

การคำนวณค่าสถิติเบื้องต้น

หลังจากที่เก็บรวบรวมข้อมูลและสร้างแฟ้มข้อมูลเรียบร้อยแล้ว ผู้วิจัยจะต้องทำการวิเคราะห์ข้อมูล เพื่อสรุปผลของการวิจัย ตามวัตถุประสงค์ของงานวิจัย ซึ่งจำเป็นต้องใช้วิธีการทางสถิติช่วยในการวิเคราะห์ ซึ่งสามารถแบ่งออกเป็น 2 ประเภท ดังนี้

1. สถิติเชิงพรรณนา (Descriptive Statistics) ซึ่งเป็นการนำเสนอข้อมูลที่เก็บรวบรวมมา โดยนำมาบรรยายถึงลักษณะของข้อมูลที่เก็บมาได้ ทั้งในรูปแบบของตาราง ข้อความ แผนภูมิ หรือกราฟต่างๆ และค่าสถิติต่างๆ

2. สถิติเชิงอนุมาน (Inference Statistics) หมายถึงการวิเคราะห์ข้อมูลตัวอย่าง แล้วนำผลการวิเคราะห์นั้นอ้างอิงถึงลักษณะที่สำคัญของประชากร โดยใช้หลักเกณฑ์ของความน่าจะเป็น และสถิติเชิงอนุมานจะประกอบด้วย การประมาณค่า การทดสอบสมมติฐานที่ใช้พารามิเตอร์ และที่ไม่ใช้พารามิเตอร์ เป็นต้น (ซึ่งเนื้อหาดังกล่าวจะกล่าวถึงรายละเอียดในหัวข้อต่อไป)

สถิติเชิงพรรณนา (Descriptive Statistics)

เป็นการนำเสนอข้อมูลที่เก็บรวบรวมมา โดยนำมาบรรยายถึงลักษณะของข้อมูลที่เก็บมาได้ ทั้งในรูปแบบของตาราง ข้อความ แผนภูมิ หรือกราฟต่างๆ หลักสำคัญของสถิติเชิงพรรณนาคือ เก็บข้อมูลชนิดใดมาได้ก็จะอธิบายได้เฉพาะข้อมูลชนิดนั้น ไม่สามารถนำไปใช้อ้างอิงถึงข้อมูลในส่วนอื่นๆ ได้ เช่น เก็บข้อมูลจากกลุ่มตัวอย่าง จำนวน 600 คน จากประชากรทั้งหมด 1,000 คน การบรรยายลักษณะของข้อมูลตลอดจนการวิเคราะห์และแปลผลจะเป็นเฉพาะในส่วนข้อมูลจากกลุ่มตัวอย่าง 600 คนเท่านั้น ผู้วิจัยไม่สามารถนำไปใช้อ้างอิงแทนประชากรทั้งหมด 1,000 คนได้ จึงเป็นการสรุปเฉพาะลักษณะที่สำคัญของข้อมูลที่ศึกษาเท่านั้น และสถิติพรรณนาประกอบด้วย เนื้อหา ดังต่อไปนี้

1. การนำเสนอข้อมูล ซึ่งประกอบด้วย

- การนำเสนอในรูปแบบบทความ เช่น สถิติของคนไทยแยกตามเพศ
- การนำเสนอในรูปแบบตาราง หรือร้อยละ ซึ่งอาจเป็นตารางจำนวนทางเดียวหรือหลายทาง
- การนำเสนอในรูปแบบกราฟ เช่น กราฟเส้น กราฟแท่ง กราฟวงกลม

2. การแจกแจงความถี่

การแจกแจงความถี่ เป็นการแสดงค่าความถี่ของข้อมูลที่เก็บมาได้ โดยแสดงเป็นจำนวน และร้อยละ (%) เช่น จำนวนและเปอร์เซ็นต์ของอาจารย์ แยกตามวุฒิการศึกษาและตามเพศ

3. การวัดแนวโน้มเข้าสู่ส่วนกลาง

เป็นการหาค่ากลางของข้อมูลเพื่อใช้เป็นตัวแทนของข้อมูลทั้งหมด เพื่อเป็นประโยชน์ในการเปรียบเทียบข้อมูลแต่ละชุด โดยไม่จำเป็นต้องพิจารณาข้อมูลทั้งหมดของแต่ละชุด สถิติของการวัดแนวโน้มเข้าสู่ส่วนกลางที่นิยมใช้ คือ

- ค่าเฉลี่ย (Mean) ประกอบด้วย ค่าเฉลี่ยเลขคณิต มัชฌิมเรขาคณิต มัชฌิมฮาร์โมนิก
- ค่ามัธยฐาน (Median)
- ค่าฐานนิยม (Mode)
- ค่าเปอร์เซ็นต์ไทล์ (Percentile)
- ค่าเดซิส์ (Decile)
- ควอไทล์ (Quatile)

4. การวัดการกระจายของข้อมูล

การวัดการกระจายของข้อมูล เป็นการอธิบายว่าข้อมูลแต่ละค่ามีค่าที่ห่างกันมากน้อยเพียงใด สถิติของการวัดการกระจายของข้อมูลที่นิยมใช้ คือ

- พิสัย (Range)
- ส่วนเบี่ยงเบนควอไทล์ (Quatile Deviation)
- ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation)
- ค่าแปรปรวน (Variance)
- สัมประสิทธิ์ความแปรผัน (Coefficient of Variance)

● สูตรที่ใช้คำนวณสถิติการแจกแจงความถี่ (Frequency)

การแจกแจงความถี่ เป็นการแสดงค่าความถี่ของข้อมูลที่เก็บมาได้ โดยแสดงเป็นจำนวนและร้อยละ (%)

ค่าร้อยละ (Percentage)

ค่าร้อยละ คือ การคำนวณหาสัดส่วนของข้อมูลในแต่ละตัวเทียบกับข้อมูลรวมทั้งหมด โดยให้ข้อมูลรวมทั้งหมดมีค่าเป็นร้อย

สูตรคำนวณ

$$\text{ร้อยละ (\%)} = \frac{X \cdot 100}{N}$$

X คือ จำนวนข้อมูล (ความถี่) ที่ต้องการนำมาหาค่าร้อยละ

N คือ จำนวนข้อมูลทั้งหมด

- สูตรที่ใช้ในการคำนวณค่าสถิติต่างๆ ของการวัดแนวโน้มเข้าสู่ส่วนกลาง

1. ค่าเฉลี่ย (Mean)

ค่าเฉลี่ยหรือค่ามัธยฐานเลขคณิต ใช้สัญลักษณ์ \bar{X} สำหรับค่าเฉลี่ยที่ได้มาจากกลุ่มตัวอย่าง และใช้สัญลักษณ์ μ สำหรับค่าเฉลี่ยที่ได้มาจากประชากรทั้งหมด การคำนวณหาค่าเฉลี่ยทำได้ 2 แบบ คือ การคำนวณหาค่าเฉลี่ยจากข้อมูลดิบที่ไม่อยู่ในรูปของตารางแจกแจงความถี่ และการคำนวณหาค่าเฉลี่ยจากข้อมูลจัดกลุ่มที่อยู่ในรูปของตารางแจกแจงความถี่ ซึ่งจะมีสูตรดังนี้

สูตรคำนวณหาค่าเฉลี่ยจากข้อมูลดิบที่ไม่อยู่ในรูปของตารางแจกแจงความถี่

$$\bar{X} = \frac{\sum_{i=1}^N X_i}{N}$$

\bar{X} คือ ค่าเฉลี่ย

$\sum_{i=1}^N X_i$ คือ ผลรวมของข้อมูลทั้งหมด

N คือ จำนวนข้อมูลทั้งหมด

สูตรคำนวณหาค่าเฉลี่ยจากข้อมูลจัดกลุ่มที่อยู่ในรูปของตารางแจกแจงความถี่

$$\bar{X} = \frac{\sum_{i=1}^N f_i X_i}{N}$$

\bar{X} คือ ค่าเฉลี่ย

$\sum_{i=1}^N f_i X_i$ คือ ผลรวมของข้อมูลทั้งหมด

N คือ จำนวนข้อมูลทั้งหมด

2. ค่ามัธยฐาน (Median)

ค่ามัธยฐาน คือ ค่าของข้อมูลที่อยู่ตรงกลางของข้อมูลทั้งหมดที่ได้นำมาเรียงลำดับจากน้อยไปหามากแล้ว ซึ่งการหาค่ามัธยฐานสามารถคำนวณหาได้ 2 แบบ ได้แก่ การคำนวณหาค่ามัธยฐานจากข้อมูลดิบที่ไม่อยู่ในรูปของตารางแจกแจงความถี่ และการคำนวณหาค่ามัธยฐานจากข้อมูลจัดกลุ่มที่อยู่ในรูปของตารางแจกแจงความถี่ ซึ่งจะมีสูตรดังนี้

สูตรคำนวณค่ามัธยฐานจากข้อมูลดิบที่ไม่อยู่ในรูปของตารางแจกแจงความถี่

$$Mdn = \frac{N+1}{2}$$

Mdn คือ ค่ามัธยฐาน

N คือ จำนวนข้อมูลทั้งหมด

สูตรคำนวณค่ามัธยฐานจากข้อมูลจัดกลุ่มที่อยู่ในรูปของตารางแจกแจงความถี่

$$Mdn = L_0 + i \left[\frac{\frac{N}{2} - cf}{f} \right]$$

Mdn คือ ค่ามัธยฐาน

L_0 คือ ขอบเขตล่างของชั้นที่มีมัธยฐานอยู่

i คือ อัตรภาคชั้น (ช่วงห่างของข้อมูลแต่ละชั้น)

N คือ จำนวนข้อมูลทั้งหมด

cf คือ ความถี่สะสมที่อยู่ก่อนถึงชั้นที่มีตำแหน่งมัธยฐาน

f คือ ความถี่ของคะแนนในชั้นที่มีมัธยฐาน

3. ค่าฐานนิยม (Mode)

ค่าฐานนิยม คือ ค่าของข้อมูลตัวที่มีค่าซ้ำกันมากที่สุดในชุดข้อมูลนั้นๆ การหาค่าฐานนิยมสามารถคำนวณหาได้ 2 แบบ ได้แก่ การคำนวณหาค่าฐานนิยมจากข้อมูลดิบที่ไม่อยู่ในรูปของตารางแจกแจงความถี่ และการคำนวณหาค่าฐานนิยมจากข้อมูลจัดกลุ่มที่อยู่ในรูปของตารางแจกแจงความถี่ ซึ่งจะมีวิธีการหาดังนี้

การคำนวณหาค่าฐานนิยมจากข้อมูลดิบที่ไม่อยู่ในรูปของตารางแจกแจงความถี่
จะพิจารณาว่าในข้อมูลชุดนั้น มีค่าใดซ้ำกันมากที่สุด ค่านั้นก็จะเป็นค่าฐานนิยม

สูตรคำนวณหาค่าฐานนิยมจากข้อมูลจัดกลุ่มที่อยู่ในรูปของตารางแจกแจงความถี่

$$Mo = L_0 + i \left[\frac{\Delta 1}{\Delta 1 + \Delta 2} \right]$$

Mo คือ ค่าฐานนิยม

L_0 คือ ขอบเขตล่างของชั้นที่มีมัธยฐานอยู่

i คือ อัตรากว้างชั้น (ช่วงห่างของข้อมูลแต่ละชั้น)

$\Delta 1$ คือ $f_1 - f_0$

$\Delta 2$ คือ $f_1 - f_2$

f_1 คือ ความถี่ของชั้นที่มีฐานนิยมอยู่

f_0 คือ ความถี่ของชั้นต่ำกว่าชั้นที่มีฐานนิยมอยู่

f_2 คือ ความถี่ของชั้นสูงกว่าชั้นที่มีฐานนิยมอยู่

4. ค่าเปอร์เซ็นต์ไทล์ (Percentile)

ถ้ากำหนดให้ $X_1, X_2, X_3, \dots, X_n$ เป็นข้อมูลในกลุ่มตัวอย่าง n จำนวน เปอร์เซ็นต์ไทล์ คือ ค่าคะแนนของตัวแปร X ซึ่งเป็นการแบ่งจำนวนข้อมูลที่มีการกระจายนี้ออกเป็น 100 ส่วน แต่ละส่วนมีค่าเท่าๆกัน ซึ่งสามารถเขียนแทนได้เป็น $P_1, P_2, P_3, \dots, P_{99}$

(Guilford and Fruchter, 1978 อ้างถึงใน อำนวย เลิศขยันดี, 2539)

ซึ่งการคำนวณค่าเปอร์เซ็นต์ไทล์สามารถแบ่งการคำนวณออกเป็น 2 แบบ ได้แก่ การคำนวณหาค่าเปอร์เซ็นต์ไทล์จากข้อมูลดิบที่ไม่อยู่ในรูปของตารางแจกแจงความถี่ และการคำนวณหาค่าเปอร์เซ็นต์ไทล์จากข้อมูลจัดกลุ่มที่อยู่ในรูปของตารางแจกแจงความถี่ ซึ่งมีวิธีการหา ดังนี้

การคำนวณหาค่าเปอร์เซ็นต์ไทล์จากข้อมูลดิบที่ไม่อยู่ในรูปของตารางแจกแจงความถี่

1. เรียงข้อมูลจากน้อยไปหามาก
2. หาดำแหน่งของค่าเปอร์เซ็นต์ไทล์

จากสูตร

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

P_r คือ ค่าตำแหน่งเปอร์เซ็นต์ไทล์ที่ต้องการหา

r คือ ตำแหน่งที่ต้องการหา

N คือ จำนวนข้อมูลทั้งหมด

3. นับตำแหน่งจากค่าที่คำนวณได้จากข้อ 2

สูตรคำนวณหาฐานนิยมจากข้อมูลจัดกลุ่มที่อยู่ในรูปของตารางแจกแจงความถี่

$$P_r = L_0 + i \left[\frac{r \frac{N}{100} - cf}{f} \right]$$

P_r คือ ค่าเปอร์เซ็นต์ไทล์

L_0 คือ ขอบเขตล่างของชั้นที่มีตำแหน่งเปอร์เซ็นต์ไทล์อยู่

i คือ อัตรากว้างชั้น (ช่วงห่างของข้อมูลแต่ละชั้น)

N คือ จำนวนข้อมูลทั้งหมด

cf คือ ความถี่สะสมที่อยู่ก่อนถึงชั้นที่มีตำแหน่งเปอร์เซ็นต์ไทล์

f คือ ความถี่ของคะแนนในชั้นที่มีตำแหน่งเปอร์เซ็นต์ไทล์

5. ค่าเดไซล์ (Decile)

ถ้ากำหนดให้ $X_1, X_2, X_3, \dots, X_n$ เป็นข้อมูลในกลุ่มตัวอย่าง n จำนวน เดไซล์ คือ ค่าคะแนนของตัวแปร X ซึ่งเป็นการแบ่งจำนวนข้อมูลที่มีการกระจายนี้ออกเป็น 10 ส่วน แต่ละส่วนมีค่าเท่าๆกัน ซึ่งสามารถเขียนแทนได้เป็น $D_1, D_2, D_3, \dots, D_9$

(Spiegel, 1961 อ้างถึงใน อำนวย เลิศขันธ์, 2539)

ซึ่งการคำนวณค่าเดไซล์สามารถแบ่งการคำนวณออกเป็น 2 แบบ ได้แก่ การคำนวณหาค่าเดไซล์จากข้อมูลดิบที่ไม่อยู่ในรูปของตารางแจกแจงความถี่ และการคำนวณหาค่าเดไซล์จากข้อมูลจัดกลุ่มที่อยู่ในรูปของตารางแจกแจงความถี่ ซึ่งมีวิธีการหาดังนี้

การคำนวณหาค่าเปอร์เซ็นต์ไทล์จากข้อมูลดิบที่ไม่อยู่ในรูปของตารางแจกแจงความถี่

1. เรียงข้อมูลจากน้อยไปหามาก
2. หาดำแหน่งของค่าเดไซล์

จากสูตร

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

D_r คือ ค่าตำแหน่งเดไซล์ที่ต้องการหา

r คือ ตำแหน่งที่ต้องการหา

N คือ จำนวนข้อมูลทั้งหมด

3. นับตำแหน่งจากค่าที่คำนวณได้จากข้อ 2

สูตรคำนวณคำนวณหาค่าฐานนิยมจากข้อมูลจัดกลุ่มที่อยู่ในรูปของตารางแจกแจงความถี่

$$D_r = L_0 + i \left[\frac{r \frac{N}{10} - cf}{f} \right]$$

D_r คือ ค่าเดไซล์

L_0 คือ ขอบเขตล่างของชั้นที่มีตำแหน่งเดไซล์อยู่

i คือ อัตรากว้างชั้น (ช่วงห่างของข้อมูลแต่ละชั้น)

N คือ จำนวนข้อมูลทั้งหมด

cf คือ ความถี่สะสมที่อยู่ก่อนถึงชั้นที่มีตำแหน่งเดไซล์

f คือ ความถี่ของคะแนนในชั้นที่มีตำแหน่งเดไซล์

6. ค่าควอไทล์ (Quatile)

ถ้ากำหนดให้ $X_1, X_2, X_3, \dots, X_n$ เป็นข้อมูลในกลุ่มตัวอย่าง n จำนวน ควอไทล์ คือ ค่าคะแนนของตัวแปร X ซึ่งเป็นการแบ่งจำนวนข้อมูลที่มีการกระจายนี้ออกเป็น 4 ส่วน แต่ละส่วนมีค่าเท่าๆกัน ซึ่งสามารถเขียนแทนได้เป็น Q_1, Q_2, Q_3

(Spiegel, 1961 อ้างถึงใน อำนวย เลิศขยันดี, 2539)

ซึ่งการคำนวณค่าควอไทล์สามารถแบ่งการคำนวณออกเป็น 2 แบบ ได้แก่ การคำนวณหาค่าควอไทล์จากข้อมูลดิบที่ไม่อยู่ในรูปของตารางแจกแจงความถี่ และการคำนวณหาค่าควอไทล์จากข้อมูลจัดกลุ่มที่อยู่ในรูปของตารางแจกแจงความถี่ ซึ่งมีวิธีการหาดังนี้

การคำนวณหาค่าควอไทล์จากข้อมูลที่ไม่อยู่ในรูปของตารางแจกแจงความถี่

- เรียงข้อมูลจากน้อยไปหามาก
- หาตำแหน่งของค่าควอไทล์

จากสูตร

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

Q_r คือ ค่าตำแหน่งควอไทล์ที่ต้องการหา

r คือ ตำแหน่งที่ต้องการหา

N คือ จำนวนข้อมูลทั้งหมด

- นับตำแหน่งจากค่าที่คำนวณได้จากข้อ 2

สูตรคำนวณหาค่าควอไทล์จากข้อมูลจัดกลุ่มที่อยู่ในรูปของตารางแจกแจงความถี่

$$Q_r = L_0 + i \left[\frac{r \frac{N}{4} - cf}{f} \right]$$

Q_r คือ ค่าควอไทล์

L_0 คือ ขอบเขตล่างของชั้นที่มีตำแหน่งควอไทล์อยู่

i คือ อัตรากว้างชั้น (ช่วงห่างของข้อมูลแต่ละชั้น)

N คือ จำนวนข้อมูลทั้งหมด

cf คือ ความถี่สะสมที่อยู่ก่อนถึงชั้นที่มีตำแหน่งควอไทล์

f คือ ความถี่ของคะแนนในชั้นที่มีตำแหน่งควอไทล์

● สูตรที่ใช้ในการคำนวณค่าสถิติต่างๆ ของการวัดการกระจายของข้อมูล

1. ค่าพิสัย (Range)

ค่าพิสัย คือ ค่าผลต่างของข้อมูลตัวที่มีค่ามากที่สุด (Max) กับตัวที่มีค่าน้อยที่สุด (Min) ค่าพิสัยจัดเป็นวิธีการวัดการกระจายของข้อมูลที่หยาบที่สุด ซึ่งมีสูตรในการคำนวณ ดังนี้

$$\text{ค่าพิสัย} = \text{ค่าสูงสุด} - \text{ค่าต่ำสุด}$$

2. ส่วนเบี่ยงเบนควอไทล์ (Qaurtile Deviation)

ส่วนเบี่ยงเบนควอไทล์ คือ ค่าที่แสดงให้เห็นว่าข้อมูลจำนวน 50% ที่อยู่ตรงกลางของข้อมูลแตกต่างกันอย่างไร สามารถหาค่าส่วนเบี่ยงเบนควอไทล์ ได้ 2 แบบคือ การคำนวณหาค่าส่วนเบี่ยงเบนควอไทล์จากข้อมูลดิบที่ไม่อยู่ในรูปของตารางแจกแจงความถี่ และการคำนวณหาค่าส่วนเบี่ยงเบนควอไทล์จากข้อมูลจัดกลุ่มที่อยู่ในรูปของตารางแจกแจงความถี่ ซึ่งมีวิธีการดังนี้

สูตรคำนวณหาค่าส่วนเบี่ยงเบนควอไทล์จากข้อมูลดิบที่ไม่อยู่ในรูปของตารางแจกแจงความถี่

$$Q.D = \frac{Q_3 - Q_1}{2}$$

โดย

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

$Q.D$ คือ ค่าส่วนเบี่ยงเบนมาตรฐาน

Q_r คือ ค่าตำแหน่งควอไทล์ที่ต้องการหา

r คือ ตำแหน่งที่ต้องการหา

N คือ จำนวนข้อมูลทั้งหมด

Q_3 คือ ค่าตัวเลขในตำแหน่งควอไทล์ที่ 3

Q_1 คือ ค่าตัวเลขในตำแหน่งควอไทล์ที่ 1

สูตรคำนวณหาค่าส่วนเบี่ยงเบนควอไทล์จากข้อมูลจัดกลุ่มที่อยู่ในรูปของตารางแจกแจงความถี่

$$Q.D = \frac{Q_3 - Q_1}{2}$$

โดย

$$Q_r = L_0 + i \left[\frac{r \frac{N}{4} - cf}{f} \right]$$

$Q.D$ คือ ค่าส่วนเบี่ยงเบนมาตรฐาน

Q_r คือ ค่าควอไทล์

L_0 คือ ขอบเขตล่างของชั้นที่มีตำแหน่งควอไทล์อยู่

i คือ อัตรากว้างชั้น (ช่วงห่างของข้อมูลแต่ละชั้น)

N คือ จำนวนข้อมูลทั้งหมด

cf คือ ความถี่สะสมที่อยู่ก่อนถึงชั้นที่มีตำแหน่งควอไทล์

f คือ ความถี่ของคะแนนในชั้นที่มีตำแหน่งควอไทล์

Q_3 คือ ค่าตัวเลขในตำแหน่งควอไทล์ที่ 3

Q_1 คือ ค่าตัวเลขในตำแหน่งควอไทล์ที่ 1

3. ค่าส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation , S.D.)

ค่าส่วนเบี่ยงเบนมาตรฐาน (S.D.) คือ ค่ารากที่สองของผลรวมของความแตกต่างระหว่างข้อมูลกับค่าเฉลี่ยยกกำลังสอง (sum of squares ของผลต่าง) หารด้วยจำนวนข้อมูลทั้งหมด
สัญลักษณ์ค่าส่วนเบี่ยงเบนมาตรฐานจะมี 2 ลักษณะ ดังนี้

σ ใช้กับข้อมูล que เก็บมาจากประชากรทั้งหมด

s ใช้กับข้อมูล que เก็บมาจากกลุ่มตัวอย่าง

ค่าส่วนเบี่ยงเบนมาตรฐานสามารถคำนวณหาได้ 2 แบบ คือ การคำนวณหาค่าส่วนเบี่ยงเบนมาตรฐานจากข้อมูลดิบที่ไม่อยู่ในรูปของตารางแจกแจงความถี่ และการคำนวณหาค่าส่วนเบี่ยงเบนมาตรฐานจากข้อมูลจัดกลุ่มที่อยู่ในรูปของตารางแจกแจงความถี่ ซึ่งมีวิธีการดังนี้

สูตรคำนวณหาค่าส่วนเบี่ยงเบนมาตรฐานจากข้อมูลดิบที่ไม่อยู่ในรูป
ของตารางแจกแจงความถี่

$$\sigma = \sqrt{\frac{\sum (X - \mu)^2}{N}} \quad (\text{สำหรับข้อมูลที่ได้จากประชากร})$$

$$s = \sqrt{\frac{\sum (X - \bar{X})^2}{N - 1}} \quad (\text{สำหรับข้อมูลที่ได้จากกลุ่มตัวอย่าง})$$

σ หรือ s คือ ค่าส่วนเบี่ยงเบนมาตรฐาน

X คือ ข้อมูลแต่ละจำนวน

μ หรือ \bar{X} คือ ค่าเฉลี่ยของข้อมูลชุดนั้น

N คือ จำนวนข้อมูลจากประชากรทั้งหมด

n คือ จำนวนข้อมูลจากกลุ่มตัวอย่าง

สูตรคำนวณค่าส่วนเบี่ยงเบนมาตรฐานจากข้อมูลจัดกลุ่มที่อยู่ในรูป
ของตารางแจกแจงความถี่

$$\sigma = i \sqrt{\frac{\sum fd^2}{N} - \left(\frac{\sum fd}{N}\right)^2} \quad (\text{สำหรับข้อมูลที่ได้จากประชากร})$$

$$s = i \sqrt{\frac{n \sum fd^2 - (\sum fd)^2}{n(n-1)}} \quad (\text{สำหรับข้อมูลที่ได้จากกลุ่มตัวอย่าง})$$

σ หรือ s คือ ค่าส่วนเบี่ยงเบนมาตรฐาน

N คือ จำนวนข้อมูลจากประชากรทั้งหมด

n คือ จำนวนข้อมูลจากกลุ่มตัวอย่าง

i คือ อัตรากว้างชั้น (ช่วงห่างของข้อมูลแต่ละชั้น)

f คือ ความถี่ของคะแนน

d คือ ค่ากึ่งกลางอัตรากว้างชั้น

4. ความแปรปรวน (Variance)

ความแปรปรวน คือ ค่าส่วนเบี่ยงเบนมาตรฐานยกกำลังสอง ซึ่งจะทำให้ความแปรปรวนมีหน่วยเป็นข้อมูลยกกำลังสอง เช่น ถ้าข้อมูลมีหน่วยเป็นบาท ความแปรปรวนก็จะมีหน่วยเป็น (บาท)² ซึ่งการคำนวณค่าความแปรปรวนสามารถแบ่งได้เป็น 2 กรณีคือ การคำนวณหาความแปรปรวนจากข้อมูลดิบที่ไม่อยู่ในรูปของตารางแจกแจงความถี่ และการคำนวณหาความแปรปรวนจากข้อมูลจัดกลุ่มที่อยู่ในรูปของตารางแจกแจงความถี่ ซึ่งจะมีวิธีการดังนี้

สูตรคำนวณหาความแปรปรวนจากข้อมูลดิบที่ไม่อยู่ในรูป
ของตารางแจกแจงความถี่

$$\sigma^2 = \frac{\sum (X - \mu)^2}{N} \quad (\text{สำหรับข้อมูลที่ได้จากประชากร})$$

$$s^2 = \frac{\sum (X - \bar{X})^2}{N} \quad (\text{สำหรับข้อมูลที่ได้จากกลุ่มตัวอย่าง})$$

σ^2 หรือ s^2 คือ ค่าส่วนเบี่ยงเบนมาตรฐาน

X คือ ข้อมูลแต่ละจำนวน

μ หรือ \bar{X} คือ ค่าเฉลี่ยของข้อมูลชุดนั้น

N คือ จำนวนข้อมูลจากประชากรทั้งหมด

n คือ จำนวนข้อมูลจากกลุ่มตัวอย่าง

สูตรคำนวณหาค่าความแปรปรวนจากข้อมูลจัดกลุ่มที่อยู่ในรูป
ของตารางแจกแจงความถี่

$$\sigma^2 = i \left(\frac{\sum fd^2}{N} - \left(\frac{\sum fd}{N} \right)^2 \right) \quad (\text{สำหรับข้อมูลที่ได้จากประชากร})$$

$$s^2 = i \left(\frac{n \sum fd^2 - (\sum fd)^2}{n(n-1)} \right) \quad (\text{สำหรับข้อมูลที่ได้จากกลุ่มตัวอย่าง})$$

σ^2 หรือ s^2 คือ ค่าส่วนเบี่ยงเบนมาตรฐาน

N คือ จำนวนข้อมูลจากประชากรทั้งหมด

n คือ จำนวนข้อมูลจากกลุ่มตัวอย่าง

i คือ อัตรภาคชั้น (ช่วงห่างของข้อมูลแต่ละชั้น)

f คือ ความถี่ของคะแนน

d คือ ค่ากึ่งกลางอัตรภาคชั้น

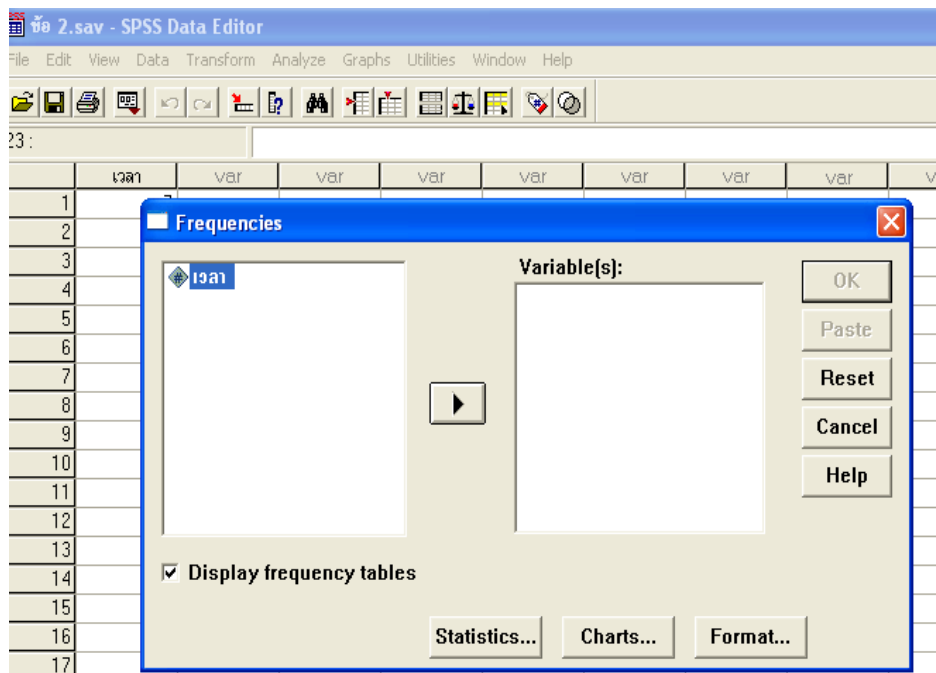
5. สัมประสิทธิ์ความแปรผัน (Coefficient of Variance, C.V.)

สัมประสิทธิ์ความแปรผัน เป็นค่าที่ใช้เปรียบเทียบการกระจายของข้อมูล 2 ชุดขึ้นไป ซึ่งค่า C.V. จะไม่มีหน่วย และข้อมูลชุดที่มีค่า C.V. มากจะมีการกระจายมากกว่าข้อมูลชุดที่มีค่า C.V. ต่ำ โดยมีสูตรในการคำนวณหา ดังนี้

$$C.V. = \frac{\sigma}{\mu} \quad (\text{ของประชากร})$$

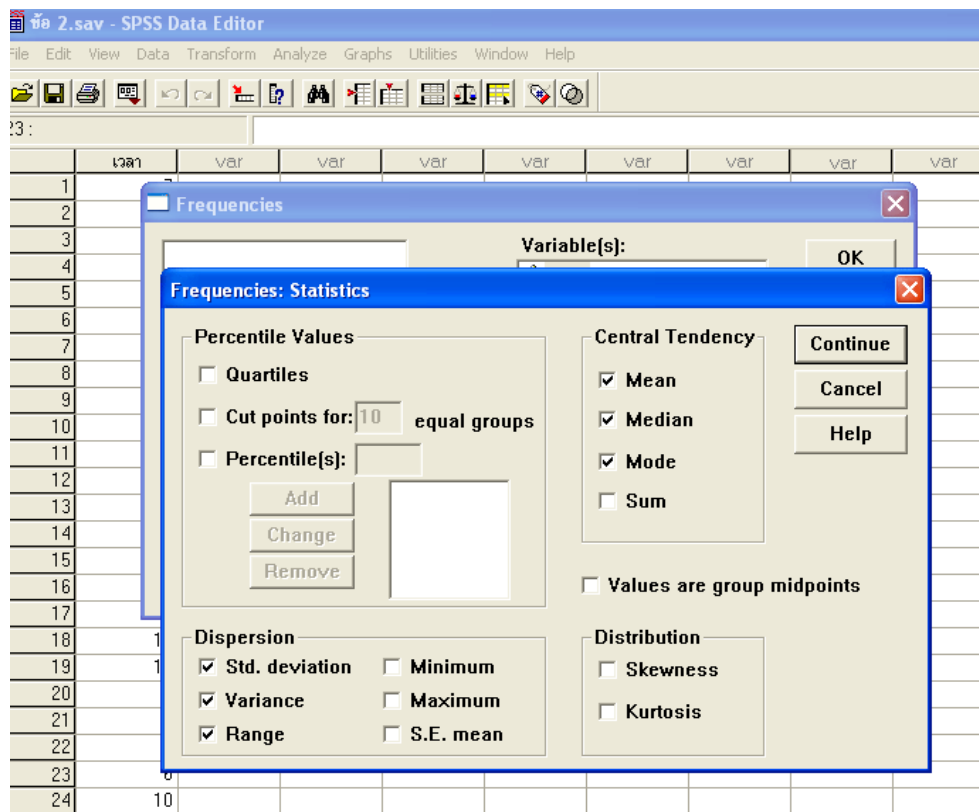
$$C.V. = \frac{s}{\bar{X}} \quad (\text{ของกลุ่มตัวอย่าง})$$

จะปรากฏหน้าจอ ดังนี้



2). เลือกตัวแปรที่เราสนใจจะศึกษาเข้าสู่ช่อง Variable

3). เลือก Statistics เพื่อใช้ในการวิเคราะห์ค่าสถิติต่าง ๆ



จากรูป จะประกอบด้วยส่วนต่างๆ ดังนี้

Percentile Values : ผู้ใช้สามารถเลือกทางเลือกต่อไปนี้ได้หลายทางเลือก คือ

- Quatile ซึ่งจะแสดงเป็นค่าของเปอร์เซ็นต์ไทล์ที่ 25 , 50 , 75
- Cut points for equal group จะแสดงเปอร์เซ็นต์ไทล์ซึ่งเป็นค่าที่แบ่งข้อมูลออกเป็นกลุ่มย่อยๆ โดยแต่ละกลุ่มย่อยมีขนาดเท่ากัน (จำนวน case เท่ากัน) ผู้ใช้สามารถระบุค่าในช่องสี่เหลี่ยมเป็นเลขจำนวนเต็มบวกที่มีค่าระหว่าง 2 และ 100
- Percentile (s) ผู้ใช้สามารถระบุค่าเปอร์เซ็นต์ไทล์ที่ต้องการได้หลายค่า ซึ่งมีค่าระหว่าง 0 ถึง 100 โดยจะต้องระบุทีละค่าแล้วเลือก Add ถ้าต้องการลบใช้แถบค่าเปอร์เซ็นต์ไทล์ที่ต้องการลบแล้วทำการ Remove ออก แต่ถ้าต้องการเปลี่ยนแปลงให้เลือก Change

Central Tendency : เป็นสถิติที่แสดงค่าการของข้อมูล ประกอบด้วย

- Mean ค่าเฉลี่ย
- Median ค่ามัธยฐาน
- Mode ค่าฐานนิยม
- Sum ผลบวกของข้อมูล

Dispersion : เป็นสถิติที่ใช้วัดการกระจายของข้อมูล ประกอบด้วย

- Std.Deviation ค่าส่วนเบี่ยงเบนมาตรฐานของข้อมูล
- Variance ค่าความแปรปรวน
- Range พิสัย
- Minimum ค่าต่ำสุดของข้อมูล
- Maximum ค่าสูงสุดของข้อมูล
- S.E.mean ค่าความคลาดเคลื่อนมาตรฐานกับค่าเฉลี่ยตัวอย่าง

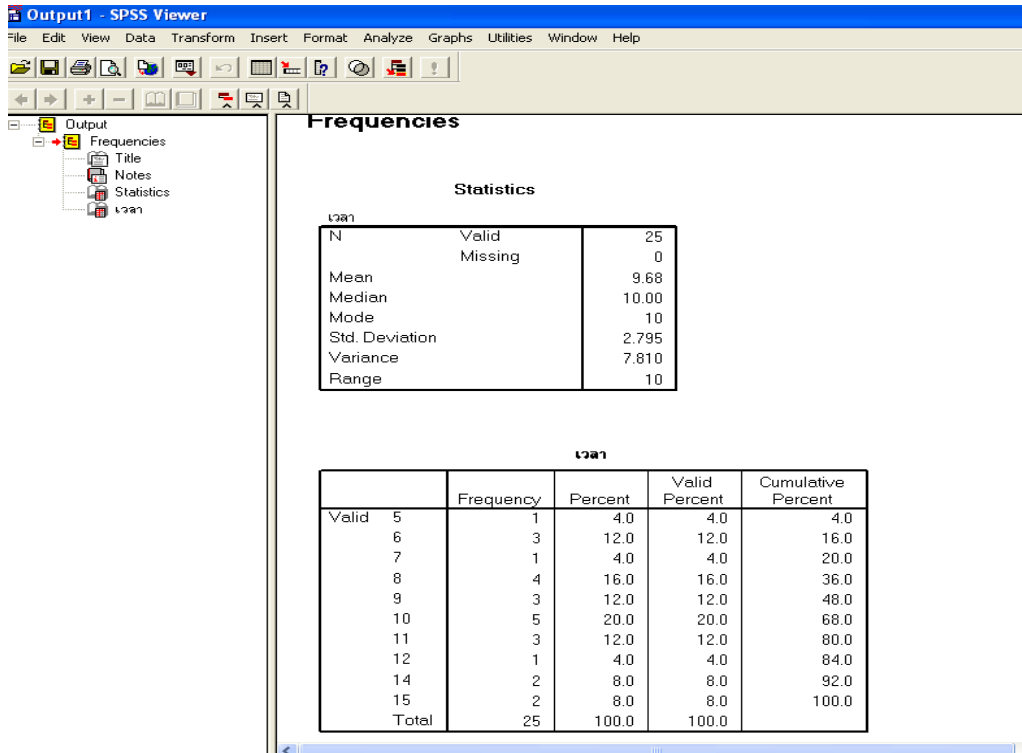
Distribution : ค่าสถิติที่แสดงการแจกแจงของข้อมูล

- Skewness ค่าความเน้นของการแจกแจงที่ไม่สมมาตร
- Kurtosis ค่าความโด่งของการแจกแจงของข้อมูล

ถ้าผู้ใช้สนใจจะศึกษาสถิติใด ก็ให้เลือกในช่องข้างหน้าของสถิติในแต่ละตัว แต่สำหรับเบื้องต้นนี้ ทางผู้จัดทำขอเลือกใช้สถิติ Mean Median Mode Std.Deviation Variance Range

4). เลือก Continue และ OK

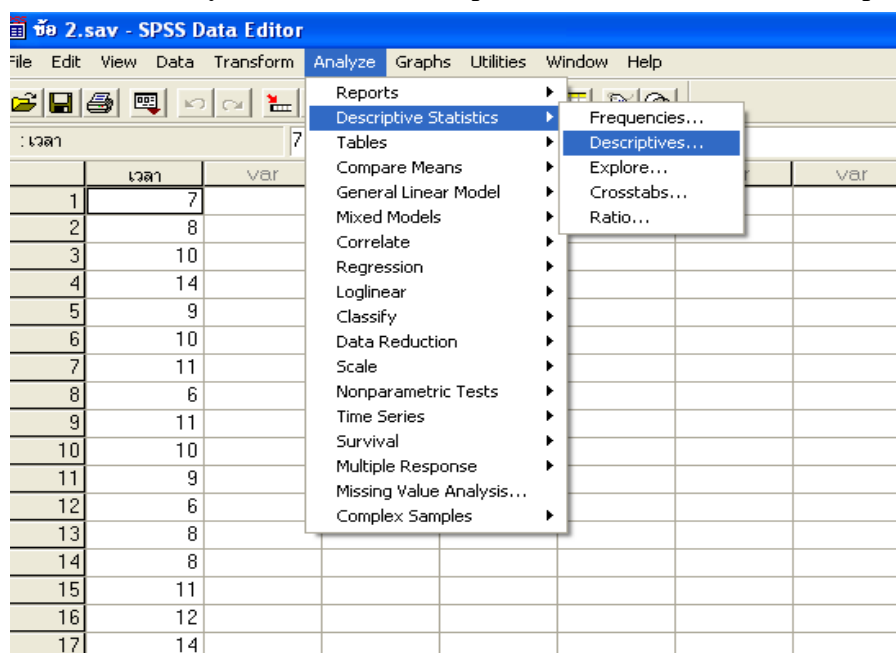
5). แล้วหน้าจอ Output จะปรากฏ



2. คำสั่ง Descriptive ซึ่งมีขั้นตอนการหาดังนี้

1). เปิดเพิ่มข้อมูล

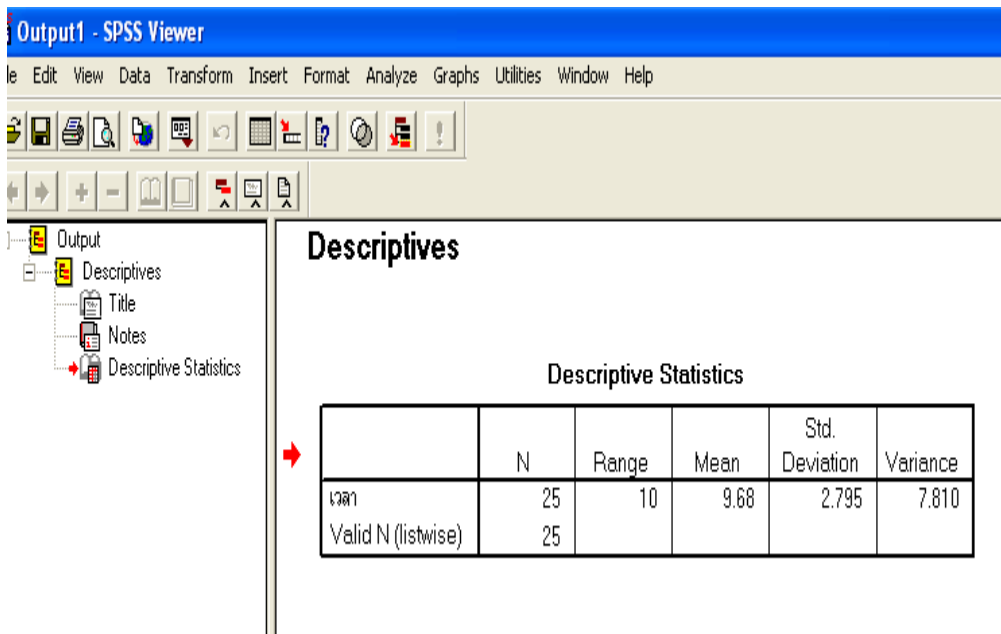
เลือก menu Analyze → Descriptive Statistics → Descriptive



ถ้าผู้ใช้สนใจจะศึกษาสถิติใด ก็ให้เลือกในช่องข้างหน้าของสถิติในแต่ละตัว แต่สำหรับเบื้องต้นนี้ ทางผู้จัดทำขอเลือกใช้สถิติ Mean Std.Deviation Variance Range

4). เลือก Continue และ OK

5). แล้วหน้าจอ Output จะปรากฏ



The screenshot shows the SPSS Output window titled "Output1 - SPSS Viewer". The main content area displays "Descriptives" with a sub-table titled "Descriptive Statistics". The table has columns for N, Range, Mean, Std. Deviation, and Variance. The first row is for "เวลา" (Time) with values: N=25, Range=10, Mean=9.68, Std. Deviation=2.795, and Variance=7.810. The second row is for "Valid N (listwise)" with N=25. A red arrow points to the "เวลา" row.

	N	Range	Mean	Std. Deviation	Variance
เวลา	25	10	9.68	2.795	7.810
Valid N (listwise)	25				

- การเลือกใช้ค่าการวัดแนวโน้มเข้าสู่ส่วนกลางและการวัดการกระจายที่เหมาะสมในการอธิบายข้อมูล

มีปัจจัยควรพิจารณาดังนี้

1. ระดับการวัดตัวของข้อมูล

- ระดับนามบัญญัติ (Nominal scale) นิยมใช้ฐานนิยม

- ระดับลำดับขั้น (Ordinal scale) นิยมใช้ มัชฎฐาน พิสัย และส่วนเบี่ยงเบนควอไทล์

- ระดับอันตรภาค (Interval scale) และระดับอัตราภาค (Ratio scale) นิยมใช้ ค่าเฉลี่ย

ค่าส่วนเบี่ยงเบนมาตรฐาน และความแปรปรวน

2. รูปแบบการกระจายของข้อมูล

3. วัตถุประสงค์ของการวิจัย

• ตัวอย่างแบบฝึกหัดและการคำนวณค่าสถิติพื้นฐาน โดยใช้ Program SPSS 13.0

1. ความสูง (เซนติเมตร) ของผู้เข้าร่วมอบรมการใช้โปรแกรม SPSS จำนวน 45 คน

เป็นดังนี้คือ

146	148	140	150	152	156	148	149	149
150	163	142	180	160	160	140	160	168
160	152	141	150	180	128	132	162	158
162	146	139	148	172	138	138	150	147
182	147	144	136	150	158	150	130	143

จงหา

1.1 ค่าเฉลี่ยเลขคณิต มัชยฐาน ฐานนิยม

1.2 ส่วนเบี่ยงเบนมาตรฐาน ส่วนเบี่ยงเบนควอไทล์ พิสัย ความแปรปรวน

วิธีทำ

1. เปิดโปรแกรม SPSS 13.0

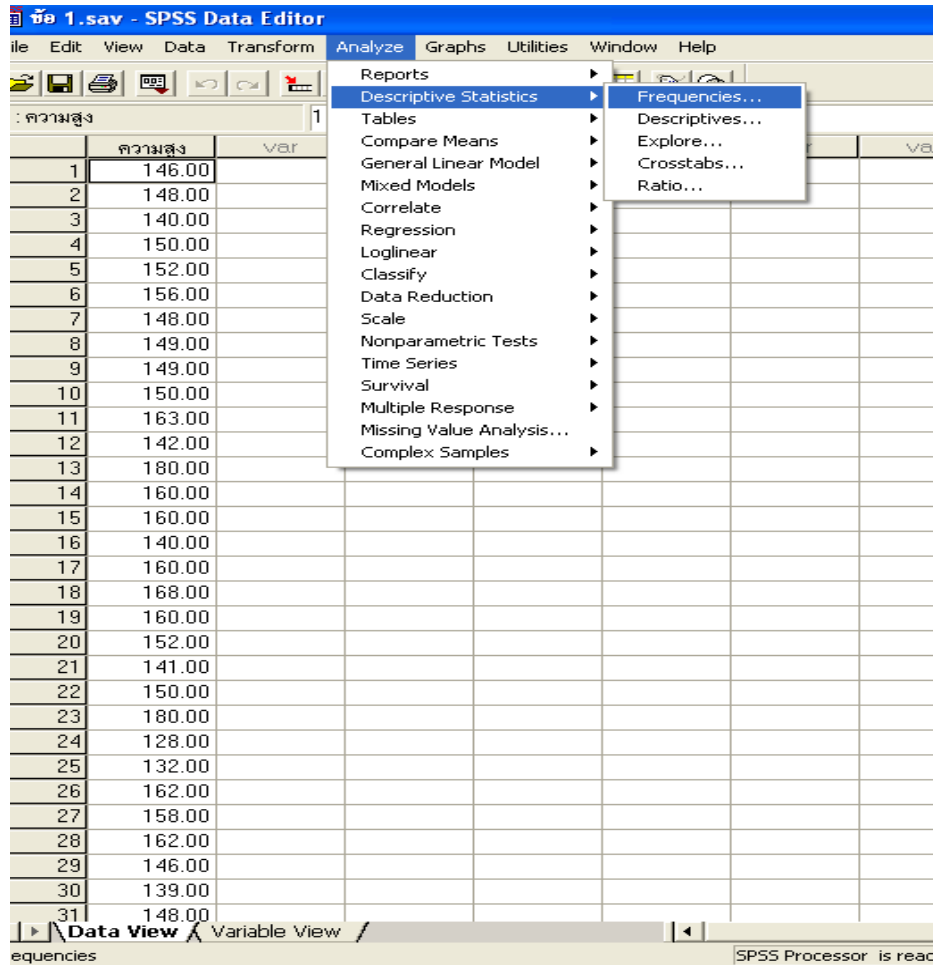
2. เลื่อนหน้าจอ Variable View และประกาศตัวแปร ดังรูป

The screenshot shows the SPSS Data Editor window with the Variable View tab selected. The variable 'ความสูง' (Height) is defined with the following properties:

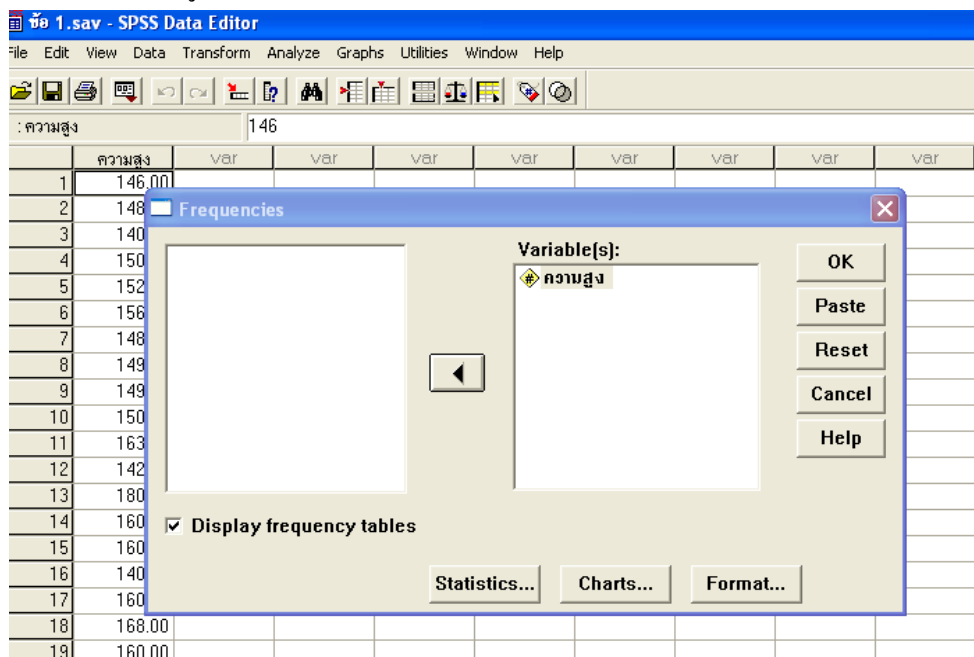
Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
ความสูง	Numeric	8	2		None	None	8	Right	Scale

The interface includes a menu bar (File, Edit, View, Data, Transform, Analyze, Graphs, Utilities, Window, Help) and a toolbar with various icons for file operations and data manipulation. The status bar at the bottom indicates 'SPSS Processor is ready'.

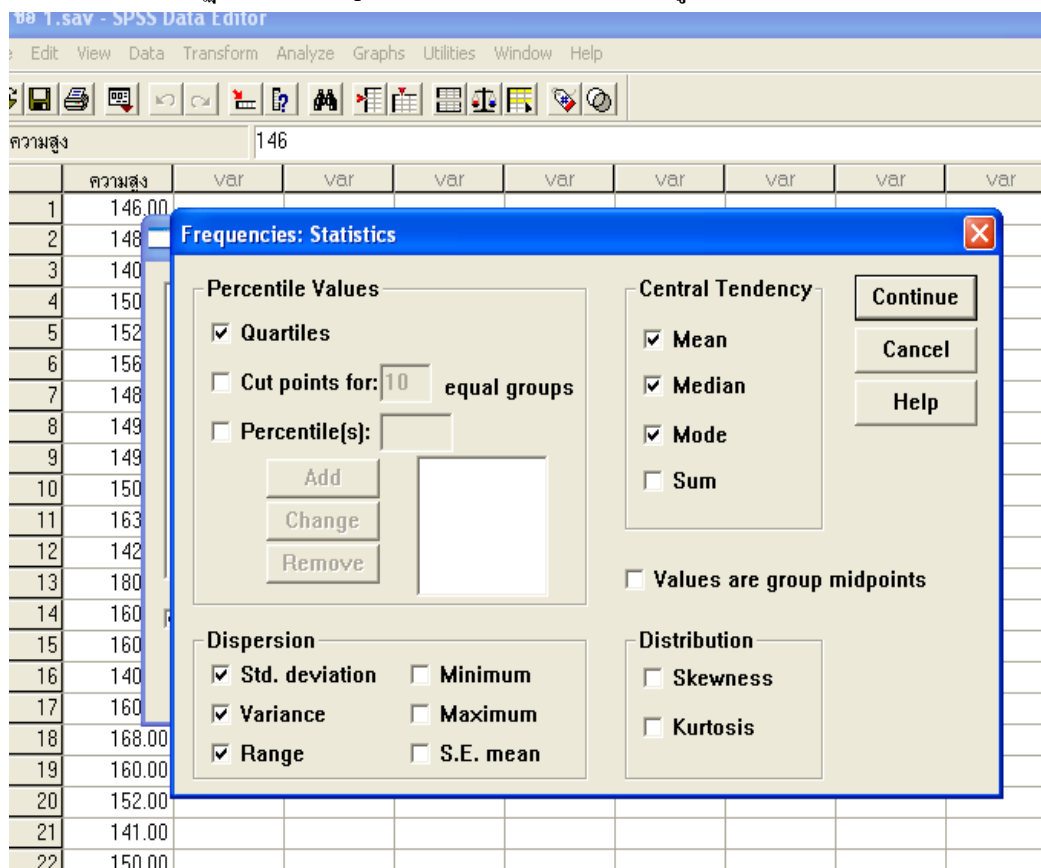
- เลือกหน้าจอ Data View ทำการใส่ข้อมูลความสูงลงไป และเลือกเมนู Analyze เพื่อวิเคราะห์ข้อมูล ดังรูป



- จะปรากฏหน้าจอ Frequency ทำการเลือกตัวแปรที่เราสนใจ เข้าสู่ช่อง Variable(s) แล้วเลือก เมนู Statistics เพื่อเลือกค่าสถิติที่จะวิเคราะห์



5. จะปรากฏหน้าจอ Frequencies : Statistics แล้ว ให้ผู้ใช้ทำการเลือกค่าสถิติที่ต้องการ



6. เลือก Continue และกด O.K จะปรากฏหน้าจอ Out-Put ดังรูป

Statistics

ความสูง		
N	Valid	45
	Missing	0
Mean		151.2000
Median		150.0000
Mode		150.00
Std. Deviation		12.52561
Variance		156.891
Range		54.00
Percentiles	25	142.5000
	50	150.0000
	75	160.0000

ความสูง

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	128.00	1	2.2	2.2	2.2
	130.00	1	2.2	2.2	4.4
	132.00	1	2.2	2.2	6.7
	136.00	1	2.2	2.2	8.9
	138.00	2	4.4	4.4	13.3
	139.00	1	2.2	2.2	15.6
	140.00	2	4.4	4.4	20.0
	141.00	1	2.2	2.2	22.2
	142.00	1	2.2	2.2	24.4
	143.00	1	2.2	2.2	26.7
	144.00	1	2.2	2.2	28.9
	146.00	2	4.4	4.4	33.3
	147.00	2	4.4	4.4	37.8
	148.00	3	6.7	6.7	44.4
	149.00	2	4.4	4.4	48.9
	150.00	6	13.3	13.3	62.2
	152.00	2	4.4	4.4	66.7
	156.00	1	2.2	2.2	68.9
	158.00	2	4.4	4.4	73.3
	160.00	4	8.9	8.9	82.2
	162.00	2	4.4	4.4	86.7
	163.00	1	2.2	2.2	88.9
	168.00	1	2.2	2.2	91.1
	172.00	1	2.2	2.2	93.3
	180.00	2	4.4	4.4	97.8
	182.00	1	2.2	2.2	100.0
	Total	45	100.0	100.0	

จากตาราง Out-Put สามารถตอบคำถามจากโจทย์ คือ ผู้เข้าสัมมนาจะมีความสูงเฉลี่ย ค่ามัธยฐาน ค่าฐานนิยม เท่ากับ 151.20 , 150.00 , 150.00 cm ตามลำดับ และจะมีค่าส่วนเบี่ยงเบนมาตรฐาน ค่าพิสัย ค่าความแปรปรวน เท่ากับ 12.53 , 54 , 156.90 ตามลำดับ แต่สำหรับค่าส่วนเบี่ยงเบนควอไทล์นั้น สามารถคำนวณได้จาก
$$Q.D = \frac{Q_3 - Q_1}{2} = \frac{160.00 - 142.50}{2} = 8.75$$

2. ข้อมูลการใช้บริการศูนย์คอมพิวเตอร์ของนักศึกษาระดับบัณฑิตศึกษา
จำนวน 160 คน ในภาคการศึกษาที่ 2 / 2549

การใช้บริการ (ครั้ง)	จำนวนนักศึกษา
20	10
18	15
17	12
12	20
10	13
8	4
7	7
5	30
2	27
1	12

จงหา

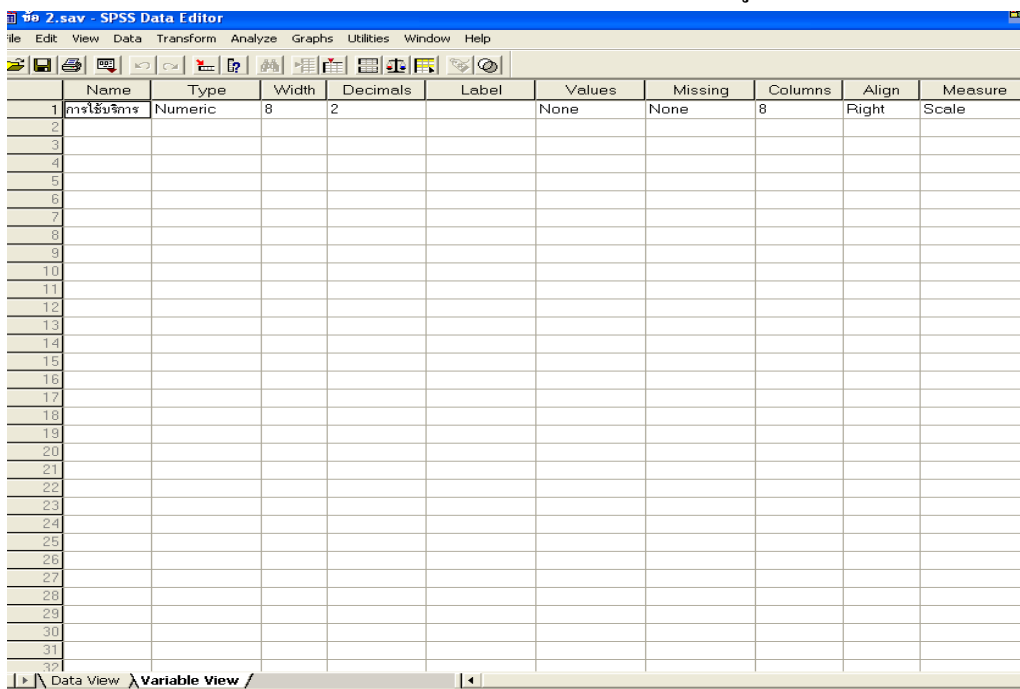
2.1 ค่าเฉลี่ยเลขคณิต มัธยฐาน ฐานนิยม

2.2 ส่วนเบี่ยงเบนมาตรฐาน ส่วนเบี่ยงเบนควอไทล์ ความแปรปรวน

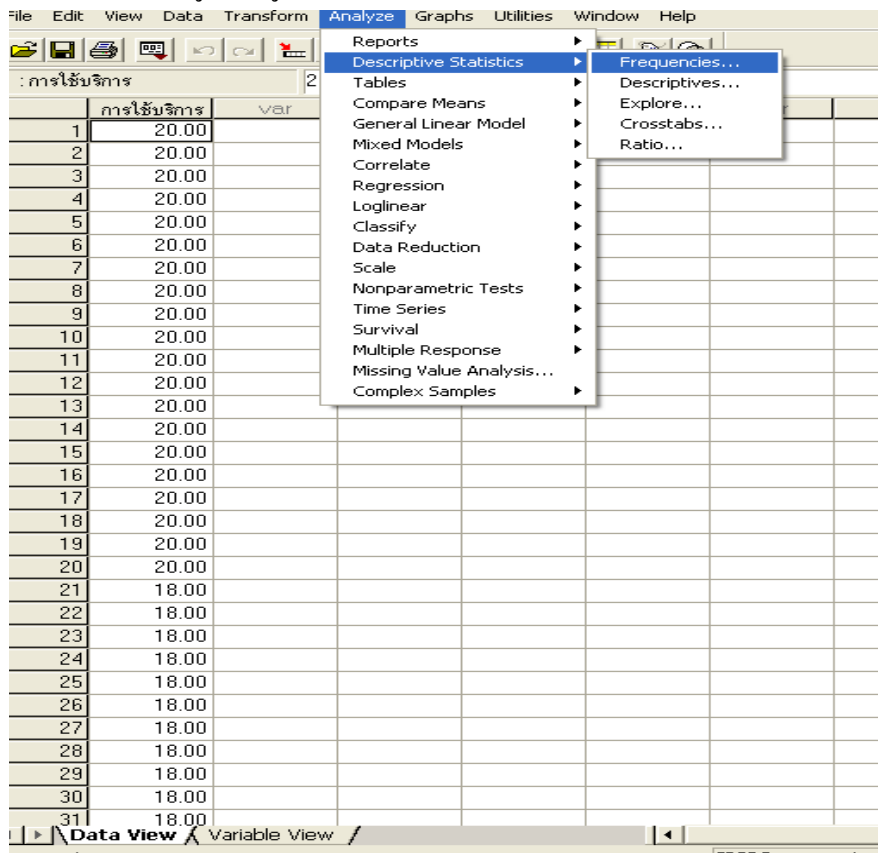
วิธีทำ

1. เปิดโปรแกรม SPSS 13.0

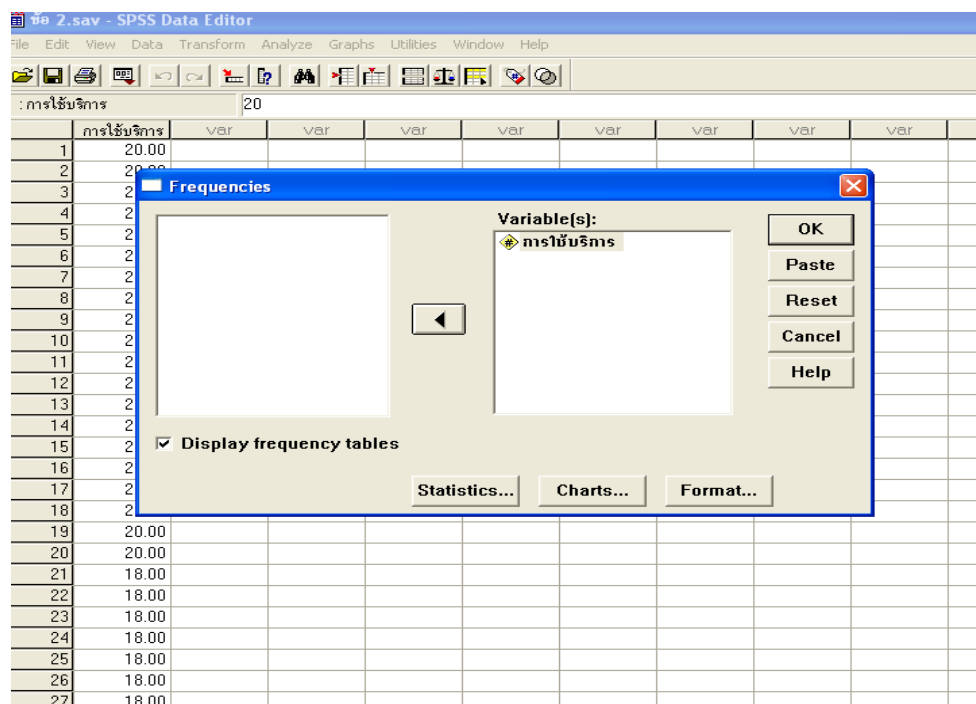
2. เลือกหน้าจอ Variable View และประกาศตัวแปร ดังรูป



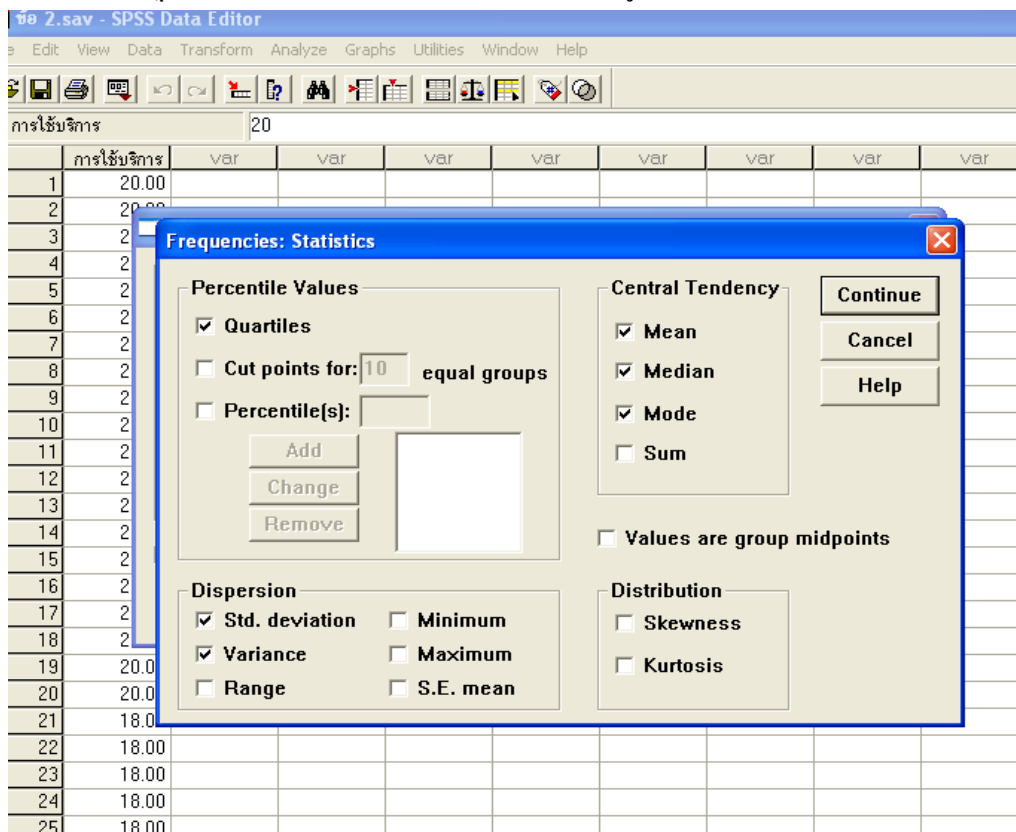
- เลือกหน้าจอ Data View ทำการใส่ข้อมูลการใช้บริการลงไป และเลือกเมนู Analyze เพื่อวิเคราะห์ข้อมูล ดังรูป



- จะปรากฏหน้าจอ Frequency ทำการเลือกตัวแปรที่เราสนใจ เข้าสู่ช่อง Variable(s) แล้วเลือก เมนู Statistics เพื่อเลือกค่าสถิติที่จะวิเคราะห์



5. จะปรากฏหน้าจอ Frequencies : Statistics แล้ว ให้ผู้ใช้ทำการเลือกค่าสถิติที่ต้องการ



6. เลือก Continue และกด O.K จะปรากฏหน้าจอ Out-Put ดังรูป

Statistics

การใช้บริการ

N	Valid	160
	Missing	0
Mean		9.6438
Median		9.0000
Mode		5.00
Std. Deviation		6.69210
Variance		44.784
Percentiles	25	5.0000
	50	9.0000
	75	17.0000

การใช้บริการ

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	12	7.5	7.5	7.5
	2.00	27	16.9	16.9	24.4
	5.00	29	18.1	18.1	42.5
	7.00	8	5.0	5.0	47.5
	8.00	4	2.5	2.5	50.0
	10.00	13	8.1	8.1	58.1
	12.00	20	12.5	12.5	70.6
	17.00	12	7.5	7.5	78.1
	18.00	15	9.4	9.4	87.5
	20.00	20	12.5	12.5	100.0
	Total	160	100.0	100.0	

จากตาราง Out-Put สามารถตอบคำถามจากโจทย์ คือ ข้อมูลการใช้บริการศูนย์คอมพิวเตอร์ของนักศึกษามีค่าเฉลี่ย ค่ามัธยฐาน ค่าฐานนิยม เท่ากับ 9.64 , 9.00 , 5.00 ครั้งตามลำดับ และจะมีค่าส่วนเบี่ยงเบนมาตรฐาน ค่าความแปรปรวน เท่ากับ 6.69 , 44.78 ตามลำดับ แต่สำหรับค่าส่วนเบี่ยงเบนควอไทล์นั้น สามารถคำนวณได้จาก

$$Q.D = \frac{Q_3 - Q_1}{2} = \frac{17.00 - 5.00}{2} = 6.00$$

3. ผลสัมฤทธิ์ทางการเรียนในการทดสอบประจำท่เรื่อง ปริมาณสารสัมพันธ์ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5 ของนักเรียนโรงเรียนแห่งหนึ่ง

คะแนน	จำนวน (คน)
15 – 19	7
20 – 24	30
25 – 29	43
30 – 35	18
36 – 39	2

3.1 จงหาแนวโน้มเข้าสู่ส่วนกลางของข้อมูลชุดนี้ โดยใช้

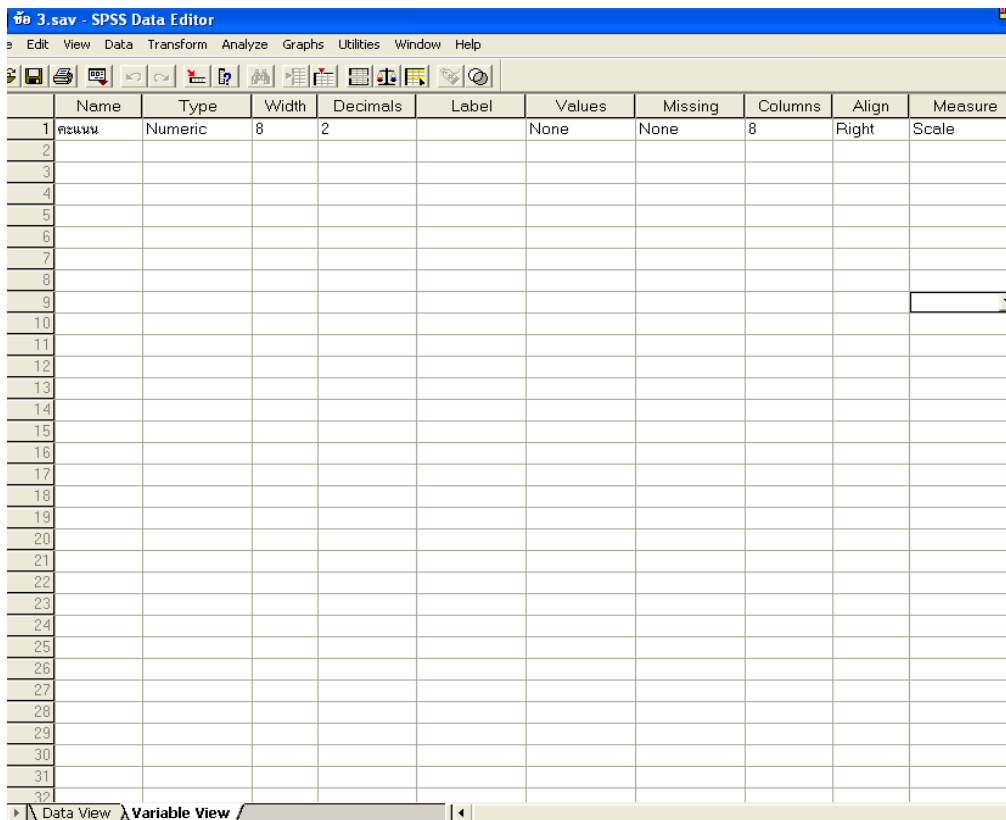
ค่าเฉลี่ยเลขคณิต มัธยฐาน ฐานนิยม

3.2 ส่วนเบี่ยงเบนมาตรฐาน ส่วนเบี่ยงเบนควอไทล์ ความแปรปรวน

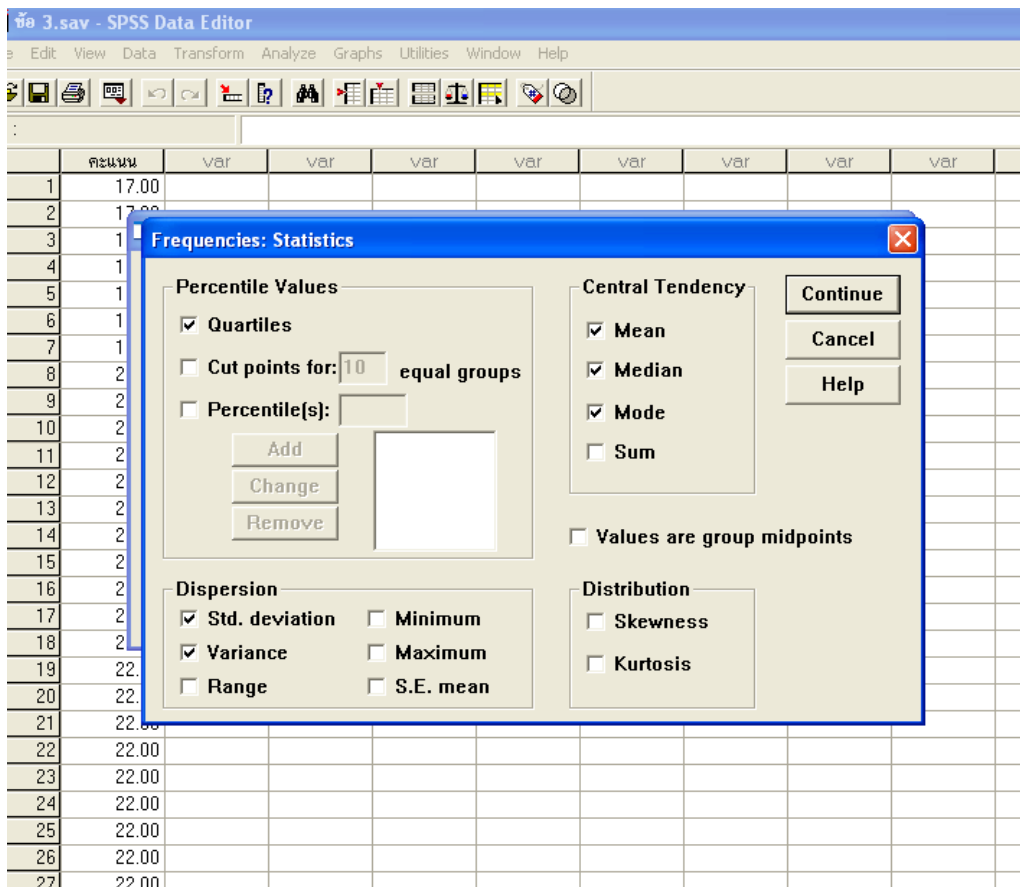
วิธีทำ

1. เปิดโปรแกรม SPSS 13.0

2. เลือกหน้าจอ Variable View และประกาศตัวแปร ดังรูป



5. จะปรากฏหน้าจอ Frequencies : Statistics แล้ว ให้ผู้ใช้ทำการเลือกค่าสถิติที่ต้องการ



6. เลือก Continue และกด O.K จะปรากฏหน้าจอ Out-Put ดังรูป

Statistics

คะแนน		
N	Valid	100
	Missing	0
Mean		25.9000
Median		27.0000
Mode		27.00
Std. Deviation		4.47101
Variance		19.990
Percentiles	25	22.0000
	50	27.0000
	75	27.0000

คะแนน

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	17.00	7	7.0	7.0	7.0
	22.00	30	30.0	30.0	37.0
	27.00	43	43.0	43.0	80.0
	32.00	18	18.0	18.0	98.0
	37.00	2	2.0	2.0	100.0
	Total	100	100.0	100.0	

จากตาราง Out-Put สามารถตอบคำถามจากโจทย์ คือ ผลสัมฤทธิ์ทางการเรียนในการทดสอบเรื่องปริมาณสารสัมพันธ์ มีค่าเฉลี่ย ค่ามัธยฐาน ค่าฐานนิยม เท่ากับ 25.90 , 27.00 , 27.00 คะแนน ตามลำดับ และจะมีค่าส่วนเบี่ยงเบนมาตรฐาน ค่าความแปรปรวน เท่ากับ 4.47 , 19.99 ตามลำดับ แต่สำหรับค่าส่วนเบี่ยงเบนควอไทล์นั้น สามารถคำนวณได้จาก

$$Q.D = \frac{Q_3 - Q_1}{2} = \frac{27.00 - 22.00}{2} = 2.5$$

4. ผลของการทดสอบในการว่ายน้ำระยะทาง 200 เมตร ของนักเรียน
ชั้นมัธยมศึกษาปีที่ 6 ของโรงเรียนแห่งหนึ่งจำนวน 3 ห้อง (นาที)

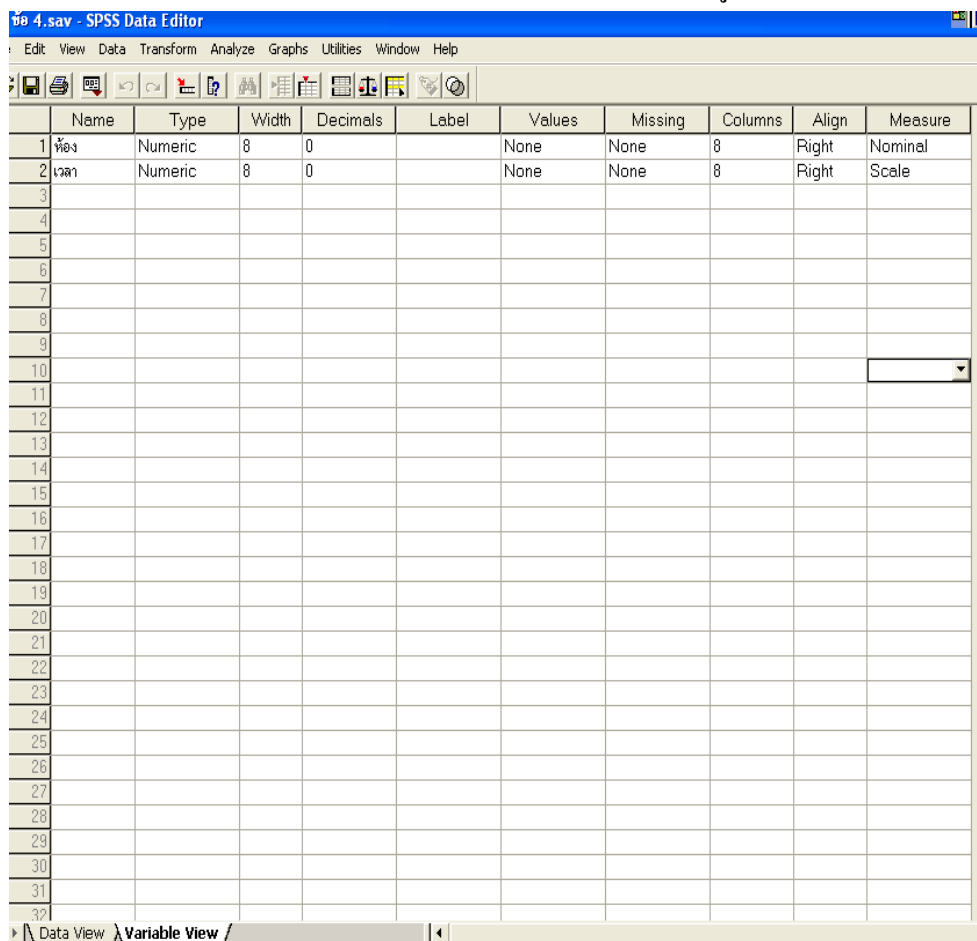
ห้องที่ 1	8	12	13	7	7	12	7	9
ห้องที่ 2	10	8	9	10	15	8	7	11
	8	15	10	12	15	8	10	
ห้องที่ 3	12	10	8	15	10	10	11	9
	9	7	13	10	12	14	10	

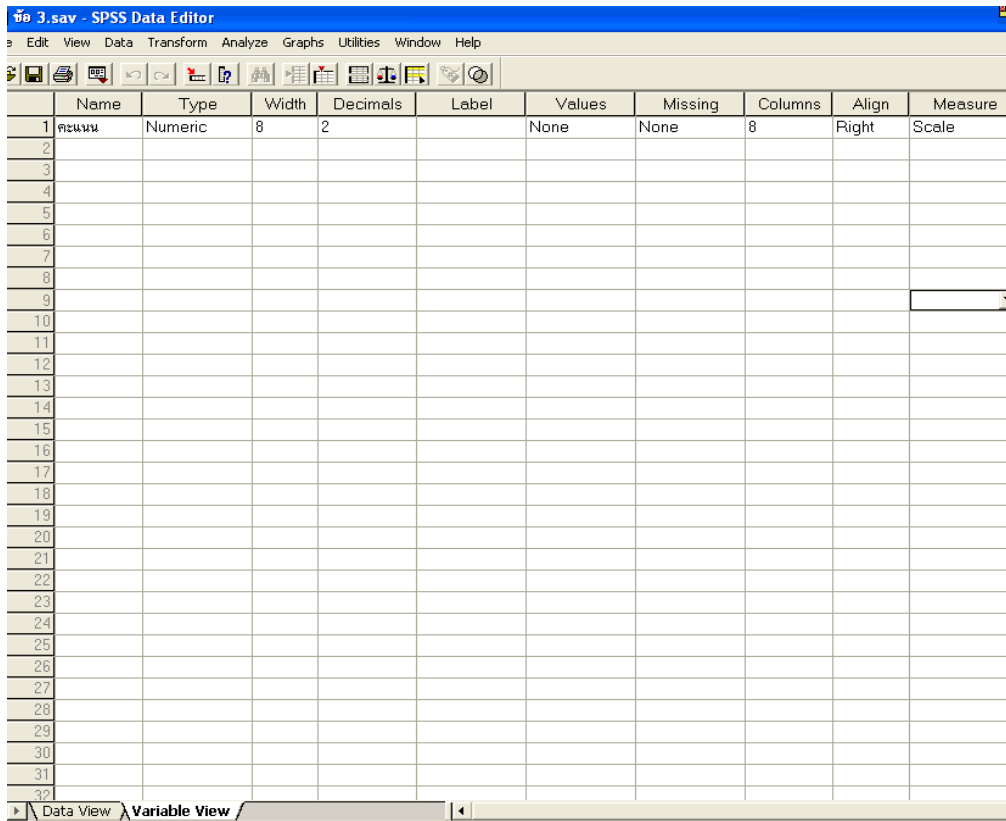
4.1 จงหาเวลาเฉลี่ยของแต่ละห้อง และเวลาเฉลี่ยรวมของการว่ายน้ำครั้งนี้

4.2 จงหาการกระจายของคะแนนสอบแต่ละห้องโดยใช้ ส่วนเบี่ยงเบน
มาตรฐาน (Standard Deviation)

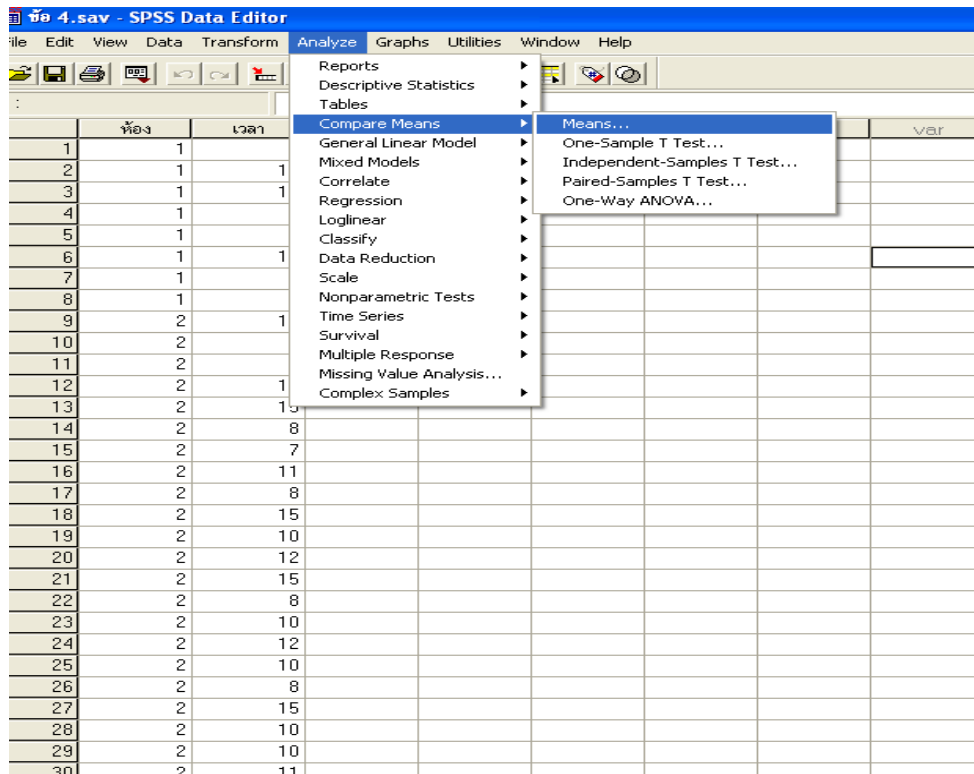
วิธีทำ

1. เปิดโปรแกรม SPSS 13.0
2. เลื่อนหน้าจอ Variable View และประกาศตัวแปร ดังรูป

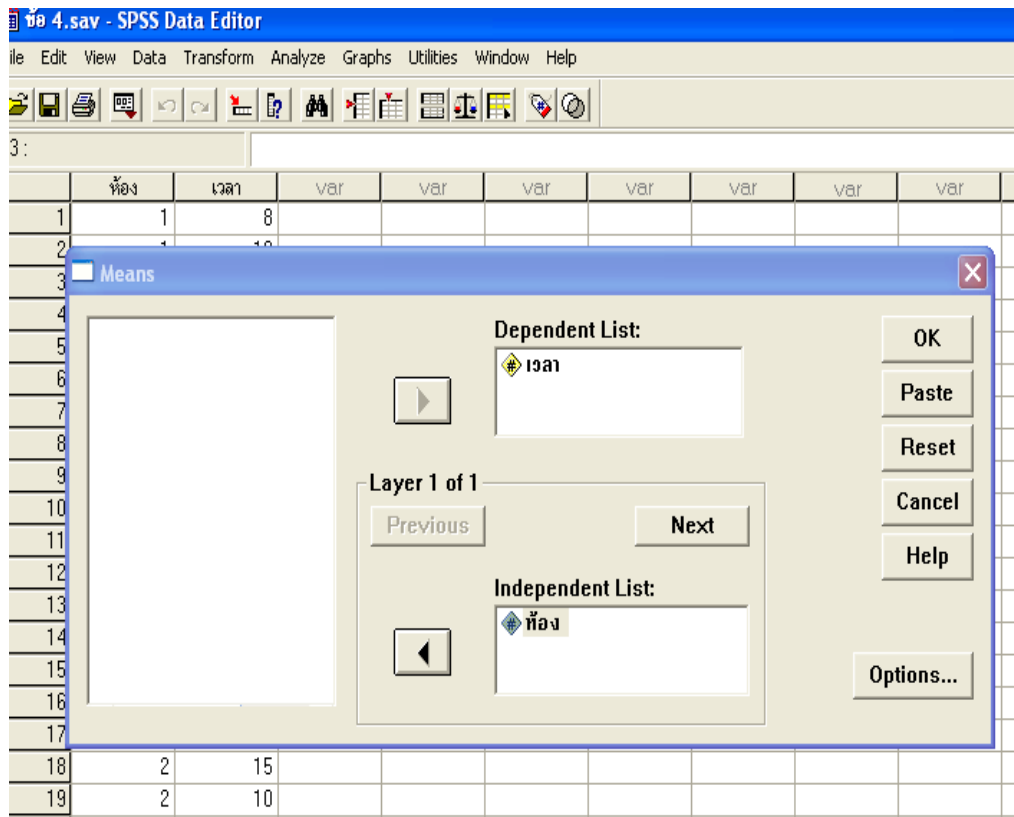




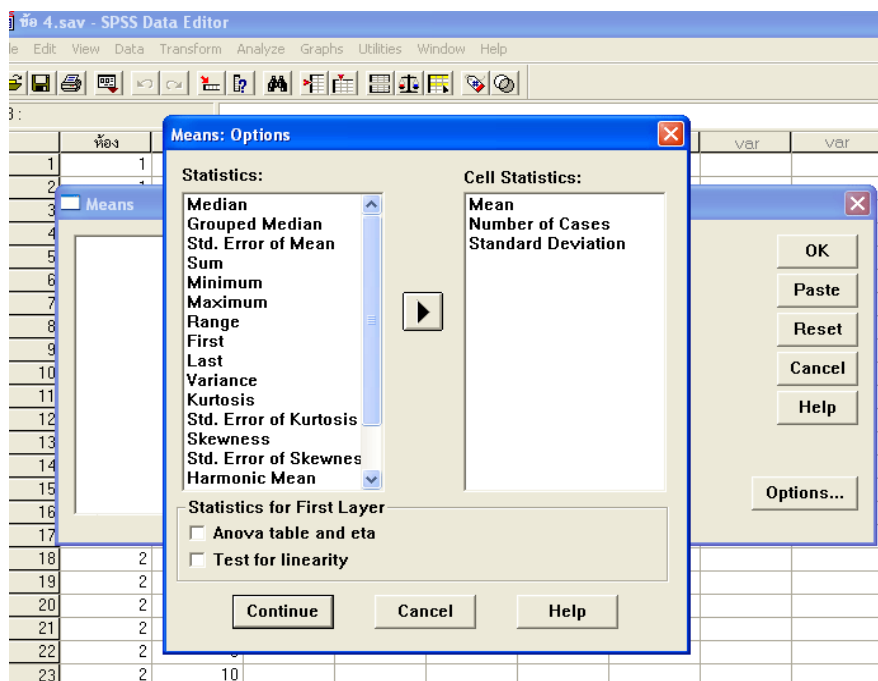
3. เลือกหน้าจอ Data View ทำการใส่ข้อมูลเวลาที่ใช้ในการว่ายน้ำลงไป เลือกเมนู Analyze ตามด้วย เมนู Compare Mean และ Mean เพราะจะเป็นการวิเคราะห์ข้อมูลภายในกลุ่มและระหว่างกลุ่ม ดังรูป



4. จะปรากฏหน้าจอ Mean ทำการเลือกตัวแปรโดยตัวแปรห้อง จะเป็นตัวแปรอิสระ (Independent List :) และตัวแปรเวลา จะเป็นตัวแปรตาม (Dependent List:) แล้วเลือกเมนู Options เพื่อเลือกค่าสถิติที่จะวิเคราะห์



5. จะปรากฏหน้าจอ Means : Options แล้ว ให้ผู้ใช้ทำการเลือกค่าสถิติที่ต้องการ



6. เลือก Continue และกด O.K จะปรากฏหน้าจอ Out-Put ดังรูป

Case Processing Summary

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
เวลา *ห้อง	38	100.0%	0	.0%	38	100.0%

Report

เวลา

ห้อง	Mean	N	Std. Deviation
1	9.38	8	2.560
2	10.55	22	2.521
3	10.50	8	2.330
Total	10.29	38	2.470

จากตาราง Out-Put สามารถตอบคำถามจากโจทย์ คือ ค่าเฉลี่ยของเวลาว่ายนํ้าในแต่ละห้อง เท่ากับ 9.38 , 10.55 , 10.50 นาที ตามลำดับ และค่าเฉลี่ยของเวลาว่ายนํ้ารวมเท่ากับ 10.29 นาที และจะมีค่าส่วนเบี่ยงเบนมาตรฐานของแต่ละห้อง เท่ากับ 2.560 , 2.521 , 2.330 วินาที ตามลำดับ