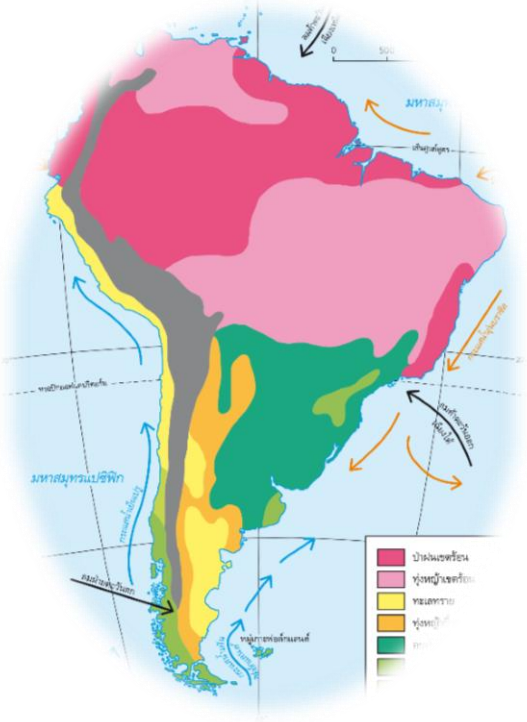
สถิติวิเคราะห์ทางภูมิศาสตร์



ผศ.รินทร์หทัย กิตติ์ธนารุจน์

บทที่ 1

หลักการเบื้องต้นทางสถิติและสถิติพรรณนาเชิงพื้นที่

เนื้อหาประกอบด้วย

1. แหล่งที่มาของข้อมูล
2. มาตราการวัดข้อมูล
3. การหาค่าร้อยละ
4. การหาค่าเปอร์เซนไทล์ , เดไซด์ และ ควอไทล์
5. การวัดแนวโน้มเข้าสู่ส่วนกลางและการวัดการกระจาย
6. การใช้โปรแกรมสำเร็จรูปวิเคราะห์ค่าสถิติพรรณนา

1. แหล่งที่มาของข้อมูล

ข้อมูล คือ เรื่องราว ข้อเท็จจริง หรือรายละเอียดของสิ่งต่าง ๆ ที่อยู่รอบตัว ซึ่งสามารถรับรู้ได้จากแหล่งข้อมูลต่าง ๆ แล้วนำข้อมูลไปใช้ ซึ่งในชีวิตประจำวันจะเกี่ยวข้องกับข้อมูลตลอดเวลา เช่น เมื่อตื่นนอนอาบน้ำก็จะได้กลิ่นของสบู่ รับประทานอาหารเช้าก็จะได้รสชาติของอาหาร ขณะเดินทาง ไปเรียนหรือไปทำงานก็จะเห็นสิ่งต่าง ๆ รอบตัว เช่น ป้ายประกาศ รถยนต์ยี่ห้อต่าง ๆ ขณะที่นั่งดู Facebook หรือ ค้นหาใน google ก็จะได้ข้อมูลที่เป็นตัวอักษร ตัวเลข และภาพจากการอ่าน

ข้อมูลสามารถจำแนกตามแหล่งข้อมูล ต้นกำเนิดหรือที่มาของข้อมูลต่าง ๆ (Source of Data) เช่น หนังสือ อินเทอร์เน็ต บุคคล และอาจรวมไปถึงสถานที่ต่าง ๆ ที่สามารถให้ข้อมูลแก่เราได้ เช่น พิพิธภัณฑ์ แหล่งท่องเที่ยว เป็นต้น ได้เป็น 2 ประเภท ดังนี้

1. แหล่งข้อมูลปฐมภูมิ (Primary Data) หมายถึง แหล่งข้อมูลที่ให้ข้อมูลแก่เราโดยตรง อาจเกิดจากการที่เราพบเห็นสิ่งต่าง ๆ ด้วยตนเอง แล้วจดบันทึกข้อมูลไว้ การสัมภาษณ์บุคคลต่าง ๆ หรือการสังเกตจากแหล่งข้อมูลโดยตรง รวมไปถึงการสำรวจข้อมูลโดยใช้แบบสอบถาม แบบสัมภาษณ์ แบบบันทึกการทดลอง หรือ แบบทดสอบ

ข้อดีของการเก็บรวบรวมข้อมูลด้วยตนเองคือ การได้ข้อมูลตรงตามความต้องการหรือวัตถุประสงค์ของการวิจัยแต่มีข้อเสียคือต้องใช้เวลาและเสียเวลาในการเก็บรวบรวมข้อมูล

2 แหล่งข้อมูลทุติยภูมิ (Secondary Data) หมายถึง แหล่งข้อมูลที่มีผู้รวบรวมไว้แล้วหรือหน่วยงานได้ทำการเก็บรวบรวมข้อมูลไว้แล้วสามารถนำข้อมูลเหล่านั้นมาใช้ได้ทันทีโดยไม่ต้องสอบถามหรือสังเกตจากแหล่งข้อมูลด้วยตนเอง เช่น ข้อมูลจากหนังสือ วิทยุ โทรทัศน์ อินเทอร์เน็ต ข้อมูลรายงานหรือสถิติของหน่วยงานต่าง ๆ เป็นต้น

ข้อดีของการใช้ข้อมูลทุติยภูมิคือผู้วิจัยไม่ต้องเสียเวลาในการจัดเก็บข้อมูลและประหยัดค่าใช้จ่าย แต่มีข้อเสียคือ ผู้วิจัยอาจได้ข้อมูลไม่ครบหรือไม่ตรงตามความต้องการของงานวิจัย

2. มาตราการวัดข้อมูล

1. มาตรานามบัญญัติ (Nominal Scale) เป็นการจัดคัดประเภทของข้อมูล เช่น การศึกษาสีของดิน อาจแยกออกเป็นดินสีแดง สีขาว สีดำ ดินร่วนกำแพงแสน ดินเหนียวองครักษ์ หมู่บ้านอาจแบ่งตามลักษณะเชื้อชาติ เช่น หมู่บ้านมอญ หมู่บ้านกะเหรี่ยง หมู่บ้านแม้ว หมู่บ้านเป็นขนาดใหญ่ ขนาดเล็ก หรือดูจากเชื้อชาติ

ในการหาค่าพื้นฐานทางสถิติจะหาได้เพียง **ความถี่ ร้อยละ ฐานนิยม** เท่านั้น

2. มาตราเรียงอันดับ (Ordinal Scale) เป็นการวัดที่ทราบว่าปรากฏการณ์หนึ่งต่างจากปรากฏการณ์หนึ่งไปยังอีกปรากฏการณ์หนึ่ง แต่เราไม่ทราบว่าปรากฏการณ์จะมากกว่าปรากฏการณ์หนึ่งเป็นกี่เท่า เช่น ที่ราบ ที่ราบสูง หรือ หมู่บ้านขนาดเล็ก ขนาดกลาง ขนาดใหญ่

สถิติพื้นฐานที่ใช้กับการวัดในระดับนี้ ได้แก่ **ฐานนิยม มัธยฐาน พิสัย เปอร์เซนต์ไทล์**

3. มาตราอันตรภาค (Interval Scale) เป็นมาตรวัดที่มีความละเอียดหน่วยการวัดคงที่ สามารถระบุความแตกต่างระหว่างสิ่งต่างๆได้ เป็นมาตรวัดที่มีความหมายเชิงปริมาณแต่ไม่ทราบอย่างแน่นอนว่าแต่ละช่วงห่างกันจริงๆเท่ากันหรือไม่ เช่นการวัดขนาดของเทศบาลตามขนาดของประชากร 10,000 คน 20,000 คน 30,000 คน 40,000 คน เป็นต้น

สถิติพื้นฐานที่ใช้กับการวัดในระดับนี้ ได้แก่
ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน

4. มาตราอัตราส่วน (Ratio Scale) เป็นมาตรวัดในระดับที่มีความละเอียดมากที่สุดเท่าที่มีอยู่ มีศูนย์ที่แท้จริงซึ่งเป็นจุดเริ่มต้น

สถิติพื้นฐานที่ใช้กับมาตราการวัดในระดับนี้สามารถใช้ได้ทุกวิธีที่มีอยู่

3. การหาค่าร้อยละ

ร้อยละ" หรือ "เปอร์เซนต์" คือ การเปรียบเทียบจำนวนที่ต้องการหากับจำนวนทั้งหมด โดยกำหนดให้จำนวนทั้งหมดเป็น 100 หรืออธิบายสั้น ๆ ก็คือ การเปรียบเทียบของปริมาณใดปริมาณหนึ่ง ต่อหนึ่งร้อย ถ้าเขียนเป็นเศษส่วนก็คือเศษส่วนที่มีส่วนเป็น 100 เสมอนั่นเอง และใช้ตัวสัญลักษณ์ % หรือเรียกอีกอย่างว่า "เปอร์เซนต์"

สูตรการหาค่าร้อยละ

$$P = \frac{F \times 100}{n}$$

เมื่อ

P แทน ร้อยละ

F แทน ความถี่ที่ต้องการแปลค่าให้เป็นร้อยละ

n แทน จำนวนความถี่ทั้งหมด

ตัวอย่างที่ 1

โรงเรียนแห่งหนึ่งมีนักเรียนทั้งหมด 400 คน ซึ่งแบ่งเป็นนักเรียนชายทั้งหมด 180 คน จงหาว่ามีนักเรียนชายทั้งหมดกี่เปอร์เซนต์

วิธีคิด

$$P = \frac{F \times 100}{n}$$

$$= \frac{180 \times 100}{400}$$

$$= 45$$

ดังนั้น มีนักเรียนชายทั้งหมด 45 เปอร์เซนต์

ตัวอย่างที่ 2

มีผลไม้ในสวนทั้งหมด 3,400 ผล เป็นแอปเปิล 1,904 ผล เป็นส้ม 1,088 ผล และเป็นมังคุด 408 ผล จงหาว่ามีผลไม้อย่างละกี่เปอร์เซนต์

วิธีคิด

4. การหาค่าเปอร์เซนไทล์ , เดไซด์ และ ควอไทล์

เปอร์เซ็นต์ไทล์ (Percentile) หมายถึงตำแหน่งที่แสดงให้ทราบว่ามีจำนวนร้อยละเท่าไหร่ของจำนวนคะแนนที่ค่าต่ำกว่าคะแนนที่ตำแหน่งนั้น เช่น นักศึกษาคนหนึ่งสอบวิชาสถิติได้คะแนน 54 คะแนน และคะแนน 54 นี้อยู่ตำแหน่งเปอร์เซนต์ไทล์ที่ 60 หมายความว่าร้อยละ 60 ของนักศึกษากลุ่มนั้นได้คะแนนวิชาสถิติต่ำกว่า 54

ขั้นตอนการคำนวณหา**เปอร์เซ็นต์ไทล์** (P_r) มีดังนี้คือ

ขั้นตอนที่ 1 เรียงข้อมูลจากน้อยไปหามาก

ขั้นตอนที่ 2 หาตำแหน่งของข้อมูลซึ่งตำแหน่งของ $P_r = \frac{r}{100}(n+1)$

เดไซล์ (Decile) หมายถึงตำแหน่งที่แสดงให้ทราบว่ามีจำนวนข้อมูลเท่าไรใน 10 ส่วนของจำนวนข้อมูลทั้งหมดที่มีค่าต่ำกว่าข้อมูลที่ตำแหน่งนั้น เช่น นักเรียนคนหนึ่งสอบวิชาภาษาไทยได้คะแนน 55 คะแนนและคะแนน 55 นี้ อยู่ในตำแหน่งเดไซล์ที่ 7(D7) หมายความว่า 7 ใน 10 ของนักศึกษากลุ่มนั้นได้คะแนนวิชาภาษาไทยต่ำกว่า 55 คะแนน และ 3 ในการ 10 ของนักศึกษากลุ่มนั้นได้คะแนนวิชาภาษาไทยสูงกว่า 55 คะแนน

โดยขั้นตอนการคำนวณหา D_r มีดังนี้คือ

<u>ขั้นตอนที่ 1</u> เรียงข้อมูลจากน้อยไปหามาก

<u>ขั้นตอนที่ 2</u> หาตำแหน่งของข้อมูลซึ่งตำแหน่งของ $D_r = \frac{r}{10}(N+1)$

ควอไทล์ (Quartile) หมายถึงตำแหน่งที่แสดงให้ทราบว่ามีจำนวนข้อมูลเท่าไรใน 4 ส่วนของจำนวนข้อมูลทั้งที่มีคำตากว่าข้อมูลที่ตำแหน่งนั้น เช่น นักศึกษาคนหนึ่งสอบวิชาภาษาไทยได้ 72 คะแนนและคะแนน 72 นั้นอยู่ในตำแหน่งควอรไทลที่ 3 (Q s) หมายความว่า 3 ใน 4 ของศึกษากลุ่มนั้นได้คะแนนวิชาไม่ต่ำกว่า 72 คะแนน และ 1 ใน 4 ของนักศึกษากลุ่มนั้นได้คะแนนวิชาภาษาไทยสูงกว่า 72 คะแนน

โดยขั้นตอนการคำนวณหา Q_r มีดังนี้คือ

<u>ขั้นตอนที่ 1</u> เรียงข้อมูลจากน้อยไปหามาก

<u>ขั้นตอนที่ 2</u> หาตำแหน่งของข้อมูลซึ่งตำแหน่งของ $Q_r = \frac{r}{4}(N+1)$

ตัวอย่างที่ 3 ในการวัดปริมาณสารคาร์บอนที่ต้นยางกักเก็บไว้ในพื้นที่แห่งหนึ่งเป็นดังนี้

15.77, 7.98, 8.32, 13.24, 11.46, 15.44, 21.23, 9.25 และ 22.39 เมตริกตัน/ไร่

จงหา P_{30} , D_4 , Q_2

วิธีคิด เรียงข้อมูลจากน้อยไปหามาก ดังนี้

7.98, 8.32, 9.25, 11.46, 13.24, 15.44, 15.77, 21.23, 22.39

(1) ตำแหน่งของ $P_r = \frac{r}{100}(N+1)$

$$P_{30} = \frac{30}{100}(9+1) = 3$$

ข้อมูลที่ตรงกับตำแหน่งที่ 3 คือ 9.25

ดังนั้น P_{30} เท่ากับ 9.25

(2) ตำแหน่งของ $D_r = \frac{r}{10}(N+1)$

$$D_4 = \frac{4}{10}(9+1) = 4$$

ข้อมูลที่ตรงกับตำแหน่งที่ 4 คือ 11.46

ดังนั้น D_4 เท่ากับ 11.46

(3) ตำแหน่งของ $Q_r = \frac{r}{4}(N+1)$

$$Q_2 = \frac{2}{4}(9+1) = 5$$

ข้อมูลที่ตรงกับตำแหน่งที่ 5 คือ 13.24

ดังนั้น Q_2 เท่ากับ 13.24

ตัวอย่างที่ 4 จากการวัดความชื้นสัมพัทธ์ในจังหวัดบุรีรัมย์ระหว่างวันที่ 1 – 14 ตุลาคม 2566 ได้ข้อมูลดังนี้

94, 98, 87, 72, 88, 96, 73, 84, 92, 90, 80, 77, 65 และ 78 เปอร์เซนต์

จงหา P_{40} , D_6 , Q_2

วิธีคิด

5. การวัดแนวโน้มเข้าสู่ส่วนกลางและการวัดการกระจาย

การวัดแนวโน้มเข้าสู่ส่วนกลาง

การวัดแนวโน้มเข้าสู่ส่วนกลางเป็นวิธีการทางสถิติเชิงพรรณนาที่ใช้ในการหาค่ากลางเพื่อใช้เป็นตัวแทนแสดงขนาดและลักษณะของข้อมูลชุดนั้น ประโยชน์ของการวัดแนวโน้มเข้าสู่ส่วนกลางคือทำให้ได้ตัวแทนของข้อมูลที่เป็นตัวเลขจำนวนเดียวที่แทนค่าทั้งหมดของข้อมูลชุดนั้นมาเสนอรายงาน โดยไม่จำเป็นต้องนำข้อมูลทั้งชุดมาพิจารณา ค่าวัดแนวโน้มเข้าสู่ส่วนกลางมีอยู่หลายค่าด้วยกัน เช่น ค่าเฉลี่ย ค่ามัธยฐาน และค่าฐาน นิยม เป็นต้น

ซึ่งการคำนวณหาค่ากลางทั้งสามชนิดนี้โดยทั่วไปแบ่งออกได้เป็น 2 กรณีใหญ่ ๆ คือ

1. การหาค่ากลางของข้อมูลที่ไม่ได้แจกแจงความถี่ (Ungrouped data) ซึ่งค่าที่ได้เป็นค่ากลางที่ถูกต้องแน่นอนของข้อมูลชุดนั้น

2. การหาค่ากลางของข้อมูลที่แจกแจงความถี่แล้ว (Grouped Data) ซึ่งค่าที่ได้เป็นค่ากลางโดยประมาณของข้อมูลชุดนั้น

แต่ในบทนี้จะขอกล่าวถึงเพียงการหาค่ากลางของข้อมูลที่ไม่ได้แจกแจงความถี่เท่านั้น

1. ค่าเฉลี่ย

ค่าเฉลี่ยเลขคณิต (Mean)

ประชากร แทนด้วย μ

$$\mu = \frac{\sum_{i=1}^{N} X_i}{N} = \frac{\left(X_1 + X_2 + \ldots + X_N\right)}{N}$$

ตัวอย่าง แทนด้วย $\bar{X}$

$$\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{\left(x_1 + x_2 + \ldots + x_n\right)}{n}$$

ค่าเฉลี่ยถ่วงน้ำหนัก (weighted mean)

ถ้าข้อมูลแต่ละค่ามีความสำคัญไม่เท่ากันหรือมีค่าถ่วงน้ำหนักไม่เท่ากัน จะต้องนำน้ำหนักมาคิดในการหาค่าเฉลี่ยด้วยเสมอ โดยการนำน้ำหนักของข้อมูลแต่ละค่ามาคูณกับข้อมูลตัวนั้น จากนั้นนำมารวมกันทุกตัวแล้วหาร ด้วยผลรวมของน้ำหนักทั้งหมด สามารถคำนวณได้จากสูตรต่อไปนี้

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i W_i}{\sum_{i=1}^{n} W_i}$$

เมื่อ W_i คือ น้ำหนักของ X_i

2. มัธยฐาน (Median , Md)

มัธยฐาน เป็นค่าที่อยู่กึ่งกลางของข้อมูล เมื่อเรียงลำดับข้อมูลจากน้อยไปหามาก และแบ่งครึ่งข้อมูลออก เป็นสองส่วนเท่ากัน ค่ามัธยฐานอาจจะเป็นนค่าใดค่าหนึ่งของชุดข้อมูล หรือมีค่าอยู่ระหว่างสองค่าใด ๆ ก็ได้

3. ฐานนิยม (Mode , M_o)

ฐานนิยม คือ ค่าของข้อมูลที่เกิดขึ้นบ่อยที่สุดหรือมีความถี่สูงสุด

ตัวอย่างที่ 5 ในการทดสอบทักษะกระบวนการทางภูมิสารสนเทศของนักศึกษาคนหนึ่งซึ่งมีคะแนนการทดสอบและความสำคัญของคะแนนทั้งหมดรวม 5 ด้าน จากคะแนนเต็ม 100 คะแนน ดังข้อมูลในตารางจงหาคะแนนเฉลี่ยของการทดสอบทักษะกระบวนการทางคณิตศาสตร์ของนักเรียนคนนี้

ด้านที่	ทักษะกระบวนการ	คะแนนที่สอบได้	ความสำคัญของคะแนน
1	การแก้ปัญหา	54	30
2	การใช้เหตุผล	65	20
3	การสื่อสาร การนำเสนอ	70	15
4	การเชื่อมโยงความรู้ทางภูมิสารสนเทศ	55	20
5	ความคิดริเริ่มสร้างสรรค์	75	15
			100

วิธีคิด จากสูตร

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i W_i}{\sum_{i=1}^{n} W_i}$$

$$= \frac{30(54)+20(65)+15(70)+20(55)+15(75)}{30+20+15+20+15}$$

$$= \frac{6,195}{100} = 61.95$$

การวัดการกระจาย (Measures of Dispersion)

การวัดการกระจาย เป็นสถิติประเภทหนึ่งที่คำนวณออกมาเป็นตัวเลข เพื่อใช้อธิบายลักษณะการกระจายของข้อมูล การที่ข้อมูลชุดหนึ่ง ๆ ประกอบด้วยคะแนนที่มีค่าต่าง ๆ กันเราเรียกว่า เป็นข้อมูลที่มีการกระจาย ถ้าข้อมูลชุดนั้นประกอบด้วยคะแนนที่มีค่าต่างกันมาก เรียกว่า เป็นข้อมูลที่มีกระจายมาก ถ้าข้อมูลชุดนั้นประกอบด้วยคะแนนที่มีค่าต่างกันน้อย เรียกว่า เป็นข้อมูลที่มีการกระจายน้อย และถ้าข้อมูลชุดนั้นประกอบด้วยคะแนนที่มีค่าเท่ากันหมด เรียกว่า เป็นข้อมูลที่ไม่มีการกระจาย

สถิติสำหรับการคำนวณหาค่าการกระจายมีอยู่หลายวิธีดังนี้

1. พิสัย หมายถึง การหาการกระจายของข้อมูลโดยนำข้อมูลที่มีค่าสูงที่สุด ลบกับข้อมูลที่มีค่าต่ำที่สุด เพื่อให้ได้ค่าที่เป็นช่วงของการกระจาย ซึ่งสามารถบอกถึงความกว้างของข้อมูลชุดนั้นๆ สำหรับสูตรที่ใช้ในการหาพิสัยคือ พิสัย (R) = $X_{max} - X_{min}$

2. ความแปรปรวน เป็นค่าที่นิยมนำมาใช้วัดการกระจายของข้อมูลมากที่สุด ซึ่งค่านี้คำนวณมาจากค่าแตกต่างระหว่างค่าของข้อมูลแต่ละค่ากับค่าเฉลี่ย ซึ่งถ้าค่าที่ได้เป็นศูนย์ แสดงว่าข้อมูลชนิดนั้นไม่มีการกระจายเลยหรือทุกค่ามีค่าเท่ากับค่ากลางนั่นเอง

จากสูตร

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

3. ค่าส่วนเบี่ยงเบนมาตรฐาน (standard deviation: SD) คือค่ารากที่ 2 ของความแปรปรวน

ตัวอย่างที่ 6 ในการวัดปริมาณน้ำฝนประจำปี 2555 – 2566 วัดเป็นมิลลิเมตร ที่จังหวัดบุรีรัมย์และสุรินทร์เป็นดังนี้

ปี พ.ศ.	บุรีรัมย์	สุรินทร์
2555	1,061	1,540
2556	1,115	1,864
2557	981	1,703
2558	865	1,600
2559	1,354	876
2560	1,260	1,320
2561	1,169	1,136
2562	1,563	1,885
2563	1,473	1,600
2564	1,364	1,516
2565	2,032	1,577
2566	981	1,501

จงหาค่าเฉลี่ย ค่ามัธยฐาน ค่าฐานนิยม และค่าเบี่ยงเบนมาตรฐานของปริมาณน้ำฝนในจังหวัดบุรีรัมย์และสุรินทร์

วิธีคิด ปริมาณน้ำฝนจังหวัดบุรีรัมย์

ค่าเฉลี่ย จากสูตร

$$\overline{X} = \frac{\sum_{i=1}^{N} X_i}{n} = \frac{\left(X_1 + X_2 + \ldots + X_n\right)}{n}$$

$$= \frac{\left(1{,}061 + 1{,}115 + 981 + \ldots + 981\right)}{12} = \frac{15{,}218}{12} = 1{,}268.17$$

ดังนั้น ปริมาณน้ำฝนเฉลี่ยจังหวัดบุรีรัมย์เท่ากับ 1,268.17 มิลลิเมตร

ค่ามัธยฐาน

วิธีคิด เรียงข้อมูลจากน้อยไปหามาก ดังนี้

865, 981, 981, 1,061, 1,115, **1,169, 1,260,** 1,354, 1,364, 1,473, 1563, 2,032

ดังนั้น ค่ามัธยฐานคือ

$$\frac{1,169+1,260}{2}$$

$$=\frac{2,429}{2}=1,214.5$$

ค่าฐานนิยม มีค่าเท่ากับ 981

ค่าเบี่ยงเบนมาตรฐาน

จากสูตร

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

$$= \sqrt{\frac{(865-1,268.17)^2 + (981-1,268.17)^2 + (981-1,268.17)^2 + \ldots + (2,032-1,268.17)^2}{12-1}}$$

$$= \sqrt{\frac{42,919.41 + 23,461.05 + 82,466.61 + \ldots + 82,466.81}{11}}$$

$$= \sqrt{\frac{1,132,627.67}{11}}$$

$$= \sqrt{102,966.15}$$

$$= 320.88$$

ดังนั้นจะได้ค่าเบี่ยงเบนมาตรฐาน (S.D.) เท่ากับ 320.88

****ให้นักศึกษาลองคำนวณของจังหวัดสุรินทร์****

6. การใช้โปรแกรมสำเร็จรูปวิเคราะห์ค่าสถิติพรรณนา

การวิเคราะห์ข้อมูลผ่านโปรแกรม SPSS จะประกอบด้วย 2 ส่วนดังนี้

1. การสร้างแฟ้มข้อมูล
2. การประมวลผลข้อมูล

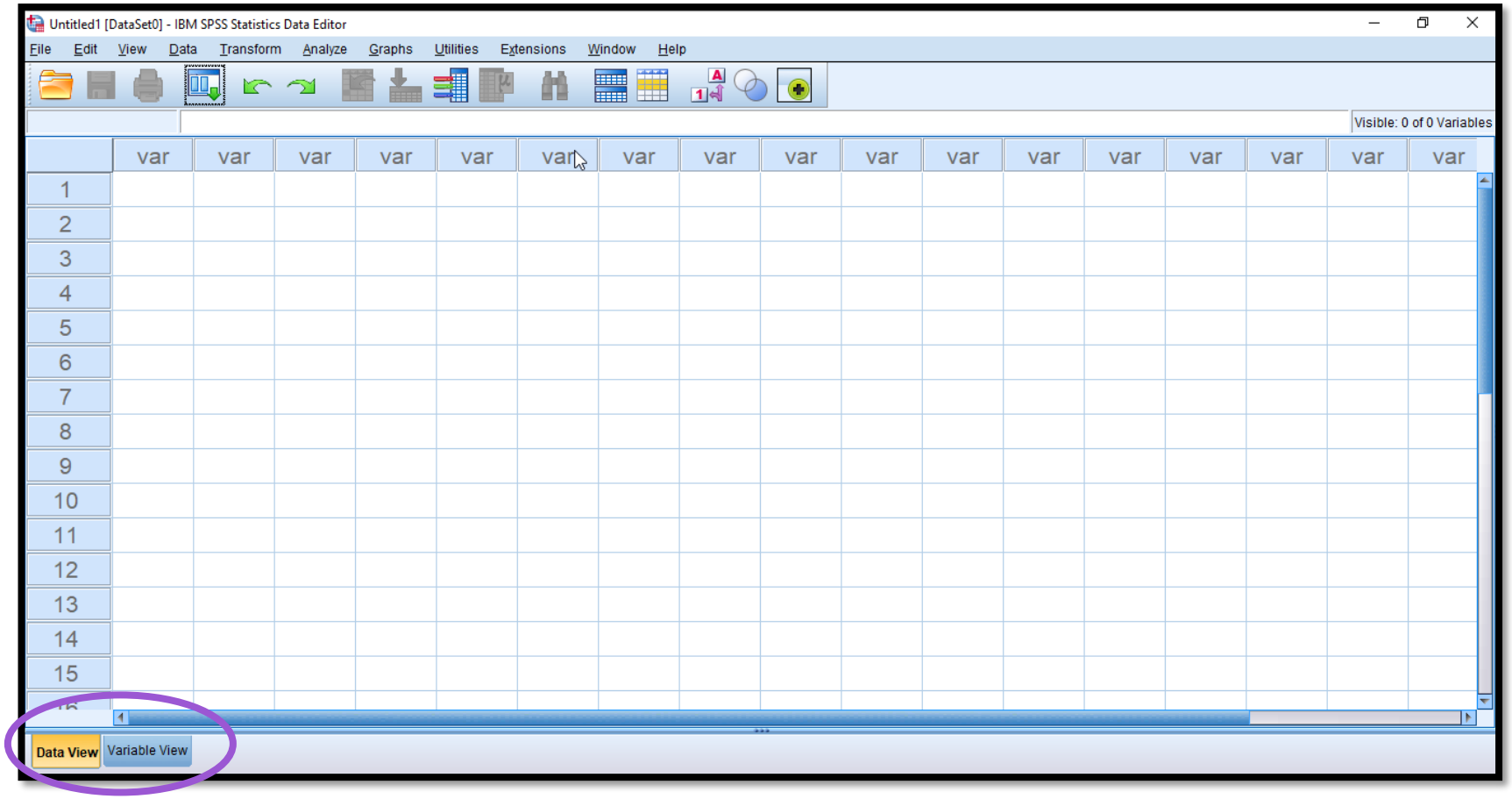
1. การสร้างแฟ้มข้อมูล

เปิดโปรแกรม Start > program > SPSS > SPSS for window

เมื่อเปิดโปรแกรมขึ้นมาแล้วจะพบกับหน้าต่างของ DATA EDITOR โดยด้านล่างจะประกอบด้วย

- Data View สำหรับทำการลงข้อมูล

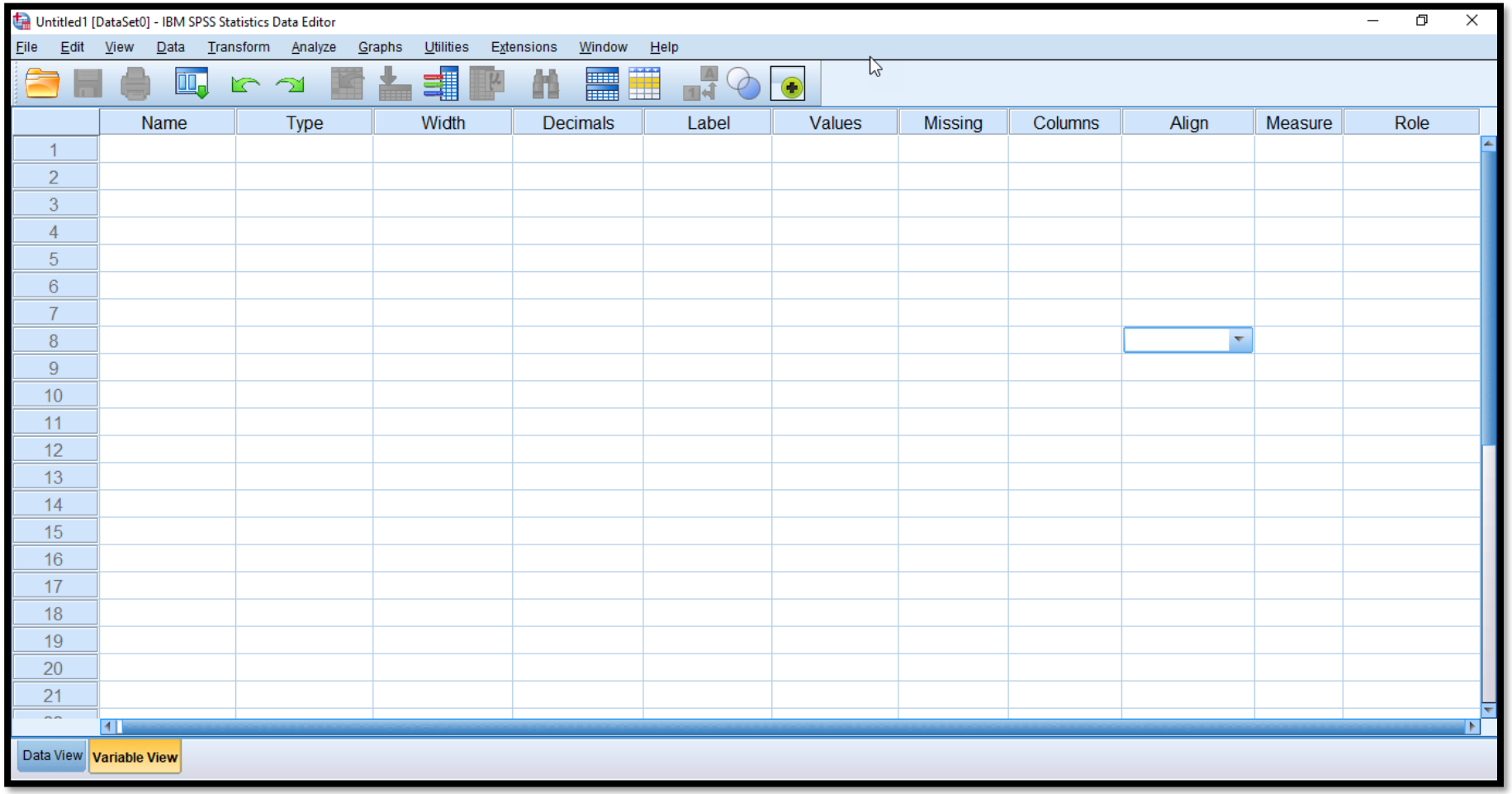
- Variable View สำหรับการกำหนดลักษณะตัวแปรของข้อมูล



ขั้นตอนการสร้างแฟ้มข้อมูล ประกอบด้วย

- การกำหนดคุณลักษณะต่าง ๆ ของตัวแปรใน Spread sheet ของ Variable View
- เมื่อกำหนดคุณลักษณะต่าง ๆ เรียบร้อยแล้ว จึงทำการกรอกข้อมูลของตัวแปรแต่ละตัวลงใน Spread sheet ของ Data View

ความหมายของ Menu bar ใน Spread sheet ของ Variable View มีรายละเอียดดังนี้



1) “Name” สำหรับการก าหนดชื่อของตัวแปร ความยาวไม่เกิน 7-10 ตัวอักษร

2) “Type” สำหรับกำหนดประเภทของตัวแปรหลัก

- Numeric สำหรับข้อมูลที่เป็นประเภทตัวเลข
- String สำหรับข้อมูลที่เป็นประเภทข้อความ

3) “Width” สำหรับกำหนดความกว้างของตัวแปรหรือ จำนวนตัวอักษรที่ต้องการให้ใส่ได้ใน “Values”

4) “Decimals” สำหรับกำหนดจำนวนทศนิยม

5) “Label” สำหรับกำหนดคำอธิบายตัวแปรหรือชื่อเต็ม ของตัวแปรนั้น ๆ จะใช้ในกรณีที่ผู้วิจัยกำหนดชื่อตัวแปร ใน Column “Name” เป็นอักษรย่อ แล้วต้องการอธิบายหรือขยายความไว้ เช่น Name ระบุเป็น ID ใน Label จะระบุเป็นลำดับที่ เป็นต้น

6) “Values” สำหรับกำหนดค่าตัวแปร เช่น ตัวแปร “เพศ” กำหนดให้ 1 แทน ชาย ให้ 2 แทนหญิง เป็นต้น

7) “Missing” สำหรับกำหนดค่าของข้อมูลที่แสดงว่าผู้ตอบแบบสอบถามไม่ได้ตอบในข้อคำถามนั้น เช่น ตัวแปร ‘เพศ’ กำหนดค่า Missing = 9 นั่นคือ ถ้ามีผู้ที่ไม่ระบุเพศในแบบสอบถาม จะบันทึกข้อมูลเป็น 9 โดยทั่วไปจะกำหนดค่า Missing ให้ไม่ซ้ำกับค่าของตัวแปร “Values”

8) “Columns” สำหรับกำหนดความกว้างของคอลัมน์ ในหน้า Data View

9) “Align” สำหรับจัดตำแหน่งการวางข้อความ

10) “Measure” สำหรับกำหนดมาตรวัดของตัวแปร >> Scale , Ordinal หรือ Nominal

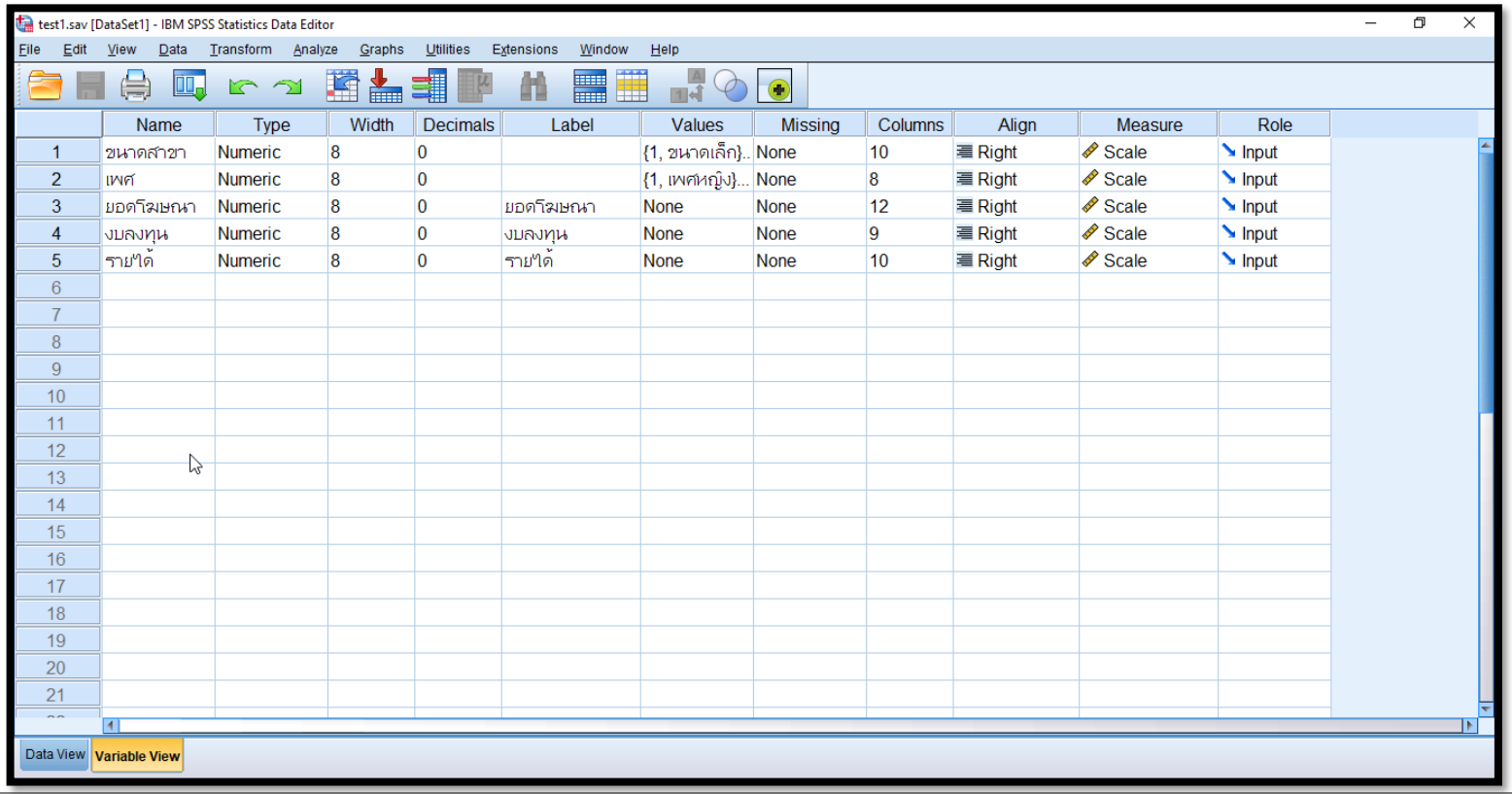
>> ข้อมูลเชิงคุณภาพ (Qualitative data) คือ ข้อมูลที่วัดค่าออกมาเป็นตัวเลขโดยตรงไม่ได้ จำแนกเป็น

- Nominal scale เป็นข้อมูลที่ถูกจัดแบ่งออกเป็นกลุ่ม ๆ เช่น เพศ-ชาย/หญิง , สถานภาพสมรส – โสด/ แต่งงาน /หย่า และหม้าย ซึ่งค่าข้อมูลยังไม่สามารถนำมาเรียงลำดับความสำคัญได้

- Ordinal scale เป็นมาตรวัดของข้อมูลที่ถูกจัดเป็นกลุ่ม ๆ โดยสามารถเรียงลำดับความสำคัญของแต่ละกลุ่มได้ แต่ยังไม่สามารถนำค่าข้อมูลมาเปรียบวัดขนาดความแตกต่างได้

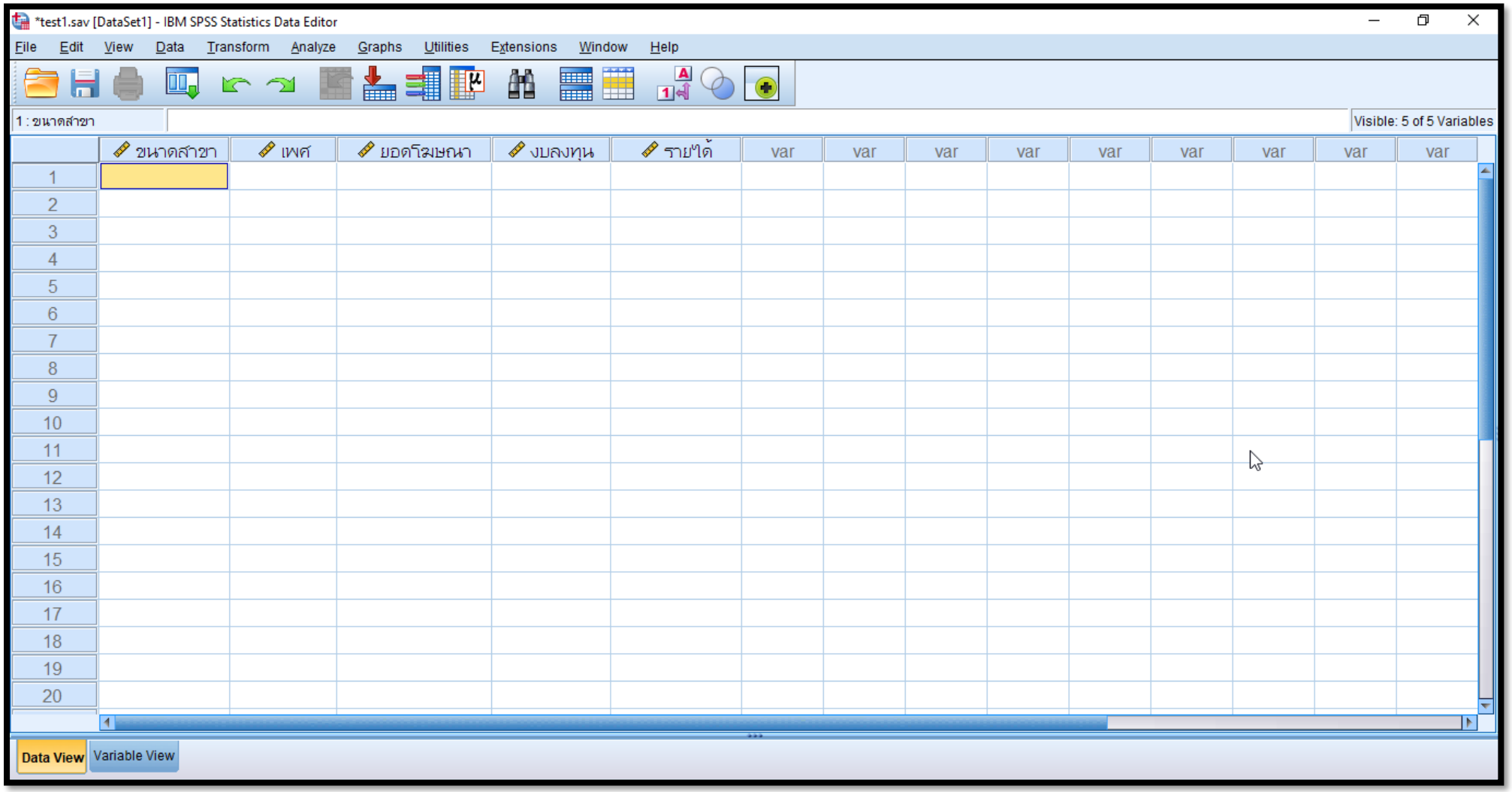
>> ข้อมูลเชิงปริมาณ (Quantitative data) เป็นข้อมูลที่วัดค่าเป็นตัวเลขได้ เช่น น้ำหนัก เงินเดือนที่ได้รับ เป็นต้น หากข้อมูลในแบบสอบถามเป็นข้อมูลเชิงปริมาณดังกล่าว ให้เลือก Scale

ตัวอย่างการกำหนดลักษณะตัวแปรใน Spread sheet ของ Variable View



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	ขนาดสาขา	Numeric	8	0		{1, ขนาดเล็ก}..	None	10	Right	Scale	Input
2	เพศ	Numeric	8	0		{1, เพศหญิง}...	None	8	Right	Scale	Input
3	ยอดโฆษณา	Numeric	8	0	ยอดโฆษณา	None	None	12	Right	Scale	Input
4	งบลงทุน	Numeric	8	0	งบลงทุน	None	None	9	Right	Scale	Input
5	รายได้	Numeric	8	0	รายได้	None	None	10	Right	Scale	Input
6											
7											
8											
9											
10											
11											
12											
13											
14											
15											
16											
17											
18											
19											
20											
21											

Spread sheet ของ Data View สำหรับทำการกรอกข้อมูล

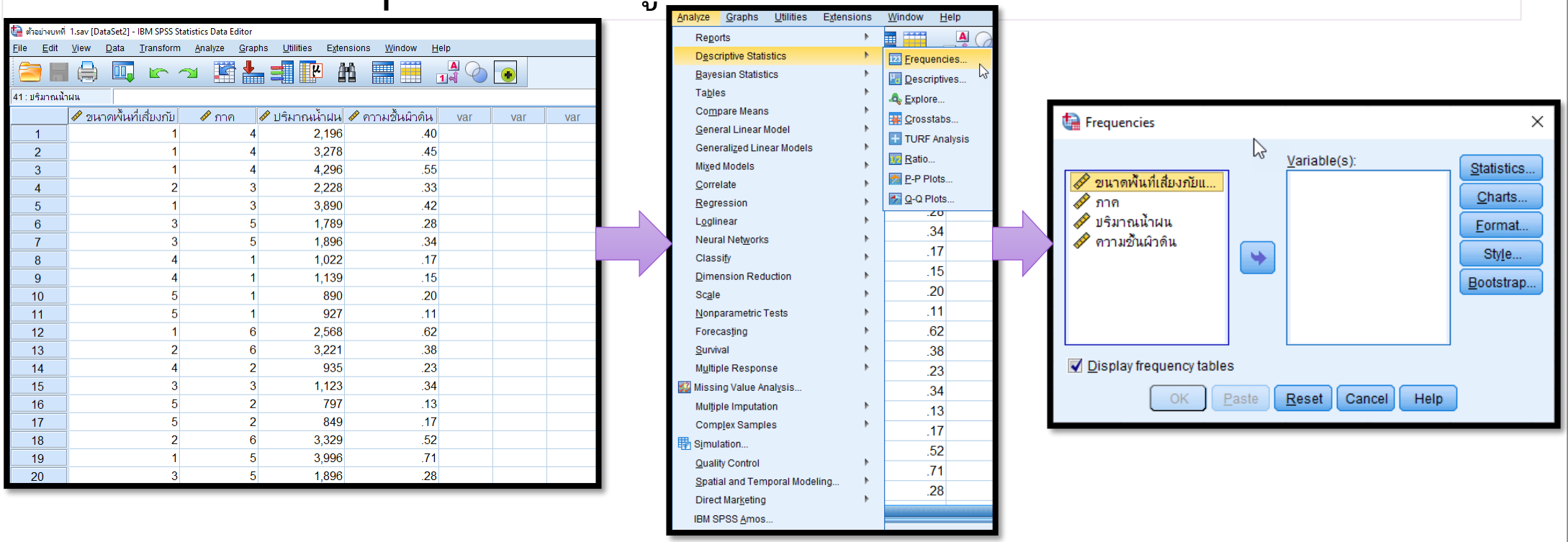


การประมวลผลของข้อมูล โดยใช้การวิเคราะห์ข้อมูลทางสถิติ

การหาค่าความถี่ ร้อยละ (%)

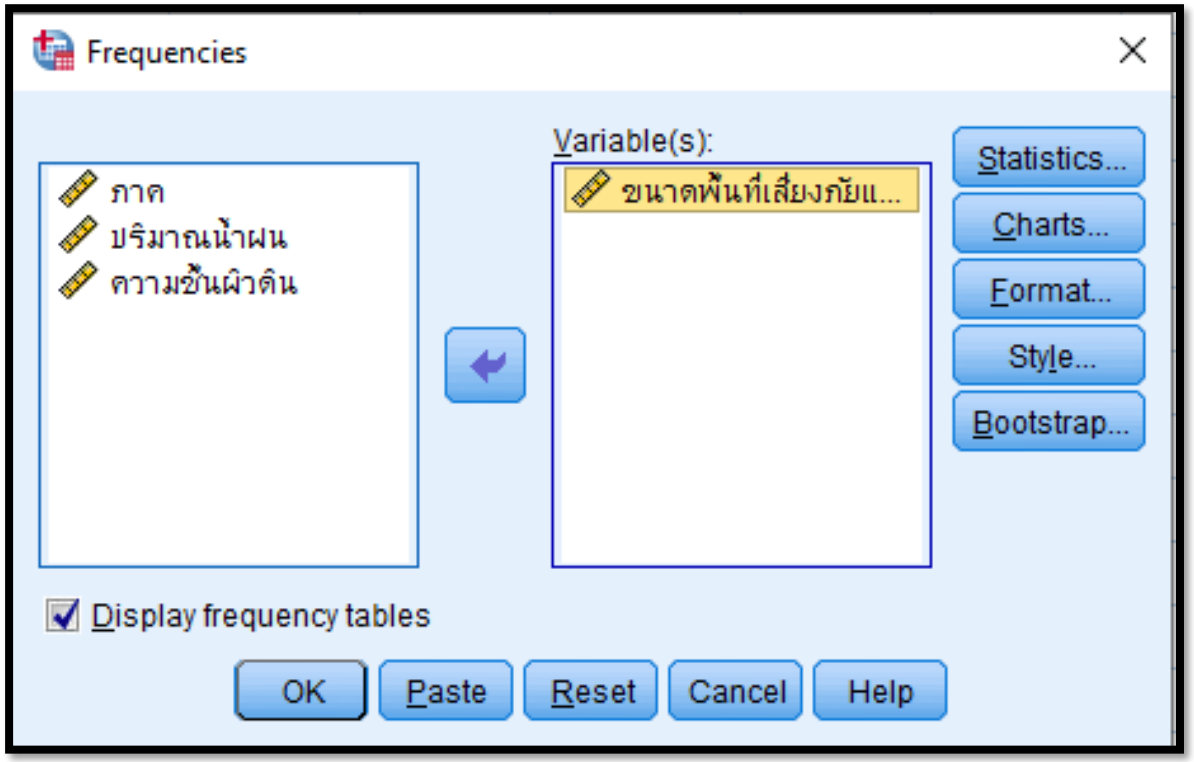
การหาค่าความถี่ ร้อยละ (%) เลือกเมนู Analyze > Descriptive

Statistics > Frequencies ดังรูป



- เลือกชุดตัวแปรที่อยู่ทางด้านซ้ายมาใส่ไว้ในกรอบขวา (สามารถเลือกได้มากกว่า 1 ตัวแปร) > OK

ในที่นี้เลือกตัวแปร ขนาดพื้นที่เสี่ยงภัย ดังรูป



คลิก Ok จะได้ผลลัพธ์ ดังนี้

Frequencies

[DataSet2] D:\เอกสารประกอบการสอน\สถิติวิเคราะห์ทางภูมิศาสตร์\ตัวอย่างบทที่ 1.sav

Statistics

ขนาดพื้นที่เสี่ยงภัยแล้ง

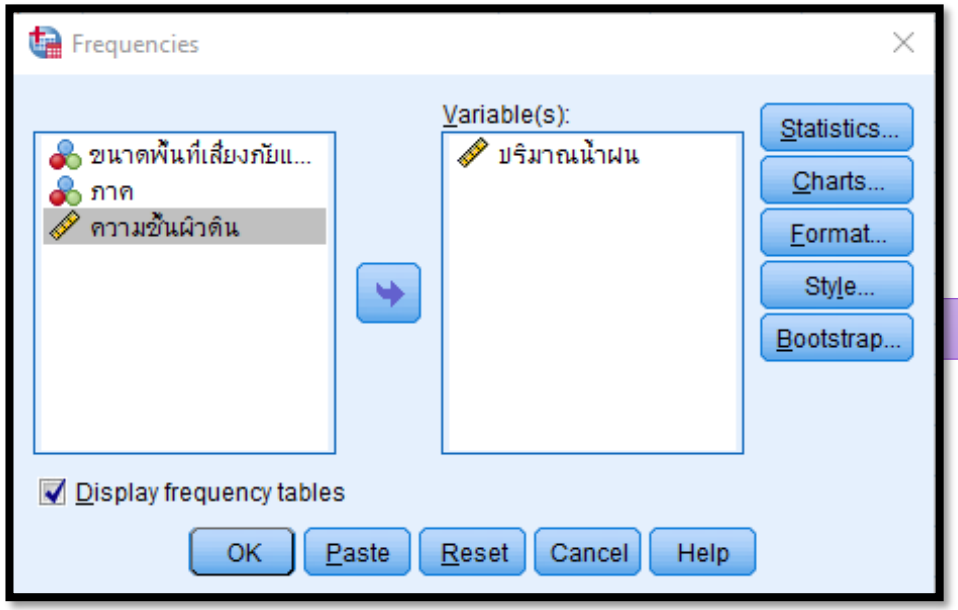
N	Valid	40
	Missing	0

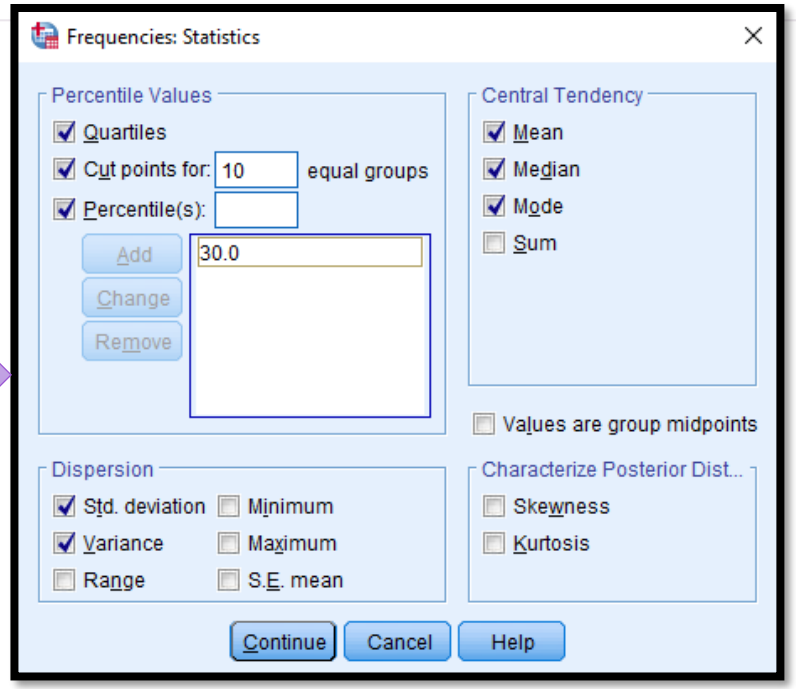
ขนาดพื้นที่เสี่ยงภัยแล้ง

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	ความเสี่ยงต่ำมาก	7	17.5	17.5	17.5
	ความเสี่ยงต่ำ	4	10.0	10.0	27.5
	ความเสี่ยงปานกลาง	14	35.0	35.0	62.5
	ความเสี่ยงสูง	7	17.5	17.5	80.0
	ความเสี่ยงสูงมาก	8	20.0	20.0	100.0
	Total	40	100.0	100.0	

การหาค่าเปอร์เซนไทล์ , เดไซด์, ควอไทล์, การวัดแนวโน้มเข้าสู่ส่วนกลางและการวัดการกระจาย

จากคำสั่ง Frequencies หากต้องการหาค่าสถิติอื่น ๆ ให้เลือกเมนู Statistics ในที่นี้จะเลือกใช้ตัวแปร ปริมาณน้ำฝน ซึ่งเป็นข้อมูลเชิงปริมาณ จะมีค่าสถิติพื้นฐานให้เลือก ดังนี้





คลิก Continue ⇒ Ok จะได้ผลลัพธ์ ดังนี้

Frequencies

Statistics

ปริมาณน้ำฝน

N	Valid	40
	Missing	0
Mean		1,700.35
Median		1,290.50
Mode		998[a]
Std. Deviation		1,013.572
Variance		1027328.387
Percentiles	10	802.20
	20	928.60
	25	960.50
	30	1,005.20
	40	1,179.40
	50	1,290.50
	60	1,426.60
	70	1,863.90
	75	2,121.00
	80	2,500.00
	90	3,506.30

a. Multiple modes exist. The smallest value is shown

สามารถสรุปค่าต่าง ๆ ได้ดังนี้

- ปริมาณน้ำฝนเฉลี่ย 1,700.35 มิลลิเมตร
- มีค่ามัธยฐาน เท่ากับ 1,290.50 มิลลิเมตร
- มีค่าฐานนิยม เท่ากับ 998 มิลลิเมตร
- มีค่าเบี่ยงเบนมาตฐาน (S.D.) เท่ากับ 1,013.572
- มีค่าความแปรปรวน เท่ากับ 1027328.387
- เปอร์เซนไทล์ที่ 50 เท่ากับ 1,290.50

Q&A