

หน่วยการเรียนรู้ที่

1

การวัดตำแหน่งที่ของข้อมูล



รายวิชาคณิตศาสตร์พื้นฐาน 6

รหัสวิชา (ค33104)

ชั้นมัธยมศึกษาปีที่ 6

สอนโดย

ครูดรชนี ดอกดวง

กลุ่มสาระการเรียนรู้คณิตศาสตร์

โรงเรียนปทุมพิทยาคม อำเภอเมือง จังหวัดอุบลราชธานี

สังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา อุบลราชธานี อำนาจเจริญ

ตัวชี้วัด

- เข้าใจและใช้ความรู้ทางสถิติในการนำเสนอข้อมูล และแปลความหมายของค่าสถิติเพื่อประกอบการตัดสินใจ (ค 3.1 ม.6/1)

การวัดตำแหน่งที่ของข้อมูล

ตัวชี้วัด

- เข้าใจและใช้ความรู้ทางสถิติในการนำเสนอข้อมูล และแปลความหมายของค่าสถิติเพื่อประกอบการตัดสินใจ (ค 3.1 ม.6/1)

ในการสอบแข่งขันวิชาคณิตศาสตร์ระหว่างโรงเรียน A+ และโรงเรียนอักษรทอง ซึ่งมีกลุ่มตัวแทนของแต่ละโรงเรียน ในการแข่งขันครั้งนี้ผลปรากฏว่า โรงเรียน A+ และโรงเรียนอักษรทอง มีคะแนนสอบเฉลี่ยเท่ากัน ดังนี้

โรงเรียน A⁺

คะแนนสอบเฉลี่ย

65.4

คะแนน

โรงเรียน อักษรทอง

คะแนนสอบเฉลี่ย

65.4

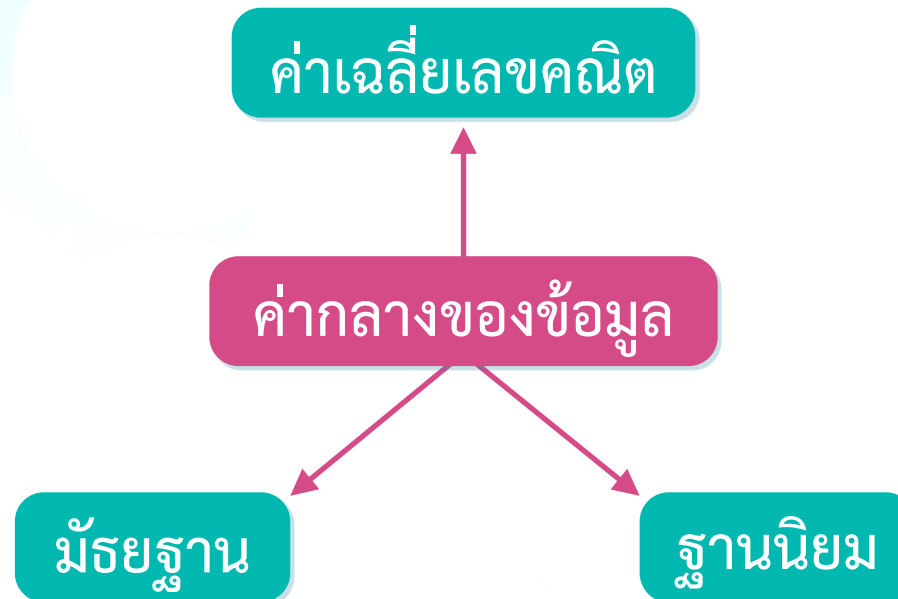
คะแนน

จากคะแนนสอบเฉลี่ยทั้งสองห้อง นักเรียนสามารถบอกได้หรือไม่ว่า กลุ่มตัวแทนจากโรงเรียนใดมีความสามารถในการสอบแข่งขันที่แตกต่างมากกว่ากัน

การวัดค่ากลางของข้อมูล

ค่ากลางของข้อมูล เป็นตัวแทนของข้อมูลในการสรุปผลและตีความหมายเกี่ยวกับข้อมูลนั้น ๆ มีด้วยกันหลายวิธี ขึ้นอยู่กับลักษณะของข้อมูลและวัตถุประสงค์ของผู้ใช้ข้อมูล ค่ากลางที่เป็นตัวแทนของข้อมูลที่นิยมใช้มี 3 ชนิด คือ **ค่าเฉลี่ยเลขคณิต มัธยฐาน และฐานนิยม**

ค่ากลางที่เป็นตัวแทนของข้อมูลที่นิยมใช้มี 3 ชนิด คือ **ค่าเฉลี่ยเลขคณิต มัธยฐาน และฐานนิยม**



ค่าเฉลี่ยเลขคณิต

ค่าเฉลี่ยเลขคณิต เป็นค่าที่ได้จากการเฉลี่ยของข้อมูลทั้งหมด ข้อมูลที่เหมาะสมนำมาหาค่าเฉลี่ยจะเป็นข้อมูลที่ไม่มีค่าสังเกตค่าใดค่าหนึ่งหรือหลาย ๆ ค่า ที่สูงหรือต่ำกว่าค่าสังเกตอื่นอย่างผิดปกติ

ค่าเฉลี่ยเลขคณิต แบ่งออกเป็น

ค่าเฉลี่ยเลขคณิตของข้อมูลที่ไม่ได้แจกแจงความถี่

ค่าเฉลี่ยเลขคณิตของข้อมูลที่ได้แจกแจงความถี่

ค่าเฉลี่ยเลขคณิตถ่วงน้ำหนัก

ค่าเฉลี่ยเลขคณิตรวม



ค่าเฉลี่ยเลขคณิตของข้อมูลที่ไม่ได้แจกแจงความถี่

ข้อมูลที่ไม่ได้แจกแจงความถี่ สามารถคำนวณค่าเฉลี่ยได้จากการนำข้อมูลทุกค่ามาบวกกันแล้วหารด้วยจำนวนข้อมูลทั้งหมด

ค่าเฉลี่ยเลขคณิตของประชากร

ให้ $x_1, x_2, x_3, \dots, x_N$ เป็นข้อมูลของประชากร และ N เป็นจำนวนของประชากร

ค่าเฉลี่ยเลขคณิตของประชากร คือ $\mu = \frac{x_1 + x_2 + x_3 + \dots + x_N}{N}$

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

ค่าเฉลี่ยเลขคณิตของตัวอย่าง

ให้ $x_1, x_2, x_3, \dots, x_n$ เป็นข้อมูลของประชากร และ n เป็นจำนวนของประชากร

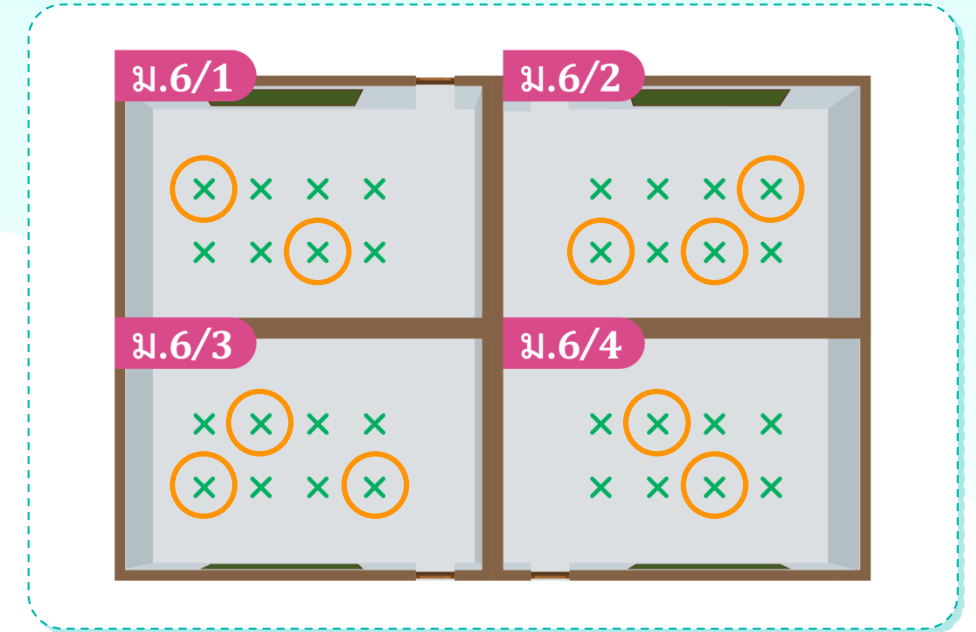
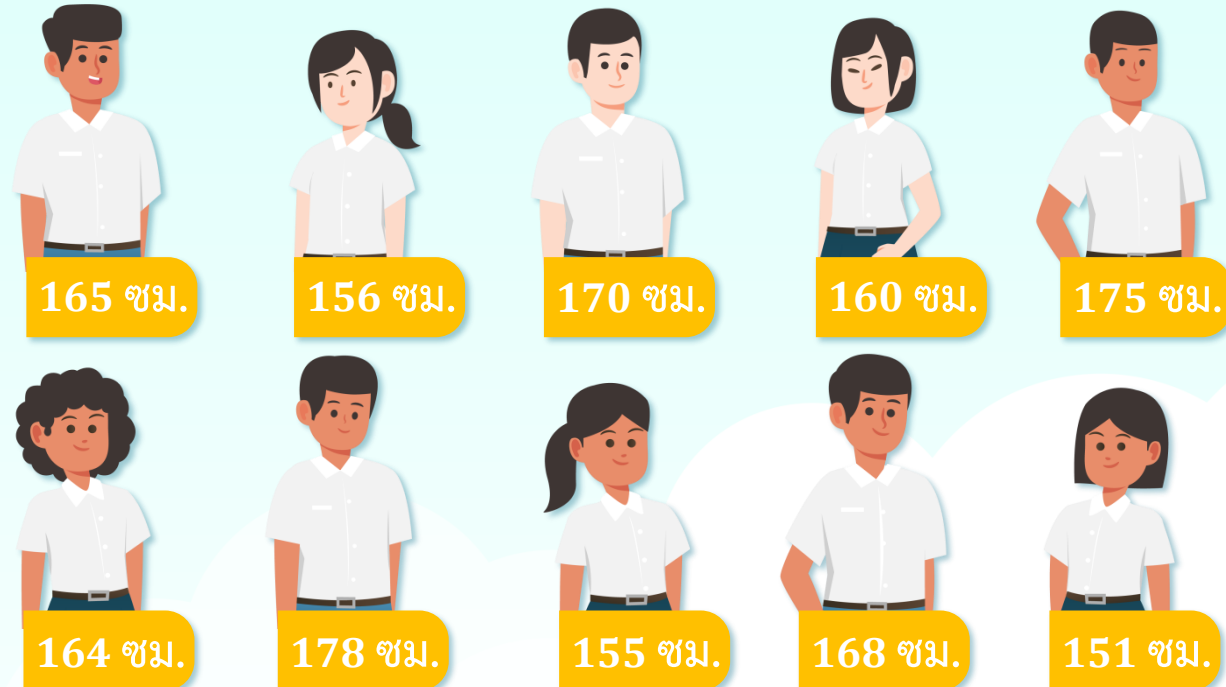
ค่าเฉลี่ยเลขคณิตของตัวอย่าง คือ $\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n}$

$$\text{หรือ } \bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$



ตัวอย่าง

ให้หาส่วนสูงเฉลี่ย (เซนติเมตร) ของนักเรียน 10 คน ที่เลือกมาเป็นตัวอย่างจากนักเรียนทั้งหมด 4 ห้อง



วิธีทำ ส่วนสูงเฉลี่ยของนักเรียนกลุ่มนี้

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$
$$\bar{x} = \frac{165 + 156 + 170 + 160 + 175 + 164 + 178 + 155 + 168 + 151}{10}$$

$$\bar{x} = \frac{1,642}{10}$$
$$= 164.2$$

ดังนั้น ส่วนสูงเฉลี่ยของนักเรียนกลุ่มนี้ คือ 164.2 เซนติเมตร



ค่าเฉลี่ยเลขคณิตของข้อมูลที่แจกแจงความถี่

ให้ $f_1, f_2, f_3, \dots, f_k$ เป็นความถี่ของค่าสังเกต $x_1, x_2, x_3, \dots, x_k$ และ n เป็นค่าสังเกตทั้งหมด

ความถี่	ค่าสังเกต	ความถี่ $\times$ ค่าสังเกต
f_1	x_1	$f_1 x_1$
f_2	x_2	$f_2 x_2$
f_3	x_3	$f_3 x_3$
$\vdots$	$\vdots$	$\vdots$
f_k	x_k	$f_k x_k$

ค่าเฉลี่ยเลขคณิตของประชากร คือ $\bar{x} = \frac{f_1 x_1 + f_2 x_2 + f_3 x_3 + \dots + f_k x_k}{f_1 + f_2 + f_3 + \dots + f_k}$

$$= \frac{\sum_{i=1}^k f_i x_i}{\sum_{i=1}^k f_i}$$

หรือ $\bar{x} = \frac{\sum_{i=1}^k f_i x_i}{n}$



ตัวอย่าง

ให้หาส่วนสูงเฉลี่ย (เซนติเมตร) ของนักเรียนกลุ่มหนึ่งที่**เลือกมา**เป็น**ตัวอย่าง**จากนักเรียนทั้งหมด 4 ห้อง

ส่วนสูง (x)	ความถี่ (f)
160	4
165	5
170	8
175	2
180	1

ข้อมูลชุดนี้มีการแจกแจงความถี่
ของข้อมูลแต่ละค่าดังตาราง
จึงหาค่าเฉลี่ยเลขคณิตได้จากสูตร

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

วิธีทำ ส่วนสูงเฉลี่ยของนักเรียนกลุ่มนี้

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

$$\bar{x} = \frac{(160)(4) + (165)(5) + (170)(8) + (175)(2) + (180)(1)}{20}$$

$$\begin{aligned}\bar{x} &= \frac{3,355}{20} \\ &= 167.75\end{aligned}$$

ดังนั้น ส่วนสูงเฉลี่ยของนักเรียนกลุ่มนี้ คือ 167.75 เซนติเมตร



ข้อมูลที่มีการแจกแจงความถี่ของข้อมูลแบบจัดกลุ่ม หรือเป็นอันตรภาคชั้น เช่น

ข้อมูลแต่ละชั้น
มีค่าสังเกตหลายค่า

จำนวนชั่วโมง	จำนวนพนักงาน (คน)
1 – 5	8
6 – 10	11
11 – 15	14
16 – 20	27
21 – 25	21
26 – 30	19

จำนวนชั่วโมง	จุดกึ่งกลาง (x_i)	ความถี่ (f_i)	$f_i x_i$
1 – 5	3	8	24
6 – 10	8	11	88
11 – 15	13	14	182
16 – 20	18	27	486
21 – 25	23	21	483
26 – 30	28	19	532

ดังนั้น เราจะใช้จุดกึ่งกลางของอันตรภาคชั้น
เป็นตัวแทนของค่าสังเกตของแต่ละอันตรภาคชั้น
ในการหาค่าเฉลี่ยเลขคณิต โดยใช้สูตร

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

จาก $\bar{x} = \frac{\sum_{i=1}^6 f_i x_i}{\sum_{i=1}^6 f_i}$

จะได้ $\bar{x} = \frac{1,795}{100}$
 $= 17.95$





ข้อมูลที่มีการแจกแจงความถี่ของข้อมูลแบบจัดกลุ่ม สามารถหาค่าเฉลี่ยเลขคณิตโดยใช้วิธีทอนค่า มีขั้นตอนดังนี้

ขั้นตอนที่ 1

หาจุดกึ่งกลางของแต่ละอันตรภาคชั้น
โดยกำหนด A เป็น จุดกึ่งกลางของชั้นที่มีความถี่มากที่สุด

ขั้นตอนที่ 2

หาค่าแตกต่าง (d_i) ระหว่างจุดกึ่งกลางของแต่ละ
อันตรภาคชั้น (x_i) กับค่า A เมื่อ $d_i = x_i - A$

ขั้นตอนที่ 3

หาค่า $f_i d_i$

ขั้นตอนที่ 4

หาค่าเฉลี่ยเลขคณิตจากสูตร $\bar{x} = A + \frac{\sum_{i=1}^k f_i d_i}{\sum_{i=1}^k f_i}$

	1		2	3
อันตรภาคชั้น	จุดกึ่งกลาง (x)	ความถี่ (f)	d_i	$f_i d_i$
1 – 5	3	8	-15	$(8)(-15) = -120$
6 – 10	8	11	-10	$(11)(-10) = -110$
11 – 15	13	14	-5	$(14)(-5) = -70$
16 – 20	18 (A)	27	0	$(27)(0) = 0$
21 – 25	23	21	5	$(21)(5) = 105$
26 – 30	28	19	10	$(19)(10) = 190$
	รวม	$\sum_{i=1}^6 f_i = 100$		$\sum_{i=1}^6 f_i d_i = -5$

4

$$\bar{x} = A + \frac{\sum_{i=1}^k f_i d_i}{\sum_{i=1}^k f_i}$$

$$\bar{x} = A + \frac{\sum_{i=1}^6 f_i d_i}{\sum_{i=1}^6 f_i}$$

$$\bar{x} = 18 + \left(\frac{-5}{100} \right)$$

$$= 17.95$$



ค่าเฉลี่ยเลขคณิตถ่วงน้ำหนัก

ค่าเฉลี่ยเลขคณิตถ่วงน้ำหนัก ใช้สำหรับข้อมูลที่ค่าสังเกตแต่ละค่ามีความสำคัญหรือน้ำหนักไม่เท่ากัน
ให้ $w_1, w_2, w_3, \dots, w_n$ เป็นความสำคัญหรือถ่วงน้ำหนักของค่าสังเกต $x_1, x_2, x_3, \dots, x_n$ ตามลำดับ

ค่าเฉลี่ยเลขคณิตถ่วงน้ำหนัก คือ $\bar{x} = \frac{w_1x_1 + w_2x_2 + w_3x_3 + \dots + w_nx_n}{w_1 + w_2 + w_3 + \dots + w_n}$

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



ค่าเฉลี่ยเลขคณิตถ่วงน้ำหนัก

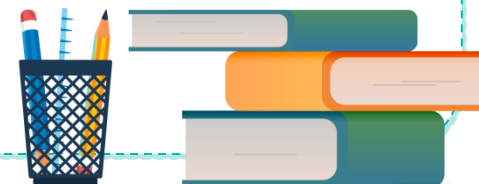
ตัวอย่าง

นายตะวัน ต้องการคิดเกรดเฉลี่ยของสี่วิชาตามที่ปรากฏในระเบียนแสดงผลการเรียนดังนี้

ระเบียนแสดงผลการเรียน		
รายวิชา	หน่วยกิต	ผลการเรียน
คณิตศาสตร์	1.0	4
ประวัติศาสตร์	0.5	3.5
ฟิสิกส์	2.0	2.5
เคมี	1.5	3

วิธีทำ ค่าสังเกตมี 4 ค่า คือ $x_1 = 4$, $x_2 = 3.5$, $x_3 = 2.5$ และ $x_4 = 3$
ความสำคัญของแต่ละวิชา คือ $w_1 = 1$, $w_2 = 0.5$, $w_3 = 2$ และ $w_4 = 1.5$

$$\begin{aligned}\text{เกรดเฉลี่ยทั้งสี่วิชา} &= \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i} \\ &= \frac{(1)(4) + (0.5)(3.5) + (2)(2.5) + (1.5)(3)}{1 + 0.5 + 2 + 1.5} \\ &= \frac{15.25}{5} \\ &= 3.05\end{aligned}$$



ค่าเฉลี่ยเลขคณิตรวม

ค่าเฉลี่ยเลขคณิตรวม ใช้หาค่าเฉลี่ยเลขคณิตของข้อมูลหลาย ๆ ชุด ที่หาค่าเฉลี่ยไว้แล้ว
ให้ $\bar{x}_1, \bar{x}_2, \bar{x}_3, \dots, \bar{x}_k$ เป็นค่าเฉลี่ยเลขคณิตของข้อมูลชุดที่ 1, 2, 3, ..., k
และ $n_1, n_2, n_3, \dots, n_k$ เป็นจำนวนค่าสังเกตในข้อมูลชุดที่ 1, 2, 3, ..., k ตามลำดับ

$$\begin{aligned}\text{ค่าเฉลี่ยเลขคณิตรวม} &= \frac{n_1\bar{x}_1 + n_2\bar{x}_2 + n_3\bar{x}_3 + \dots + n_k\bar{x}_k}{n_1 + n_2 + n_3 + \dots + n_k} \\ &= \frac{\sum_{i=1}^k n_i x_i}{\sum_{i=1}^k n_i}\end{aligned}$$

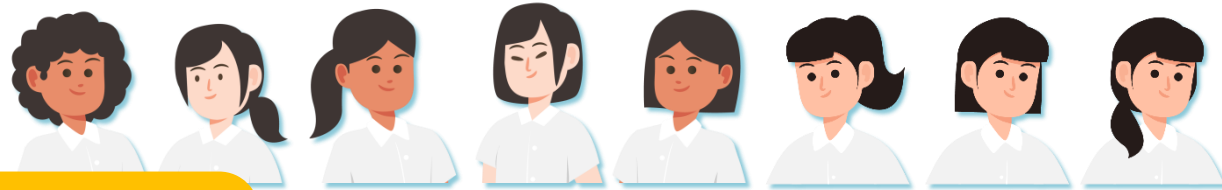


ค่าเฉลี่ยเลขคณิตรวม

ตัวอย่าง



นักเรียนชาย



นักเรียนหญิง

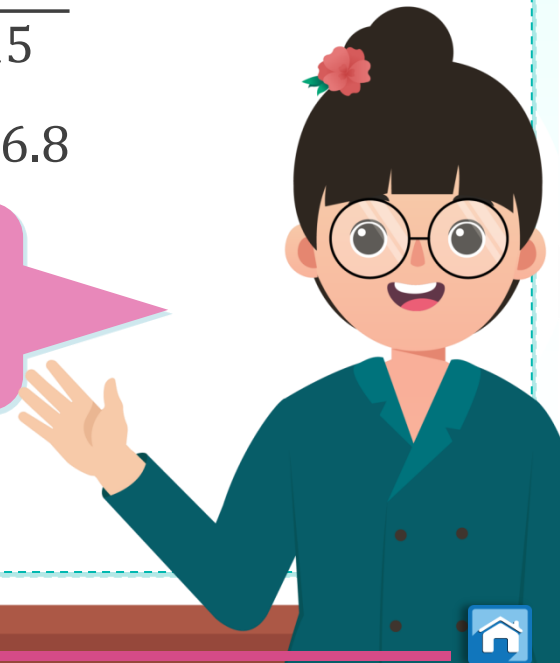
นักเรียนกลุ่มหนึ่งที่เลือกมาเป็นตัวอย่าง มีนักเรียนชาย 7 คน
และนักเรียนหญิง 8 คน ถ้านักเรียนชายมีส่วนสูงเฉลี่ย 178 เซนติเมตร
และนักเรียนหญิงมีส่วนสูงเฉลี่ย 157 เซนติเมตร ให้หาส่วนสูงเฉลี่ยของนักเรียนกลุ่มนี้

วิธีทำ ให้ $\bar{x}_1 = 178$ และ $\bar{x}_2 = 157$

$$n_1 = 7 \quad \text{และ} \quad n_2 = 8$$

$$\begin{aligned} \text{ค่าเฉลี่ยเลขคณิตรวม } \bar{x} &= \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2} \\ &= \frac{(7)(178) + (8)(157)}{7 + 8} \\ &= \frac{2,502}{15} \\ &= 166.8 \end{aligned}$$

นักเรียนกลุ่มนี้
มีส่วนสูงเฉลี่ย 166.8 เซนติเมตร

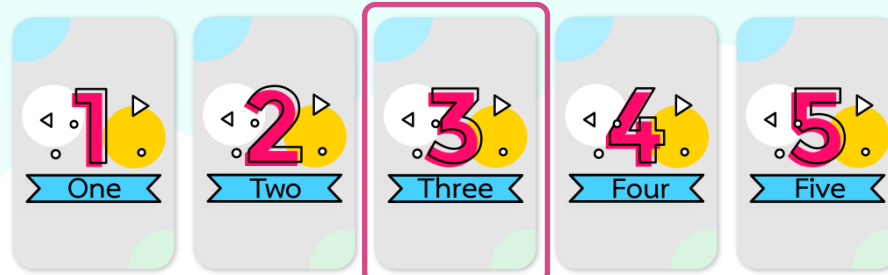


มัธยฐาน

มัธยฐาน เป็นค่ากลางที่เหมาะสมกับข้อมูลชุดใดชุดหนึ่งที่มีค่าบางค่าที่มากกว่าหรือน้อยกว่าข้อมูลอื่นอย่างผิดปกติ

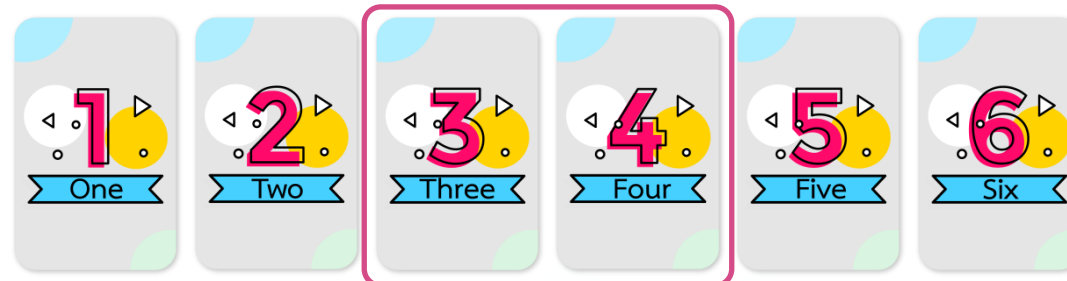
การหามัธยฐานของข้อมูลชุดหนึ่งจะต้องเรียงข้อมูลจากน้อยไปมาก หรือเรียงข้อมูลจากมากไปน้อย โดยทั่วไป ถ้าจำนวนข้อมูลมีทั้งหมด N ค่า แล้วมัธยฐานจะอยู่ในตำแหน่งที่ $\frac{N+1}{2}$

- ▶ ถ้าจำนวนข้อมูลทั้งหมดเป็น **จำนวนคี่** แล้วมัธยฐานจะเป็น **ค่าที่อยู่ตำแหน่งตรงกลางของข้อมูลทั้งหมด**



มัธยฐาน คือ ข้อมูลที่อยู่ในตำแหน่งที่ 3

- ▶ ถ้าจำนวนข้อมูลทั้งหมดเป็น **จำนวนคู่** แล้วมัธยฐานจะเป็น **ค่าเฉลี่ยของข้อมูลสองค่าที่อยู่ระหว่างกลางของข้อมูลทั้งหมด**

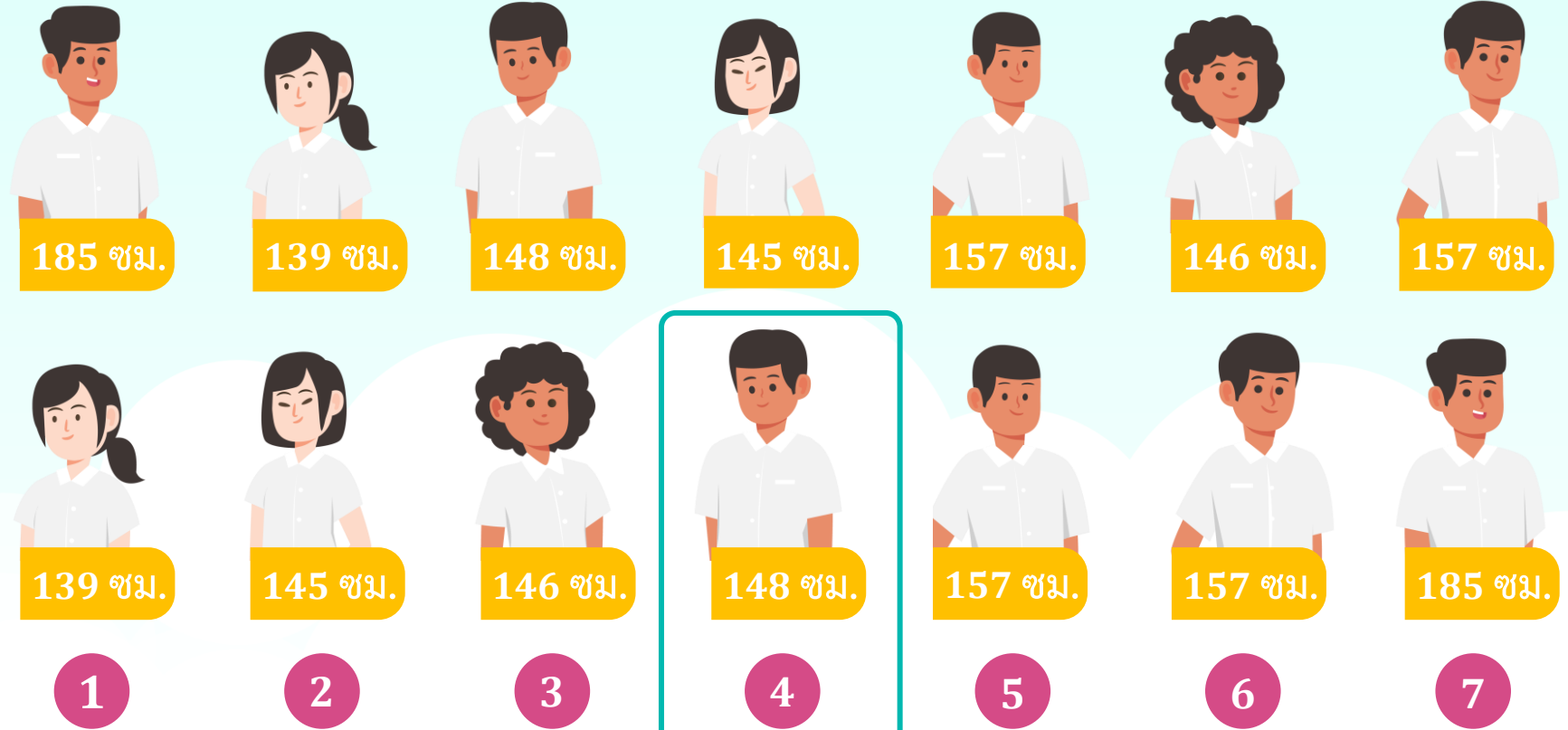


มัธยฐาน คือ ข้อมูลที่อยู่ในตำแหน่งที่ $\frac{3 + 4}{2} = 3.5$



ตัวอย่าง

ให้หามัธยฐานของส่วนสูงของนักเรียนกลุ่มหนึ่งที่เลือกสุ่มมาในโรงเรียน



เรียงข้อมูลจากน้อยไปมาก

หาตำแหน่งตรงกลางของข้อมูล

$$\frac{N + 1}{2} = \frac{7 + 1}{2} = 4$$

มัธยฐานของข้อมูลชุดนี้ คือ 148 เซนติเมตร

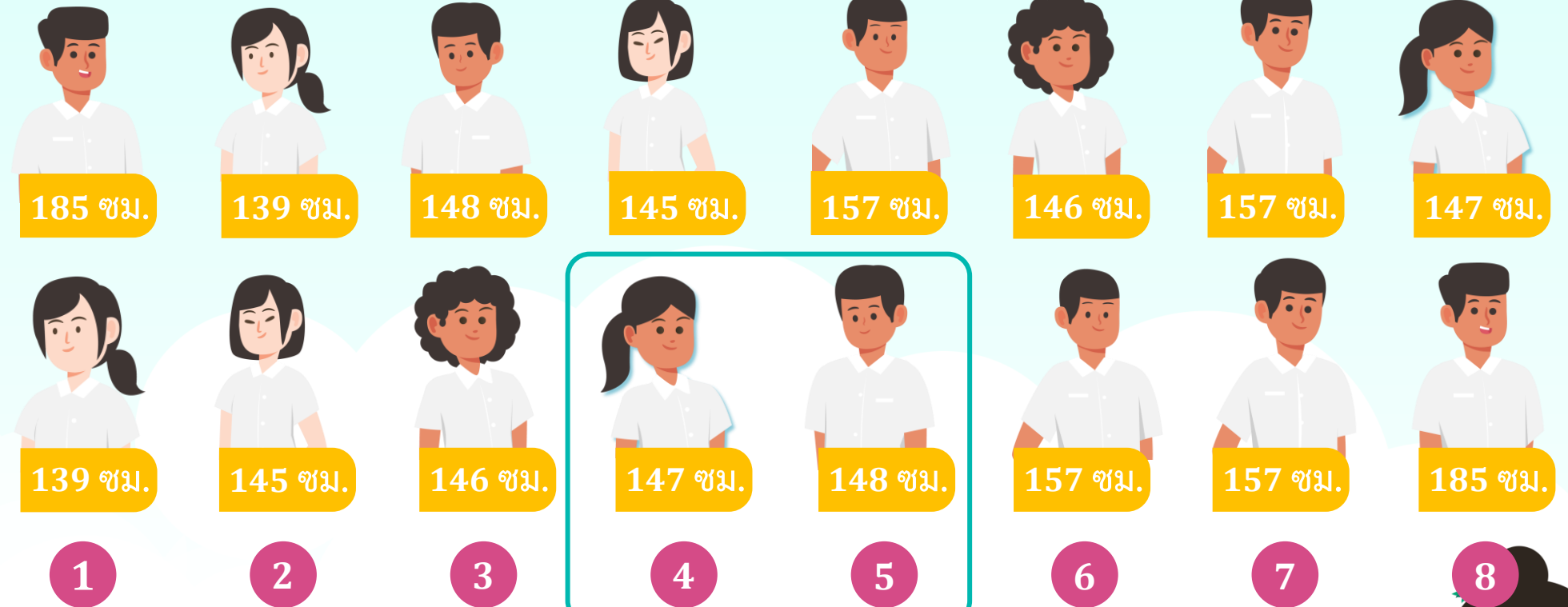
มัธยฐาน = 148 เซนติเมตร

นักเรียนจะเห็นว่าข้อมูลชุดนี้มีจำนวนข้อมูลทั้งหมดเป็นจำนวนคี่
ดังนั้น มัธยฐานจะเป็นค่าที่อยู่ตำแหน่งตรงกลางของข้อมูลทั้งหมด



ตัวอย่าง

ถ้าต้องการสູมนักเรียนเพิ่มอีกหนึ่งคน ให้หามัธยฐานของนักเรียนกลุ่มนี้



เรียงข้อมูลจากน้อยไปมาก

หาตำแหน่งตรงกลางของข้อมูล

$$\frac{N+1}{2} = \frac{8+1}{2} = 4.5$$

มัธยฐานของข้อมูลชุดนี้ คือ 147.5 เซนติเมตร

$$\text{มัธยฐาน} = \frac{147 + 148}{2} = 147.5 \text{ เซนติเมตร}$$

นักเรียนจะเห็นว่าข้อมูลชุดนี้มีจำนวนข้อมูลทั้งหมดเป็นจำนวนคู่
ดังนั้น มัธยฐานจะเป็นค่าเฉลี่ยของข้อมูลสองค่าที่อยู่ระหว่างกลางของข้อมูลทั้งหมด



ฐานนิยม

ฐานนิยม เป็นค่ากลางของข้อมูลที่มี**จำนวนซ้ำกันมากที่สุด**หรือ**มีความถี่สูงสุด** ข้อมูลนั้นจะเป็นฐานนิยมของข้อมูลชุดนั้น ซึ่งใช้ได้ทั้งข้อมูลเชิงคุณภาพและข้อมูลเชิงปริมาณ

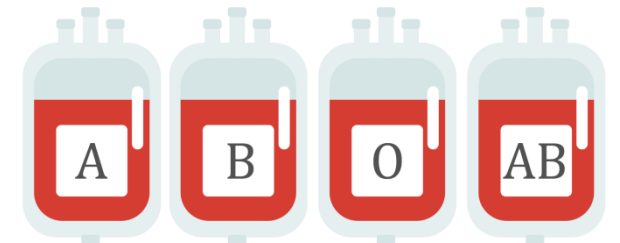
ตัวอย่างข้อมูลเชิงคุณภาพที่ใช้ฐานนิยม เช่น ขนาดของรองเท้า ยี่ห้อน้ำอัดลม กลุ่มเลือด



ขนาดรองเท้า



ยี่ห้อน้ำอัดลม



กลุ่มเลือด

ข้อมูลบางชุดอาจมีฐานนิยมมากกว่า 1 ค่า หรือไม่มีฐานนิยมเลย

- ▶ ถ้าข้อมูลชุดหนึ่งมีความถี่สูงสุด**เท่ากัน 2 ค่า** แล้วข้อมูลชุดนั้นจะมี**ฐานนิยม 2 ค่า**
- ▶ ถ้าข้อมูลชุดหนึ่งมีความถี่สูงสุดเท่ากัน**มากกว่า 2 ค่า** แล้วข้อมูลชุดนั้นจะ**ไม่มีฐานนิยม**



ตัวอย่าง

ให้หาฐานนิยมของส่วนสูงของนักเรียนกลุ่มนี้



นับข้อมูลที่มีจำนวนซ้ำกันมากที่สุด



ฐานนิยมของข้อมูลชุดนี้ คือ 155 ซม.



ถ้ามีนักเรียนเพิ่มอีก 1 คน ที่มีส่วนสูง 169 ซม.
จะเห็นว่า มีข้อมูลที่มีความถี่สูงสุดเท่ากัน 2 ค่า
ฐานนิยมของนักเรียนกลุ่มนี้ คือ 155 ซม. และ 169 ซม.



ถ้ามีนักเรียนเพิ่มอีก 1 คน ที่มีส่วนสูง 180 ซม.
จะเห็นว่า มีข้อมูลที่มีความถี่สูงสุดเท่ากันมากกว่า 2 ค่า
ข้อมูลชุดนี้ไม่มีฐานนิยม



145 ซม.



147 ซม.



155 ซม.



166 ซม.



169 ซม.



180 ซม.



ข้อสังเกตและหลักเกณฑ์ที่สำคัญในการใช้ค่ากลางชนิดต่าง ๆ



Mr.Mean

ค่าเฉลี่ยเลขคณิต

▶ เป็นค่ากลางที่ใช้ได้กับข้อมูลเชิงปริมาณเท่านั้น ซึ่งเป็นที่นิยมใช้และน่าเชื่อถือมากกว่ามัธยฐาน และฐานนิยม เนื่องจากใช้ข้อมูลทุกตัวมาคำนวณหาค่า แต่ข้อมูลชุดนั้นจะต้องไม่มีข้อมูลที่ต่ำกว่า หรือสูงกว่าข้อมูลอื่นอย่างผิดปกติ



Mr.Median

มัธยฐาน

▶ เป็นค่ากลางที่ใช้กับข้อมูลเชิงปริมาณเท่านั้น สามารถใช้กับข้อมูลที่มีค่าต่ำกว่าหรือสูงกว่าข้อมูลอื่นอย่างผิดปกติ ถ้าข้อมูลเป็นจำนวนคี่ แล้วมัธยฐานที่ได้จะเป็นค่าใดค่าหนึ่งของข้อมูลชุดนั้น แต่ถ้าข้อมูลเป็นจำนวนคู่ แล้วมัธยฐานที่ได้ อาจจะไม่ใช่ค่าใดค่าหนึ่งของข้อมูลชุดนั้น



Mr.Mode

ฐานนิยม

▶ เป็นค่ากลางที่ใช้ได้กับข้อมูลเชิงปริมาณและข้อมูลเชิงคุณภาพ ซึ่งฐานนิยมของข้อมูลบางชุดอาจมีมากกว่าหนึ่งค่า หรือข้อมูลบางชุดอาจไม่มีฐานนิยม



ปัญหาชวนคิด

นักเรียนคิดว่า Mr.Mean , Mr.Median และ Mr.Mode เหมาะสมกับสถานการณ์ใด

1

นายตะวัน อยากทราบว่าเงินเดือนเฉลี่ยของพนักงานบริษัทแห่งหนึ่งมีค่าเป็นเท่าใด ถ้าอัตราเงินเดือนของบริษัทมีค่าที่แตกต่างกันมาก



Mr.Mean



Mr.Median



Mr.Mode

2

นายอนันต์ อยากทราบว่าคะแนนสอบเฉลี่ยของวิชาคณิตศาสตร์ของนักเรียน 10 คน มีค่าเป็นเท่าใด ถ้านักเรียนกลุ่มนั้นมีคะแนนสอบใกล้เคียงกัน



Mr.Mean



Mr.Median



Mr.Mode

3

นางสาวใจดี อยากทราบว่าเพื่อนร่วมชั้นส่วนใหญ่ชอบรับประทานไอศกรีมรสชาติใดมากที่สุด



Mr.Mean



Mr.Median



Mr.Mode



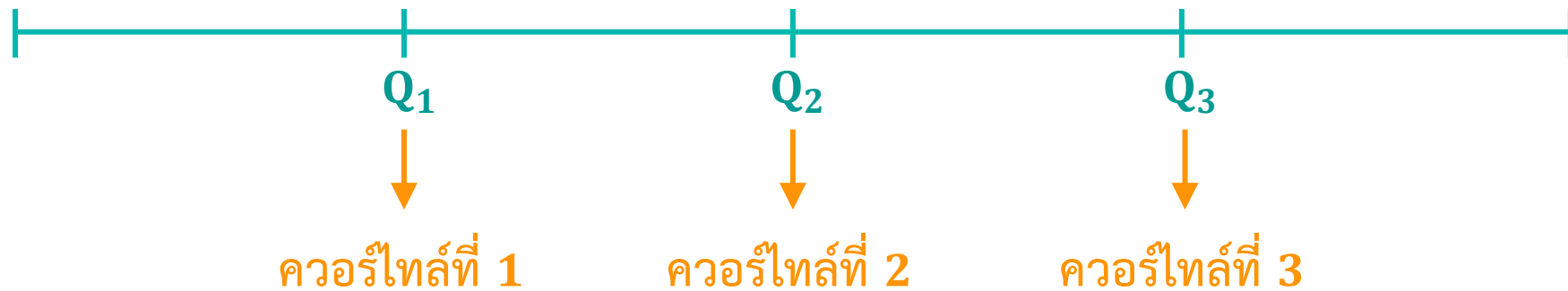
การวัดตำแหน่งที่ของข้อมูล



การวัดตำแหน่งที่ของข้อมูล สามารถบอกได้ว่าข้อมูลที่น่าสนใจอยู่ตรงตำแหน่งใดของข้อมูลทั้งหมด โดยมีข้อมูลอื่นประกอบ และสามารถบอกได้ว่าตำแหน่งนั้นเป็นตำแหน่งที่ดีหรือไม่เพียงใดในกลุ่ม จึงมีวิธีการบอกตำแหน่งโดยใช้ **ควอร์ไทล์** **เดไซล์** และ **เปอร์เซ็นต์ไทล์**

ควอร์ไทล์

ควอร์ไทล์ เป็นการแบ่งข้อมูลที่เรียงจากน้อยไปมากออกเป็น 4 ส่วน โดยที่แต่ละส่วนมีจำนวนข้อมูลเท่า ๆ กัน



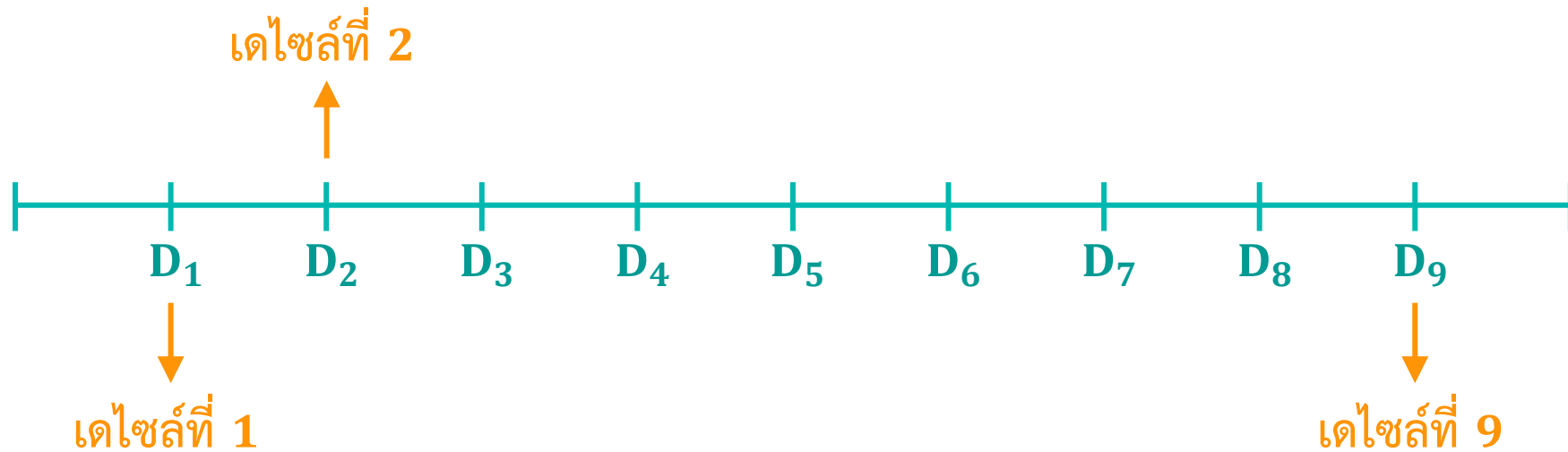
- ▶ **ควอร์ไทล์ที่ 1** คือ ค่าที่มีจำนวนข้อมูลน้อยกว่าค่านี้อยู่ประมาณ **หนึ่งในสี่** ของจำนวนข้อมูลทั้งหมด หรือ **ร้อยละ 25**
- ▶ **ควอร์ไทล์ที่ 2** คือ ค่าที่มีจำนวนข้อมูลน้อยกว่าค่านี้อยู่ประมาณ **หนึ่งในสอง** ของจำนวนข้อมูลทั้งหมด หรือ **ร้อยละ 50**
- ▶ **ควอร์ไทล์ที่ 3** คือ ค่าที่มีจำนวนข้อมูลน้อยกว่าค่านี้อยู่ประมาณ **สามในสี่** ของจำนวนข้อมูลทั้งหมด หรือ **ร้อยละ 75**



การวัดตำแหน่งที่ของข้อมูล

เดซิล์

เดซิล์ เป็นการแบ่งข้อมูลที่เรียงจากน้อยไปมากออกเป็น 10 ส่วน โดยที่แต่ละส่วนมีจำนวนข้อมูลเท่า ๆ กัน



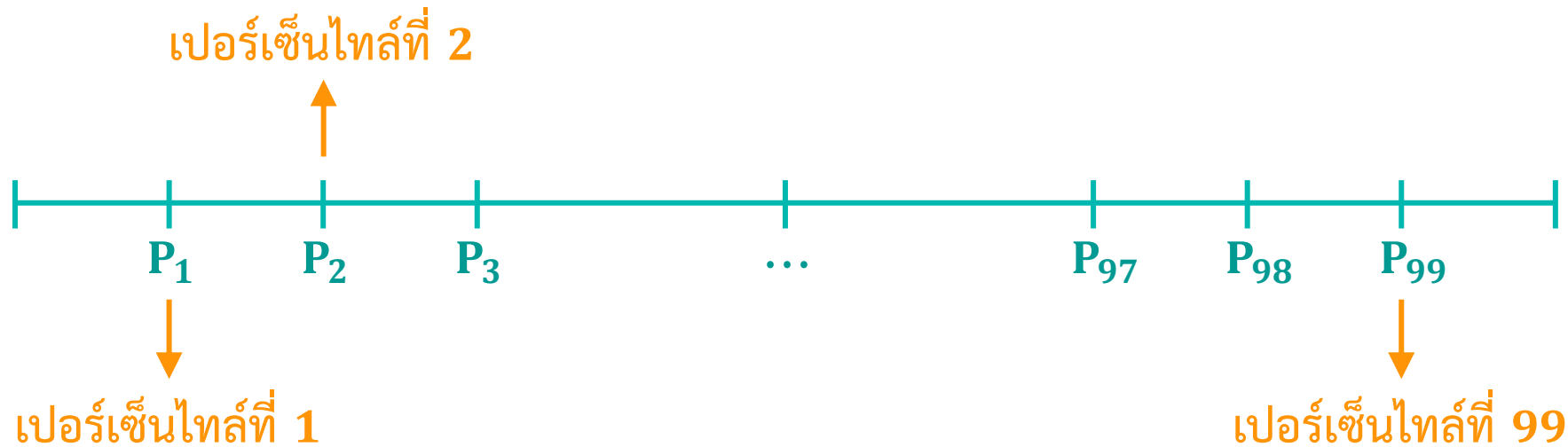
- ▶ **เดซิล์ที่ 1** คือ ค่าที่มีจำนวนข้อมูลน้อยกว่าค่านี้อยู่ประมาณ**หนึ่งในสิบ**ของจำนวนข้อมูลทั้งหมด
- ▶ **เดซิล์ที่ 2** คือ ค่าที่มีจำนวนข้อมูลน้อยกว่าค่านี้อยู่ประมาณ**สองในสิบ**ของจำนวนข้อมูลทั้งหมด
- ▶ **เดซิล์ที่ 9** คือ ค่าที่มีจำนวนข้อมูลน้อยกว่าค่านี้อยู่ประมาณ**เก้าในสิบ**ของจำนวนข้อมูลทั้งหมด



การวัดตำแหน่งที่ของข้อมูล

เปอร์เซ็นต์ไทล์

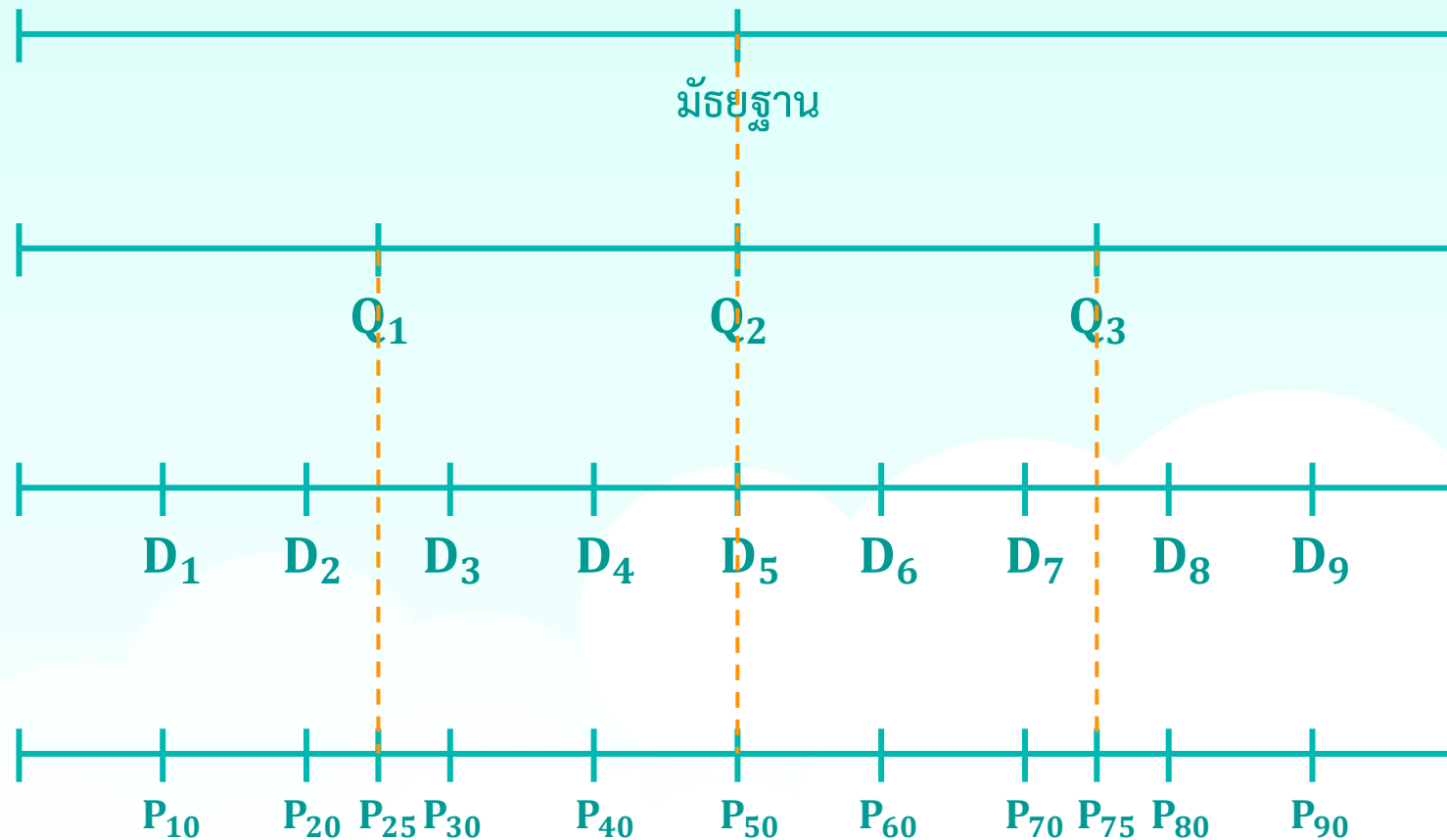
เปอร์เซ็นต์ไทล์ เป็นการแบ่งข้อมูลที่เรียงจากน้อยไปมากออกเป็น 100 ส่วน โดยที่แต่ละส่วนมีจำนวนข้อมูลเท่า ๆ กัน



- ▶ เปอร์เซ็นต์ไทล์ที่ 1 คือ ค่าที่มีจำนวนข้อมูลน้อยกว่าค่านี้อยู่ประมาณหนึ่งในร้อยของจำนวนข้อมูลทั้งหมด
- ▶ เปอร์เซ็นต์ไทล์ที่ 2 คือ ค่าที่มีจำนวนข้อมูลน้อยกว่าค่านี้อยู่ประมาณสองในร้อยของจำนวนข้อมูลทั้งหมด
- ⋮
- ▶ เปอร์เซ็นต์ไทล์ที่ 99 คือ ค่าที่มีจำนวนข้อมูลน้อยกว่าค่านี้อยู่ประมาณเก้าสิบเก้าในร้อยของจำนวนข้อมูลทั้งหมด



ความสัมพันธ์ของมัธยฐาน ควอร์ไทล์ เดซิส์ และเปอร์เซ็นต์ไทล์



แบ่งข้อมูลออกเป็น **2** ส่วนเท่า ๆ กัน

แบ่งข้อมูลออกเป็น **4** ส่วนเท่า ๆ กัน

แบ่งข้อมูลออกเป็น **10** ส่วนเท่า ๆ กัน

แบ่งข้อมูลออกเป็น **100** ส่วนเท่า ๆ กัน

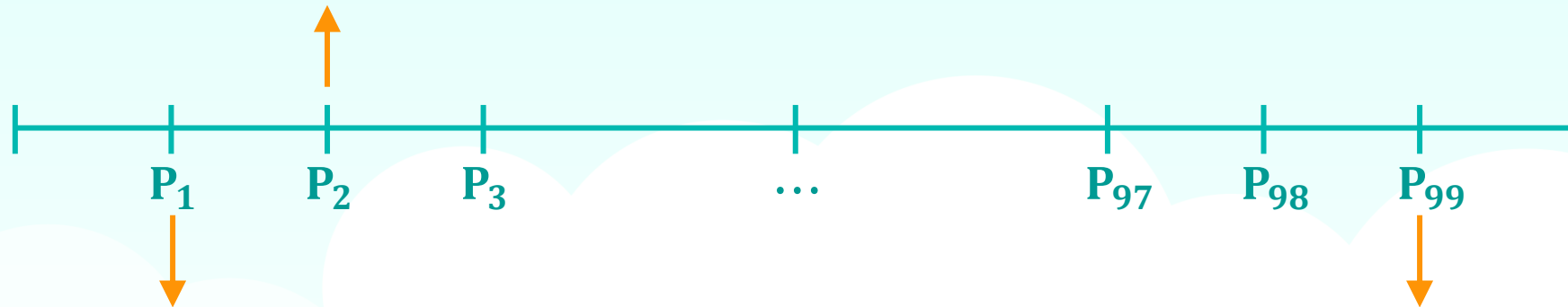
จากความสัมพันธ์ข้างต้น จะเห็นว่า $P_{25} = Q_1$
 $P_{50} = Q_2 = D_5 = \text{มัธยฐาน}$
 $P_{75} = Q_3$



การหาเปอร์เซ็นต์ไทล์ของข้อมูลที่ไม่ได้แจกแจงความถี่

การหาเปอร์เซ็นต์ไทล์ของข้อมูลที่ไม่ได้แจกแจงความถี่ จะต้องเรียงข้อมูลจากน้อยไปมาก แล้วหาตำแหน่งต่าง ๆ ของเปอร์เซ็นต์ไทล์

P_2 อยู่ในตำแหน่งที่ $\frac{2(N+1)}{100}$



P_1 อยู่ในตำแหน่งที่ $\frac{(N+1)}{100}$

P_{99} อยู่ในตำแหน่งที่ $\frac{99(N+1)}{100}$

ในกรณีทั่วไป จะหาตำแหน่งของเปอร์เซ็นต์ไทล์ได้จาก

ตำแหน่งของ P_r คือ $\frac{r(N+1)}{100}$ เมื่อ $r \in \{1, 2, \dots, 99\}$

เมื่อ r แทนตำแหน่งของเปอร์เซ็นต์ไทล์ และ N แทนจำนวนข้อมูลทั้งหมด



ตัวอย่าง

คะแนนสอบวิชาสถิติของนักเรียนกลุ่มหนึ่ง เป็นดังนี้

35 20 16 31 28 20 33 35 29

ให้หา P_{30} , P_{55} และ P_{78}

วิธีทำ ข้อมูลทั้งหมด 9 จำนวน จะได้ $N = 9$

เรียงข้อมูลจากน้อยไปมากได้ ดังนี้

16 20 20 28 29 31 33 35 35



ตัวอย่าง

คะแนนสอบวิชาสถิติของนักเรียนกลุ่มหนึ่ง เป็นดังนี้

35 20 16 31 28 20 33 35 29

ให้หา P_{30} , P_{55} และ P_{78}

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

16 20 20 28 29 31 33 35 35
1 2 3

ดังนั้น P_{30} คือ 20

ตำแหน่งของ P_{30} คือตำแหน่งที่ 3 ของข้อมูลชุดนี้



ตัวอย่าง

คะแนนสอบวิชาสถิติของนักเรียนกลุ่มหนึ่ง เป็นดังนี้

35 20 16 31 28 20 33 35 29

ให้หา P_{30} , P_{55} และ P_{78}

วิธีทำ หาตำแหน่งของ $P_{55} = \frac{55(9+1)}{100} = 5.5$

16 20 20 28 29 31 33 35 35
1 2 3 4 5 6
↓
5.5

ดังนั้น P_{55} คือ $\frac{29+31}{2} = 30$

ตำแหน่งที่ 5.5 เป็นตำแหน่งที่อยู่กึ่งกลางระหว่าง
ตำแหน่งที่ 5 คือ 29 และตำแหน่งที่ 6 คือ 31
ดังนั้น P_{55} จึงเป็นค่าเฉลี่ยระหว่าง 29 และ 31



ตัวอย่าง

คะแนนสอบวิชาสถิติของนักเรียนกลุ่มหนึ่ง เป็นดังนี้

35 20 16 31 28 20 33 35 29

ให้หา P_{30} , P_{55} และ P_{78}

ตำแหน่งที่ 7.8 เป็นตำแหน่งที่อยู่ระหว่าง
ตำแหน่งที่ 7 คือ 33 และตำแหน่งที่ 8 คือ 35

วิธีทำ หาตำแหน่งของ $P_{78} = \frac{78(9+1)}{100} = 7.8$

16 20 20 28 29 31 33 35 35

1 2 3 4 5 6 7 8

7.8

ตำแหน่งที่ต่างกัน $8 - 7 = 1$

ข้อมูลต่างกัน $35 - 33 = 2$

ตำแหน่งที่ต่างกัน $7.8 - 7 = 0.8$

ข้อมูลต่างกัน $\frac{2}{1} \times 0.8 = 1.6$

ดังนั้น P_{78} คือ $33 + 1.6 = 34.6$



การวัดการกระจายของข้อมูล

การวัดการกระจายของข้อมูล เป็นค่าสถิติที่ใช้อธิบายลักษณะการกระจายของข้อมูล เพื่อให้เห็นลักษณะของข้อมูล ได้ชัดเจนมากขึ้น ซึ่งในที่นี้จะกล่าวถึงการวัดการกระจายสัมบูรณ์ ได้แก่ **พิสัย** และ **ส่วนเบี่ยงเบนมาตรฐาน**

การวัดการกระจายของข้อมูล แบ่งออกเป็น

พิสัย

ส่วนเบี่ยงเบนมาตรฐาน



พิสัย

ข้อมูลที่ไม่ได้มีการแจกแจงความถี่

พิสัย คือ ค่าที่ใช้วัดการกระจายที่ได้จากผลต่างระหว่างข้อมูลที่มีค่าสูงสุดและข้อมูลที่มีค่าต่ำสุด ถ้า $x_1, x_2, x_3, \dots, x_n$ เป็นข้อมูลชุดหนึ่ง แล้วพิสัยของข้อมูลชุดนี้เท่ากับ $x_{\max} - x_{\min}$ เมื่อ $x_{\max}$ เป็นค่าสูงสุดของข้อมูล และ $x_{\min}$ เป็นค่าต่ำสุดของข้อมูล

ตัวอย่าง

ให้หาพิสัยของข้อมูล 39, 12, 40, 35, 47, 22, 36, 30

วิธีทำ จากข้อมูล 39 12 40 35 47 22 36 30

↓ ↓

$x_{\min}$ $x_{\max}$

จะได้ พิสัย = $x_{\max} - x_{\min}$
 $= 47 - 12$
 $= 35$

ดังนั้น พิสัยข้อมูลชุดนี้เท่ากับ 35



พิสัย

ข้อมูลที่มีการแจกแจงความถี่โดยแบ่งอันตรภาคชั้น

พิสัย คือ ผลต่างระหว่างขอบบนของอันตรภาคชั้นของข้อมูลที่มีค่าสูงสุดและขอบล่างของอันตรภาคชั้นของข้อมูลที่มีค่าต่ำสุด

ตัวอย่าง

ให้หาพิสัยของข้อมูลชุดนี้

ข้อมูล	ความถี่
11 – 20	4
21 – 30	7
31 – 40	14
41 – 50	2
51 – 60	3

วิธีทำ เนื่องจากขอบบนของอันตรภาคชั้นของข้อมูลที่มีค่าสูงสุดเท่ากับ 60.5 และขอบล่างของอันตรภาคชั้นของข้อมูลที่มีค่าต่ำสุดเท่ากับ 10.5
จะได้ $\text{พิสัย} = 60.5 - 10.5 = 50$

ดังนั้น พิสัยของข้อมูลชุดนี้เท่ากับ 50

ข้อมูล	ขอบล่าง-ขอบบน	ความถี่
11 – 20	10.5 – 20.5	4
21 – 30	20.5 – 30.5	7
31 – 40	30.5 – 40.5	14
41 – 50	40.05 – 50.5	2
51 – 60	50.5 – 60.5	3

ขอบล่างของข้อมูลที่มีค่าต่ำสุด

ขอบบนของข้อมูลที่มีค่าสูงสุด



ส่วนเบี่ยงเบนมาตรฐาน

การวัดการกระจายของข้อมูลโดยใช้ส่วนเบี่ยงเบนมาตรฐาน เป็นวิธีที่ใช้วัดการกระจายได้ดีที่สุด เนื่องจากการวัดการกระจายวิธีนี้จะใช้ข้อมูลทุก ๆ ค่าหรือนำตัวแทนของข้อมูลทุกค่ามาคำนวณ

ส่วนเบี่ยงเบนมาตรฐานของประชากร

ให้ $x_1, x_2, x_3, \dots, x_N$ เป็นข้อมูล N จำนวนของประชากร และ k แทนจำนวนอันตรภาคชั้น

▶ ข้อมูลที่ไม่ได้แจกแจงความถี่

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

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

▶ ข้อมูลที่มีการแจกแจงความถี่

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

หรือ $\sigma = \sqrt{\frac{\sum_{i=1}^k f_i x_i^2}{N} - \mu^2}$

ส่วนเบี่ยงเบนมาตรฐานของตัวอย่าง

ให้ $x_1, x_2, x_3, \dots, x_n$ เป็นข้อมูล n จำนวนของตัวอย่าง และ k แทนจำนวนอันตรภาคชั้น

▶ ข้อมูลที่ไม่ได้แจกแจงความถี่

ส่วนเบี่ยงเบนมาตรฐานของตัวอย่าง คือ $s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$

หรือ $s = \sqrt{\frac{\sum_{i=1}^n x_i^2 - n\bar{x}^2}{n - 1}}$

▶ ข้อมูลที่มีการแจกแจงความถี่

ส่วนเบี่ยงเบนมาตรฐานของตัวอย่าง คือ $s = \sqrt{\frac{\sum_{i=1}^k f_i (x_i - \bar{x})^2}{n - 1}}$

หรือ $s = \sqrt{\frac{\sum_{i=1}^k f_i x_i^2 - n\bar{x}^2}{n - 1}}$

ตัวอย่าง

ให้หาส่วนเบี่ยงเบนมาตรฐานของคะแนนสอบของนักเรียนกลุ่มหนึ่งที่เลือกมาเป็นตัวอย่าง เป็นดังนี้

54 55 52 60 64 70 53 72

วิธีทำ หาค่า $\bar{x}$ และ x_i^2

x_i	x_i^2
54	$54^2 = 2,916$
55	$55^2 = 3,025$
52	$52^2 = 2,704$
60	$60^2 = 3,600$
64	$64^2 = 4,096$
70	$70^2 = 4,900$
53	$53^2 = 2,809$
72	$72^2 = 5,184$
$\sum_{i=1}^8 x_i = 480$	$\sum_{i=1}^8 x_i^2 = 29,234$

$$\begin{aligned}\text{จะได้ } \bar{x} &= \frac{\sum_{i=1}^8 x_i}{8} \\ &= \frac{480}{8} \\ &= 60\end{aligned}$$

$$\begin{aligned}\text{และ } s &= \sqrt{\frac{\sum_{i=1}^8 x_i^2 - n\bar{x}^2}{n-1}} \\ &= \sqrt{\frac{29,234 - (8)(60)^2}{8-1}} \\ &= \sqrt{62} \\ &\approx 7.87\end{aligned}$$

จากข้อมูล จะเห็นว่าเป็นข้อมูลตัวอย่าง และเป็นข้อมูลที่ไม่ได้แจกแจงความถี่

ดังนั้น ส่วนเบี่ยงเบนมาตรฐานของคะแนนสอบของนักเรียนกลุ่มนี้มีค่าประมาณ 7.87 คะแนน



ส่วนเบี่ยงเบนมาตรฐาน

ในกรณีที่มีการแจกแจงความถี่ของข้อมูลแบบจัดกลุ่ม สามารถหาส่วนเบี่ยงเบนมาตรฐานได้ โดยใช้จุดกึ่งกลางของแต่ละอันตรภาคชั้นแทนข้อมูลของอันตรภาคชั้นนั้น ๆ ซึ่งใช้วิธีการคำนวณวิธีเดียวกันกับการหาส่วนเบี่ยงเบนมาตรฐานของข้อมูลที่ไม่มีการแจกแจงความถี่

ตัวอย่าง

ให้หาส่วนเบี่ยงเบนมาตรฐานของข้อมูลตัวอย่างชุดนี้

ข้อมูล	ความถี่
1 – 10	5
11 – 20	9
21 – 30	15
31 – 40	8
41 – 50	3

จากข้อมูล จะเห็นว่าเป็นข้อมูลตัวอย่าง และเป็นข้อมูลที่มีแจกแจงความถี่ของข้อมูลแบบจัดกลุ่ม



วิธีทำ

ข้อมูล	ความถี่ (f_i)	จุดกึ่งกลาง (x_i)	x_i^2	$f_i x_i$	$f_i x_i^2$
1 – 10	5	3	9	15	45
11 – 20	9	8	64	72	576
21 – 30	15	13	169	195	2,535
31 – 40	8	18	324	144	2,592
41 – 50	3	23	529	69	1,587
รวม	$\sum_{i=1}^5 f_i = 40$			$\sum_{i=1}^5 f_i x_i = 495$	$\sum_{i=1}^5 f_i x_i^2 = 7,335$

$$\begin{aligned}\text{จะได้ } \bar{x} &= \frac{\sum_{i=1}^5 f_i x_i}{n} \\ &= \frac{495}{40}\end{aligned}$$

$$\begin{aligned}s &= \sqrt{\frac{\sum_{i=1}^5 f_i x_i^2 - n\bar{x}^2}{n - 1}} \\ &= \sqrt{\frac{7,335 - (40)\left(\frac{495}{40}\right)^2}{40 - 1}} \\ &\approx 5.57\end{aligned}$$

ดังนั้น ส่วนเบี่ยงเบนมาตรฐาน
ของข้อมูลชุดนี้มีค่าประมาณ **5.57**



ส่วนเบี่ยงเบนมาตรฐาน



ถ้าส่วนเบี่ยงเบนมาตรฐานมี**ค่ามาก** แสดงว่าข้อมูลมี**การกระจายมาก**
ถ้าส่วนเบี่ยงเบนมาตรฐานมี**ค่าน้อย** แสดงว่าข้อมูลมี**การกระจายน้อย**

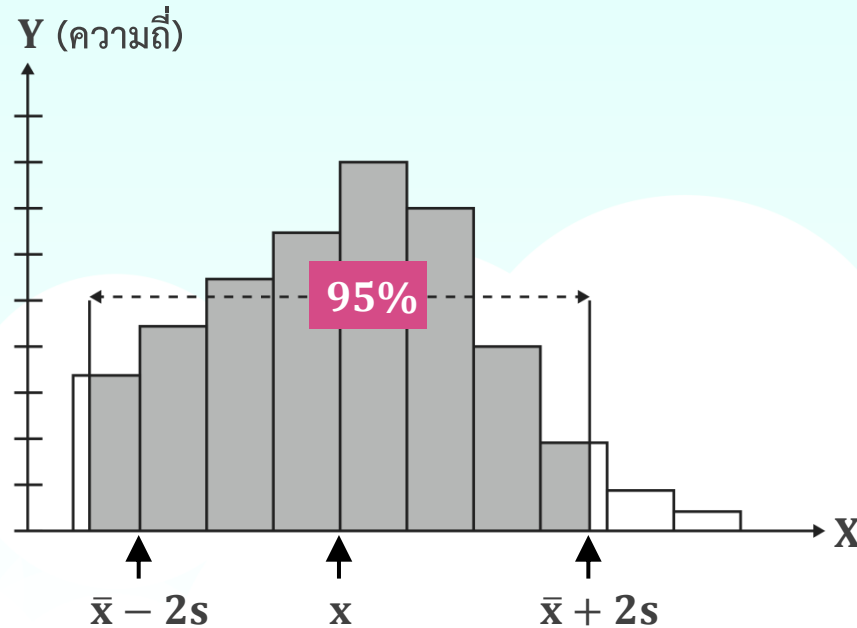
สมบัติของส่วนเบี่ยงเบนมาตรฐาน

- ▶ หน่วยของส่วนเบี่ยงเบนมาตรฐานเป็นหน่วยเดียวกันกับข้อมูล
- ▶ ส่วนเบี่ยงเบนมาตรฐานมีค่ามากกว่าหรือเท่ากับ 0
- ▶ ถ้าทุกค่าของข้อมูลเท่ากัน แล้วส่วนเบี่ยงเบนมาตรฐานมีค่าเป็น 0 แสดงว่าข้อมูลนั้นไม่มีการกระจายของข้อมูล



ส่วนเบี่ยงเบนมาตรฐาน

จำนวนข้อมูลที่อยู่ในช่วง $(\bar{x} - 2s, \bar{x} + 2s)$ หรือ $(\mu - 2\sigma, \mu + 2\sigma)$ มีค่าประมาณร้อยละ 95 ของจำนวนข้อมูลทั้งหมด ซึ่งเป็นตามกฎเกณฑ์ที่เรียกว่า “**The 95% Rule**” ได้กล่าวว่า โดยทั่วไปไม่ว่าข้อมูลจะกระจายในลักษณะใด จะมีข้อมูลประมาณ 95% ที่อยู่ในช่วง $(\bar{x} - 2s, \bar{x} + 2s)$ หรือ $(\mu - 2\sigma, \mu + 2\sigma)$



ถ้าข้อมูลมีลักษณะเป็นระฆังคว่ำและมีลักษณะโดยประมาณสมมาตร
แล้วจะประมาณส่วนเบี่ยงเบนมาตรฐานได้จากความสัมพันธ์

$$R \approx 4s \text{ หรือ } s \approx \frac{R}{4}$$

เมื่อ R คือ พิสัย และ s คือ ส่วนเบี่ยงเบนมาตรฐาน



ความแปรปรวน

ความแปรปรวน คือ กำลังสองของส่วนเบี่ยงเบนมาตรฐาน

ความแปรปรวนของประชากร

ให้ $x_1, x_2, x_3, \dots, x_N$ เป็นข้อมูล N จำนวนของประชากร และ k แทนจำนวนอันตรภาคชั้น

▶ ข้อมูลที่ไม่ได้แจกแจงความถี่

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

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

▶ ข้อมูลที่มีการแจกแจงความถี่

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

$$\text{หรือ } \sigma^2 = \frac{\sum_{i=1}^k f_i x_i^2}{N} - \mu^2$$

ความแปรปรวนของตัวอย่าง

ให้ $x_1, x_2, x_3, \dots, x_n$ เป็นข้อมูล n จำนวนของตัวอย่าง และ k แทนจำนวนอันตรภาคชั้น

▶ ข้อมูลที่ไม่ได้แจกแจงความถี่

ความแปรปรวนของตัวอย่าง คือ $s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$

$$\text{หรือ } s^2 = \frac{\sum_{i=1}^n x_i^2 - n\bar{x}^2}{n - 1}$$

▶ ข้อมูลที่มีการแจกแจงความถี่

ความแปรปรวนของตัวอย่าง คือ $s^2 = \frac{\sum_{i=1}^k f_i (x_i - \bar{x})^2}{n - 1}$

$$\text{หรือ } s^2 = \frac{\sum_{i=1}^k f_i x_i^2 - n\bar{x}^2}{n - 1}$$

ตัวอย่าง

ในการสุ่มคะแนนสอบวิชาคณิตศาสตร์ของนักเรียน 6 คน จากนักเรียนกลุ่มหนึ่ง เป็นดังนี้

17 8 11 15 9 12

- 1) ให้หาความแปรปรวนของคะแนนสอบของนักเรียน 6 คนนี้
- 2) ถ้าคะแนนสอบของนักเรียนแต่ละคนเพิ่มขึ้นคนละ 5 คะแนน แล้วความแปรปรวนของคะแนนสอบชุดใหม่จะเป็นเท่าใด

วิธีทำ 1)

$$\begin{aligned}\bar{x} &= \frac{17 + 8 + 11 + 15 + 9 + 12}{6} \\ &= \frac{72}{6} \\ &= 12\end{aligned}$$

$$\begin{aligned}s^2 &= \frac{\sum_{i=1}^N x_i^2 - n\bar{x}^2}{n - 1} \\ &= \frac{(17^2 + 8^2 + 11^2 + 15^2 + 9^2 + 12^2) - 6(12)^2}{6 - 1} \\ &= \frac{924 - 864}{5} \\ &= 12\end{aligned}$$

ดังนั้น ความแปรปรวนของคะแนนสอบของนักเรียน 6 คนนี้ เท่ากับ 12



ตัวอย่าง

ในการสุ่มคะแนนสอบวิชาคณิตศาสตร์ของนักเรียน 6 คน จากนักเรียนกลุ่มหนึ่ง เป็นดังนี้

17 8 11 15 9 12

1) ให้หาความแปรปรวนของคะแนนสอบของนักเรียน 6 คนนี้

2) ถ้าคะแนนสอบของนักเรียนแต่ละคนเพิ่มขึ้นคนละ 5 คะแนน แล้วความแปรปรวนของคะแนนสอบชุดใหม่จะเป็นเท่าใด

จากข้อ 2) นักเรียนจะเห็นว่า ข้อมูลชุดใหม่จะมีความแปรปรวนเท่ากับข้อมูลชุดเดิม กล่าวคือ ถ้านำจำนวนจริง c ไปบวกหรือลบกับข้อมูลทุก ๆ ตัวแล้ว ความแปรปรวนของข้อมูลชุดใหม่จะมีค่าเท่ากับความแปรปรวนของข้อมูลชุดเดิม

วิธีทำ 2)

ข้อมูลชุดเดิม

17 8 11 15 9 12

+5
→

ข้อมูลชุดใหม่

22 13 16 20 14 17

หาค่าเฉลี่ยเลขคณิตของคะแนนสอบชุดใหม่ของนักเรียน 6 คนนี้

$$\begin{aligned}\text{จะได้ } \bar{x} &= \frac{22 + 13 + 16 + 20 + 14 + 17}{6} \\ &= \frac{104}{6} \\ &= 17\end{aligned}$$

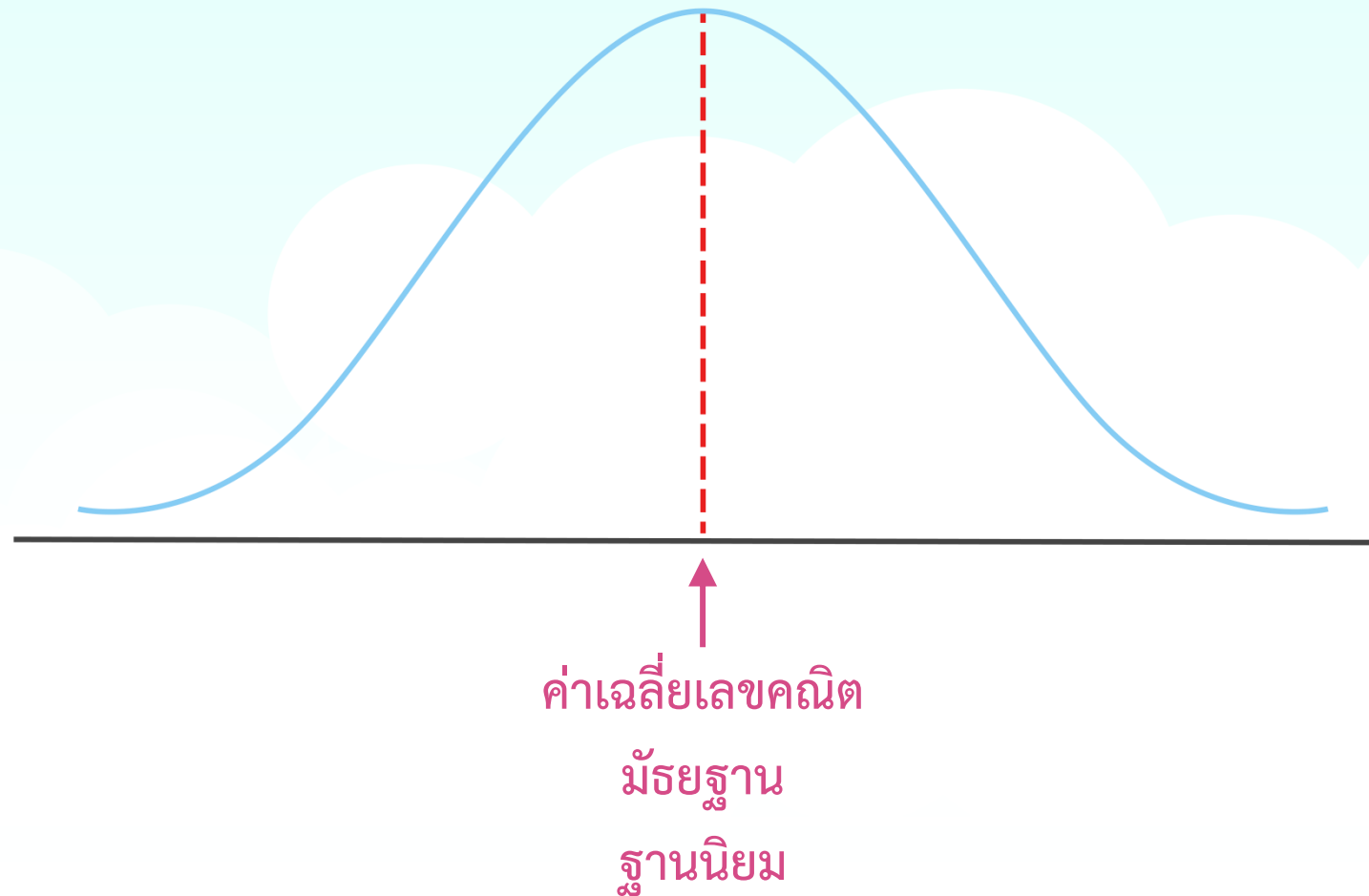
$$\begin{aligned}s^2 &= \frac{\sum_{i=1}^N x_i^2 - n\bar{x}^2}{n - 1} \\ &= \frac{(22^2 + 13^2 + 16^2 + 20^2 + 14^2 + 17^2) - 6(17)^2}{6 - 1} \\ &= \frac{1,794 - 1,734}{5} \\ &= 12\end{aligned}$$

ดังนั้น ความแปรปรวนของคะแนนสอบชุดใหม่เท่ากับ 12



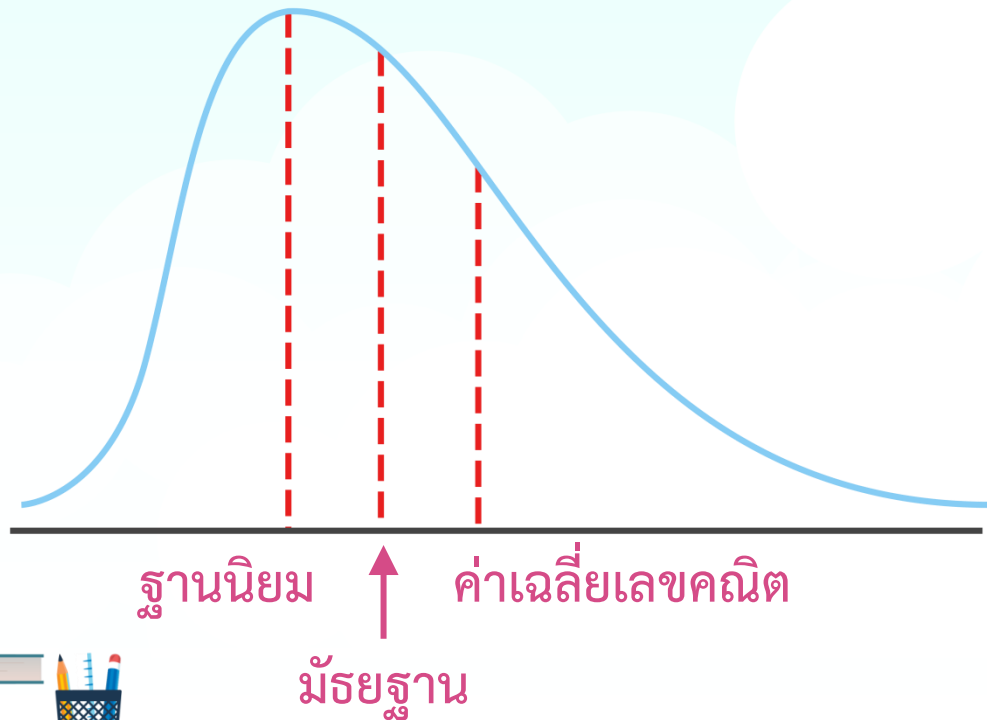
● ความสัมพันธ์ระหว่างการแจกแจงความถี่ ค่ากลาง และค่าการกระจายของข้อมูล ●

ถ้าข้อมูลชุดใดมี **ค่ากลางเท่ากันทั้ง 3 ค่า** (ค่าเฉลี่ยเลขคณิต มัธยฐาน และฐานนิยม มีค่าเท่ากัน)
แล้วข้อมูลชุดนั้นจะมีลักษณะการกระจายของข้อมูลเป็นแบบสมมาตร (symmetric distribution)

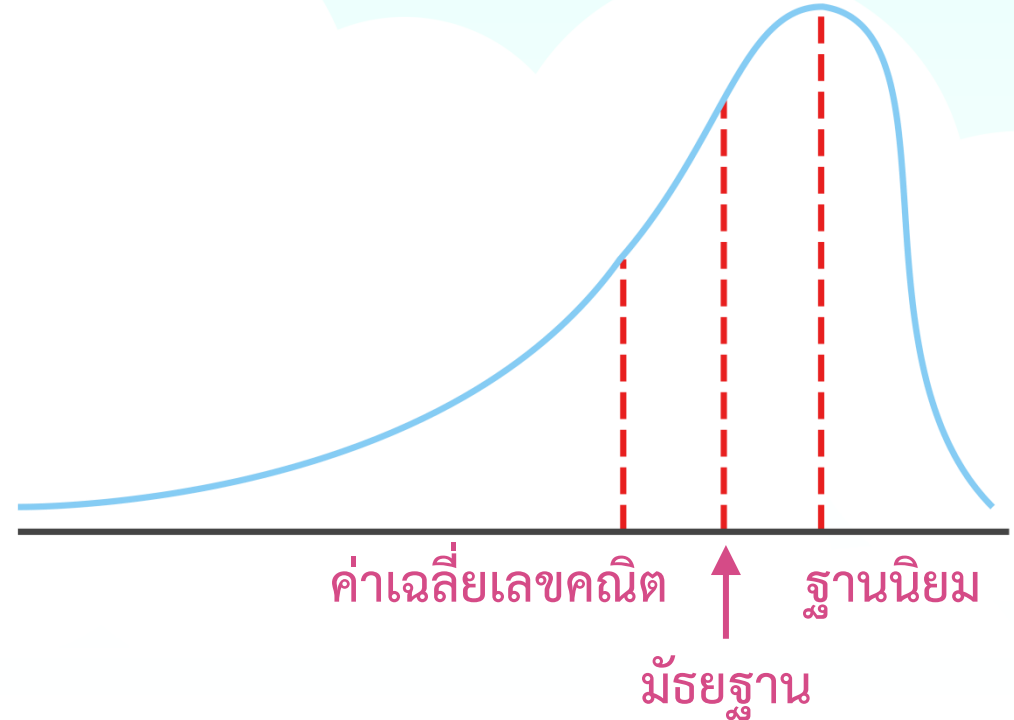


ความสัมพันธ์ระหว่างการแจกแจงความถี่ ค่ากลาง และค่าการกระจายของข้อมูล

ถ้าข้อมูลชุดใดมีค่ากลางทั้ง 3 ค่าไม่เท่ากัน โดยที่
 $\text{ฐานนิยม} < \text{มัธยฐาน} < \text{ค่าเฉลี่ยเลขคณิต}$
แล้วข้อมูลชุดนั้น จะมีลักษณะการกระจาย
ที่เบ้ทางขวา (right-skewed distribution)



ถ้าข้อมูลชุดใดมีค่ากลางทั้ง 3 ค่าไม่เท่ากัน โดยที่
 $\text{ค่าเฉลี่ยเลขคณิต} < \text{มัธยฐาน} < \text{ฐานนิยม}$
แล้วข้อมูลชุดนั้น จะมีลักษณะการกระจาย
ที่เบ้ทางซ้าย (left-skewed distribution)



แผนภาพกล่อง

นอกจากการวัดการกระจายของข้อมูลโดยใช้พิสัยและส่วนเบี่ยงเบนมาตรฐาน
ยังสามารถใช้แผนภาพกล่องเพื่อให้ทราบถึงลักษณะการกระจายของข้อมูลจากมัธยฐาน

ถ้านำข้อมูลชุดหนึ่งซึ่งเรียงจากน้อยไปมากมาเขียนแผนภาพกล่องจะได้ ดังนี้

12 15 19 23 27 33 37 39 42 44 45 48 50 53 55

หาข้อมูลที่มีค่าต่ำสุด, ค่าสูงสุด, Q_1 , Q_2 และ Q_3

จะได้ ข้อมูลที่มีค่าต่ำสุด คือ 12 ข้อมูลที่มีค่าสูงสุด คือ 55

$$Q_1 \text{ อยู่ในตำแหน่ง } \frac{N+1}{4} = \frac{15+1}{4} = 4$$

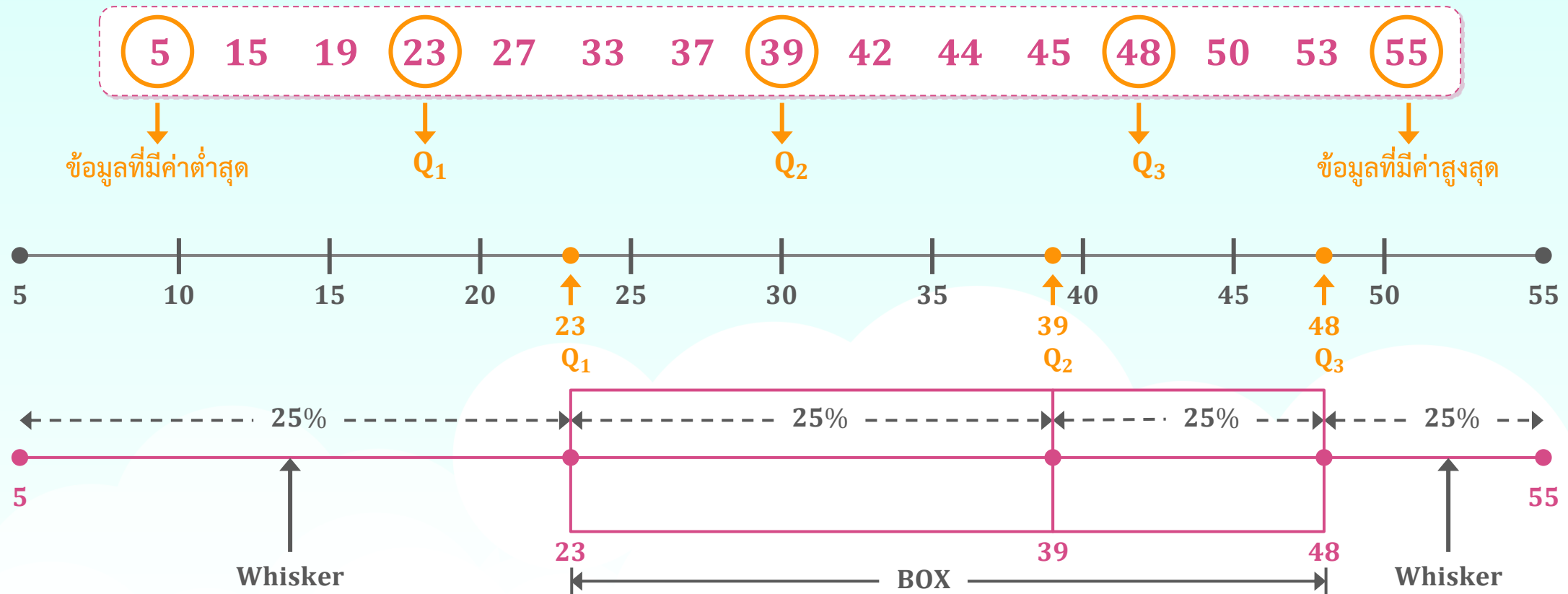
$$Q_2 \text{ อยู่ในตำแหน่ง } \frac{2(N+1)}{4} = \frac{2(15+1)}{4} = 8$$

$$Q_3 \text{ อยู่ในตำแหน่ง } \frac{3(N+1)}{4} = \frac{3(15+1)}{4} = 12$$



แผนภาพกล่อง

จากข้อมูลข้างต้น จะสร้างแผนภาพกล่องได้ ดังนี้



จากแผนภาพ จะเห็นว่า

- ▶ ข้อมูลที่มีค่าอยู่ระหว่าง 5 ถึง 23 และ 48 ถึง 55 มีจำนวนเท่ากัน คือ ประมาณ 25% ของจำนวนข้อมูลทั้งหมด
- ▶ ข้อมูลที่มีค่าอยู่ระหว่าง 23 ถึง 48 และ 39 ถึง 55 มีประมาณ 50% ของจำนวนข้อมูลทั้งหมด
- ▶ ข้อมูลทั้งหมด (100%) มีค่าอยู่ระหว่าง 5 ถึง 55
- ▶ ข้อมูลที่มีค่าอยู่ระหว่าง Q_1 และ Q_2 จะมีการกระจายมากกว่าข้อมูลที่อยู่ระหว่าง Q_2 และ Q_3 และมีการกระจายมากกว่าข้อมูลที่อยู่ระหว่าง Q_3 ถึงค่าที่มากที่สุด





นักเรียนจำคำถามในตอนต้นได้ไหมว่า
“จากคะแนนสอบเฉลี่ยทั้งสองห้อง นักเรียนสามารถบอกได้หรือไม่ว่า
กลุ่มตัวแทนจากโรงเรียนใดที่มีความสามารถในการสอบแข่งขันที่แตกต่างมากกว่ากัน”

เราสามารถหาคำตอบนี้ได้ โดยนำข้อมูลจากกลุ่มตัวแทนทั้งสองโรงเรียน
มาหาค่าส่วนเบี่ยงเบนมาตรฐาน ดังตัวอย่างต่อไปนี้



กำหนดให้

คะแนนสอบวิชาคณิตศาสตร์ของนักเรียนทั้งสองกลุ่ม ดังนี้

โรงเรียน A ⁺	โรงเรียนอักษรทอง
60	43
65	58
62	89
70	64
70	73
$\bar{x} = 65.4$	$\bar{x} = 65.4$

วิธีทำ หาค่า x_i^2 ของข้อมูลแต่ละกลุ่ม ดังนี้



โรงเรียน A ⁺	x_i^2
60	3,600
65	4,225
62	3,844
70	4,900
70	4,900
$\bar{x} = 65.4$	21,469

โรงเรียนอักษรทอง	x_i^2
43	1,849
58	3,364
89	7,921
64	4,096
73	5,329
$\bar{x} = 65.4$	22,559

$$\begin{aligned}
 s &= \sqrt{\frac{\sum_{i=1}^5 x_i^2 - n\bar{x}^2}{n-1}} \\
 &= \sqrt{\frac{21,469 - (5)(65.4)^2}{5-1}} \\
 &= \sqrt{20.8} \\
 &\approx 4.56
 \end{aligned}$$

$$\begin{aligned}
 s &= \sqrt{\frac{\sum_{i=1}^5 x_i^2 - n\bar{x}^2}{n-1}} \\
 &= \sqrt{\frac{22,559 - (5)(65.4)^2}{5-1}} \\
 &= \sqrt{293.3} \\
 &\approx 17.13
 \end{aligned}$$



สรุปเป็นตารางได้ ดังนี้

	ค่าเฉลี่ยเลขคณิต	ส่วนเบี่ยงเบนมาตรฐาน
โรงเรียน A ⁺	65.4	4.56
โรงเรียนอักษรทอง	65.4	17.13



เมื่อพิจารณาสวนเบี่ยงเบนมาตรฐานจากทั้งสองโรงเรียน พบว่า ส่วนเบี่ยงเบนมาตรฐานของคะแนนสอบจากโรงเรียนอักษรทองมีค่ามากกว่า หมายความว่า คะแนนสอบของนักเรียนกลุ่มนี้มีการกระจายมากกว่า นั่นคือความสามารถในการสอบของนักเรียนแต่ละคนในกลุ่มมีความแตกต่างกันมาก

ในทางกลับกัน ส่วนเบี่ยงเบนมาตรฐานของคะแนนสอบจากโรงเรียน A+ มีค่าน้อยกว่า หมายความว่า คะแนนสอบของนักเรียนกลุ่มนี้มีการกระจายน้อยกว่า นั่นคือ ความสามารถในการสอบของนักเรียนแต่ละคนในกลุ่มมีความแตกต่างกันน้อย

ดังนั้น การวิเคราะห์ค่าเฉลี่ยเลขคณิตเพียงอย่างเดียวไม่สามารถบอกความแตกต่างของข้อมูลชุดนั้นได้ จึงต้องใช้ส่วนเบี่ยงเบนมาตรฐานเป็นตัวบอกว่าข้อมูลชุดนั้นมีการกระจายของข้อมูลมากน้อยเพียงใด



เอกสารอ้างอิง

หนังสือเรียน รายวิชาพื้นฐาน คณิตศาสตร์ ระดับชั้น ม.5 ฉบับอนุญาต (อญ.)

ตรงตามหลักสูตรแกนกลาง '51 (ฉบับปรับปรุง '60) อจท.

หนังสือเรียน รายวิชาเพิ่มเติม คณิตศาสตร์ ระดับชั้น ม.5 ฉบับอนุญาต (อญ.)

ตรงตามหลักสูตรแกนกลาง '51 (ฉบับปรับปรุง '60) อจท.

แบบฝึกหัด รายวิชาเพิ่มเติม คณิตศาสตร์ ระดับชั้น ม.5 เล่ม 1 ตรงตามหลักสูตร
แกนกลาง '51 (ฉบับปรับปรุง '60) อจท.

