียงลำดับได้ เช่น ชื่อ ที่อยู่ เบอร์โทรศัพท์ เลขประจำตัวประชาชน หมายเลขโทรศัพท์ เป็นต้น
- ข้อมูลเสียง (Audio Data) หมายถึง ข้อมูลที่เกิดจากการได้ยิน เช่น เสียงคนพูด เสียงสัตว์ร้อง เสียงจากปรากฏการณ์ธรรมชาติ หรือเสียงจากอุปกรณ์ต่าง ๆ ที่สามารถแสดงผลข้อมูลในรูปแบบเสียงได้ เช่น แผ่นซีดี โทรศัพท์ วิทยุ เป็นต้น
- ข้อมูลภาพ (Images Data) หมายถึง ข้อมูลที่เป็นภาพในลักษณะต่าง ๆ ที่มองเห็น ซึ่งอาจเป็นภาพนิ่ง เช่น ภาพวาด ภาพถ่าย เป็นต้น หรือภาพเคลื่อนไหว เช่น ภาพจากโทรทัศน์ ภาพจากวีดิทัศน์ ภาพจากคอมพิวเตอร์ เป็นต้น
- ข้อมูลภาพเคลื่อนไหว (Video Data) หมายถึง ข้อมูลที่เป็นภาพเคลื่อนไหว หรือข้อมูลที่มีทั้งภาพ เสียง ข้อความปนกัน เช่น ภาพเคลื่อนไหวที่ถ่ายด้วยกล้องวิดีโอ หรือภาพที่ทำจากโปรแกรมต่าง ๆ เป็นต้น

### 2.3 Clustering<sup>3</sup>

Clustering Model คือ Machine Learning Model ประเภท Unsupervised ที่ไม่มี Target หรือ ไม่มีต้นแบบของผลลัพธ์ ซึ่งเป็น Model ที่นำไปใช้ในการจัดกลุ่มของข้อมูลที่ไม่เคยมีการจัดกลุ่มมาก่อน โดยจะแบ่งกลุ่มข้อมูลจากความคล้าย เช่น การจัดกลุ่มลูกค้าจากพฤติกรรมการซื้อสินค้าของลูกค้าที่มีลักษณะคล้ายกันจะเป็นกลุ่มลูกค้าประเภทเดียวกัน

Algorithms ที่ใช้ในการทำ Clustering Model ได้แก่

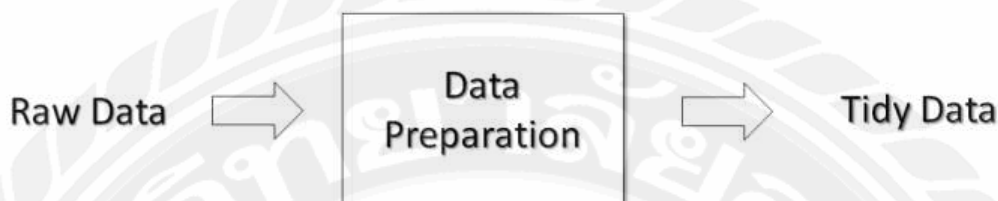
- 2.3.1 K-Mean Clustering เป็นการจัดกลุ่มโดยต้องมีการกำหนดจำนวนกลุ่มที่ต้องการจัดกลุ่มข้อมูลก่อน เป็นการวิเคราะห์แบบไม่เป็นขั้นตอน

<sup>3</sup> อ้างอิง [https://medium.com/tni-university/การทำ-machine-learning-ด้วย-clustering-model-](https://medium.com/tni-university/การทำ-machine-learning-ด้วย-clustering-model-2a3c392e7faa)

2.3.2 Hierarchical Clustering เป็นการจัดกลุ่มโดยไม่ต้องมีการกำหนดจำนวนกลุ่มที่ต้องการจัดกลุ่มข้อมูลก่อน เป็นการวิเคราะห์แบบเป็นขั้นตอน

## 2.4 Data Preparation <sup>4</sup>

Data Preparation หรือ การเตรียมข้อมูล หมายถึงกระบวนการใดๆ ที่จำเป็นต้องทำกับข้อมูลดิบ (Raw Data) ที่ได้รับมา เพื่อปรับเปลี่ยนข้อมูลให้อยู่ในรูปแบบที่เหมาะสม ที่จะนำไปโหลดเข้าฐานข้อมูล หรือนำไปวิเคราะห์ต่อไป



รูปที่ 2.1 กระบวนการเตรียมข้อมูล (Data Preparation)

การเตรียมข้อมูลที่มีประสิทธิภาพ ควรมีลักษณะสำคัญดังนี้

- ให้ผลลัพธ์ที่ครบถ้วนสมบูรณ์
- ให้ความสำคัญกับนิยามข้อมูล
- จัดบันทึกขั้นตอนการเตรียมข้อมูลโดยละเอียด
- ปรับกระบวนการให้เป็นอัตโนมัติให้มากที่สุด

### 2.4.1 วิธีการจัดการข้อมูลสูญหาย (Missing Data Management)

- วิธีการลบข้อมูล (Listwise Deletion or Complete Case Analysis) กรณีที่ข้อมูลสูญหายเกิดขึ้นหลายตัวแปร แต่ปริมาณการสูญหายต้องไม่เกิน 5% ของข้อมูลทั้งหมด ก็จะตัดหรือลบข้อมูลส่วนที่สูญหายออกทั้ง record
- วิธีการประมาณค่าข้อมูลสูญหาย (Imputation Methods) เป็นวิธีการประมาณค่าสูญหายโดยเอาหลักการทางคณิตศาสตร์ มาเติมเต็มค่าที่สูญหายไป ซึ่งมีหลากหลายวิธีมากเช่น ประมาณค่าด้วยค่าเฉลี่ย (Mean Imputation: MI) , วิธีการถดถอย (Regression Imputation: RI) เป็นต้น

<sup>4</sup> อ้างอิง <https://bzinsight.wordpress.com/2014/06/11/การทำ-data-preparation-อย่างมืออาชีพ/>

<sup>4</sup> อ้างอิง <https://www.roots.tech/blog/products-services-1/how-to-manage-missing-data-25>

<sup>4</sup> อ้างอิง <https://ichi.pro/th/kar-cadkar-kab-outliers-ni-chud-khxmml-60750975160994>

- วิธีการขั้นสูง (Advanced Methods) เป็นลักษณะของการประมาณค่าด้วยวิธีการคำนวณที่มีความซับซ้อนมากขึ้น ตัวอย่างเช่น
  - 1) วิธีการประมาณค่าด้วยวิธีเอ็มไอ (Multiple Imputation)
  - 2) วิธีการประมาณค่าโดยวิธีอีเอ็ม (Expectation Maximization Algorithm: EM Algorithm)
  - 3) วิธีการประมาณค่าโดยวิธีเพื่อนบ้านใกล้เคียง (K-Nearest Neighbor: KNN)

#### 2.4.2 วิธีการจัดการกับ Outliers

- 1) วิธีการเบี่ยงเบนมาตรฐาน
- 2) วิธีช่วงระหว่างควอไทล์ (IQR)
- 3) การตรวจจับค่าผิดปกติโดยอัตโนมัติ

### 2.5 Data Visualization <sup>5</sup>

Data Visualization คือ การสรุปและแสดงข้อมูลออกมาให้อยู่ในรูปของแผนภาพ กราฟ หรือวิดีโอที่อธิบายเกี่ยวกับข้อมูลที่คุณต้องการจะแสดง ซึ่ง Data Visualization นั้นมีมานานมากแล้ว แต่ได้หิบนำมาใช้กันมากขึ้นในช่วงหลังมานี้ เนื่องจากการทำงานในปัจจุบัน ข้อมูลที่มีนั้นไม่ได้เป็นเพียงข้อความอย่างเดียว แต่มีข้อมูลที่หลากหลายมากขึ้น ดังนั้น จึงต้องมีการรองรับข้อมูลประมวลผล วิเคราะห์ข้อมูลออกมาให้เห็นเป็นภาพที่ชัดเจนและเข้าใจง่ายกว่าเดิม

ซึ่งกระบวนการทำงาน Data Visualization แบบเข้าใจง่ายก็คือการนำข้อมูลดิบ มาวิเคราะห์ และแสดงผลข้อมูลออกมาในรูปแบบของกราฟ แผนภูมิ ที่ช่วยให้คุณ ได้ข้อมูลเชิงลึกจากข้อมูลดิบเหล่านั้น ทำให้คุณเห็นคุณค่าของข้อมูล พร้อมทั้งแนวโน้มการตลาดในอนาคตจากข้อมูลเหล่านี้จริง ๆ คุณอาจไม่เคยรู้เลยว่าข้อมูลของคุณนั้นมีอะไรอยู่บ้าง สามารถนำข้อมูลออกมาใช้ได้มากน้อยแค่ไหน แต่การใช้ Data Visualization จะทำให้คุณรู้ว่าข้อมูลในที่ที่น่าสนใจ คุณอาจจะเจอข้อมูลที่ทำให้คุณเจอแนวทางใหม่ ๆ ในการทำการตลาดได้ด้วย

Data Visualization ก็เหมือนกับการเล่าเรื่องด้วยรูปภาพ ซึ่งแน่นอนว่าภาพที่อยากนำไปถ่ายทอดให้กับคนอ่านที่มีหลากหลายรูปแบบ แต่ยังคงคำนึงถึงจุดเด่นของภาพ หรือจุดที่ควรโฟกัสว่าข้อมูลไหน คือส่วนที่อยากสื่อสารไปยังผู้อ่านมากที่สุด ดังนั้น การใช้รูปแบบภาพในรูปทรงต่าง ๆ และเพิ่มจุดเด่นลงไป จะเป็นการช่วยให้เห็นภาพได้ชัดขึ้น และเกิดความเข้าใจได้ง่ายขึ้น

<sup>5</sup> อ้างอิง <https://blog.ourgreenfish.com/get-to-know-data-visualization-quick-and-easy-to-understand>

<sup>3</sup> อ้างอิง <https://stepstraining.co/trendy/data-visualization-trends-in-2021>

### 2.5.1 รูปแบบของข้อมูลที่จะแสดงถึงขนาดที่ต่างกัน หรือ จัดลำดับ<sup>6</sup>

แผนภูมิแท่ง (Bar Chart) เป็น แผนภาพที่ง่ายที่สุด ที่สามารถใช้เปรียบเทียบข้อมูลในแต่ละหมวดหมู่ สามารถแสดงถึงลำดับและขนาดได้ชัดเจน

แผนภูมิรูปภาพ (Pictogram) ในบางกรณี แผนภูมิรูปภาพอาจจะเป็นทางเลือกที่สวยงามและเข้าใจง่าย ควรใช้สำหรับกรณีที่น่าเสนอเป็นจำนวนเต็มที่ไม่เยอะมาก

### 2.5.2 รูปแบบของข้อมูลที่ต้องการเห็นความสัมพันธ์

แผนภาพการกระจาย (Scatter Plot) เป็นวิธีที่ดีที่สุดใน การแสดงความสัมพันธ์ระหว่างตัวแปรประเภทตัวเลขสองตัวแปร สามารถใส่เส้นประกอบเพื่อแสดงให้เห็นถึงแนวโน้มได้

แผนภูมิฟอง (Bubble chart) หากมีตัวแปรที่สามที่ต้องการแสดงเป็นมิติเพิ่มเติมจากแผนภาพการกระจายทั่วไป อาจนำตัวแปรนั้นมาแสดงโดยใช้นาฬิกาหรือพื้นที่ของวงกลมได้

ตารางแผนภาพความร้อน (Heatmap) หากเป็นข้อมูลหมวดหมู่ แต่มีลำดับชั้น (Ordinal) และอยากแสดงถึงความสัมพันธ์ของตัวแปรเหล่านี้กับปริมาณที่สนใจ อาจจะใช้ตารางแผนภาพความร้อนหรือ หลายคนรู้จักกันในชื่อ Heatmap โดยใช้สีแสดงถึงปริมาณที่สนใจ

### 2.5.3 รูปแบบของข้อมูลที่ต้องการเห็นการกระจายตัว

ฮิสโตแกรม (Histogram) ใช้แสดงการกระจายตัวได้ดีและละเอียด หากต้องการแสดงจำนวนของตัวแปรประเภทตัวเลข

Boxplot ใช้แสดงการกระจายตัวของข้อมูลเช่นเดียวกับฮิสโตแกรมแต่จะมีความกระชับกว่า และสามารถใช้ในการเปรียบเทียบของแต่ละหมวดหมู่ได้

### 2.5.4 รูปแบบของข้อมูลที่ต้องการเห็นการเปลี่ยนแปลงตามช่วงเวลา

แผนภูมิเส้น/แผนภูมิแท่ง จะแสดงให้เห็นถึงการเชื่อมโยงของแต่ละช่วงเวลาได้ดีกว่ากราฟแท่ง แต่โดยทั่วไปสามารถใช้ได้ทั้งสองอย่าง หรือ อาจจะผสมกันถ้าต้องการแสดงสองปริมาณพร้อมกันและต้องการให้เห็นความแตกต่าง

แผนภูมิพื้นที่ (Area Chart) หากต้องการแสดงให้เห็นถึงองค์ประกอบในแต่ละช่วงเวลาประกอบกับการเปลี่ยนแปลงตามช่วงเวลาอาจเลือกใช้แผนภูมิพื้นที่นำเสนอข้อมูลชุดนั้น

<sup>6</sup> อ้างอิง <https://bigdata.go.th/big-data-101/picking-chart-for-data-visualization/>

### 2.5.5 รูปแบบของข้อมูลที่ต้องการแสดงถึงองค์ประกอบ

แผนภูมิวงกลม (Pie Chart) จะใช้กับตัวแปรประเภทหมวดหมู่ว่าเพื่อแสดงให้เห็นว่าแต่ละหมวดหมู่มีจำนวนหรือปริมาณที่เกี่ยวข้องเท่าใด โดยใช้ขนาดมุมหรือพื้นที่เป็นตัวบ่งชี้ ทั้งนี้ไม่ควรใช้ในกรณีที่มีหมวดหมู่มากเกินไป ปริมาณที่ใช้ควรรวมเป็น 100% ไม่ควรใช้ในกรณีที่ปริมาณนั้นไม่สามารถนำมาบวกกันได้ เช่น ค่าเฉลี่ยของยอดขายในแต่ละประเภทผลิตภัณฑ์

แผนภูมิโดนัท (Donut Chart) เหมือน Pie Chart แต่มีพื้นที่ตรงกลางไว้ที่สามารถใช้แสดงข้อความ หรือ ปริมาณรวมได้

แผนที่ต้นไม้ (Tree map) คล้าย Pie Chart แต่สามารถใช้แบ่งหมวดหมู่เป็นลำดับชั้นได้ และหากมีหลายหมวดหมู่ก็จะดูรู้เรื่องกว่า โดยจะใช้พื้นที่ที่สัมพันธ์ในการแสดงถึงปริมาณ และสื่ออาจจะใช้ในการแบ่งหมวดหมู่หรือจัดลำดับได้

### 2.5.6 รูปแบบของข้อมูลที่ต้องการแสดงถึงข้อมูลเชิงพื้นที่

แผนที่โคโรเพลท (Choropleth Map) คือการใช้สีในการแสดงปริมาณต่าง ๆ บนพื้นที่นั้น มีข้อควรระวังคือหากพื้นที่มีขนาดใหญ่ บริเวณที่มีสีก็จะใหญ่ตามไปด้วยซึ่งอาจจะทำให้ผู้อ่านเข้าใจผิด ดังนั้นปริมาณที่นำมาใช้แสดงโดยสีใน Choropleth map ควรจะเป็นปริมาณที่ไม่ขึ้นกับขนาดพื้นที่ เช่น ความหนาแน่นประชากรต่อพื้นที่ มากกว่าปริมาณที่อาจจะขึ้นกับพื้นที่ เช่น จำนวนประชากรรวม

แผนที่แสดงด้วยสัญลักษณ์สัดส่วน (Proportional Symbol Map) ใช้ขนาดของสัญลักษณ์ เช่น วงกลม แสดงถึงปริมาณ การใช้แผนภาพลักษณะนี้จะทำให้แสดงข้อมูลได้โดยไม่ต้องกังวลถึงขนาดของพื้นที่ ตัวอย่าง เช่น จำนวนผู้ติดเชื้อในแต่ละประเทศมักจะใช้แผนภาพแบบนี้ แทนที่จะเป็น Choropleth Map

แผนที่แสดงข้อมูลด้วยจุด (Dot Map) / แผนที่ความร้อน (Heatmap) หากข้อมูลมีความละเอียดถึงขั้นละติจูด ลองจิจูด อาจแสดงข้อมูลเป็นจุดตามแผนที่ แต่หากข้อมูลมีจำนวนมากสามารถใช้ Heatmap ในการแสดงถึงความหนาแน่นของข้อมูลบนพื้นที่แทน

## 2.6 Tableau <sup>7</sup>

Tableau เป็นซอฟต์แวร์ที่สามารถนำข้อมูลจำนวนมากที่มีหลากหลายในองค์กรมาทำการวิเคราะห์เพื่อเป็นข้อมูลเชิงธุรกิจช่วยให้ผู้บริหารมีข้อมูลเชิงลึกเพิ่มความสามารถในการตัดสินใจอย่างรวดเร็วและชาญฉลาด ในรูปแบบของ Data Visualization ซึ่งเป็นการใช้ภาพเพื่อแสดงข้อมูลในเชิงปริมาณที่วัดได้ ไม่ว่าจะเป็นตัวเลข แผนภูมิ กราฟ และอื่น ๆ อีกมากมาย คำว่า Data คือ ข้อมูล ส่วน Visualization คือ การมองเห็น เมื่อนามารวมกันแล้วหมายถึง ข้อมูลที่มองเห็นได้ด้วยตา

### - Connect to Your Data

ก่อนที่คุณจะสร้างมุมมองและวิเคราะห์ข้อมูลของคุณคุณต้องเชื่อมต่อ Tableau กับข้อมูลของคุณก่อน Tableau รองรับการเชื่อมต่อกับข้อมูลที่หลากหลาย จัดเก็บในที่ต่าง ๆ ตัวอย่างเช่นข้อมูลของคุณอาจถูกเก็บไว้ในคอมพิวเตอร์ของคุณในสเปรดชีตหรือไฟล์ข้อความหรือในฐานข้อมูลขนาดใหญ่เชิงสัมพันธ์หรือคิวบ์ (หลายมิติ) บนเซิร์ฟเวอร์ในองค์กรของคุณ หรือคุณอาจเชื่อมต่อกับข้อมูลสาธารณะสมบัติที่มีอยู่บนเว็บเช่นข้อมูลของ US Census Bureau หรือไปยังแหล่งฐานข้อมูลระบบคลาวด์เช่น Google Analytics, Amazon Redshift หรือ Salesforce

### - Modern Analytics Workflow

เวิร์กโฟลว์การวิเคราะห์สมัยใหม่ช่วยให้ผู้คนสามารถถามและตอบคำถามของตนเองด้วยข้อมูลช่วยให้ผู้คนสามารถแบ่งปันข้อมูลเชิงลึกกับทีมและส่วนอื่น ๆ ในองค์กรได้อย่างง่ายดายและมั่นใจได้ว่าข้อมูลที่สนับสนุนการตัดสินใจของพวกเขาจำเป็นต้องเชื่อถือได้และปลอดภัย เวิร์กโฟลว์ประกอบด้วยขั้นตอนการดำเนินการหลัก 5 ประการ ได้แก่ การเข้าถึงและดูได้ตอบวิเคราะห์และค้นพบแบ่งปันและส่งเสริมและกำกับดูแล



รูปที่ 2.2 เวิร์กโฟลว์ของ Tableau

<sup>7</sup> อ้างอิง <https://help.tableau.com/current/pro/desktop/en-us/basicconnectoverview.htm>

<sup>7</sup> อ้างอิง [https://help.tableau.com/current/blueprint/en-us/bp\\_modern\\_analytics\\_workflow.htm](https://help.tableau.com/current/blueprint/en-us/bp_modern_analytics_workflow.htm)



## บทที่ 3

### การวิเคราะห์ข้อมูล

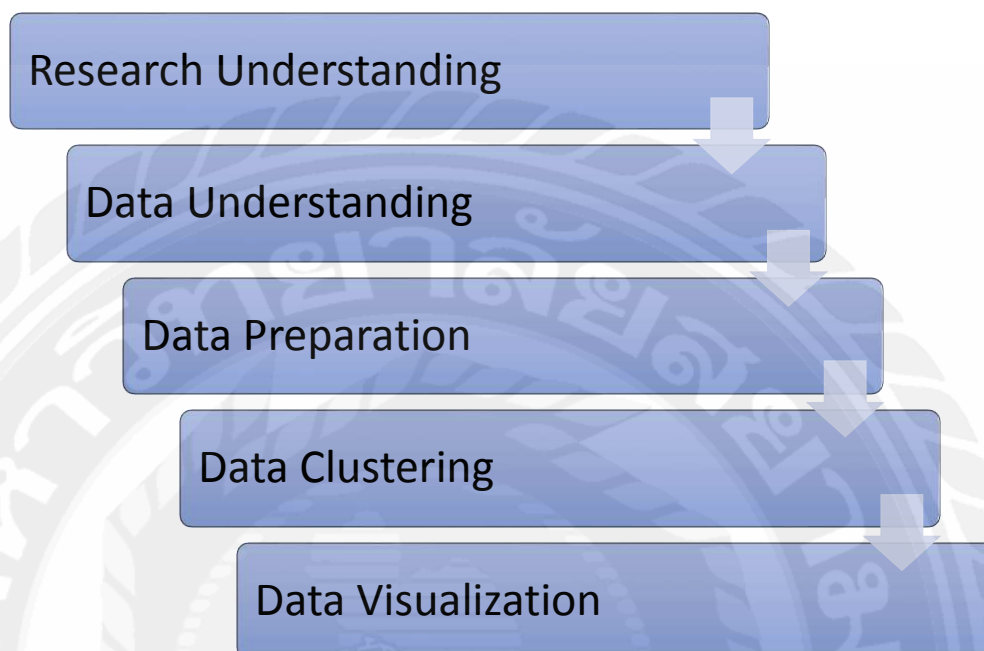
#### 3.1 รายละเอียดของปริญญาโท

ในการวิเคราะห์ปัจจัยที่ส่งผลต่อการสำเร็จการศึกษาของนักศึกษาภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยสยาม โดยนักศึกษาที่เข้าศึกษาต่อในหลักสูตรมีอัตรา การย้ายสาขาในช่วงขึ้นชั้นปีที่ 2 มีอัตราสูงถึง 50% ของนักศึกษาที่เข้าศึกษาต่อในปีการศึกษานั้นๆ อีกทั้งอัตรานักศึกษาที่สำเร็จการศึกษาตามเกณฑ์ 4 ปี มีอัตราต่ำ ทางหลักสูตรจึงต้องการหาสาเหตุ หรือปัจจัยที่ทำให้ นักศึกษามีการย้ายสาขา หรือจบการศึกษาล่าช้า เพื่อหาแนวทางในการแก้ไขและ ป้องกันต่อไป ซึ่งคณะผู้จัดทำได้ทำการรวบรวมข้อมูลจากหลายๆ แหล่งข้อมูล ทั้งจากนักศึกษา อาจารย์ประจำหลักสูตร เอกสารคำร้อง คู่มือนักศึกษา โครงสร้างหลักสูตร และระบบของสำนัก ทะเบียนและวัดผลฯ และใช้เทคนิคทั้งการสัมภาษณ์ การประชุม และรวบรวมข้อมูลจากเอกสารที่เป็น กระดาษและดิจิทัลไฟล์ ได้แก่ ทำการสัมภาษณ์นักศึกษาที่ย้ายสาขาหรือต้อออก ซึ่งนักศึกษาได้ให้ เหตุผลในการย้ายออกหรือต้อออก ประกอบด้วย เนื้อหาวิชายากเกินไปเรียนไม่ไหว ยังไม่ใช้สิ่ง ตนเองชอบ หรืออยากไปเรียนกับเพื่อน เป็นต้น รวมถึงมีการประชุมร่วมกับอาจารย์ประจำ หลักสูตรเพื่อรวบรวมข้อมูลส่วนตัวจากอาจารย์ที่ปรึกษาชั้นปี และในการหาปัจจัยที่มีผลต่อการ สำเร็จการศึกษาของนักศึกษานั้น ได้ขอข้อมูลจากระบบของสำนักทะเบียนและวัดผล ประกอบด้วย ข้อมูลวุฒิการศึกษาที่จบก่อนเข้าศึกษาต่อในหลักสูตร ข้อมูลเกรดเฉลี่ยในแต่ละภาคการศึกษา (GPA) ข้อมูลเกรดเฉลี่ยสะสมตลอดการศึกษ (GPA) ข้อมูลรายวิชาบังคับก่อนเรียนของแต่ละ รายวิชา (Prerequisite) ข้อมูลผลการศึกษตามรายวิชา และข้อมูลจำนวนหน่วยกิตที่ลงทะเบียนใน แต่ละภาคการศึกษา

หลังจากที่ได้รวบรวมข้อมูลจากแหล่งข้อมูลต่างๆ มาแล้ว จะทำการตรวจสอบคุณภาพของ ข้อมูลว่าพร้อมที่จะนำไปทำการวิเคราะห์ต่อไปหรือไม่ โดยได้ทำการหาค่าที่ขาดหายไป (Missing Value) ค่าที่ผิดปกติ (Outlier) ข้อมูลที่ซ้ำซ้อน (Duplicate) ถ้าพบข้อผิดพลาดดังกล่าวจะทำการ ปรับปรุงแก้ไขข้อมูลให้มีความถูกต้องครบถ้วนให้ได้มากที่สุด เพื่อให้การวิเคราะห์มีความแม่นยำ สูง ในการจัดกลุ่มข้อมูล (Clustering) และนำเสนอแผนภาพของข้อมูล (Data Visualization) นั้นใช้ โปรแกรม Tableau ในการจัดกลุ่มข้อมูลและนำเสนอแผนภาพของข้อมูล โดยจัดทำในรูปแบบของ Dashboard เพื่อให้ง่ายต่อการดูสารสนเทศ

### 3.2 ขั้นตอนในการวิเคราะห์ข้อมูล

ขั้นตอนในการวิเคราะห์ปัจจัยที่มีผลการสำเร็จการศึกษาของนักศึกษาภาควิชาวิทยาการคอมพิวเตอร์ มีดังนี้



รูปที่ 3.1 ขั้นตอนการวิเคราะห์ปัจจัยที่มีผลต่อการสำเร็จการศึกษาของนักศึกษาภาควิชาวิทยาการคอมพิวเตอร์

#### 3.2.1 กำหนดวัตถุประสงค์ในการวิเคราะห์ข้อมูล (Research Understanding)

ในเบื้องต้นได้ทำการคาดการณ์ถึงปัจจัยที่ส่งผลต่อการสำเร็จการศึกษาของนักศึกษาภาควิชาวิทยาการคอมพิวเตอร์ ไว้ด้วยกัน 7 ปัจจัย ได้แก่

- 1) การศึกษาที่จบก่อนเข้าศึกษาต่อในหลักสูตร
- 2) เกรดเฉลี่ยสะสม (GPA)
- 3) รายวิชาบังคับที่จะต้องเรียนก่อน (Prerequisite)
- 4) ผลการศึกษาของแต่ละรายวิชาที่ลงทะเบียนเรียน
- 5) จำนวนหน่วยกิตสะสมที่ลงทะเบียนเรียนและผ่าน
- 6) การเปิดสอนรายวิชาของภาควิชาฯ โดยเปิดสอนปีการศึกษาละ 1 ครั้ง
- 7) รายวิชาโครงการคอมพิวเตอร์

### 3.2.2 รวบรวมข้อมูล (Data Understanding)

ในการรวบรวมข้อมูลเพื่อนำมาใช้ในการวิเคราะห์ นั้นทางคณะผู้จัดทำได้ทำการรวบรวมจากแหล่งข้อมูลต่างๆ ดังนี้

- คู่มือนักศึกษาปีการศึกษา เพื่อศึกษาโครงสร้างหลักสูตรวิทยาการคอมพิวเตอร์ ถึงโครงสร้างรายวิชาที่นักศึกษาต้องเรียนตลอดการศึกษา 4 ปี รวมถึงเกณฑ์ในการจบการศึกษา
- เอกสารคำร้องขอย้ายสาขาและลาออก ของนักศึกษาภาควิชาวิทยาการคอมพิวเตอร์ เพื่อรวบรวมเหตุผลในการขอย้ายสาขาหรือลาออก
- สัมภาษณ์อาจารย์ที่ปรึกษาชั้นปี เพื่อรวบรวมข้อมูลส่วนตัวและการเรียนของนักศึกษา
- ข้อมูลจากระบบของสำนักทะเบียนและวัดผล โดยได้ทำการร้องขอข้อมูลพฤติกรรมการศึกษาที่จบก่อนเข้าศึกษาต่อในหลักสูตร ข้อมูลเกรดเฉลี่ยในแต่ละภาคการศึกษา (GPS) ข้อมูลเกรดเฉลี่ยสะสมตลอดการศึกษ (GPA) ข้อมูลรายวิชาบังคับก่อนเรียนของแต่ละรายวิชา (Prerequisite) ข้อมูลผลการศึกษารายรายวิชา และข้อมูลจำนวนหน่วยกิตที่ลงทะเบียนในแต่ละภาคการศึกษา ของนักศึกษาที่เข้าศึกษาต่อในหลักสูตรตั้งแต่ปีการศึกษา 2555 ถึง 2560 โดยข้อมูลอยู่ในรูปแบบของไฟล์ Excel

### 3.2.3 เตรียมข้อมูล (Data Preparation)

หลังจากที่ได้ข้อมูลจากระบบของสำนักทะเบียนและวัดผลซึ่งอยู่ในรูปแบบของไฟล์ Excel แล้วนั้น ทางคณะผู้จัดทำได้ทำการทำความสะอาดข้อมูล (Data Cleaning) โดยการหาค่าที่ขาดหายไป (Missing Value) และค่าที่มีความผิดปกติ (Outlier) ถ้าพบข้อมูลที่มีความผิดปกติในลักษณะดังกล่าวจะทำการพิจารณาว่าจะคงเก็บรายการข้อมูล (Record) นั้นไว้หรือลบทิ้ง โดยถ้าเป็นลักษณะของรายการซ้ำ (Duplicate) จะทำการลบรายการนั้น แต่ถ้าค่าที่ผิดปกติเป็นแอททริบิวต์ที่ไม่มีผลต่อการวิเคราะห์จะยังคงเก็บรายการข้อมูลนั้นไว้ เมื่อทำ Data Cleaning แล้วจะทำการแปลงข้อมูล (Transformation) ให้อยู่ในรูปแบบที่เหมาะสม สามารถนำไปทำการวิเคราะห์ได้ โดยมีรายละเอียดในการเตรียมข้อมูลดังนี้







### 3.2.4 จัดกลุ่มข้อมูล (Data Clustering)

นำข้อมูลที่ได้จัดเตรียมไว้มาทำการจัดกลุ่มข้อมูล (Data Clustering) โดยนำเข้าโปรแกรม Tableau และมีการใช้ฟังก์ชันในการคำนวณ ดังนี้

- COUNT นับจำนวนรายการทั้งหมด
- COUNTD (Distinct Count) นับจำนวนรายการทั้งหมด แต่ไม่นับรายการที่มีความเหมือนกัน

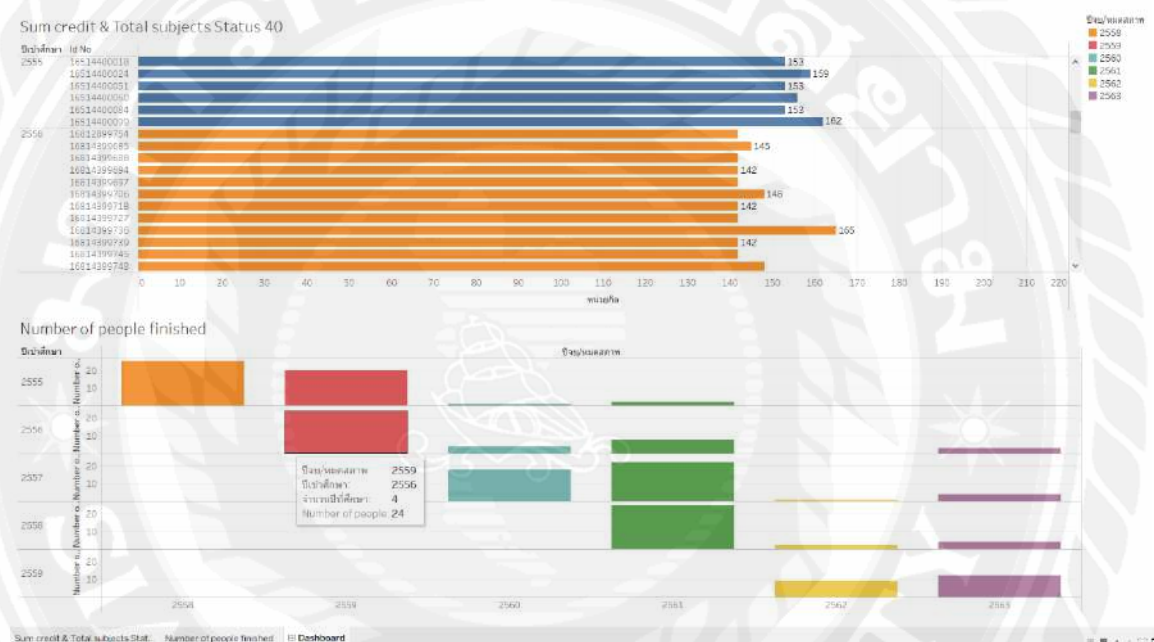
ตัวแปรที่เพิ่มเข้ามาโดยใช้ Function ในการคำนวณของ Tableau

- Group Grade คือ นับจำนวนเกรดของในแต่ละวิชา โดยใช้ COUNT([เกรด])
- Number of People คือ นับจำนวนนักศึกษา โดยใช้ COUNTD ([id No])
- Total Subject คือ นับจำนวนที่ลงทะเบียนทั้งหมดกี่ครั้ง โดยใช้ COUNT ([id No])
- วิชาที่ลงทะเบียนต่อภาค คือ นับจำนวนวิชาที่ลงทะเบียนในภาคการศึกษานั้น โดยใช้ COUNT([ภาค])
- จำนวนปีที่ศึกษา คือ เป็นการคำนวณที่บอกว่าคุณศึกษารับการศึกษาภายในกี่ปี โดยใช้  $STR(INT([ปีจบ/หมดสภาพ]) - INT([ปีเข้าศึกษา]) + 1)$  เพราะว่าข้อมูลที่สำนักทะเบียน บันทึกมาเป็น String ไม่ใช่ date

การวิเคราะห์ข้อมูลดังต่อไปนี้จะเป็นการใช้ข้อมูลของนักศึกษาที่มีสถานภาพนักศึกษาเป็น 40 (ตารางข้อมูลของนักศึกษาที่สำเร็จการศึกษาแล้ว) เชื่อมกับ ตารางสถานภาพ (ตารางความหมายรหัสสถานภาพ)

### WorkBook Name: Status40\_Number of people finished

Dashboard นี้ เป็นเป็นการแสดงข้อมูลเพื่อที่จะแสดงว่าจำนวนนักศึกษาที่เข้าศึกษาตั้งแต่ 2555-2559 ทั้งหมด ได้สำเร็จการศึกษาในปีใดและมีจำนวนลงทะเบียนหน่วยกิตสะสมทั้งหมดเท่าใด



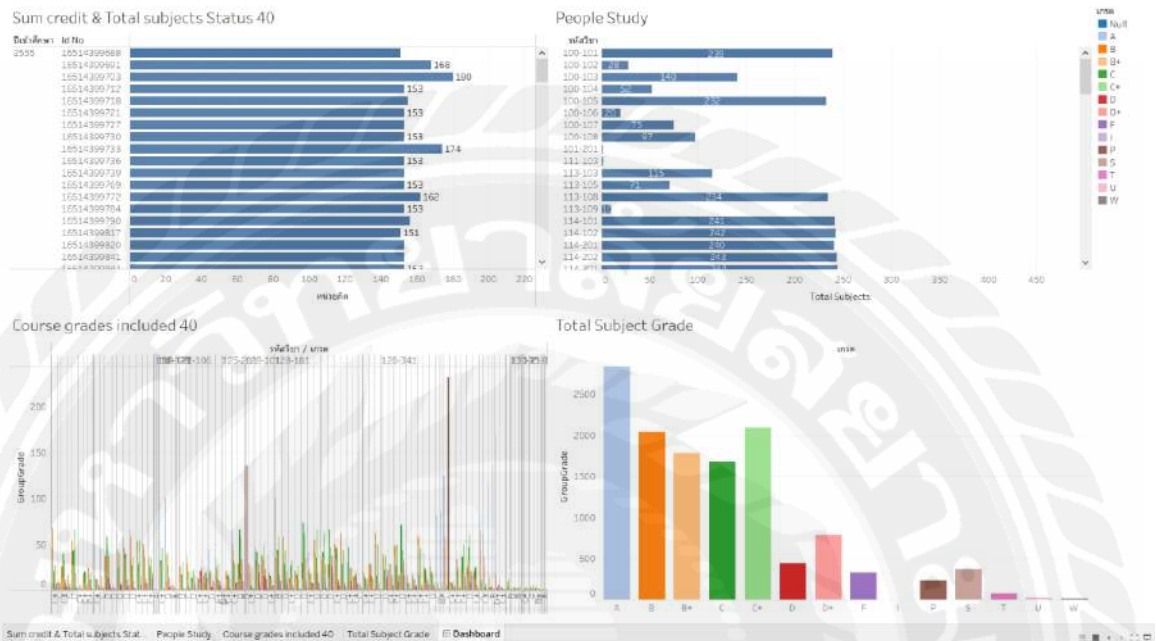
รูปที่ 3.7 WorkBook Name: Status40\_Number of people finished

- Sum Credit & Total Subject Status 40 เป็นการแสดงถึงข้อมูลของนักศึกษาแต่ละคนที่สำเร็จการศึกษาได้ลงทะเบียนหน่วยกิตสะสมทั้งหมดตลอดการศึกษาเป็นจำนวนเท่าไร (จะเป็น Main หลักสำหรับทุก workbook เนื่องจากแสดงได้เป็นรายบุคคล)
- Number of people finished นับจำนวนนักศึกษาที่จบในแต่ละปี โดยสามารถจำแนกได้ว่า นักศึกษาที่เข้าศึกษาในแต่ละปีการศึกษา สำเร็จการศึกษาปีการศึกษาใด เป็นจำนวนเท่าไร



### WorkBook Name: Status40\_People Study and Course grades included

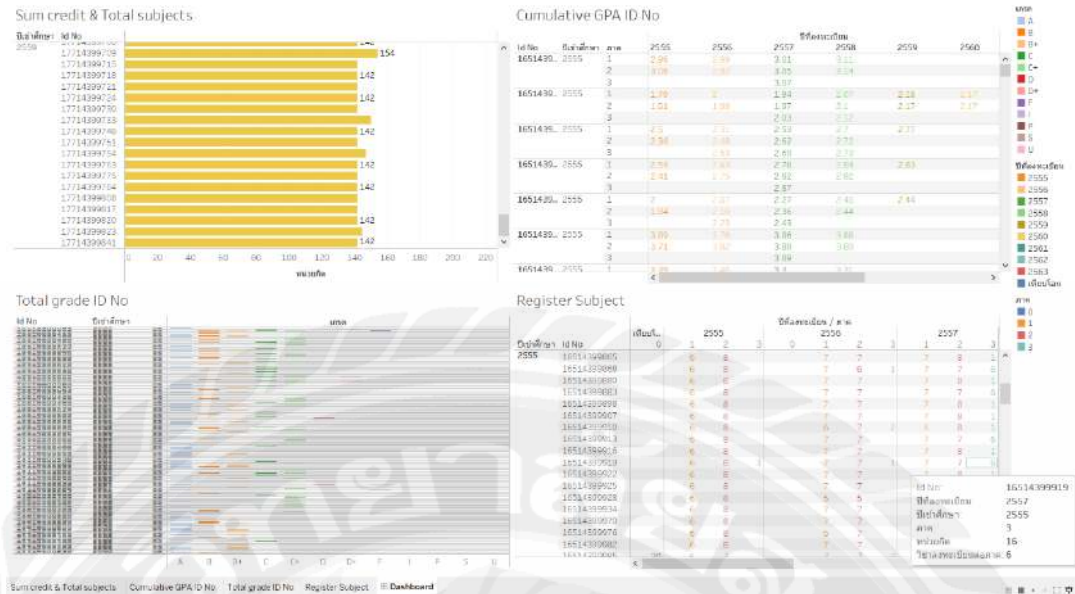
Dashboard นี้ เป็นการแสดงข้อมูลรายละเอียดของเกรดรวมทั้งหมด จำนวนครั้งที่ลงทะเบียนแต่ละรายวิชา (นักศึกษาที่สำเร็จการศึกษาแล้ว)



### รูปที่ 3.8 WorkBook Name: Status40\_People Study and Course grades included

- Total Subject Grade แสดงกราฟแท่งของการนับจำนวนเกรดทั้งหมด
- People Study แสดงกราฟแท่งของในแต่ละวิชาที่มีการลงทะเบียนทั้งหมดกี่ครั้ง
- Course grades included 40 แสดงกราฟแท่งที่แสดงเกรดทั้งหมดในแต่ละรายวิชาที่นักศึกษาได้

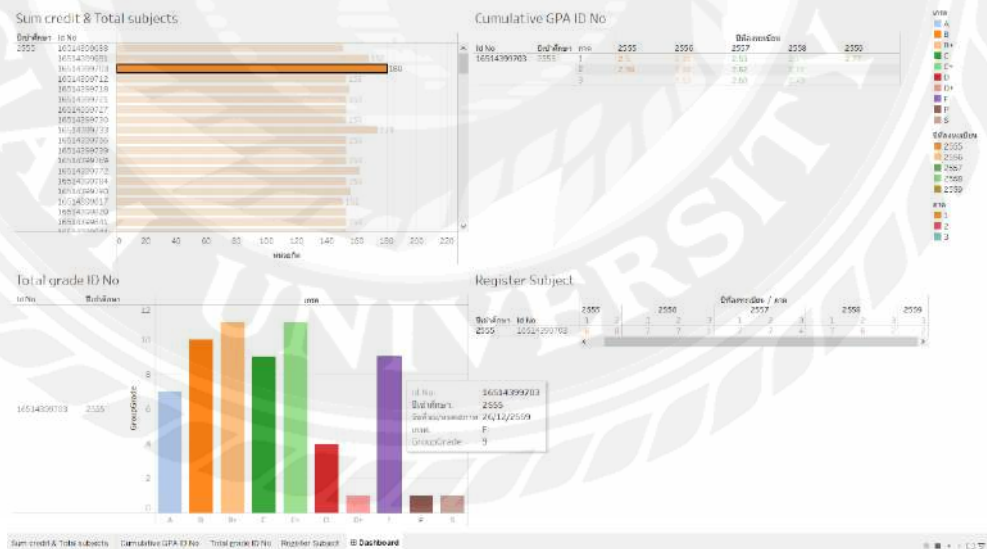
เมื่อต้องการเข้าถึงข้อมูลการศึกษาของแต่ละบุคคล สามารถคลิกที่ id No ก็จะแสดงผลเฉพาะรหัสนั้นๆ ดังรูปที่ 3.9



รูปที่ 3.9 แสดงข้อมูลการศึกษาของนักศึกษาเป็นรายบุคคล

**WorkBook Name: Status40\_Total grades and Cumulative GPA and Register Subject**

Dashboard นี้เป็นการแสดงข้อมูลรายละเอียด GPA สะสมในแต่ละภาคการศึกษาและแสดงว่าในภาคการศึกษานั้นนักศึกษาได้ลงทะเบียนไปทั้งสิ้นกี่รายวิชา



รูปที่ 3.10 WorkBook Name: Status40\_Total grades and Cumulative GPA and Register Subject

- Register Subject ตารางข้อมูลแสดงถึงรายละเอียดของแต่ละภาคการศึกษามีการลงทะเบียนเรียนที่รายวิชา เป็นจำนวนกี่หน่วยกิต

- Total grade ID No กราฟแท่งแสดงข้อมูลเกรดทั้งหมด ว่าได้ A ทั้งหมดกี่รายวิชา ได้เกรด B ทั้งหมดกี่รายวิชา เป็นต้น
- Cumulative GPA ID No ตารางแสดงรายละเอียดว่าในแต่ละภาคการศึกษานักศึกษาได้ GPA สะสมเท่าไร

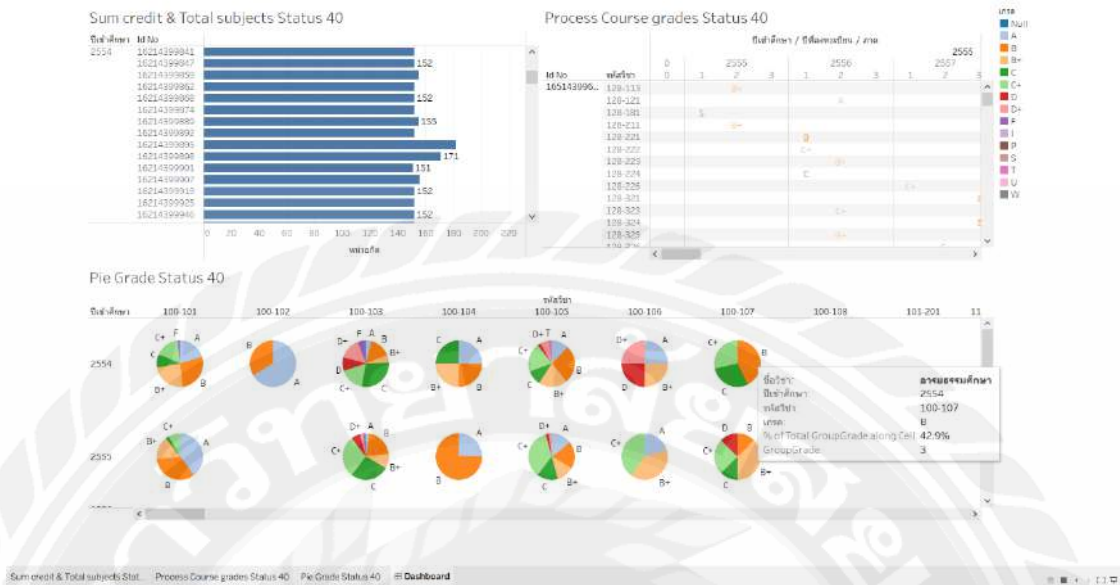
ถ้าต้องการเข้าถึงข้อมูลการศึกษาของนักศึกษาเป็นรายบุคคล ให้คลิกที่ id No ก็จะแสดงผลเฉพาะข้อมูลของรหัสนั้นๆ ดังรูปที่ 3.11



รูปที่ 3.11 แสดงข้อมูลการศึกษาของนักศึกษาเป็นรายบุคคล

**WorkBook Name: Status40\_Total grades and Cumulative GPA and Register Subject**

Dashboard นี้ เป็นการแสดงรายละเอียดข้อมูลเกรดของแต่ละรายวิชาในแต่ละปีการศึกษาว่าในปีการศึกษานั้นผลการศึกษจากจำนวนนักศึกษาทั้งหมดที่ได้ออกมาคิดเป็นกึ่งเปอร์เซ็นต์

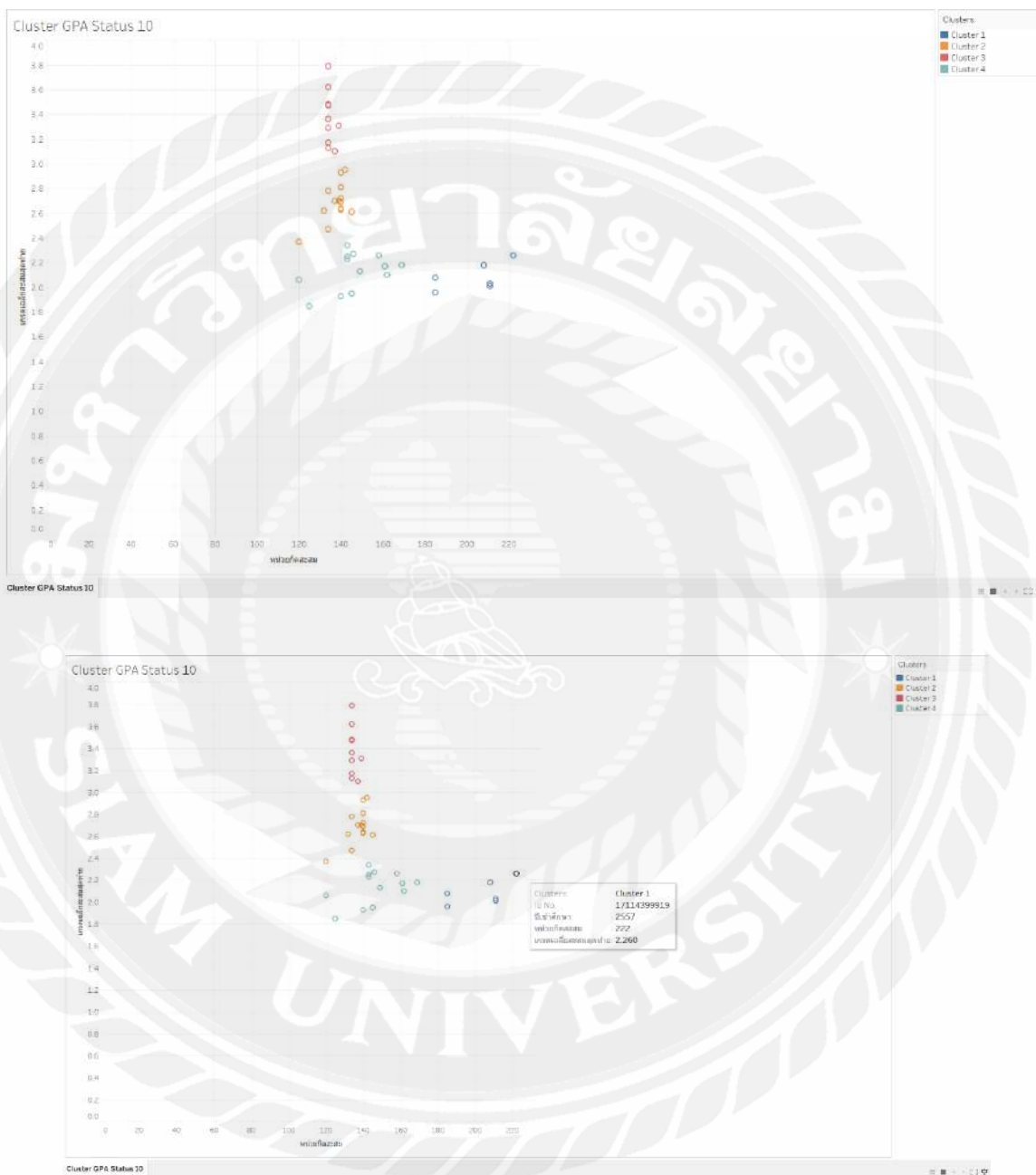


รูปที่ 3.12 Workbook Name: Status40\_Total grades and Cumulative GPA and Register Subject

- Pie Grade Status 40 Pie chart แสดงข้อมูลเกรดในแต่ละวิชา ผลการศึกษาเกรดนั้นๆ มีกี่คน คิดเป็นกี่เปอร์เซ็นต์
- Process Course grades Status 40 ตารางข้อมูลแสดงเกรดรายวิชาในแต่ละปีการศึกษา/ภาคการศึกษา ของนักศึกษาแต่ละคน

การวิเคราะห์ข้อมูลดังต่อไปนี้จะเป็นการใช้ตาราง GPA All & Credit All (10)  
(ตารางข้อมูล GPA และหน่วยกิตสะสมของนักเรียนที่กำลังศึกษาอยู่ในปัจจุบัน)

Workbooks Name: Status10\_Cluster GPA



รูปที่ 3.13 Workbooks Name: Status10\_Cluster GPA

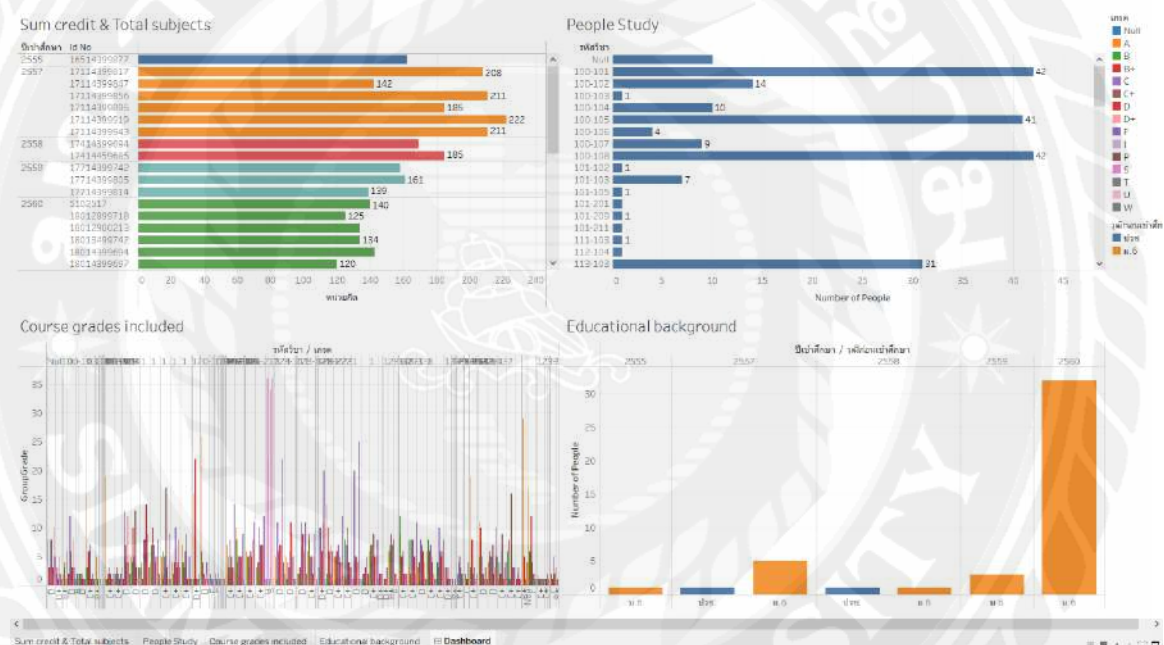
เป็นการจัดกลุ่มข้อมูลของนักศึกษาที่มีความเสี่ยงที่จะไม่สำเร็จการศึกษาตามหลักสูตร 4 ปี โดยใช้การ Clustering แบ่งเป็นจำนวน 4 กลุ่ม โดยกลุ่มที่ 1 และ 4 จะมีความเสี่ยงสูง

มากตามลำดับที่จะไม่สามารถสำเร็จการศึกษาภายในหลักสูตร 4 ปี โดยใช้ GPA สะสม และ หน่วยกิตสะสมมา Plot เป็นกราฟจุด จากนักศึกษาที่กำลังศึกษาอยู่ในปัจจุบัน จำนวน 44 คน

การวิเคราะห์ข้อมูลดังต่อไปนี้ใช้ตารางสถานภาพ 10 (ตารางข้อมูลของนักศึกษาที่กำลังการศึกษาอยู่ในปัจจุบัน ) เชื่อมกับ ตารางสถานภาพ (ตารางความหมายรหัสสถานภาพ)

**Workbooks Name: Status10\_People Study and Course grades included**

Dashboard นี้เป็นการแสดงข้อมูลรายละเอียดของหน่วยกิตสะสมรวมทั้งหมด จำนวนครั้งที่ลงทะเบียนเรียนในแต่ละรายวิชา ข้อมูลเกรดที่ได้ในแต่ละรายวิชา ข้อมูลวุฒิการศึกษาของนักศึกษาก่อนเข้าศึกษาในหลักสูตร (นักศึกษาที่กำลังศึกษาอยู่ในปัจจุบัน)



รูปที่ 3.14 Workbooks Name: Status10\_People Study and Course grades included

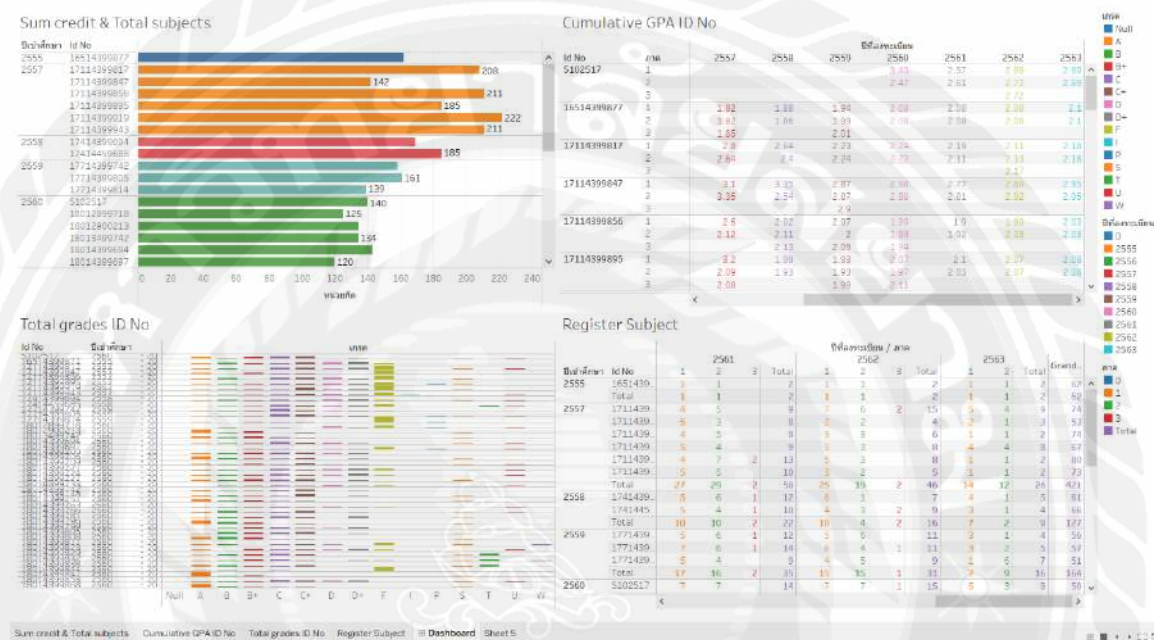
- Sum Credit & Total Subject Status เป็น sheet ที่แสดงถึง นักศึกษาแต่ละคนที่ การศึกษาอยู่ได้ลงทะเบียนมีหน่วยกิตสะสมทั้งหมดตลอดการศึกษาเป็นจำนวนเท่าใด (จะเป็น Main หลักของทุก workbook เนื่องจากแสดงได้เป็นรายบุคคล)
- People Study แสดงกราฟแท่งของแต่ละรายวิชาที่มีการลงทะเบียนทั้งหมดกี่ครั้ง
- Course grades included แสดงกราฟแท่งที่แสดงเกรดทั้งหมดในแต่ละรายวิชาที่ นักศึกษาได้
- Educational background ข้อมูลวุฒิการศึกษาก่อนเข้าศึกษาต่อในหลักสูตร



## Workbooks Name: Status10\_Total grades and Cumulative GPA and Register

### Subject

Dashboard นี้ เป็นการแสดงข้อมูลรายละเอียด GPA สะสมในแต่ละภาคการศึกษาและแสดงว่าในภาคการศึกษานั้นนักศึกษาได้ลงทะเบียนเรียนไปทั้งสิ้นกี่รายวิชา (นักศึกษาที่กำลังศึกษาอยู่ในปัจจุบัน)



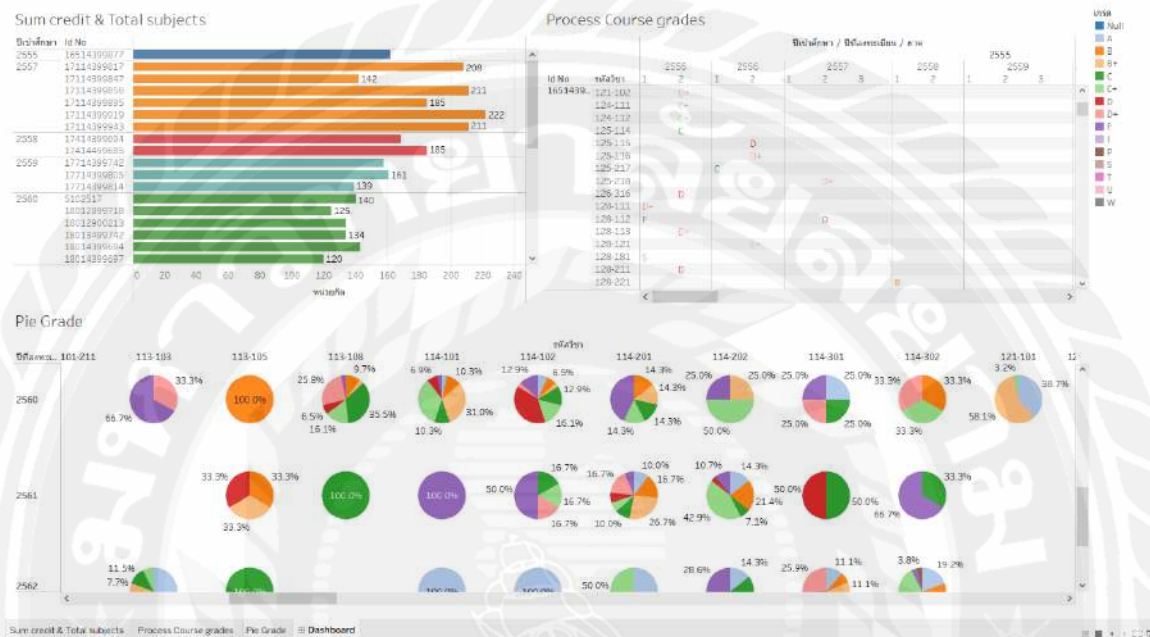
รูปที่ 3.16 Workbooks Name: Status10\_Total grades and Cumulative GPA and Register Subject

- Register Subject ตารางข้อมูลแสดงรายละเอียดในแต่ละภาคการศึกษามีการลงทะเบียนที่รายวิชา
- Total grade ID No กราฟแท่งแสดงข้อมูลเกรดทั้งหมด ว่าได้ A ทั้งหมดกี่รายวิชา B ทั้งหมดกี่รายวิชา เป็นต้น
- Cumulative GPA ID No ตารางที่แสดงรายละเอียดในแต่ละภาคการศึกษานักศึกษาได้ GPA สะสมเท่าใด



### Workbooks Name: Status10\_Process Course grades and Pie Grade

Dashboard นี้ เป็นการแสดงข้อมูลรายละเอียดข้อมูลเกรดของแต่ละรายวิชาในแต่ละปีว่า ในปีการศึกษานั้นผลการศึกษากจากจำนวนนักเรียนทั้งหมดได้ออกมาคิดเป็นกี่ปอร์เซ็นต์ (นักศึกษาปัจจุบัน)



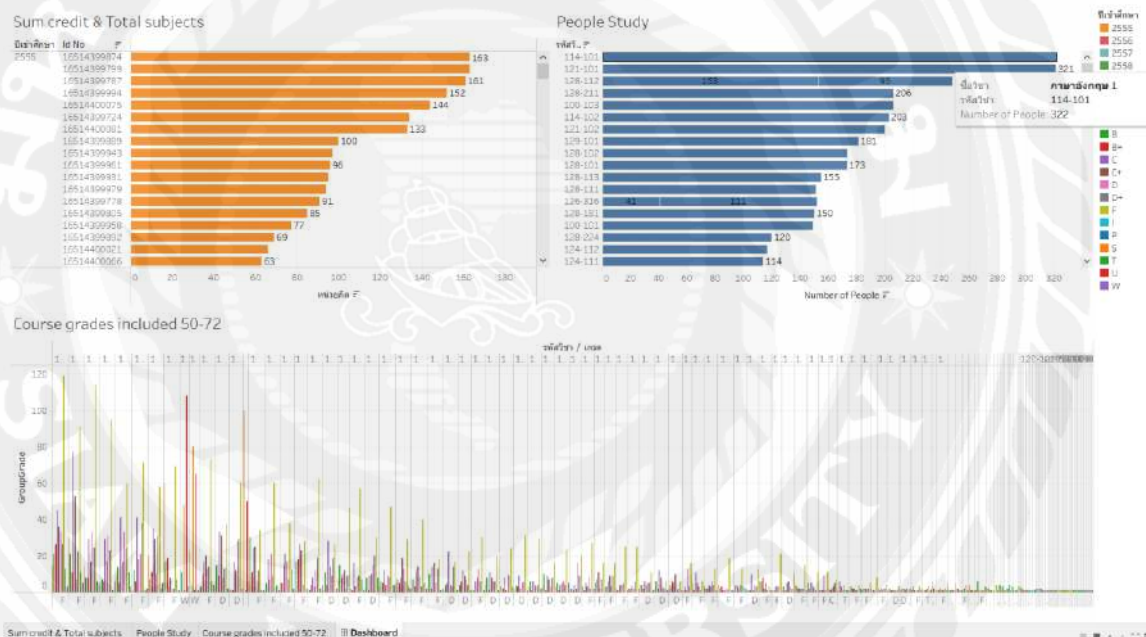
รูปที่ 3.17 Workbooks Name: Status10\_Process Course grades and Pie Grade

- Pie Grade Pie chart แสดงข้อมูลเกรดในแต่ละวิชา ผลการศึกษาเกรดนั้นๆ มีกี่คน คิดเป็นกี่ปอร์เซ็นต์ (นักศึกษาปัจจุบัน)
- Process Course ตารางข้อมูลแสดงเกรดรายวิชาในแต่ละปีการศึกษา/ภาคการศึกษา ของนักศึกษาแต่ละคน (นักศึกษาปัจจุบัน)

การวิเคราะห์ข้อมูลดังต่อไปนี้จะเป็นการใช้ตารางข้อมูลนักศึกษาที่ไม่สำเร็จการศึกษา มีข้อมูลเฉพาะนักศึกษาที่มีสถานภาพ 50 (หมดสภาพนักศึกษา), 52 (หมดระยะเวลาการศึกษา), 70 (ออกตามระเบียบการวัดผล), 71 (ออกตามคำสั่งมหาวิทยาลัย), 72 (ออกระหว่างการศึกษา) เป็นกลุ่มของนักศึกษาที่ไม่สามารถเรียนให้สำเร็จการศึกษาได้ เชื่อมกับ ตารางสถานภาพ (ตารางความหมายรหัสสถานภาพ)

#### Workbooks Name: Status50-72\_People Study and Course grades included

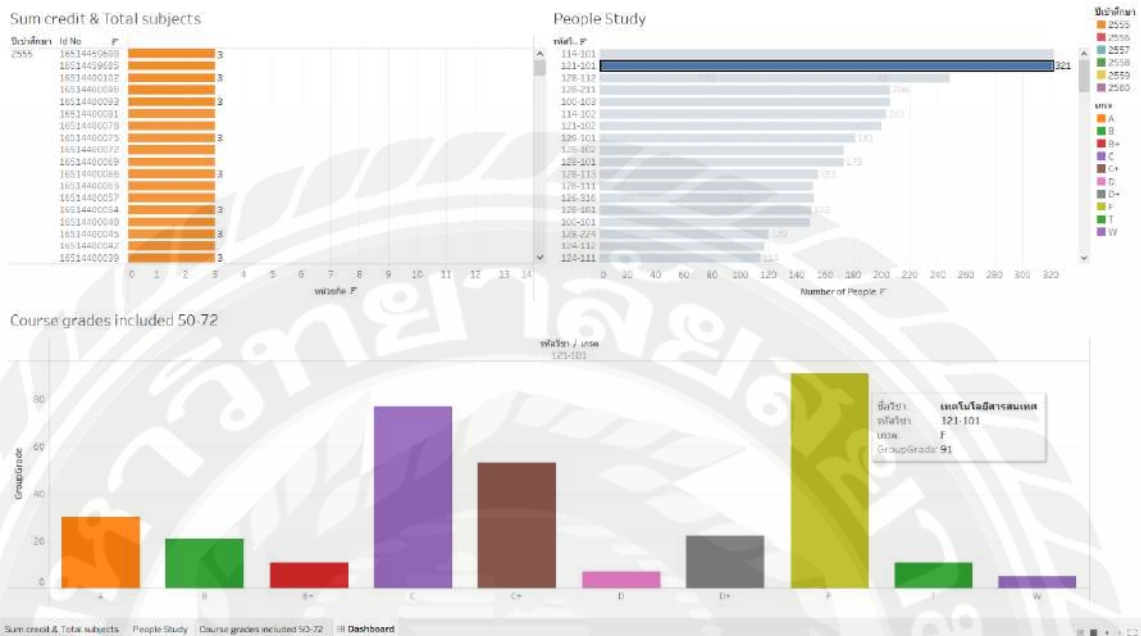
Dashboard นี้เป็นการแสดงข้อมูลรายละเอียดหน่วยกิตสะสมรวมทั้งหมด จำนวนครั้งที่ลงทะเบียนเรียนแต่ละรายวิชา ข้อมูลเกรดที่ได้ในแต่ละรายวิชา (ข้อมูลนักศึกษาที่ไม่สำเร็จการศึกษา)



รูปที่ 3.18 Workbooks Name: Status50-72\_People Study and Course grades included

- Sum Credit & Total Subjects เป็น sheet แสดงข้อมูลประวัติหน่วยกิตสะสมของนักศึกษาทั้งหมดในขณะที่กำลังศึกษาอยู่ ก่อนเปลี่ยนสถานภาพที่จะมีการต้อออก ลาออก หรือย้ายสาขา (รหัสสถานภาพตั้งแต่ 50 ถึง 72)
- People Study แสดงกราฟแท่งของแต่ละวิชาที่มีการลงทะเบียนทั้งหมดกี่ครั้ง
- Course grades included 50-72 แสดงกราฟแท่งที่แสดงเกรดทั้งหมดในแต่ละรายวิชาที่นักศึกษาได้

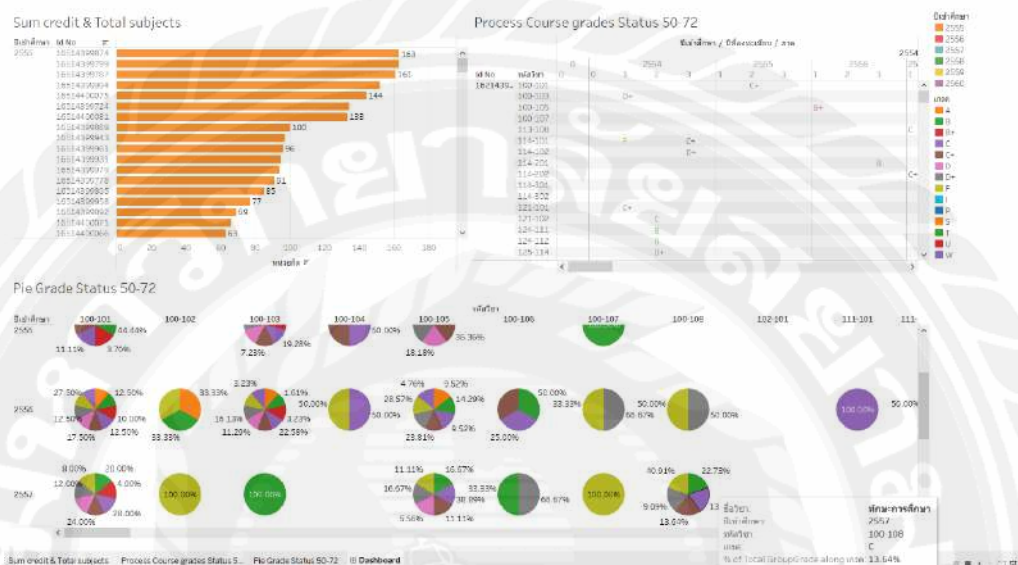
เมื่อคลิกที่ รหัสนักศึกษา หรือรหัสวิชา จะสามารถเข้าถึงข้อมูลรายละเอียดเกรดในแต่ละวิชาได้ หรือ จะเป็นเกรดทั้งหมดของนักศึกษาที่ได้



รูปที่ 3.19 แสดงข้อมูลรายละเอียดเกรดในแต่ละวิชา

### Workbooks Name: Status50-72\_Process Course grades and Pie Grade

Dashboard นี้เป็นการแสดงข้อมูลรายละเอียดข้อมูลเกรดของแต่ละรายวิชาในแต่ละปีการศึกษา โดยปีการศึกษานั้นมีผลการศึกษามาจากจำนวนนักเรียนทั้งที่ได้คิดเป็นที่เปอร์เซ็นต์ (นักศึกษาที่ไม่สำเร็จการศึกษา)

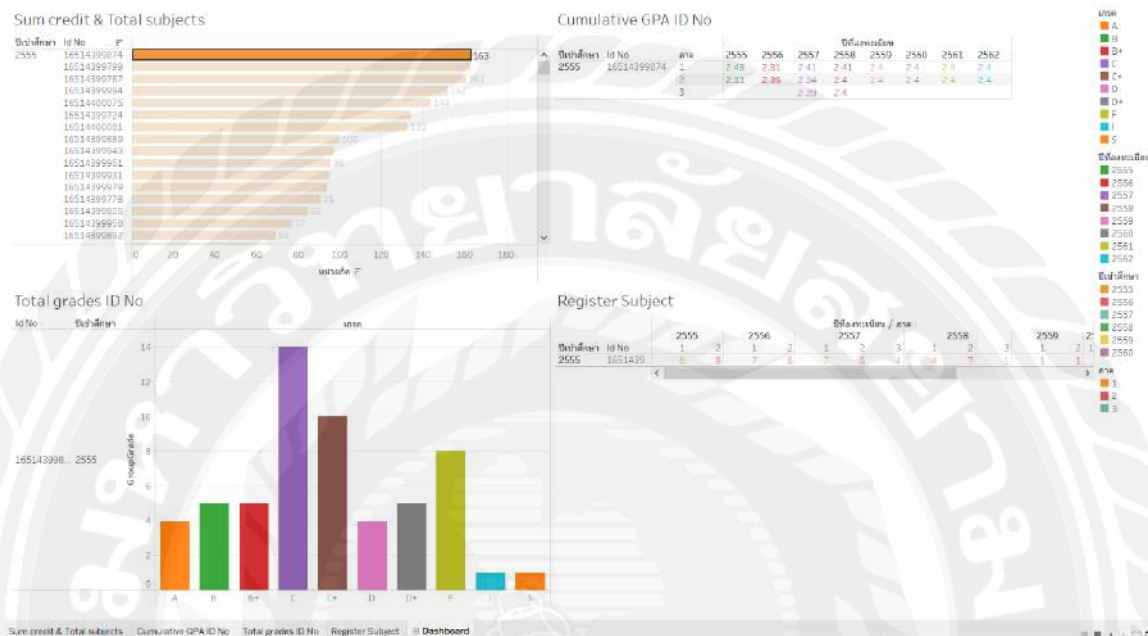


รูปที่ 3.20 Workbooks Name: Status50-72\_Process Course grades and Pie Grade

- Pie Grade Status 50-72 Pie chart แสดงข้อมูลเกรดในแต่ละวิชา ผลการศึกษาของแต่ละเกรดว่ามีนักศึกษาได้จำนวนกี่คน คิดเป็นที่เปอร์เซ็นต์ (นักศึกษาที่ไม่สำเร็จการศึกษา)
- Process Course grade Status 50-72 ตารางข้อมูลแสดงเกรดรายวิชาในแต่ละปีการศึกษา/ภาคการศึกษา ของนักศึกษาแต่ละคน (นักศึกษาที่ไม่สำเร็จการศึกษา)

### Workbooks Name: Status50-72\_Total grades and Cumulative GPA and Register Subject

Dashboard นี้ เป็นการแสดงข้อมูลรายละเอียด GPA สะสมในแต่ละภาคการศึกษาและแสดงว่าในภาคการศึกษานั้นมีนักศึกษาได้ลงทะเบียนไปทั้งหมดกี่รายวิชา (นักศึกษาที่ไม่สำเร็จการศึกษา)

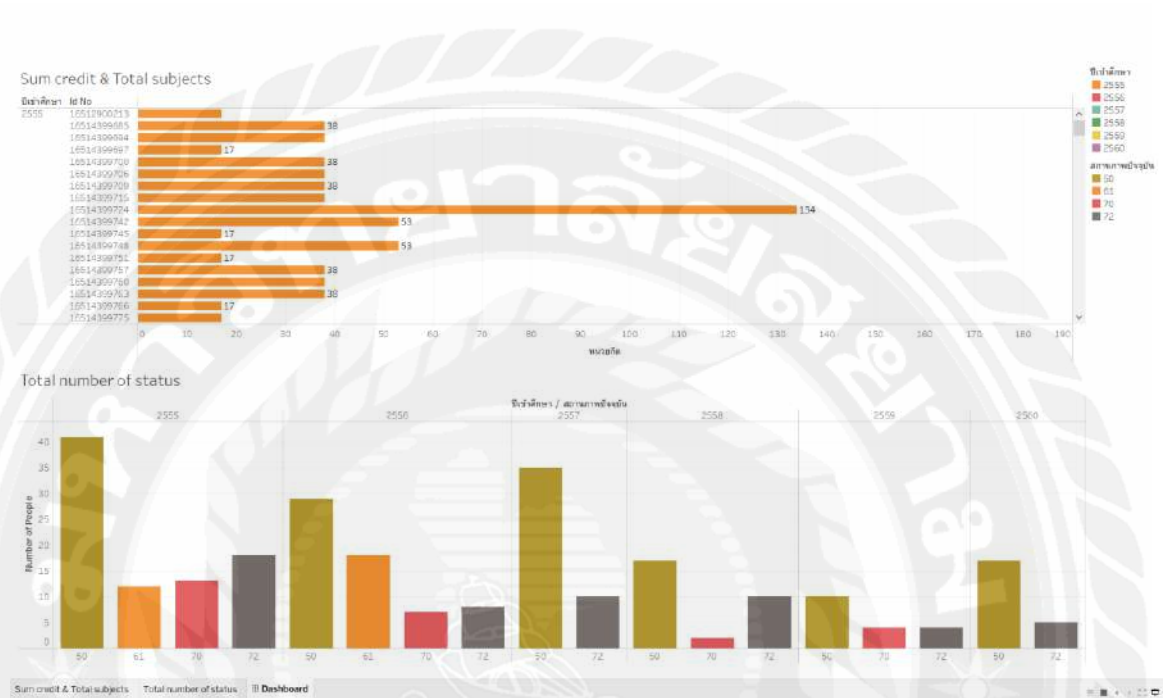


รูปที่ 3.21 Workbooks Name: Status50-72\_Total grades and Cumulative GPA and Register Subject

- Register Subject ตารางข้อมูลแสดงรายละเอียดว่าในแต่ละภาคการศึกษามีการลงทะเบียนเรียนกี่รายวิชา รวมเป็นกี่หน่วยกิต
- Total grade ID No กราฟแท่งแสดงข้อมูลเกรดทั้งหมด ว่าได้ A ทั้งหมดกี่รายวิชา B ทั้งหมดกี่รายวิชา เป็นต้น
- Cumulative GPA ID No ตารางที่บอกรายละเอียดว่าในแต่ละภาคการศึกษา นักศึกษาได้ GPA สะสมเท่าไร

**Workbooks Name: Status50-72\_Total number of status and people**

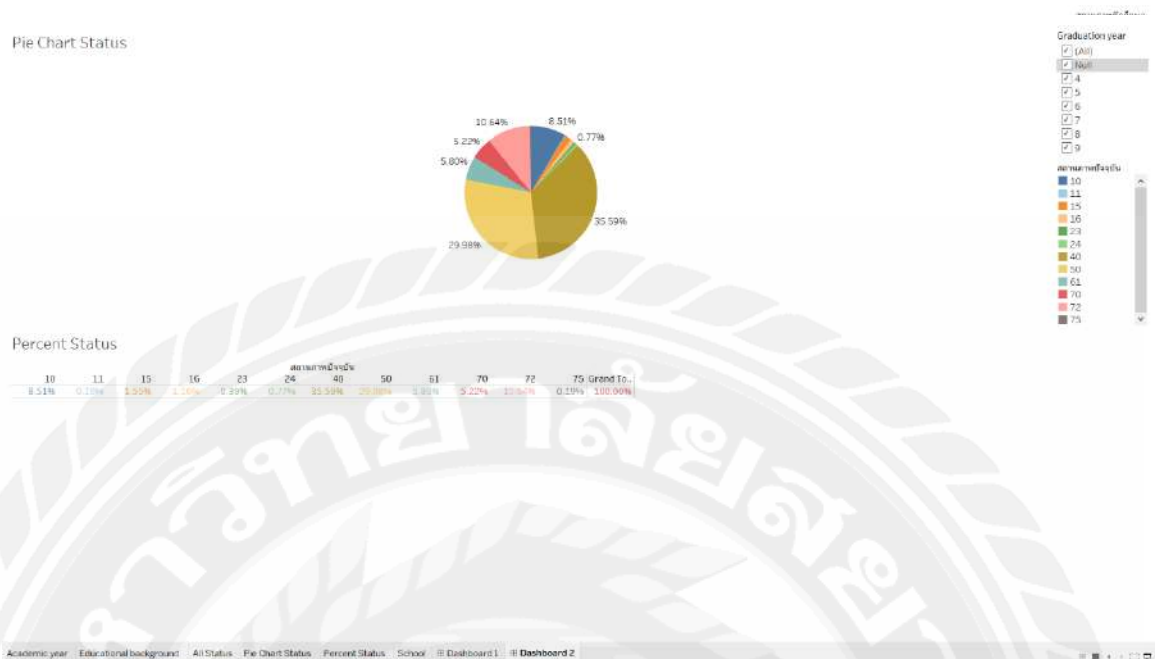
Dashboard นี้เป็นการแสดงข้อมูลรายละเอียดจำนวนนักศึกษาที่ไม่สำเร็จการศึกษาในสถานภาพต่างๆ ในแต่ละปี และสามารถบอกรายละเอียดได้ว่านักศึกษารหัสนี้เรียนสะสมมาได้กี่หน่วยกิตก่อนที่จะพ้นสภาพการเป็นนักศึกษาหรือย้ายคณะสาขาวิชา



รูปที่ 3.22 Workbooks Name: Status50-72\_Total number of status and people

- Total number of status แสดงจำนวนของนักศึกษาในแต่ละสถานภาพในแต่ละปีการศึกษา

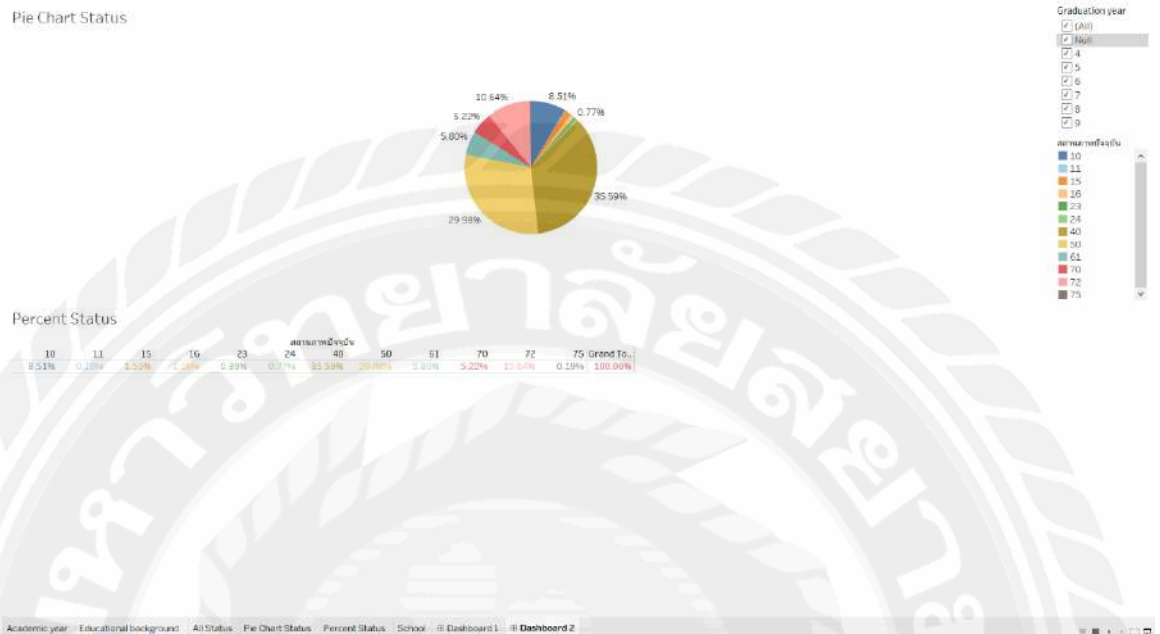




รูปที่ 3.24 Dashboard 1

- Academic year กราฟแท่งแสดงจำนวนคนที่สมัครเข้าเรียนมหาวิทยาลัยสยาม คณะวิทยาศาสตร์ สาขาวิทยาการคอมพิวเตอร์ในแต่ละปีการศึกษา
- Educational background กราฟแท่งแสดงจำนวนคนที่สมัครเข้าเรียนในหลักสูตร โดยแยกด้วยวุฒิการศึกษาที่จบมา
- All Status กราฟแท่งแสดงจำนวนนักศึกษาแต่ละสถานภาพในแต่ละปีการศึกษา





รูปที่ 3.25 Dashboard 2

- Pie Chart Status Pie Chart แสดงข้อมูลจำนวนนักศึกษาแต่ละสถานภาพโดยแสดงเป็นเปอร์เซ็นต์
- Percent Status ตารางข้อมูลจำนวนนักศึกษาแต่ละสถานภาพโดยแสดงเป็นเปอร์เซ็นต์

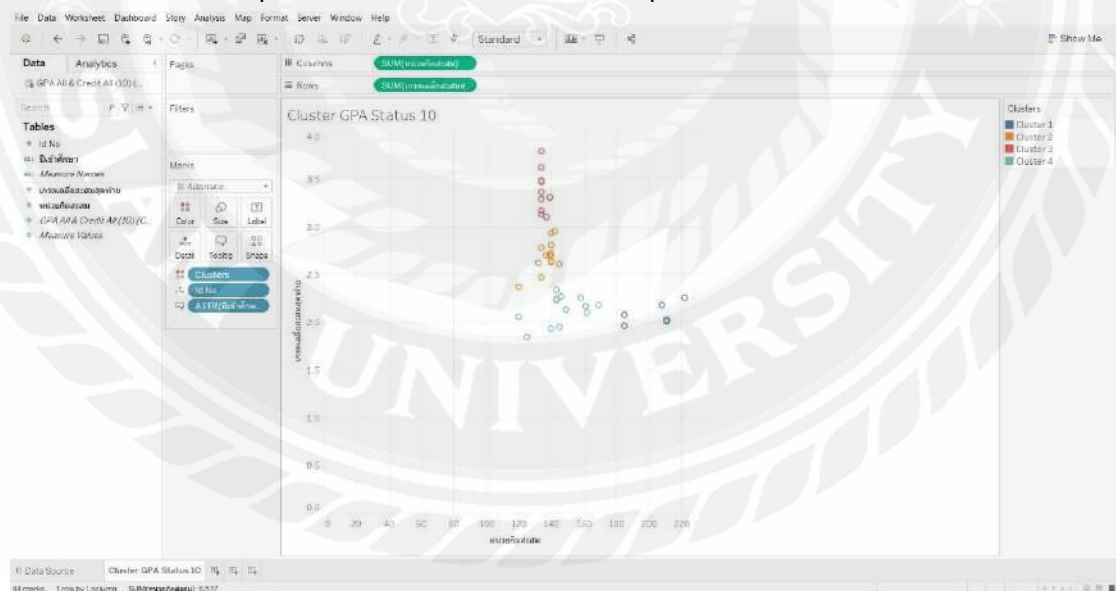
## บทที่ 4

### การนำเสนอแผนภาพของข้อมูล

หลังจากที่ได้ทำการเตรียมข้อมูล (Data Preparation) จนได้ข้อมูลที่สามารถนำไปทำการวิเคราะห์เพื่อจัดกลุ่มข้อมูล (Data Clustering) ตามเงื่อนไขที่กำหนดได้แล้ว ต่อไปจะเป็นการนำผลลัพธ์ที่ได้มานำเสนอให้ผู้ใช้เข้าใจได้ง่าย และสามารถเข้าถึงข้อมูลเชิงลึกได้ง่าย นั่นคือ การนำเสนอแผนภาพของข้อมูล (Data Visualization) โดยใช้โปรแกรม Tableau ในการจัดทำแผนภาพของข้อมูล โดยคณะผู้จัดทำได้ทำการเลือกการแสดงผลในรูปแบบของกราฟ (Graph) และแผนภูมิ (Chart) ที่เหมาะสมกับสารสนเทศที่ต้องการจะนำเสนอ เช่น ถ้าแสดงผลในเชิงแนวโน้มจะเลือกใช้กราฟเส้น (Line Graph) ถ้าแสดงผลเป็นเปอร์เซ็นต์จะเลือกใช้กราฟวงกลม (Pie Chart) ถ้าแสดงผลเปรียบเทียบหรือแสดงจำนวนจะเลือกใช้กราฟแท่ง (Bar Graph) เป็นต้น และนำมาสร้างเป็น Dashboard เพื่อให้ดูง่ายและสวยงาม สามารถนำสารสนเทศในมิติต่างๆ มาเรียงไว้ในหน้าเดียวกันได้ สามารถทำการ Drill Down เพื่อเข้าไปดูรายละเอียดของข้อมูลต่างๆ ได้

#### 4.1 การวิเคราะห์ข้อมูลแบบจัดกลุ่ม GPA ของนักศึกษาที่กำลังศึกษาอยู่

เป็นการนำข้อมูลเข้าไปวิเคราะห์เพื่อจัดกลุ่มของนักศึกษาที่กำลังศึกษาอยู่โดยแบ่งกลุ่มตามหน่วยกิตสะสมในปัจจุบันทั้งหมด และ เกรดเฉลี่ยสะสมสุดท้าย ได้ผลลัพธ์ดังนี้



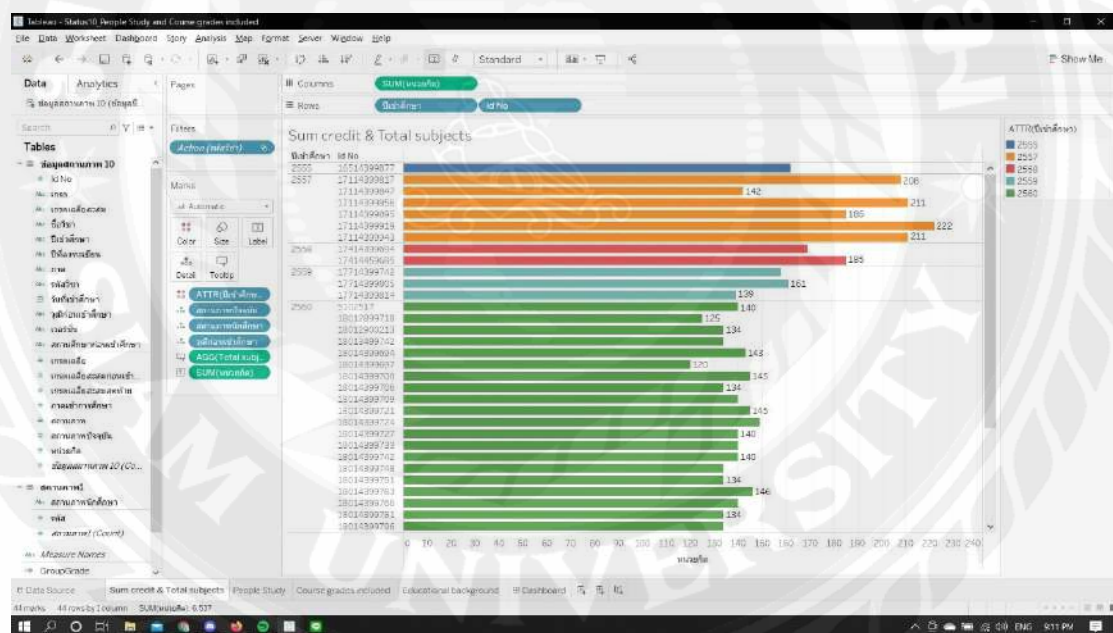
รูปที่ 4.1 แสดงผลการวิเคราะห์ข้อมูลแบบจัดกลุ่ม GPA ของนักศึกษาที่กำลังศึกษา

จากรูปที่ 4.1 จะแบ่งกลุ่มออกเป็น 4 สี สีที่ 1 สีน้ำเงิน (Cluster 1) จะเป็นกลุ่มของนักศึกษาที่มีหน่วยกิตสะสมมากคิดปกติและมีเกรดเฉลี่ยสะสมที่ไม่สามารถจบตามระยะเวลาของหลักสูตร 4

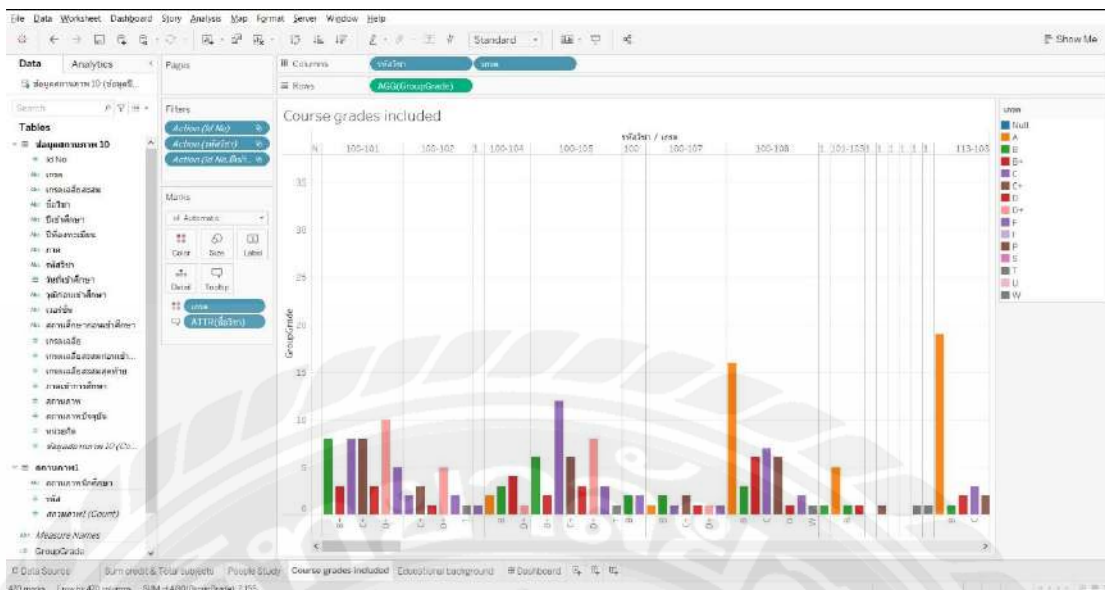
ปีได้ สี่ที่ 2 สีส้ม (Cluster 2) จะเป็นกลุ่มของนักศึกษาที่มีหน่วยกิตสะสมอยู่ในเกณฑ์ที่ผิดปกติเล็กน้อยและมีเกรดเฉลี่ยสะสมที่มีแนวโน้มต่ำที่จะไม่จบตามระยะเวลาของหลักสูตร 4 ปี สี่ที่ 3 สีแดง (Cluster 3) จะเป็นกลุ่มของนักศึกษาที่มีหน่วยกิตสะสมปกติและมีเกรดเฉลี่ยสะสมที่มีแนวโน้มสูงว่าจะจบตามระยะเวลาของหลักสูตร 4 ปี และสี่ที่ 4 สีฟ้า (Cluster 4) จะเป็นกลุ่มของนักศึกษาที่มีหน่วยกิตสะสมผิดปกติและมีเกรดเฉลี่ยสะสมที่มีแนวโน้มสูงว่าจะไม่จบตามระยะเวลาของหลักสูตร 4 ปี

#### 4.2 การวิเคราะห์ภาพรวมผลการเรียนของแต่ละรายวิชา

จะเป็นการนำข้อมูลเข้าไปวิเคราะห์เพื่อหาภาพรวมผลการเรียนของแต่ละรายวิชาที่นักศึกษา โดยจะใช้ รหัสวิชา, เกรดรายวิชาของนักศึกษาแต่ละคนเพื่อหาภาพรวมของผลการเรียนในแต่ละรายวิชาโดยแยกสีไว้อย่างชัดเจน เพื่อให้ผู้ใช้รู้ว่ารายวิชาใดที่มีปัญหาในการเรียนมากที่สุด โดยดูจากผลการเรียนที่เป็น F ได้จากกราฟดังรูปที่ 4.3 ทั้งนี้ยังใช้รหัสนักศึกษาและรหัสวิชาเพื่อหาจำนวนคนที่เรียนในแต่ละรายวิชาด้วย



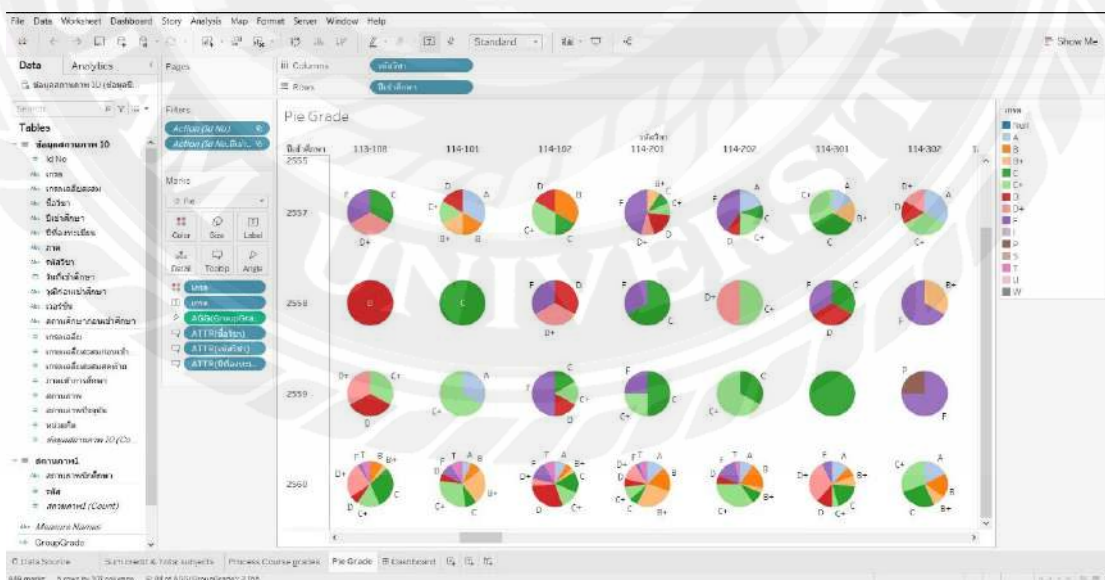
รูปที่ 4.2 แสดงภาพรวมผลการลงทะเบียนของนักศึกษาแต่ละคน



รูปที่ 4.3 แสดงภาพรวมผลการศึกษของแต่ละรายวิชา

### 4.3 การวิเคราะห์ภาพรวมผลการศึกษาแต่ละรายวิชาตามปีการศึกษา

จะเป็นการวิเคราะห์เพื่อหาภาพรวมผลการศึกษาของแต่ละรายวิชาโดยแยกตามปีการศึกษาที่นักศึกษาลงทะเบียนเรียน โดยเปลี่ยนมุมมองและเจาะจงเกรดรายวิชาตามปีการศึกษาเพื่อดูปัญหาว่าในปีการศึกษานั้น ๆ มีปัญหาที่วิชาใดในปีการศึกษาใด ทางผู้สอนรายวิชานั้นสามารถนำผลการวิเคราะห์ไปใช้ประโยชน์ในการเรียนการสอนได้

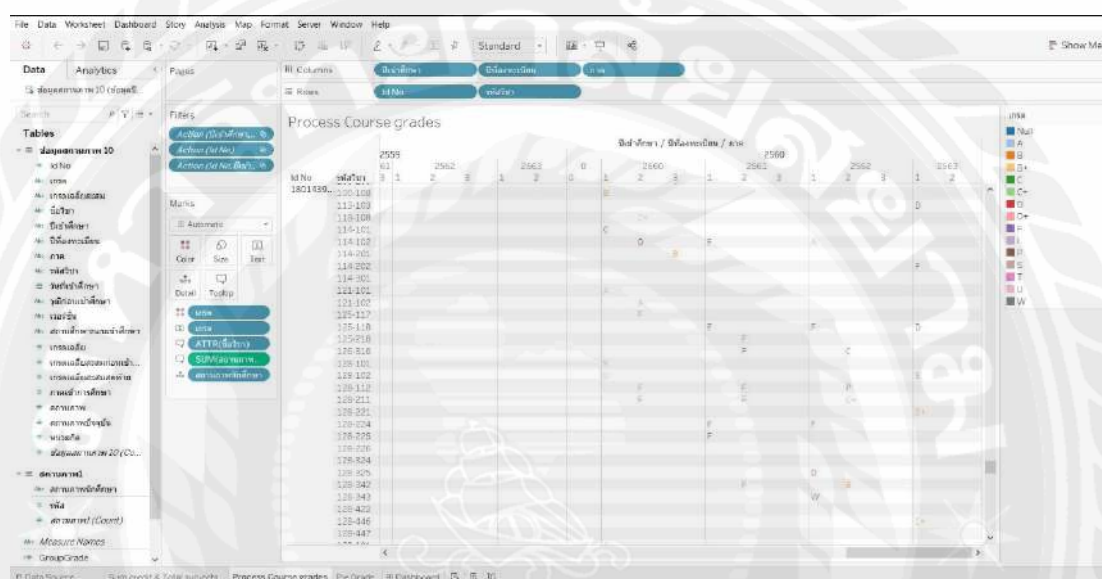


รูปที่ 4.4 ผลลัพธ์การศึกษาแต่ละรายวิชาตามปีการศึกษา

จากรูปที่ 4.4 จะใช้แอททริบิวต์รหัสวิชาเป็นตัวตั้ง (Row) และแอททริบิวต์ปีการศึกษาต่อในหลักสูตรเพื่อหาผลลัพธ์ ซึ่งจะแสดงเป็นกราฟวงกลม (Pie Chart) เพื่อแยกข้อมูลของเกรดในแต่ละรายวิชาออกเป็นเปอร์เซ็นต์ แยกตามปีการศึกษาที่เข้าศึกษาต่อในหลักสูตร

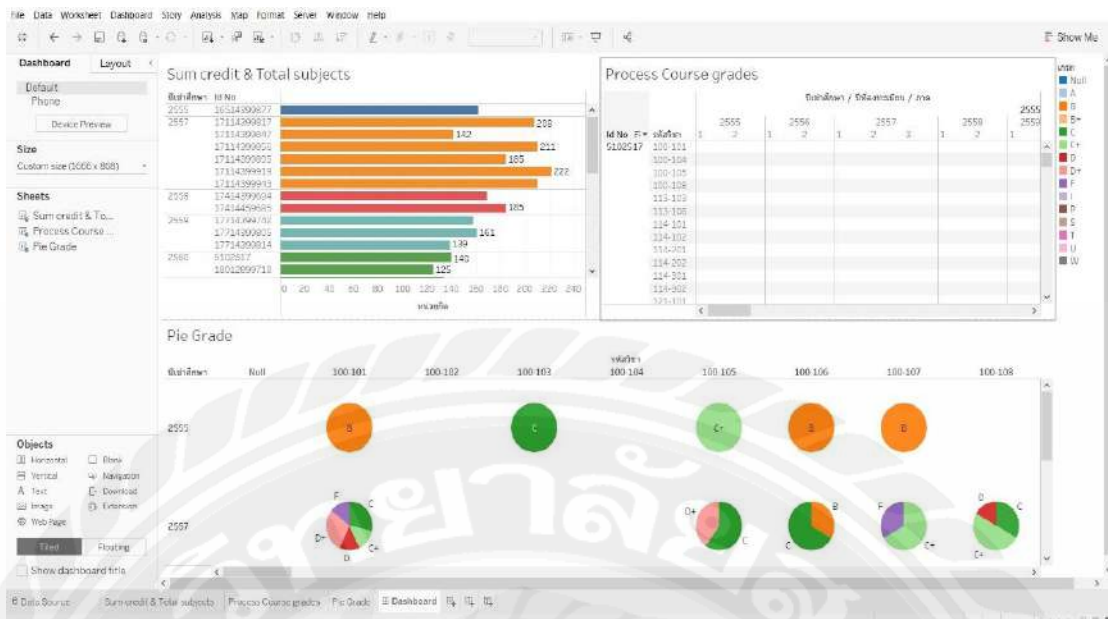
#### 4.4 การวิเคราะห์เกรดตามรายวิชาของนักศึกษาตามปีการศึกษา

จะเป็นการวิเคราะห์เพื่อหาเกรดรายวิชาของนักศึกษาตามปีการศึกษาที่ลงทะเบียนเรียน และเจาะจงลงไปตามภาคการศึกษาของแต่ละปีการศึกษา ว่านักศึกษาได้เกรดใดในรายวิชานั้น ๆ



รูปที่ 4.5 ผลลัพธ์การศึกษาและจำนวนวิชาที่ลงทะเบียนตามภาคการศึกษา

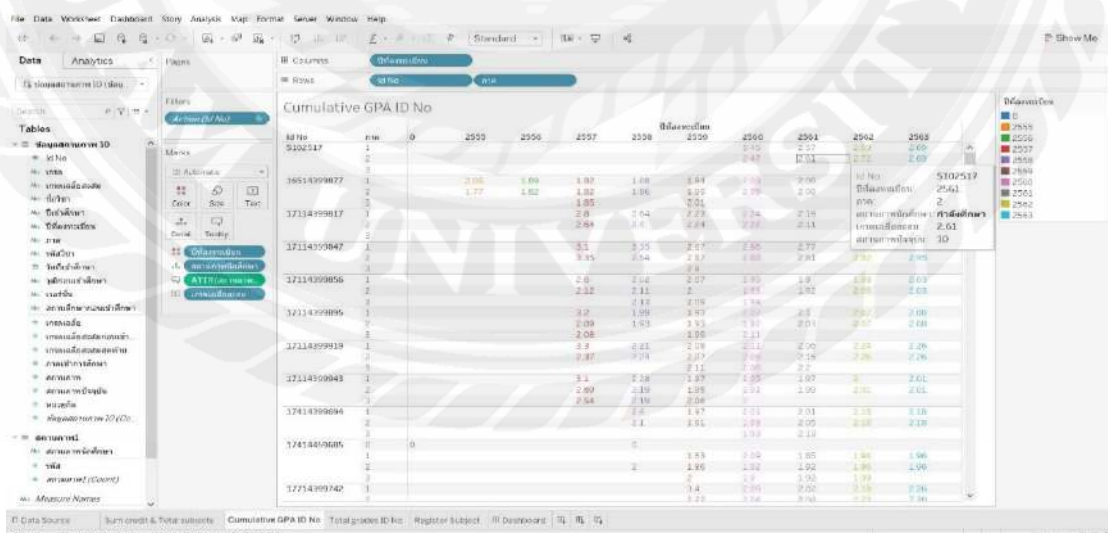
จากรูปที่ 4.5 จะใช้แอททริบิวต์รหัสนักศึกษา ปีที่ลงทะเบียนเรียนในรายวิชานั้นๆ และรหัสวิชา เป็นหลักเพื่อดูเกรดของแต่ละรายวิชาที่ลงทะเบียนเรียนทั้งหมดเป็นรายบุคคล โดยจะแยกไปตามภาคการศึกษาของแต่ละปีการศึกษา



รูปที่ 4.6 แสดงผลการวิเคราะห์ผลการศีกษาโดยรวม

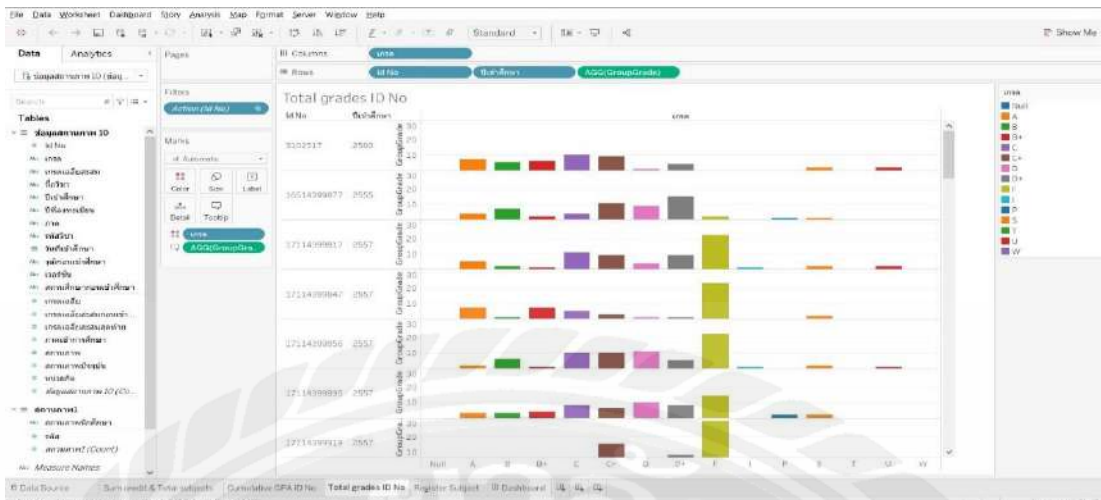
จากรูปที่ 4.6 เป็นการแสดงผลที่เป็นภาพรวมผลการศีกษาทั้งหมด โดยนำข้อมูลจำนวนหน่วยกิตทั้งหมดมาเป็นตัวกลางในการแสดงผลลัพท์ โดยจะใช้การ Use as Filter เพื่อการแสดงผลข้อมูล เกรดรายวิชาที่เคลงทะเลเบียน และ ภาพรวมรายวิชาต่อปีเป็นรายบุคคล

#### 4.5 การวิเคราะห์เกรดเฉลี่ยสะสมและจำนวนรายวิชาที่ลงทะเบียนของนักศึกษาที่กำลังศีกษา



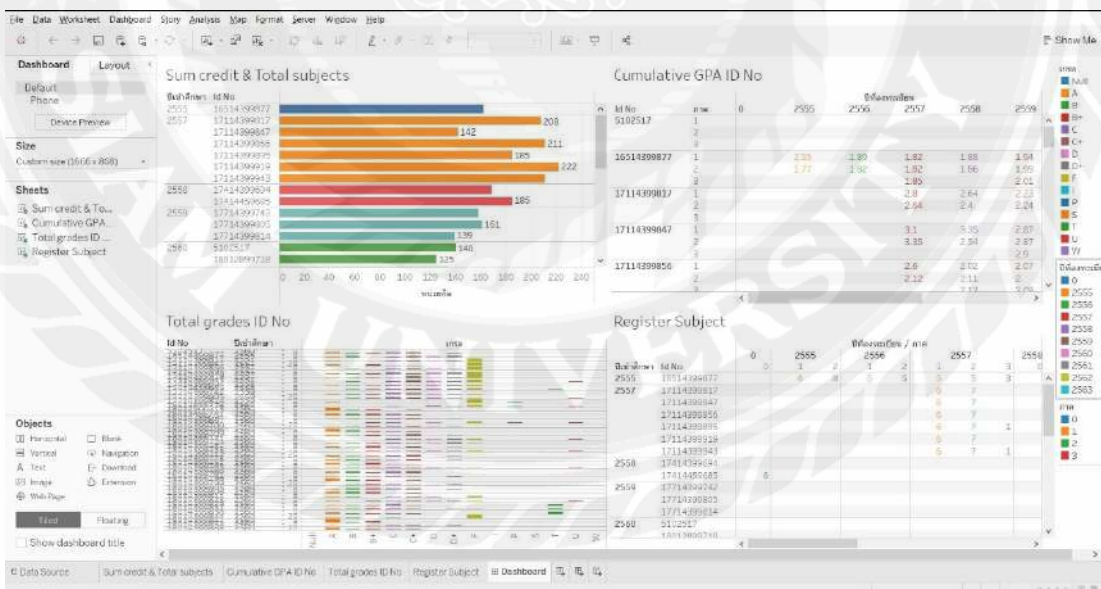
รูปที่ 4.7 แสดงภาพรวมเกรดเฉลี่ยสะสม (GPA)





รูปที่ 4.9 ผลลัพธ์การนับจำนวนเกรดที่ลงทะเบียนเรียน

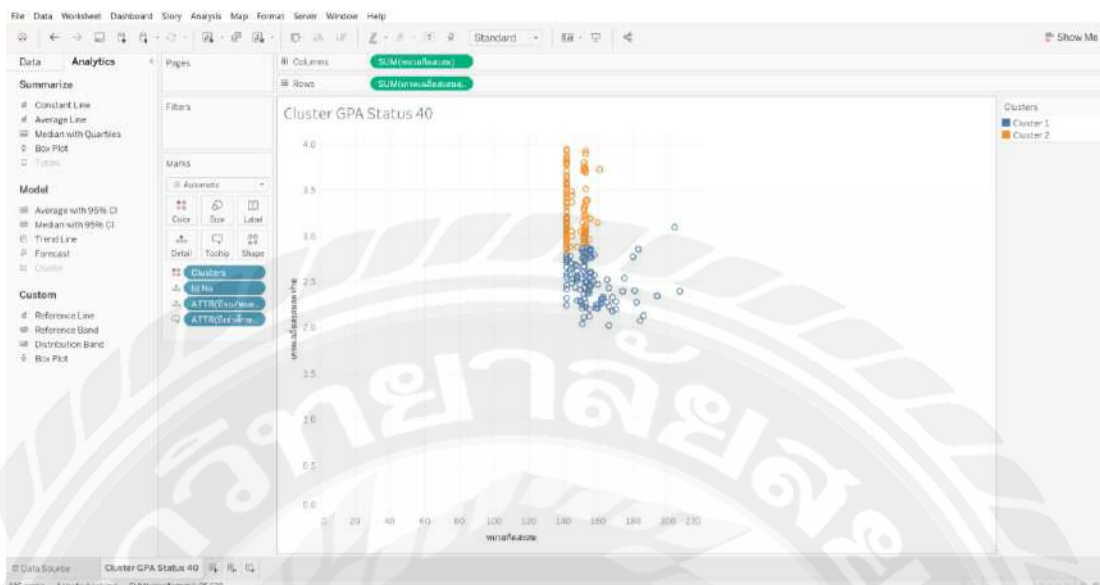
จากรูปที่ 4.9 แสดงผลรวมของเกรดแต่ละรายวิชาที่นักศึกษาได้ลงทะเบียนเรียนทั้งหมด ทั้งที่เรียนผ่านและไม่ผ่าน จะนับรวมไว้ทั้งหมดในหน้านี้ เพื่อแสดงว่านักศึกษาแต่ละคนได้เกรดอะไรบ้างในขณะที่กำลังศึกษาอยู่ และยังบอกได้ว่านักศึกษาคอนใดที่มีผลการเรียนที่เป็น F มากผิดปกติ อาทิเช่น ได้ F ในรายวิชาใดรายวิชาหนึ่งมากกว่า 3 ครั้ง ซึ่งจะต้องหาสาเหตุต่อไปว่าเพราะอะไรถึงเรียนไม่ผ่านในรายวิชานั้น เป็นเพราะอาจารย์ผู้สอน หรือตัวนักศึกษาเอง



รูปที่ 4.10 แสดงภาพรวมเกรดเฉลี่ยสะสมและจำนวนรายวิชาที่ลงทะเบียนเรียนของนักศึกษาที่กำลังศึกษาอยู่



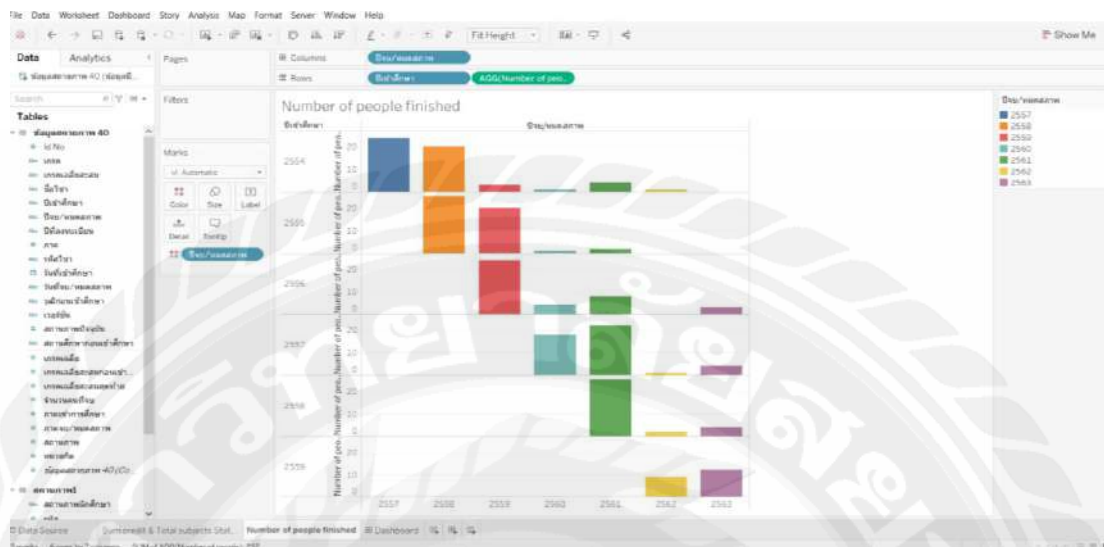
#### 4.6 การวิเคราะห์ข้อมูลแบบจัดกลุ่ม GPA ของนักศึกษาที่จบการศึกษาแล้ว



รูปที่ 4.11 แสดงการจัดกลุ่ม GPA ของนักศึกษาที่จบการศึกษาแล้ว

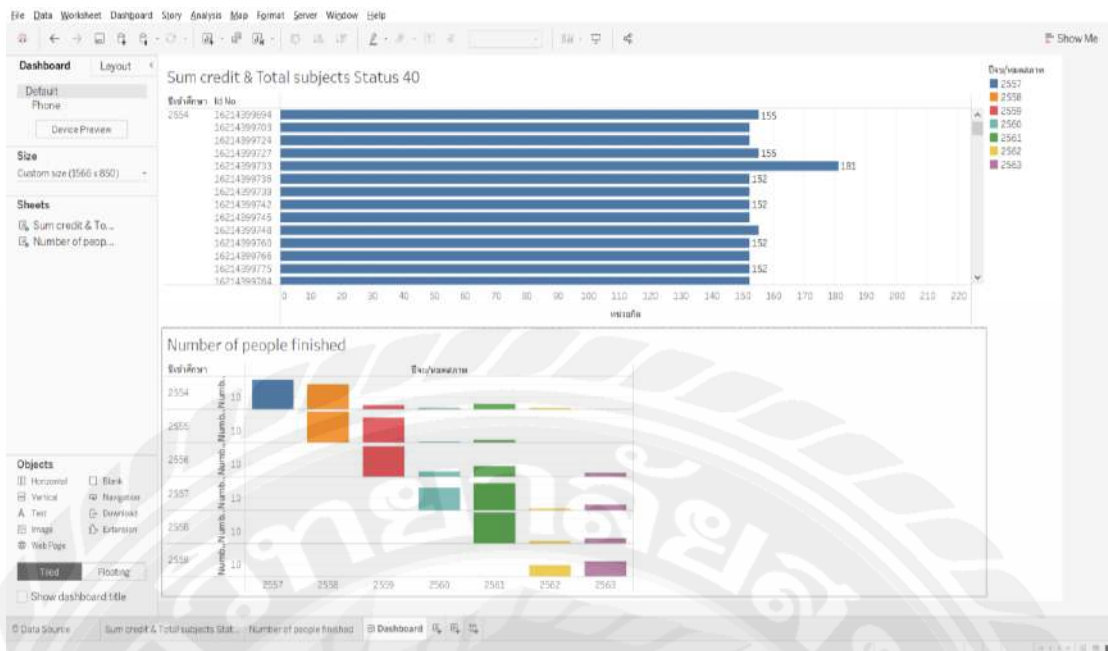
จากรูปที่ 4.11 จะเป็นการแสดงผลของการจัดกลุ่มข้อมูล โดยจะแบ่งออกมาเป็น 2 กลุ่มคือ กลุ่มที่จบการศึกษาตามระยะเวลาของหลักสูตรที่กำหนดไว้ กล่าวคือจบการศึกษาภายใน 4 ปี และกลุ่มที่จบการศึกษาไม่ตรงตามระยะเวลาของหลักสูตรที่กำหนดไว้ กล่าวคือ ใช้เวลาในการศึกษา มากกว่า 4 ปี

#### 4.7 การวิเคราะห์ภาพรวมปีการศึกษาที่นักศึกษาเข้าศึกษาต่อในหลักสูตรและปีการศึกษาที่จบการศึกษา



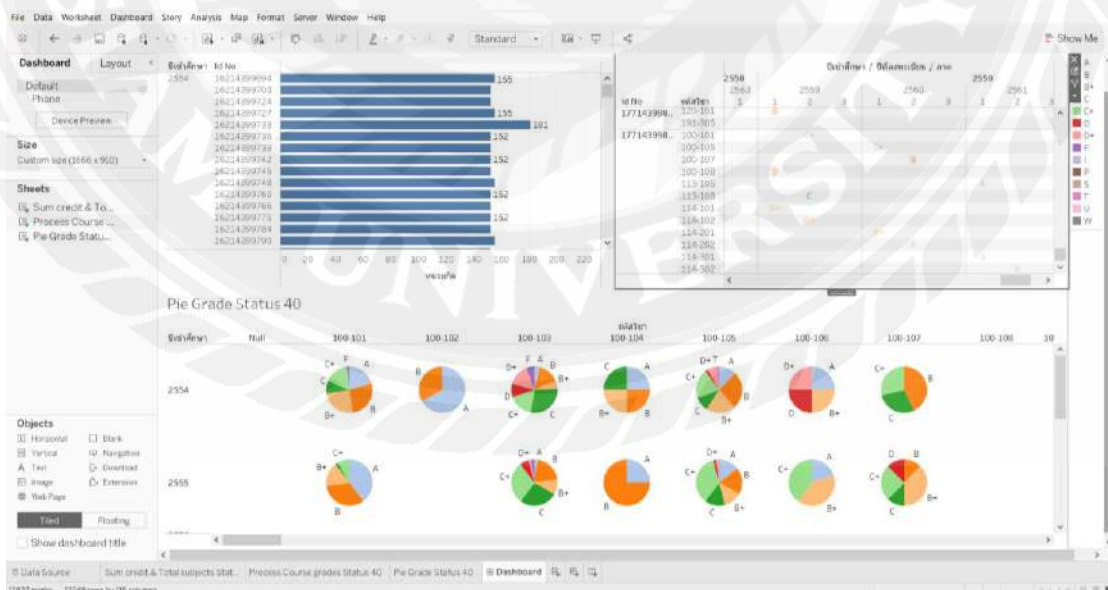
รูปที่ 4.12 แสดงจำนวนนักศึกษาที่จบการศึกษาตามเกณฑ์และจบการศึกษาเกินเกณฑ์

จากรูปที่ 4.12 เป็นการแสดงจำนวนของนักศึกษาที่จบการศึกษาแล้วในแต่ละปีการศึกษา โดยใช้เอทริบิวต์ปีที่จบการศึกษา และปีที่เข้าศึกษาต่อในหลักสูตร เพื่อนำมาหาผลลัพธ์ ว่ามีนักศึกษากี่คนที่จบตามระยะเวลาที่หลักสูตรกำหนดไว้ 4 ปี และมีจำนวนกี่คนที่จบการศึกษาโดยใช้เวลาในการเรียนมากกว่า 4 ปี



รูปที่ 4.13 แสดงภาพรวมของจำนวนนักศึกษาที่จบการศึกษา

**4.8 การวิเคราะห์เกรดของแต่ละรายวิชาตามปีการศึกษาของนักศึกษาที่จบการศึกษาแล้ว**  
 จะเป็นการวิเคราะห์หาเกรดของแต่ละรายวิชาของนักศึกษาเป็นรายบุคคล และเกรดโดยรวมของรายวิชาต่างๆ ของนักศึกษาที่จบการศึกษาแล้ว รวมทั้งที่จบการศึกษาตามเกณฑ์ 4 ปี และเกินเกณฑ์ 4 ปี เพื่อนำไปเป็นโมเดล (Model) ในการวิเคราะห์ปัจจัยที่มีผลต่อการศึกษา

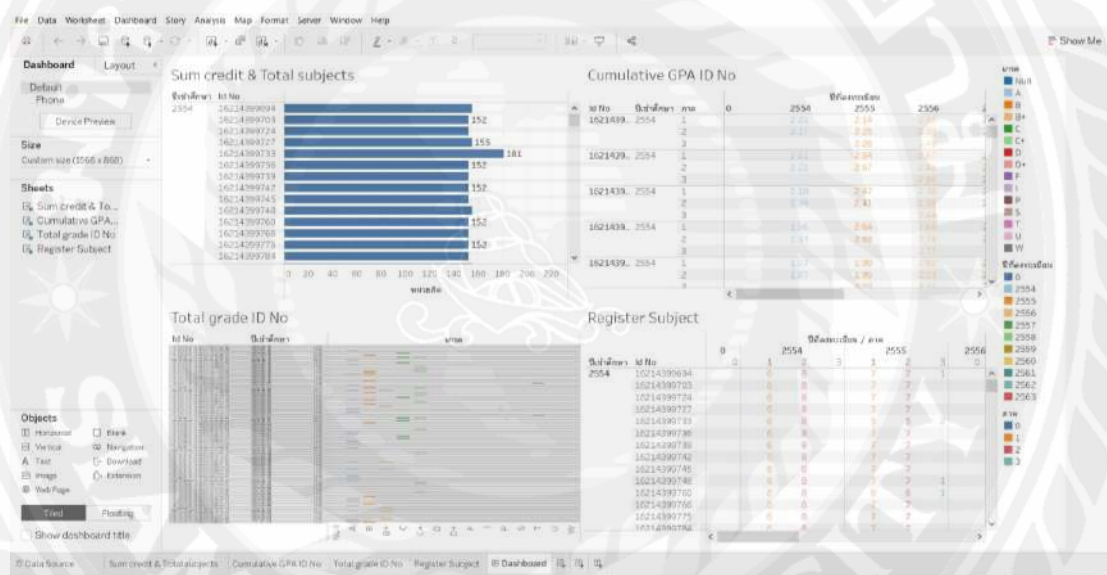


รูปที่ 4.14 แสดงผลเกรดของแต่ละรายวิชาตามปีการศึกษาของนักศึกษาที่จบการศึกษาแล้ว

จากรูปที่ 4.14 เป็นการแสดงผลโดยรวมทั้งหมด โดยนำข้อมูลจำนวนหน่วยกิตทั้งหมดมาใช้เป็นตัวกลาง โดยจะใช้การ Use as Filter เพื่อการแสดงผลข้อมูลเกรดรายวิชาที่เคลงทะเบียนเรียนและภาพรวมของรายวิชาต่อปีเป็นรายบุคคลของนักศึกษาที่จบการศึกษาไปแล้วทั้งหมด

#### 4.9 การวิเคราะห์ภาพรวมเกรดเฉลี่ยสะสมและจำนวนรายวิชาที่ลงทะเบียนเรียนของนักศึกษาที่จบการศึกษาแล้ว

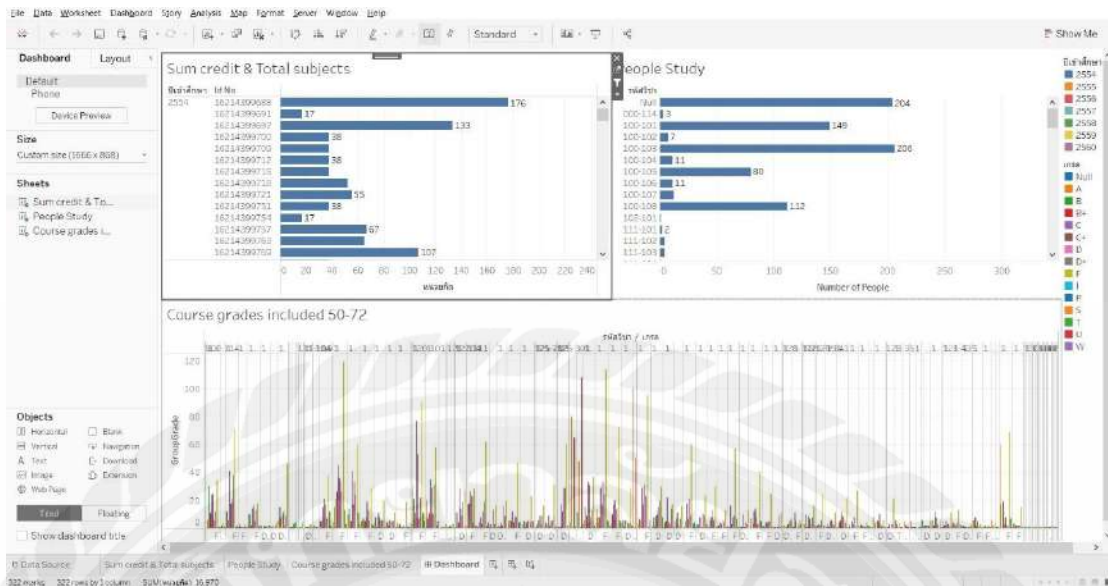
จะเป็นการวิเคราะห์ข้อมูลเพื่อหาเกรดเฉลี่ยสะสม โดยจะดูเกรดเฉลี่ยต่อภาคการศึกษาว่า มีนักศึกษาคนใดที่มีเกรดเฉลี่ยสะสมต่ำกว่า 2.00 ซึ่งจะนำไปเป็นโมเดลตัวอย่างสำหรับการวิเคราะห์ข้อมูลของนักศึกษาที่กำลังศึกษาอยู่ และยังวิเคราะห์การลงทะเบียนเรียนในแต่ละภาคการศึกษาว่าตรงตามเกณฑ์ของมหาวิทยาลัยหรือไม่ เพื่อหาปัจจัยและปัญหาที่เกิดขึ้นในการลงทะเบียนเรียนของนักศึกษาต่อไป



รูปที่ 4.15 แสดงภาพรวมเกรดเฉลี่ยสะสมและจำนวนรายวิชาที่ลงทะเบียนเรียนของนักศึกษาที่จบการศึกษาแล้ว

#### 4.10 การวิเคราะห์ภาพรวมผลการเรียนของแต่ละรายวิชาของนักศึกษาที่ตกออกหรือย้ายออก

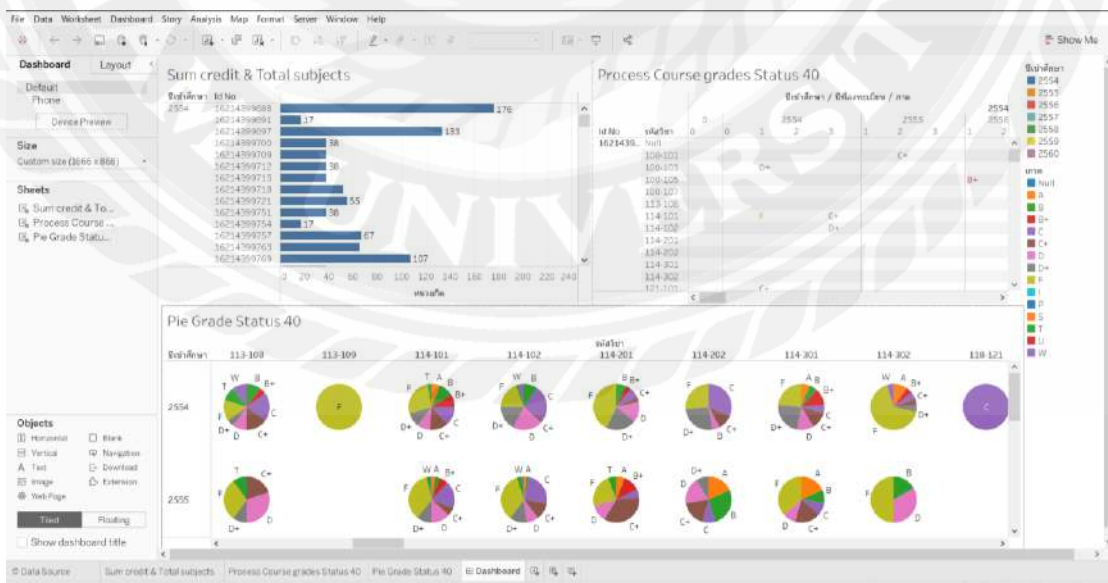
จะเป็นการวิเคราะห์เพื่อนำไปหาปัจจัยหรือสาเหตุของการตกออกและย้ายออกของนักศึกษาในระหว่างที่กำลังศึกษาอยู่ โดยดูจากเกรดของแต่ละรายวิชาที่ลงทะเบียนเรียนและจำนวนหน่วยกิตที่ได้ลงทะเบียนเรียนไปทั้งหมดตลอดระยะเวลาของการศึกษา เพื่อให้อาจารย์ที่ปรึกษาชั้นปีได้นำไปใช้ประโยชน์ในการดูแลนักศึกษาต่อไป



รูปที่ 4.16 แสดงภาพรวมผลการศึกษารายวิชาของแต่ละรายวิชาของนักศึกษาที่ตกออกหรือย้ายออก

#### 4.11 การวิเคราะห์ภาพรวมจำนวนเกรดของแต่ละรายวิชาและเกรดในแต่ละรายวิชาที่นักศึกษาลงทะเบียนเรียน

จะเป็นการวิเคราะห์เกรดในแต่ละรายวิชาของนักศึกษาที่ตกออกหรือย้ายออก ว่านักศึกษาคอนใดที่ได้เกรดใดบ้างในแต่ละรายวิชา เพื่อที่จะหาสาเหตุว่านักศึกษาคอนดังกล่าวได้ตกออกหรือย้ายออกเพราะผลการเรียนในรายวิชาที่ตนเรียนหรือไม่ และตั้งสมมติฐานความเป็นไปได้ว่าตกออกหรือย้ายออกด้วยสาเหตุใด



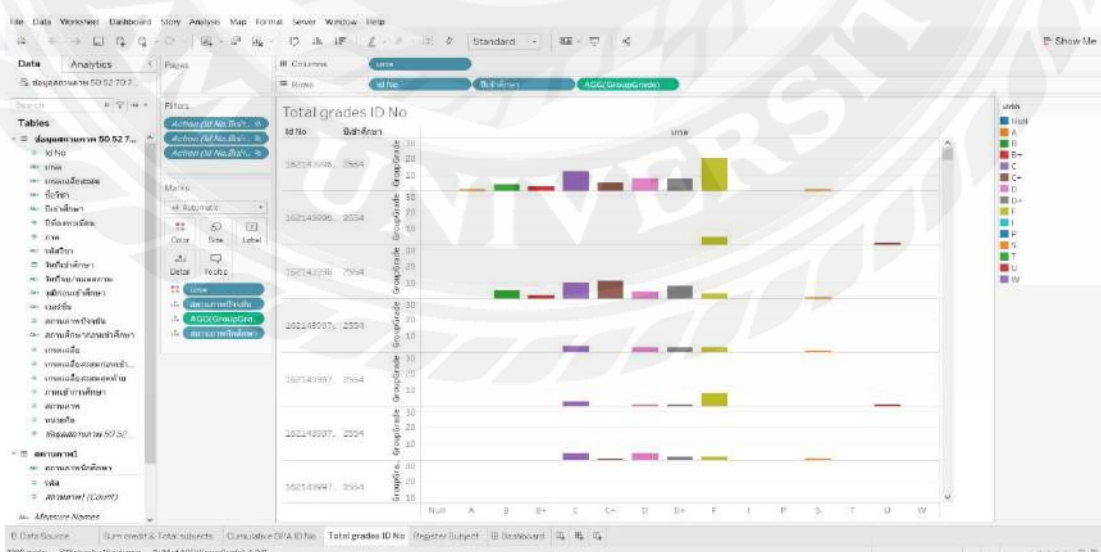
รูปที่ 4.17 แสดงภาพรวมจำนวนเกรดของแต่ละรายวิชาและเกรดในแต่ละรายวิชาที่นักศึกษาลงทะเบียนเรียน

### 4.12 การวิเคราะห์ภาพรวมของเกรดเฉลี่ยสะสมและจำนวนรายวิชาที่ลงทะเบียนเรียนของนักศึกษาที่ตบออกหรือย้ายออก

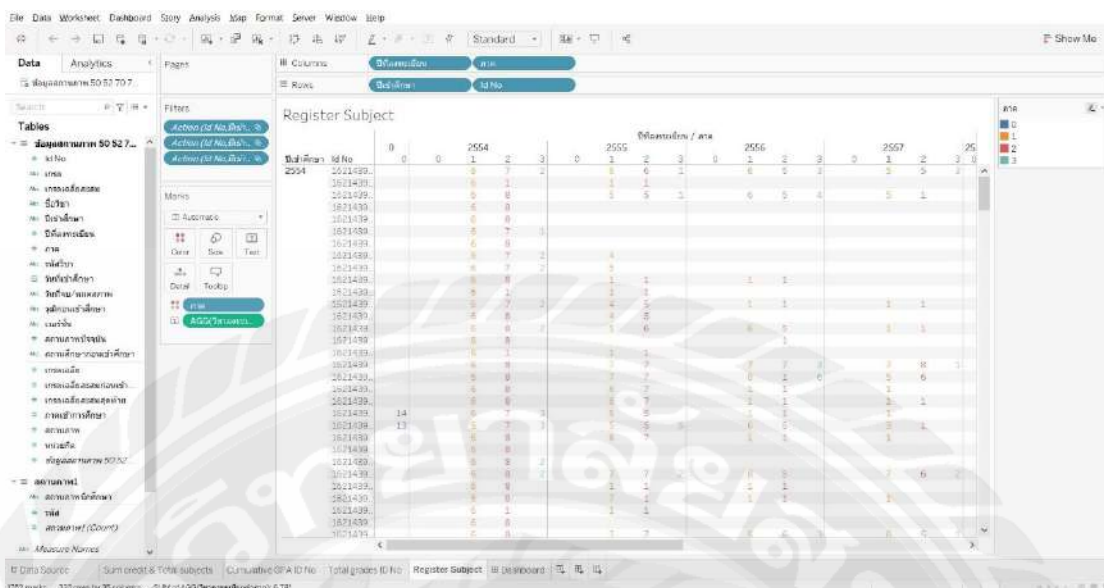
จะเป็นการวิเคราะห์และหาปัจจัยที่เกี่ยวข้องกับการตบออกหรือย้ายออกของนักศึกษาแต่ละคน ซึ่งจะใช้รหัสนักศึกษาของหัวข้อ Sum credit เป็นตัวกรองข้อมูล (Filter) เพื่อให้หน้าอื่น ๆ แยกข้อมูลได้ง่ายขึ้น การวิเคราะห์จะประกอบไปด้วย เกรดเฉลี่ยสะสมของแต่ละภาคการศึกษาแยกเป็นรายบุคคล เกรดทั้งหมดที่สอบได้ในแต่ละรายวิชา จำนวนวิชาที่ลงทะเบียนเรียนต่อภาคการศึกษา เพื่อนำมาวิเคราะห์ถึงสาเหตุที่นักศึกษตบออกหรือย้ายออก



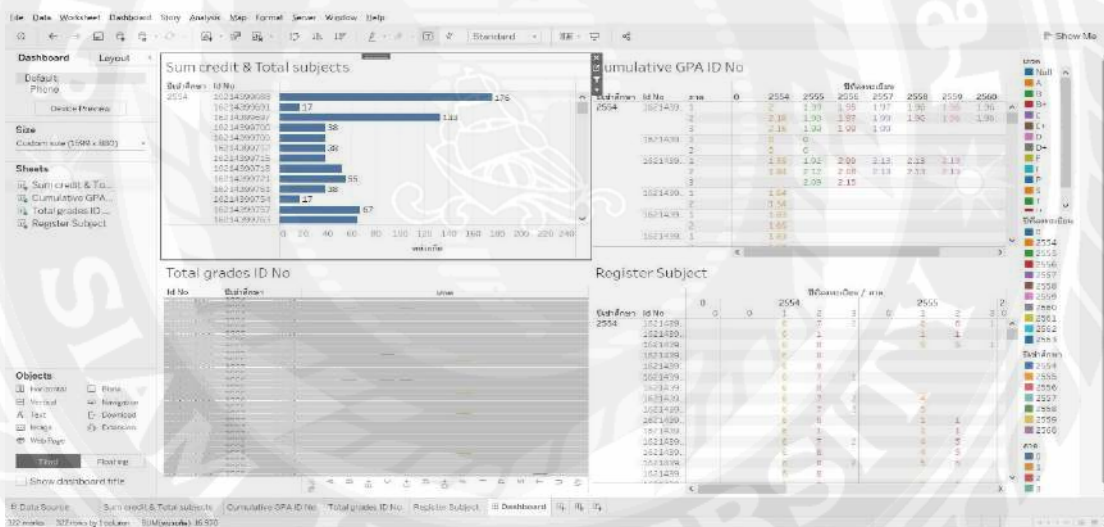
รูปที่ 4.18 แสดงผลลัพธ์เกรดเฉลี่ยสะสมต่อภาคการศึกษา



รูปที่ 4.19 แสดงผลลัพธ์ผลการเรียนเกรดโดยรวมของนักศึกษาแต่ละคน

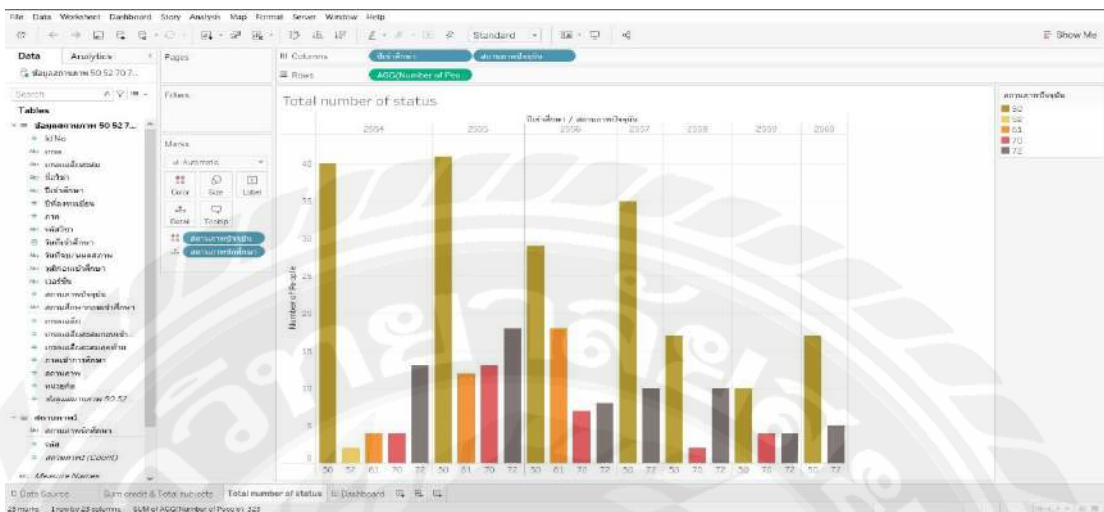


รูปที่ 4.20 แสดงผลลัพธ์จำนวนรายวิชาที่ลงทะเบียนในแต่ละภาคการศึกษา

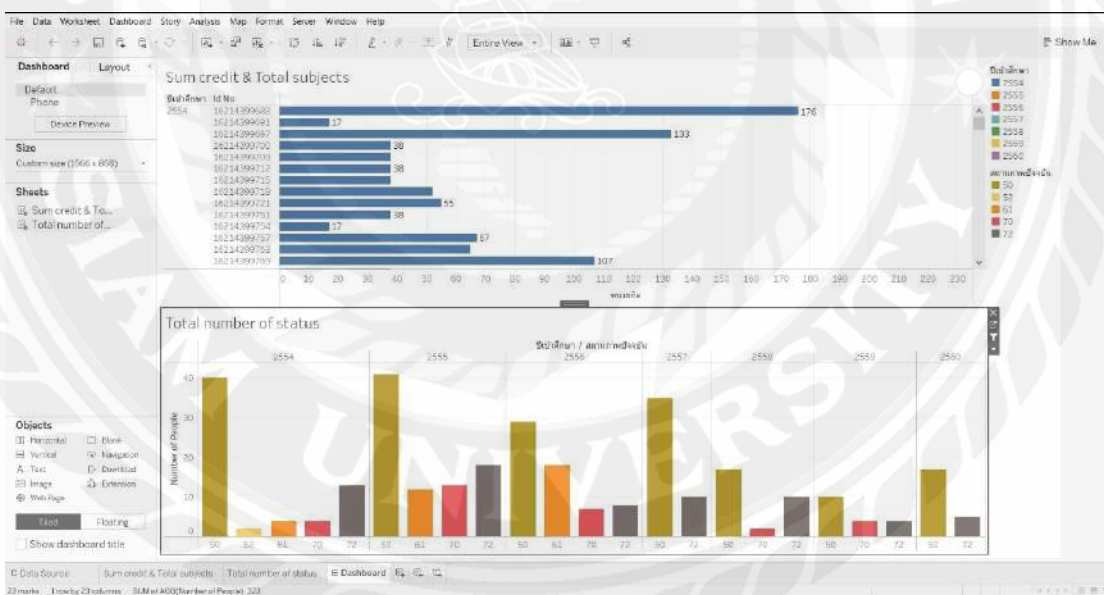


รูปที่ 4.21 แสดงภาพรวมของเกรดเฉลี่ยสะสมและจำนวนรายวิชาที่ลงทะเบียนเรียนของนักศึกษาที่ ตกลงหรือย้ายออก

### 4.13 การวิเคราะห์ภาพรวมจำนวนนักศึกษานักศึกษาตามสถานภาพของนักศึกษา ที่มีสถานภาพ นักศึกษาอยู่ระหว่าง 50 และ 72



รูปที่ 4.22 แสดงจำนวนนักศึกษานักศึกษาตามสถานภาพของนักศึกษา ที่มีสถานภาพนักศึกษาอยู่ระหว่าง 50 และ 72



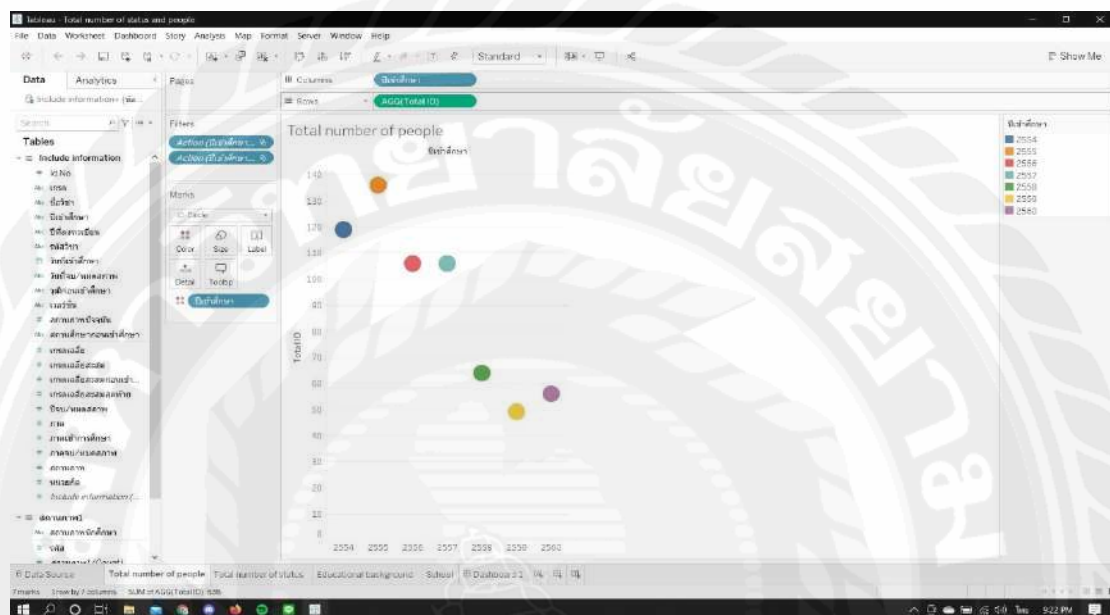
รูปที่ 4.23 แสดงผลสถานภาพนักศึกษาเป็นรายบุคคล

จากรูปที่ 4.23 เมื่อเลือกไปที่รหัสนักศึกษา กราฟทางด้านล่างจะเปลี่ยนไปโดยจะเป็นการ แสดงว่านักศึกษาคงดังกล่าวอยู่ในสถานภาพใด หรือเมื่อเลือกไปที่กราฟด้านล่างที่เป็นรหัส



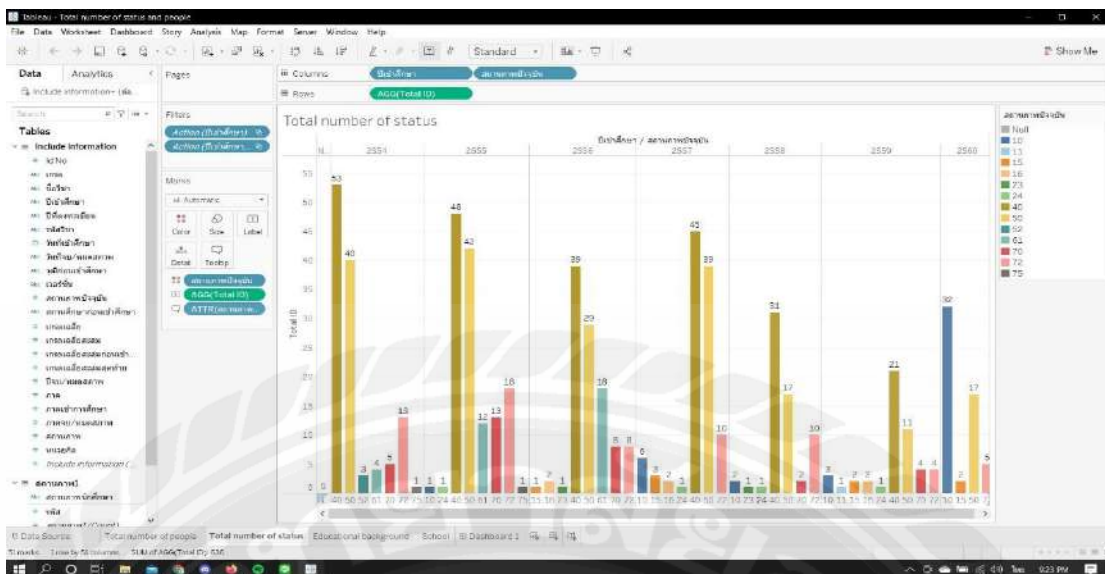
สถานภาพ จะทำการกรองรหัสนักศึกษาทั้งหมดที่เราเลือกในกราฟด้านล่าง ซึ่งกราฟด้านล่างจะแบ่งรหัสสถานภาพตามปีการศึกษาที่นักศึกษาเข้าศึกษาต่อในหลักสูตร

#### 4.14 การวิเคราะห์ภาพรวมสถานภาพต่าง ๆ ของจำนวนคนที่เข้าศึกษาต่อในแต่ละปีการศึกษา และจำนวนวุฒิการศึกษาที่จบก่อนเข้าศึกษาต่อในหลักสูตร



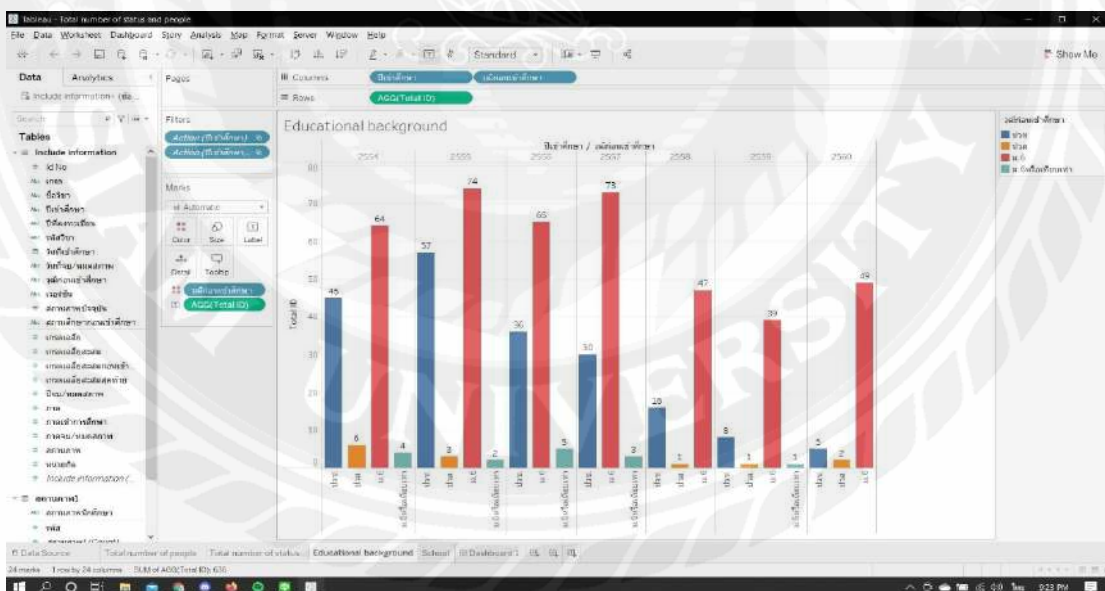
รูปที่ 4.24 แสดงจำนวนนักศึกษาที่เข้าศึกษาต่อในแต่ละปีการศึกษา

จากรูปที่ 4.24 จะเป็นการแสดงจำนวนนักศึกษาที่เข้าศึกษาต่อในหลักสูตรวิทยาการคอมพิวเตอร์ในแต่ละปีการศึกษามีจำนวนเท่าใด



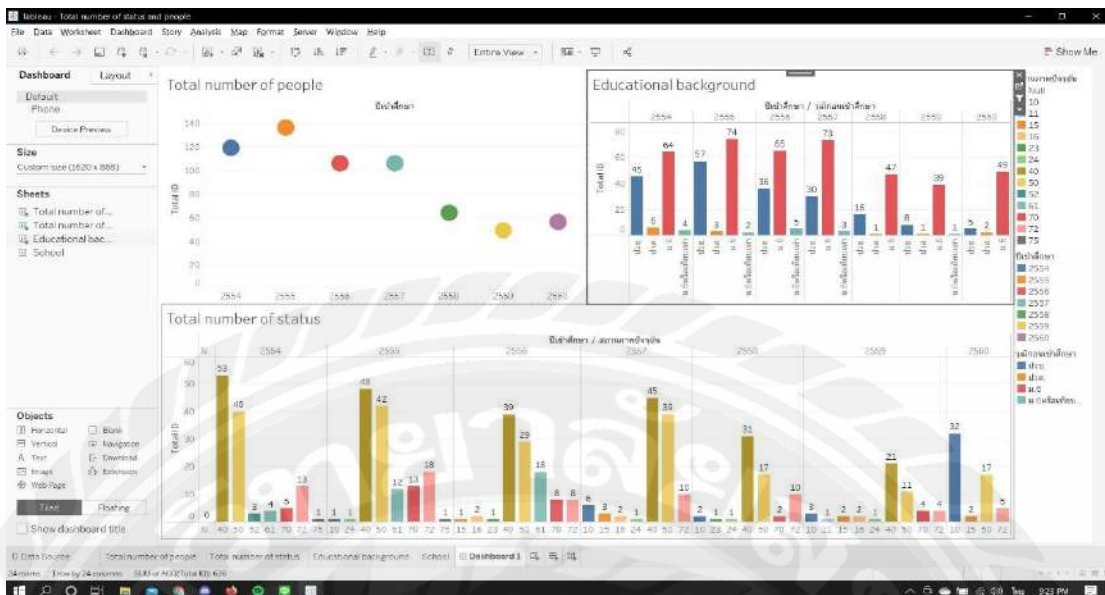
รูปที่ 4.25 แสดงจำนวนนักศึกษาแยกตามสถานภาพ

จากรูปที่ 4.25 จะเป็นการแยกข้อมูลในแต่ละสถานภาพของนักศึกษาจากรูปที่ 4.23 เพื่อที่จะแสดงผลต่างของแต่ละสถานภาพว่ามีนักศึกษาเข้าศึกษาต่อทั้งหมดจำนวนกี่คน จบการศึกษานานกี่คน ตกออกหรือย้ายออกจำนวนกี่คนในแต่ละปีการศึกษา



รูปที่ 4.26 แสดงจำนวนนักศึกษาแยกตามวุฒิการศึกษาที่จบก่อนเข้าศึกษาต่อในหลักสูตร

จากรูปที่ 4.26 จะเป็นการวิเคราะห์ปัจจัยด้านวุฒิการศึกษาที่จบมาก่อนเข้าศึกษาต่อในหลักสูตรวิทยาการคอมพิวเตอร์ว่ามีผลต่อการสำเร็จการศึกษาของนักศึกษาหรือไม่



รูปที่ 4.27 แสดงภาพรวมสถานภาพต่าง ๆ ของจำนวนคนที่เข้าศึกษาต่อในแต่ละปีการศึกษา และจำนวนวุฒิการศึกษาที่จบก่อนเข้าศึกษาต่อในหลักสูตร

จากรูปที่ 4.27 เป็นการแสดงผลโดยรวมของสถานภาพต่าง ๆ ของจำนวนคนที่เข้าศึกษาต่อในแต่ละปีการศึกษา และจำนวนวุฒิการศึกษาที่จบก่อนเข้าศึกษาต่อในหลักสูตร เพื่อแสดงให้เห็นข้อมูลที่วิเคราะห์มาข้างต้นได้ดีมากขึ้น และตอบปัญหาได้ดีมากขึ้น ในปัจจัยทางด้านวุฒิการศึกษาที่จบมาก่อนเข้าศึกษาต่อในหลักสูตร ซึ่งไม่ส่งผลต่อการสำเร็จการศึกษาของนักศึกษา

## บทที่ 5

### สรุปผลและข้อเสนอแนะ

#### 5.1 สรุปผลปริญญานิพนธ์

ผลที่ได้จากการวิเคราะห์ปัจจัยที่มีผลต่อการสำเร็จการศึกษาของนักศึกษาภาควิชาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยสยาม ด้วยการจัดกลุ่มข้อมูล (Clustering) ของนักศึกษาตามปัจจัยต่างๆ ที่ได้กำหนดทั้งหมด 5 ด้าน โดยใช้ข้อมูลย้อนหลังตั้งแต่ปีการศึกษา 2555 ถึง 2560 โดยนำเสนอด้วยแผนภาพข้อมูล (Data Visualization) นำมาจัดวางให้ดูง่ายด้วย Dashboard บนโปรแกรม Tableau ซึ่งจากผลการวิเคราะห์ได้ผลลัพธ์ดังนี้

- 1) วุฒิการศึกษาที่จบก่อนเข้าศึกษาต่อในหลักสูตรวิทยาการคอมพิวเตอร์ ไม่มีผลต่อการสำเร็จการศึกษา จากผลการวิเคราะห์อัตราการจบการศึกษาของนักศึกษาที่จบการศึกษา ปวช. หรือ ปวส. นั้นมีอัตราการจบที่สูงเมื่อเทียบกับการจำนวนนักศึกษาทั้งหมดที่เข้าศึกษาต่อในหลักสูตร
- 2) เกรดเฉลี่ยสะสม (GPA) มีผลต่อการสำเร็จการศึกษาตามหลักสูตร เนื่องจากตามข้อกำหนดของทางมหาวิทยาลัยแล้ว ถ้านักศึกษาคนใดที่มีเกรดเฉลี่ยสะสมน้อยกว่า 2.00 จะไม่สามารถลงทะเบียนเรียนได้เกิน 15 หน่วยกิต และถ้ามีเกรดเฉลี่ยสะสมสุดท้ายไม่เกิน 2.00 ทางมหาวิทยาลัยจะไม่สามารถให้นักศึกษาจบการศึกษาได้ ซึ่งจากผลการวิเคราะห์ทำให้เห็นว่านักศึกษาที่มีเกรดเฉลี่ยสะสมน้อยกว่า 2.00 ลงทะเบียนเรียนต่อภาคการศึกษาได้ต่ำกว่าแผนการเรียนที่กำหนดไว้ ทำให้สำเร็จการศึกษาล่าช้า เกินกว่า 4 ปี
- 3) รายวิชาบังคับที่ต้องเรียนก่อน (Prerequisite) ไม่มีผลต่อการสำเร็จการศึกษาตามหลักสูตร ถึงแม้ทางภาควิชาวิทยาการคอมพิวเตอร์นั้นจะเปิดสอนวิชาชีพเฉพาะด้านเพียงปีการศึกษาละ 1 ครั้ง แต่ก็ไม่ได้เป็นปัญหาต่อการสำเร็จการศึกษาตามหลักสูตร แต่จะส่งผลต่อการสำเร็จการศึกษาก็ต่อเมื่อนักศึกษาไม่สามารถเรียนแก้ไขในรายวิชาที่เป็นวิชาบังคับก่อนได้ในปีการศึกษาถัดไป
- 4) ผลการศึกษาตามรายวิชา มีผลต่อการสำเร็จการศึกษาตามหลักสูตร เนื่องจากนักศึกษามีผลการเรียนที่เป็น F หลายครั้งในรายวิชาเดิม ก็จำเป็นที่จะต้องลงทะเบียนเรียนซ้ำในรายวิชานั้นในภาคการศึกษาถัดไป หรือในปีการศึกษาถัดไป ทำให้ตารางการเรียนที่หลักสูตรกำหนดไว้นั้นมีการเปลี่ยนแปลงซึ่งอาจจะทำให้จบการศึกษา

ล่าช้าได้ โดยเฉพาะถ้ารายวิชานั้นเป็นรายวิชาบังคับที่ต้องเรียนก่อน (Prerequisite) ของรายวิชาอื่น

- 5) จำนวนหน่วยกิตสะสมที่ลงทะเบียนเรียนและผ่าน มีผลต่อการสำเร็จการศึกษาตามหลักสูตร โดยถ้านักศึกษามีการลงทะเบียนเรียนหน่วยกิตที่เกินจากที่กำหนดไว้จะทำให้รู้ว่านักศึกษาคนนั้นมีผลการเรียนที่เป็น F จนทำให้ไม่จบการศึกษาตามหลักสูตร หรือมีการลงทะเบียนเรียนหน่วยกิตที่น้อยทำให้รู้ว่าเกรดเฉลี่ยสะสมระหว่างศึกษานั้นต่ำกว่า 2.00 ทำให้ลงทะเบียนเรียนได้ไม่เกิน 15 หน่วยกิตต่อภาคการศึกษาจากปกติลงทะเบียนได้ 21 หน่วยกิตต่อภาคการศึกษา จึงส่งผลทำให้สำเร็จการศึกษาล่าช้า หรือมีความเสี่ยงที่จะไม่สำเร็จการศึกษา

## 5.2 ข้อเสนอแนะ

- 5.2.1 อาจารย์ที่ปรึกษาชั้นปีและอาจารย์ผู้สอนสามารถนำผลการวิเคราะห์นี้ไปใช้เพื่อการวางแผนการกำกับดูแลนักศึกษา เพื่อให้อัตราการต้อออกหรือย้ายสาขาลดลง และอัตราการสำเร็จการศึกษาตามเกณฑ์เพิ่มมากขึ้น รวมถึงอาจารย์ผู้สอนในแต่ละรายวิชามีการปรับเนื้อหาหรือวิธีการเรียนการสอนเพื่อให้นักศึกษาสามารถเรียนผ่านไปได้ ไม่มีผลการเรียนเป็น F เป็นจำนวนหลายครั้ง โดยมุ่งเน้นนักศึกษาจะต้องได้ความรู้ตามผลการเรียนรู้ที่ได้กำหนดไว้
- 5.2.2 ทางหลักสูตรสามารถนำผลการวิเคราะห์ไปใช้เพื่อปรับหลักสูตรให้สอดคล้องกับวัตถุประสงค์ที่ได้กำหนดไว้ในหลักสูตร เพื่อให้นักศึกษาจบการศึกษาตามเกณฑ์ในอัตราร้อยละ 100
- 5.2.3 เพื่อให้ผลการวิเคราะห์แม่นยำมากขึ้น สามารถดูข้อมูลได้ในหลากหลายมิติมากขึ้น ควรมีการรวบรวมข้อมูลอื่นๆ เพิ่มเติม ได้แก่ ข้อมูลคะแนนในแต่ละรายวิชา ซึ่งจะช่วยให้สามารถวิเคราะห์ถึงพฤติกรรมผลการเรียนของนักศึกษาแต่ละคนได้ เป็นต้น
- 5.2.4 เพื่อให้สามารถดูข้อมูลในเชิงลึกมากยิ่งขึ้น หรือเพื่อให้สามารถทำการพยากรณ์ได้ว่านักศึกษามีโอกาสจบการศึกษาตามเกณฑ์ 4 ปี มากหรือน้อย ควรมีการใช้อัลกอริทึมอื่นๆ หรือเครื่องมือที่มีความสามารถทางการพยากรณ์ได้

## บรรณานุกรม

จุฬารัตน์ งานชัยภูมิ. (2564 , 24 กรกฎาคม). Missing Data Management [เว็บบล็อก]. เข้าถึงได้จาก

<https://www.roots.tech/blog/products-services-1/how-to-manage-missing-data-25>

ดาต้าคิวเบเตอร์. (2564). *Validation Set สำคัญไฉน?*. เข้าถึงได้จาก

<https://medium.com/datacubator/validation-set-สำคัญไฉน-1abf22a68b75>

ธนาคารกสิกรไทย. (2564, 11 กันยายน). Big Data คืออะไร นำมาใช้ประโยชน์กับอะไรได้บ้าง [เว็บ

บล็อก]. เข้าถึงได้จาก <https://katalyst.kasikornbank.com/th/blog/Pages/what-is-big-data.html>

บิกเดต้าไทยแลนด์. (2564). *เลือกแผนภาพอย่างไรสำหรับการทำ Data Visualization*. เข้าถึงได้จาก

<https://bigdata.go.th/big-data-101/picking-chart-for-data-visualization/>

บิซอินไซต์. (2557, 14 มกราคม). การทำ Data Preparation อย่างมืออาชีพ [เว็บบล็อก]. เข้าถึงได้จาก

<https://bzinsight.wordpress.com/2014/06/11/การทำ-data-preparation-อย่างมืออาชีพ/>

บริษัท ริคโค จำกัด. (2564). *ประเภทของข้อมูลสำหรับการวิเคราะห์* . เข้าถึงได้จาก

<https://riccosmartdata.com/data-type/>

มัลลิกา. (2562). *การทำ Machine Learning ด้วย Clustering Model*. เข้าถึงได้จาก

<https://medium.com/tni-university/การทำ-machine-learning-ด้วย-clustering-model-2a3c392e7faa>

ศราวุธ รัตนตรัย. (2564). *Data Analytics คืออะไร*. เข้าถึงได้จาก

<https://work360.in.th/digital-marketing/data-analytics/what-is-data-analytics/>

อามเวอร์กรีนฟิช. (2563, 29 ตุลาคม). ทำความรู้จักกับ Data Visualization แบบรวดเร็วและเข้าใจ

ง่าย [เว็บบล็อก]. เข้าถึงได้จาก <https://blog.ourgreenfish.com/get-to-know-data-visualization-quick-and-easy-to-understand>

อิชิโปร. (2564). *วิธีจัดการกับ Outlier* . เข้าถึงได้จาก

<https://ichi.pro/th/kar-cadkar-kab-outliers-ni-chud-khxmul-60750975160994>

Tableau. (2564). *Tableau Analytics Workflow*. เข้าถึงได้จาก

[https://help.tableau.com/current/blueprint/en-us/bp\\_modern\\_analytics\\_workflow.htm/](https://help.tableau.com/current/blueprint/en-us/bp_modern_analytics_workflow.htm/)

Tableau. (2564). *Tableau Connect data*. เข้าถึงได้จาก

<https://help.tableau.com/current/pro/desktop/en-us/basicconnectoverview.htm>

