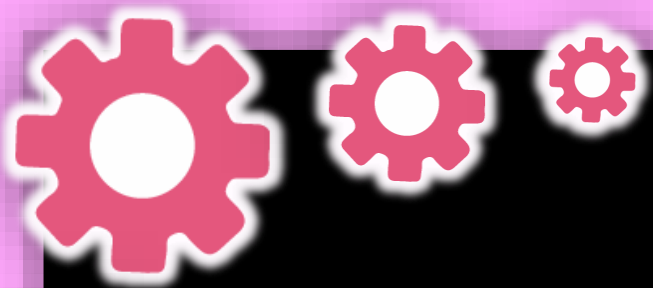


บทที่ 1

ความรู้เบื้องต้นเกี่ยวกับสถิติวิศวกรรม

(Introduction to Engineering Statistics)



## “สถิติ”

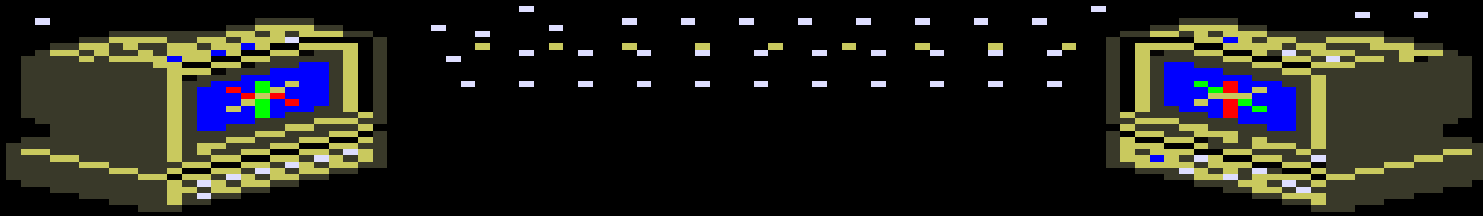
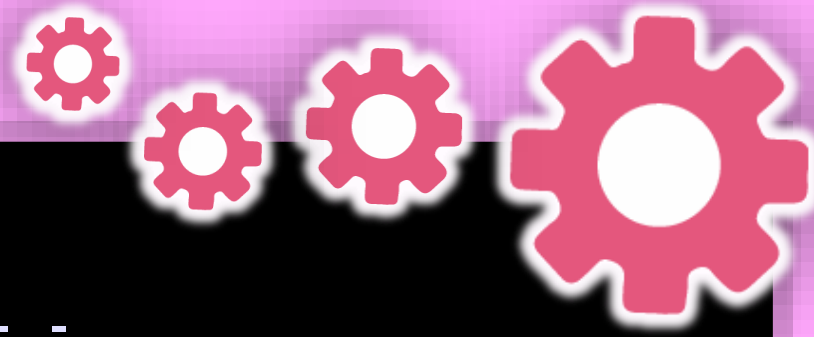


การเก็บรวบรวมข้อมูลต่าง ๆ ในอดีตหรือปัจจุบัน เพื่อการนำเสนอ  
หรือนำไปวิเคราะห์เพื่อการสรุปผล แล้วนำผลสรุปนั้นมาช่วยในการตัดสินใจ

## “สถิติวิศวกรรม”

สถิติวิศวกรรม เป็นวิธีการสำหรับการตัดสินใจ โดยอาศัยข้อมูล  
ที่ได้เก็บรวบรวมมา เพื่อประยุกต์ใช้ให้เกิดประโยชน์มากที่สุด

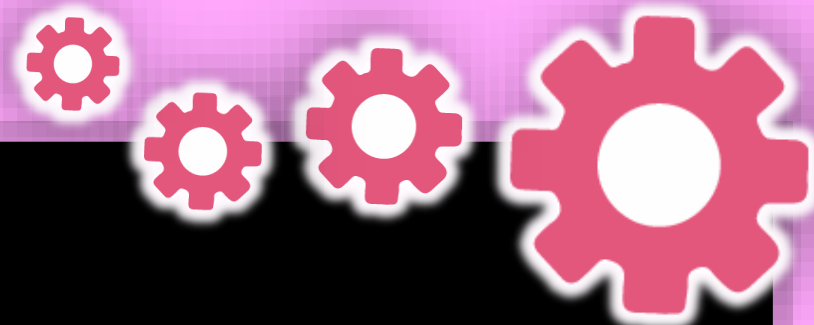




อยากทราบว่าการผลิตชิ้นงานยังมีความยาวที่ 5 นิ้ว หรือไม่



# STATISTICS



## สถิติ สามารถจำแนกออกเป็น 2 กลุ่ม

### 1. สถิติเชิงพรรณนา

จัดระเบียบเพื่อการสรุป อธิบายลักษณะต่าง ๆ ของสิ่งที่ต้องการศึกษา การนำเสนอข้อมูล เช่น ตาราง กราฟ ค่าเฉลี่ย

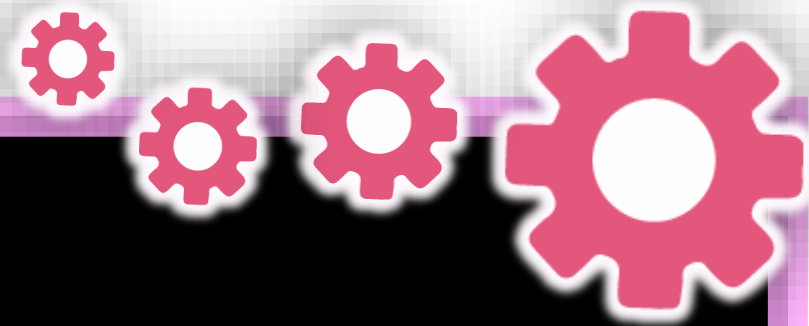
### 2. สถิติเชิงอนุมาน

วิเคราะห์วินิจฉัยกลุ่มตัวอย่าง เพื่อใช้เป็นข้อสรุปสำหรับประชากร เช่น การทดสอบสมมติฐาน การวิเคราะห์ความแปรปรวน

#### จำนวนนิสิตชั้น ปี 1 ปีการศึกษา 2566

วิศวกรรมโยธา	92
วิศวกรรมอุตสาหการ	88
วิศวกรรมเครื่องกล	103
วิศวกรรมไฟฟ้า	88
วิศวกรรมคอมพิวเตอร์	84
วิศวกรรมวัสดุ	63
วิศวกรรมสิ่งแวดล้อม	50
วิศวกรรมเคมี	71
วิศวกรรมนิวเคลียร์	20

# STATISTICS



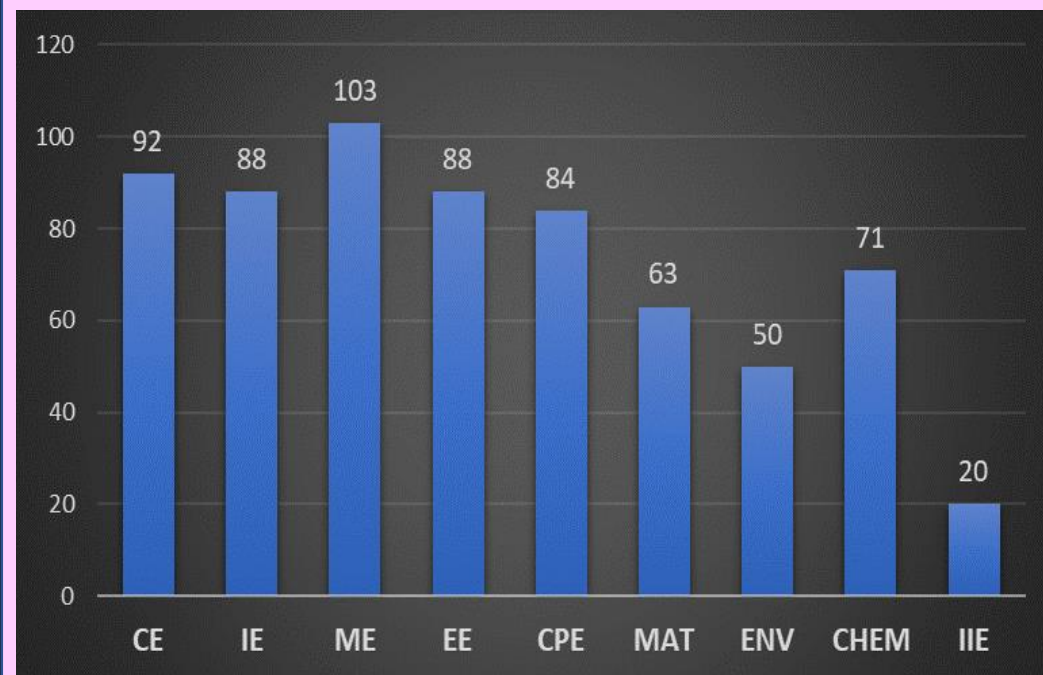
## สถิติ สามารถจำแนกออกเป็น 2 กลุ่ม

### 1. สถิติเชิงพรรณนา

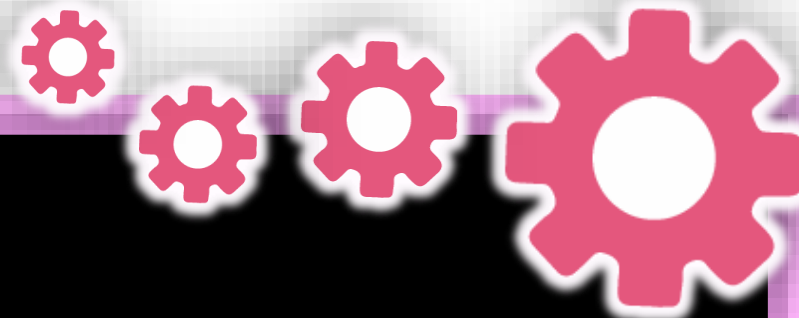
จัดระเบียบเพื่อการสรุป อธิบายลักษณะต่าง ๆ ของสิ่งที่ต้องการศึกษา การนำเสนอข้อมูล เช่น ตาราง กราฟ ค่าเฉลี่ย

### 2. สถิติเชิงอนุมาน

วิเคราะห์วินิจฉัยกลุ่มตัวอย่าง เพื่อใช้เป็นข้อสรุปสำหรับประชากร เช่น การทดสอบสมมติฐาน การวิเคราะห์ความแปรปรวน



# STATISTICS



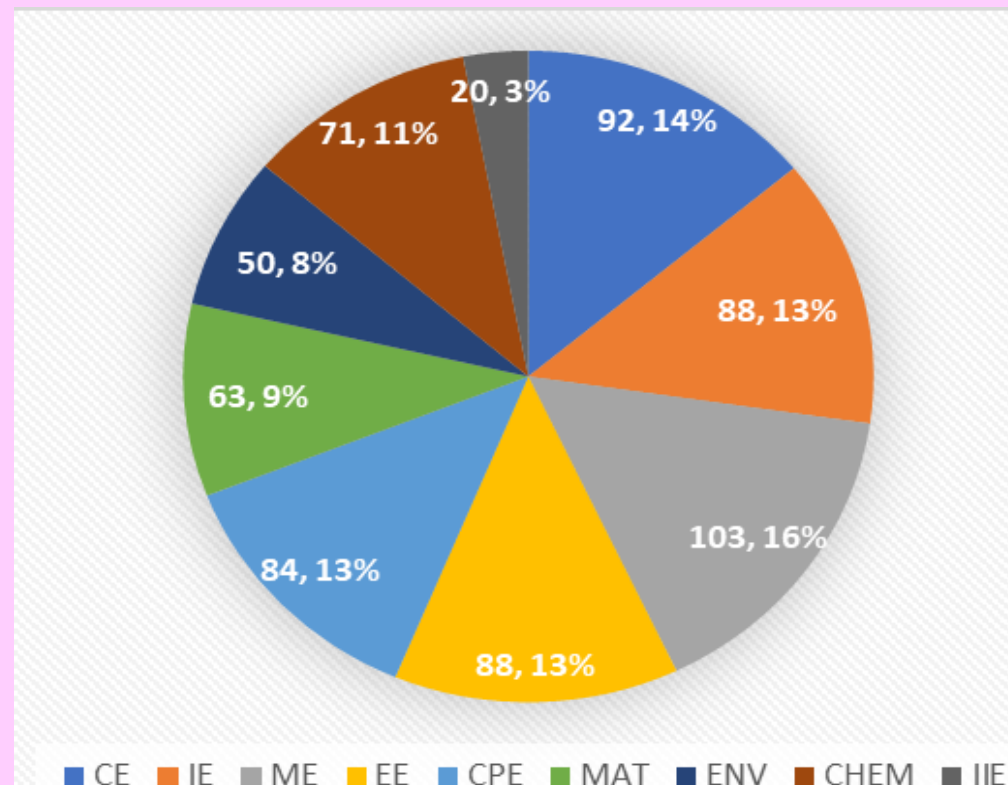
## สถิติ สามารถจำแนกออกเป็น 2 กลุ่ม

### 1. สถิติเชิงพรรณนา

จัดระเบียบเพื่อการสรุป อธิบายลักษณะต่าง ๆ ของสิ่งที่ต้องการศึกษา การนำเสนอข้อมูล เช่น ตาราง กราฟ ค่าเฉลี่ย

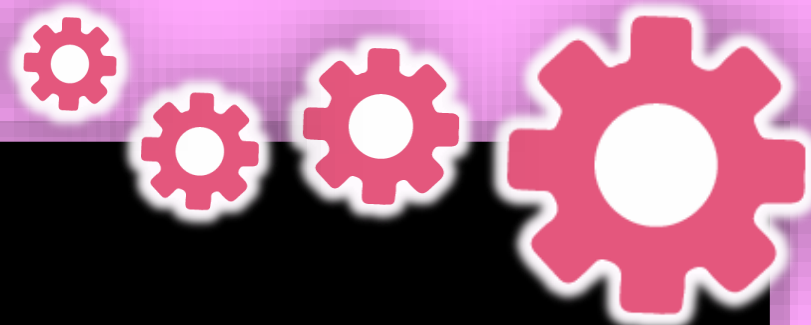
### 2. สถิติเชิงอนุมาน

วิเคราะห์วินิจฉัยกลุ่มตัวอย่าง เพื่อใช้เป็นข้อสรุปสำหรับประชากร เช่น การทดสอบสมมติฐาน การวิเคราะห์ความแปรปรวน





# STATISTICS



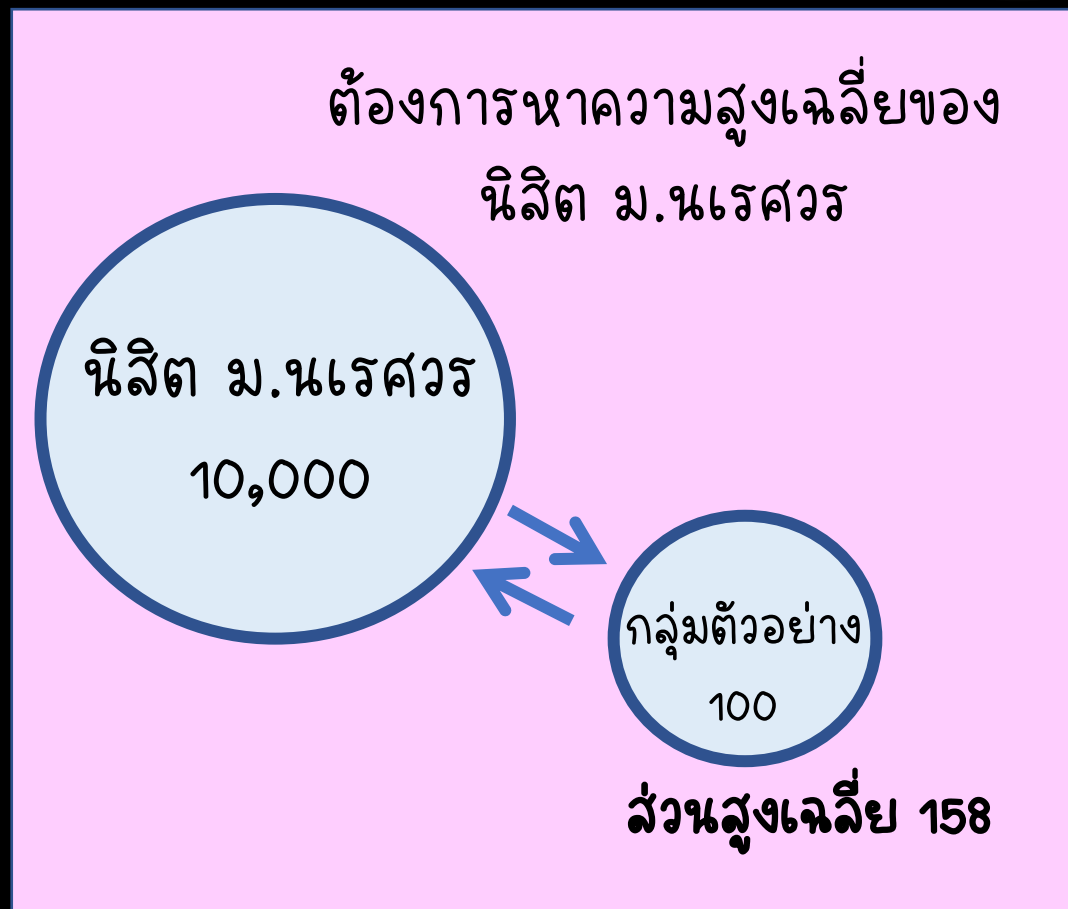
สถิติ สามารถจำแนกออกเป็น 2 กลุ่ม

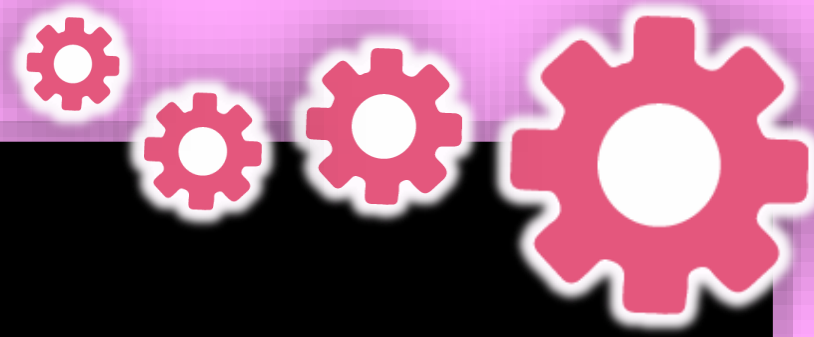
## 1. สถิติเชิงพรรณนา

จัดระเบียบเพื่อการสรุป อธิบายลักษณะต่าง ๆ ของสิ่งที่ต้องการศึกษา การนำเสนอข้อมูล เช่น ตาราง กราฟ ค่าเฉลี่ย

## 2. สถิติเชิงอนุมาน

วิเคราะห์วินิจัยกลุ่มตัวอย่าง เพื่อใช้เป็นข้อสรุป สำหรับประชากร เช่น การทดสอบสมมติฐาน การวิเคราะห์ความแปรปรวน



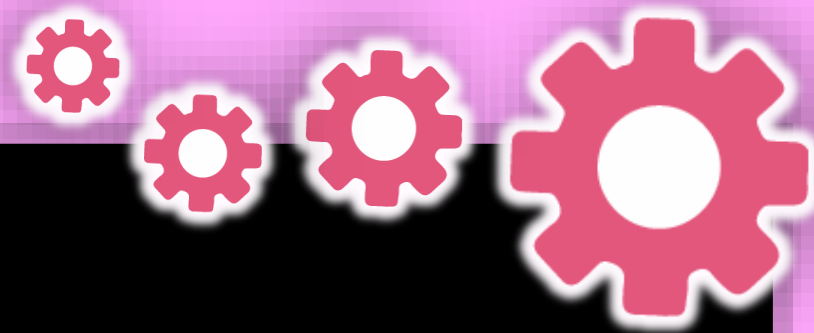


# การนำเสนอข้อมูล

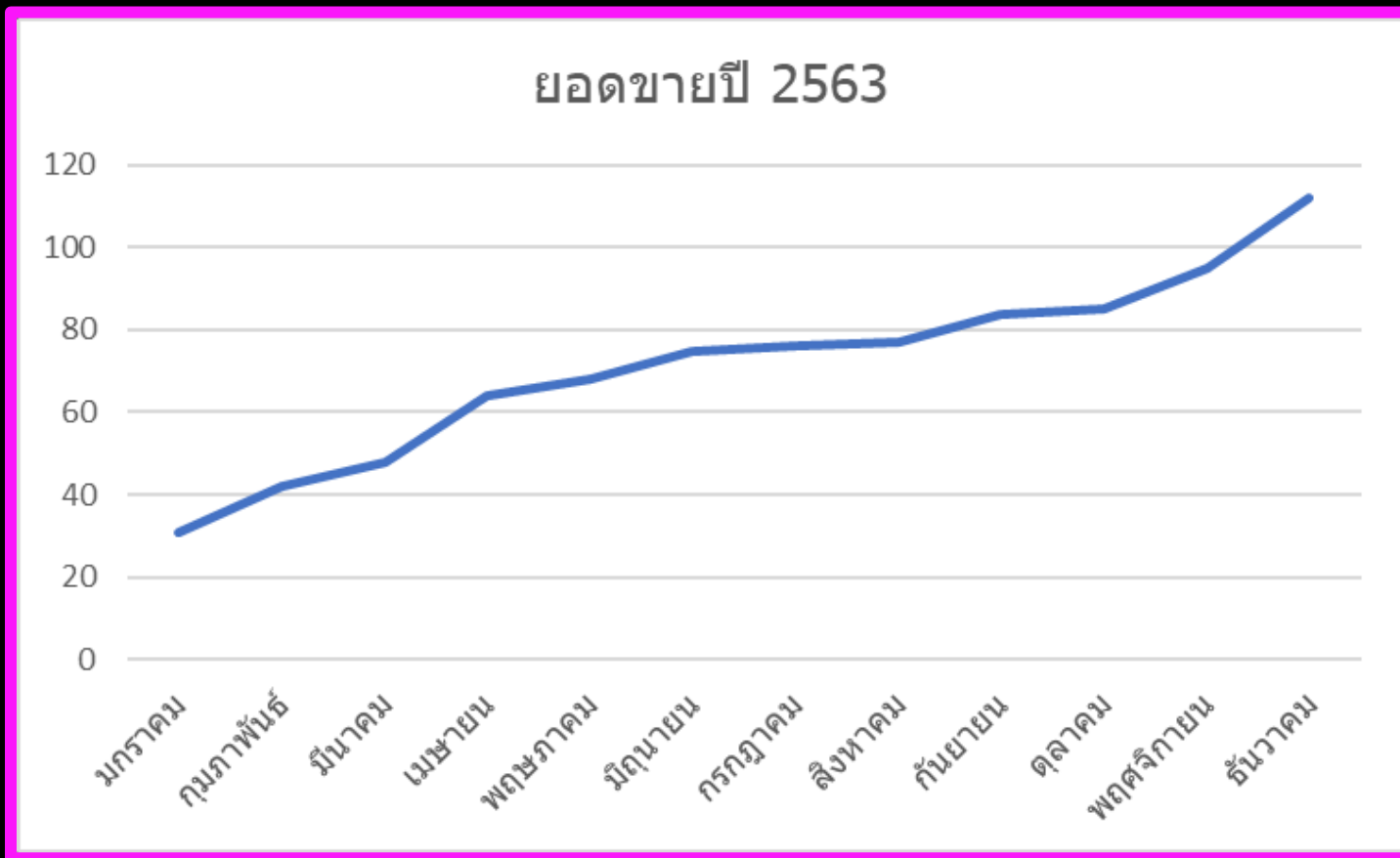




# STATISTICS

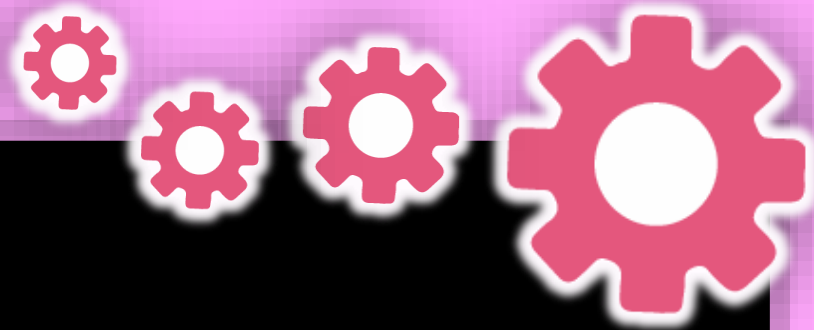


## การนำเสนอข้อมูลด้วยกราฟเส้น

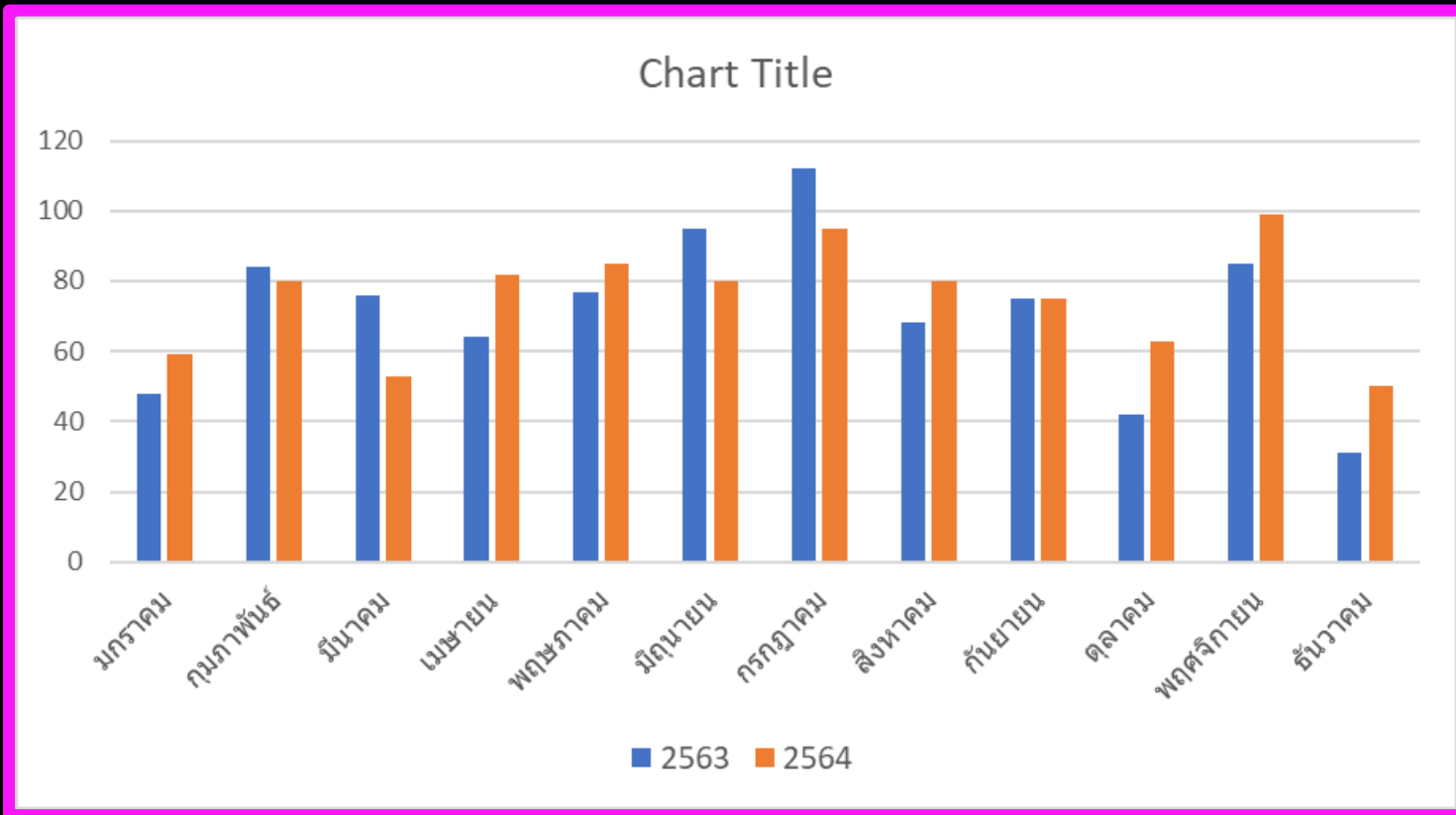


ดูแนวโน้ม

# STATISTICS

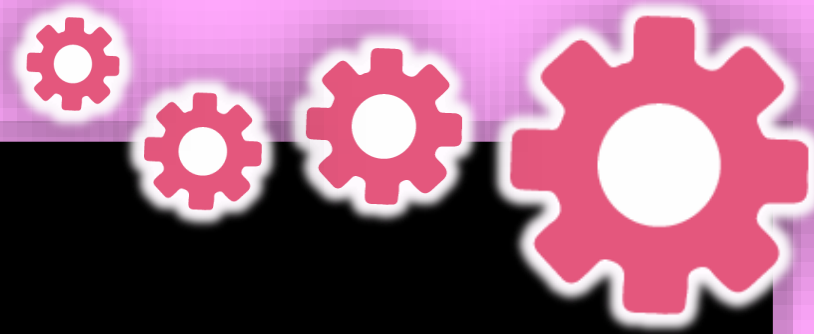


## การนำเสนอข้อมูลด้วยกราฟแท่ง

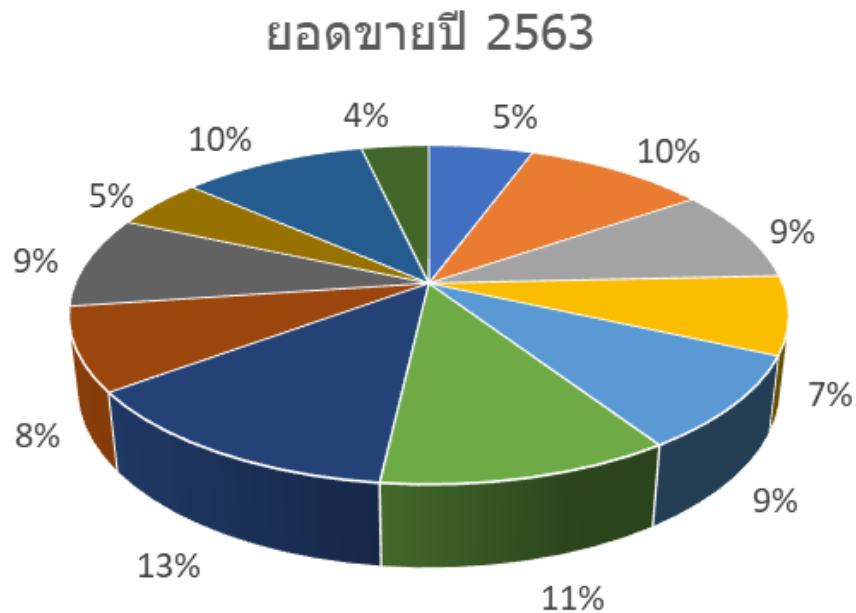


เปรียบเทียบ

# STATISTICS



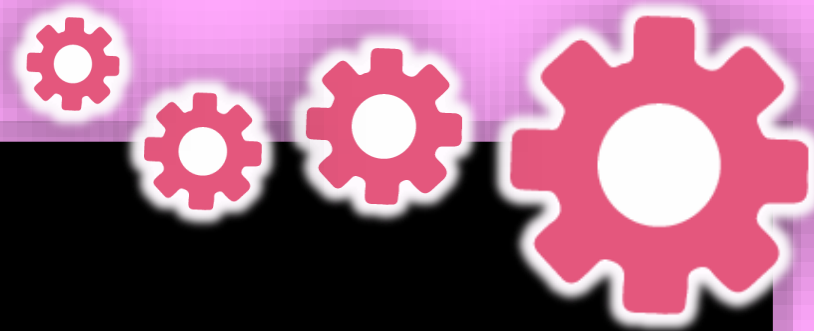
## การนำเสนอข้อมูลด้วยกราฟวงกลม



■ มกราคม ■ กุมภาพันธ์ ■ มีนาคม ■ เมษายน ■ พฤษภาคม ■ มิถุนายน  
■ กรกฎาคม ■ สิงหาคม ■ กันยายน ■ ตุลาคม ■ พฤศจิกายน ■ ธันวาคม

อัตราส่วน  
สัดส่วน

# STATISTICS



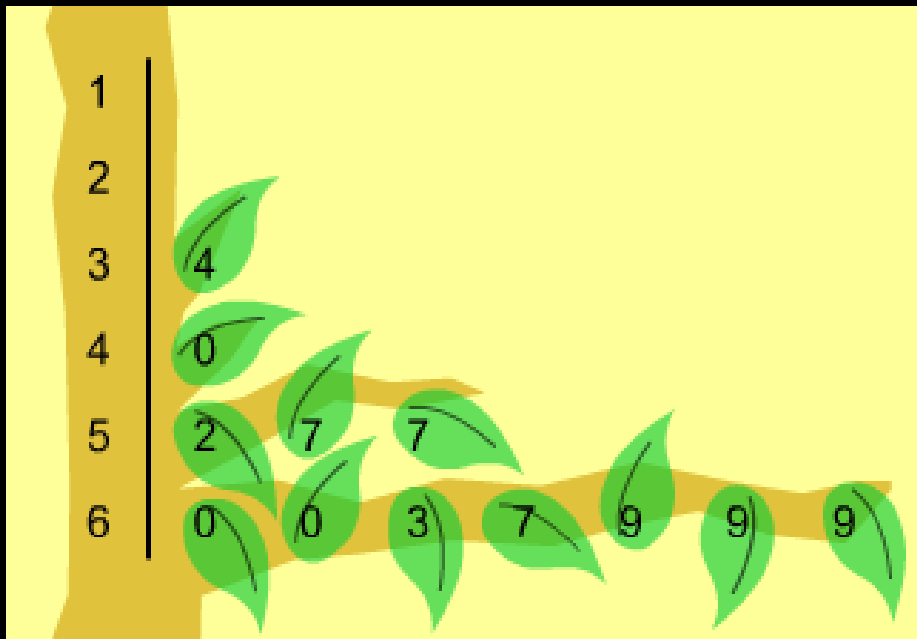
การนำเสนอข้อมูลด้วยแผนภาพก้านและใบ

ต้น	ใบ
1	9
2	3 4 4 4
3	1 3 4 5 6 6 8 8
4	2 3 4 4 7 8 9

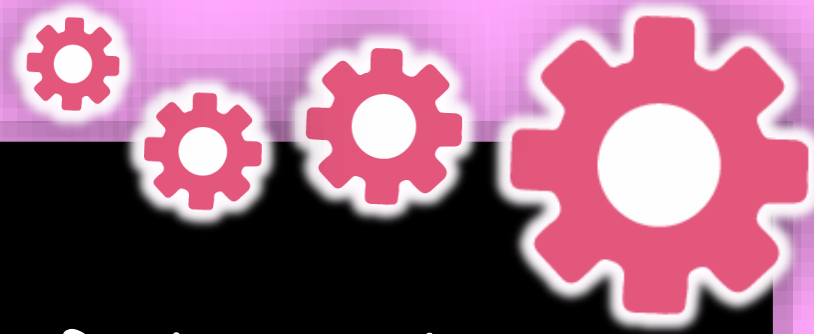
ดูการกระจาย



## แผนภูมิก้านและใบ (Stem-and-Leaf Diagrams)



เป็นวิธีแสดงข้อมูลเชิงตัวเลขที่ง่ายและทำได้รวดเร็ว มีประโยชน์สำหรับใช้เรียงและจำแนกค่าที่สังเกตได้ของข้อมูลเป็นกลุ่มย่อยจากน้อยที่สุดไปมากที่สุด ทำให้เราสามารถเห็นตำแหน่งของค่าต่าง ๆ ที่สนใจได้ง่าย



## ส่วนประกอบของแผนภูมิก้านและใบ (Stem-and-Leaf Diagrams)

1. ส่วนที่เป็นก้าน (stem) เปรียบได้กับก้านของต้นไม้ อาจมีมากกว่า 1 ก้านก็ได้ จำนวนก้านจะขึ้นอยู่กับขนาดของข้อมูล ปกติควรไม่ต่ำกว่า 5 และไม่ควรเกิน 20
2. ส่วนที่เป็นใบ (leaf) จะเป็นส่วนของตัวเลขที่มีการเปลี่ยนแปลงเร็วที่สุด

# STATISTICS



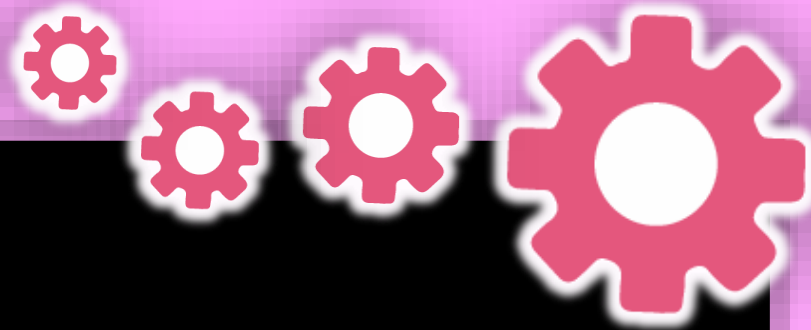
ตัวอย่างที่ 1 ตัวอย่างน้ำหนัก (กก.) ของวัตถุนิตหนึ่งจำนวน 30 ข้อมูล

ลงสร้างแผนภูมิกำนและใบ

65	77	57	50	72	69
41	69	64	65	51	63
75	79	68	76	87	52
52	84	55	83	81	64
63	40	64	56	75	90
92	43	73	68	66	76



# STATISTICS



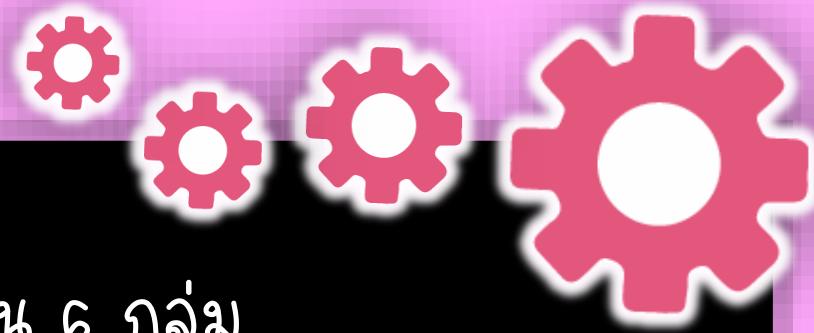
65	77	57	50	72	69
41	69	64	65	51	63
75	79	68	76	87	52
52	84	55	83	81	64
63	40	64	56	75	90
92	43	73	68	66	76

จากข้อมูลจะเห็นว่ากำหนดหลักประกอบด้วย

4 5 6 7 8 และ 9

stem	leaf
4	
5	7 0
6	5
7	7 2
8	
9	

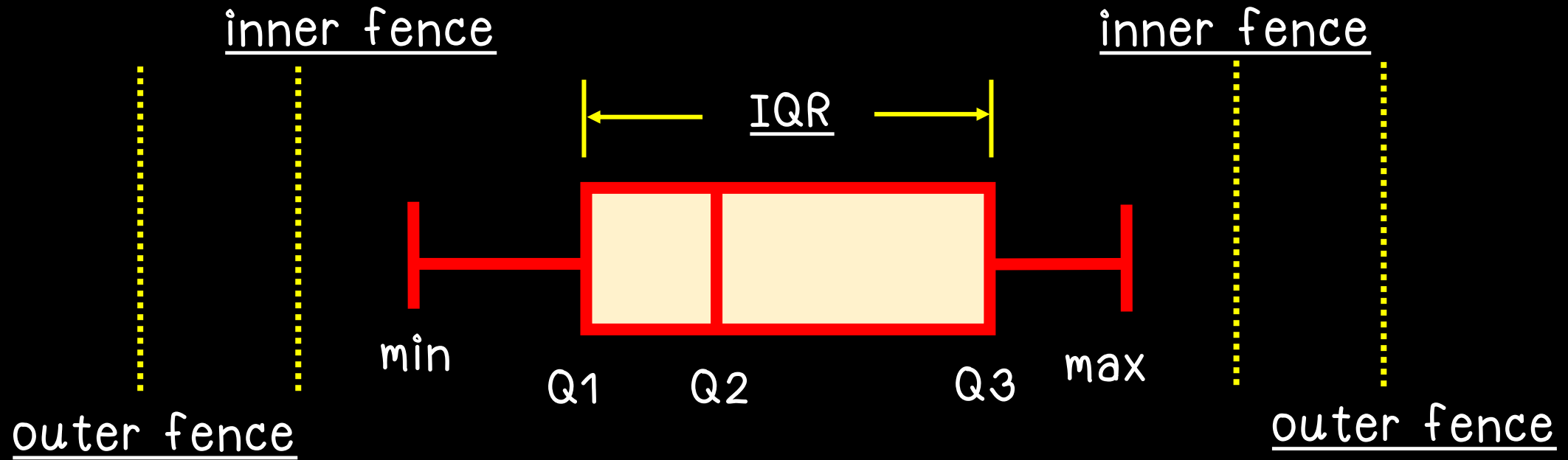
# STATISTICS



stem	leaf
4	0 1 3
5	0 1 2 2 5 6 7
6	3 3 4 4 4 5 5 6 8 8 9 9
7	2 3 5 5 6 6 7 9
8	1 3 4 7
9	0 2

1. ข้อมูลถูกแบ่งออกเป็น 6 กลุ่ม
2. ข้อมูลมากที่สุด = 92  
ข้อมูลน้อยที่สุด = 40
3. น้ำหนักของวัตถุส่วนใหญ่อยู่ในช่วง 63–69 กก.
4. น้ำหนักของวัตถุส่วนใหญ่จะเกาะกลุ่มกันในช่วง 50–79 กก ซึ่งมีถึง 27 ค่า หรือ 75 % ของจำนวนค่าสังเกตทั้งหมด

# แผนภาพกล่อง **BOX PLOT**

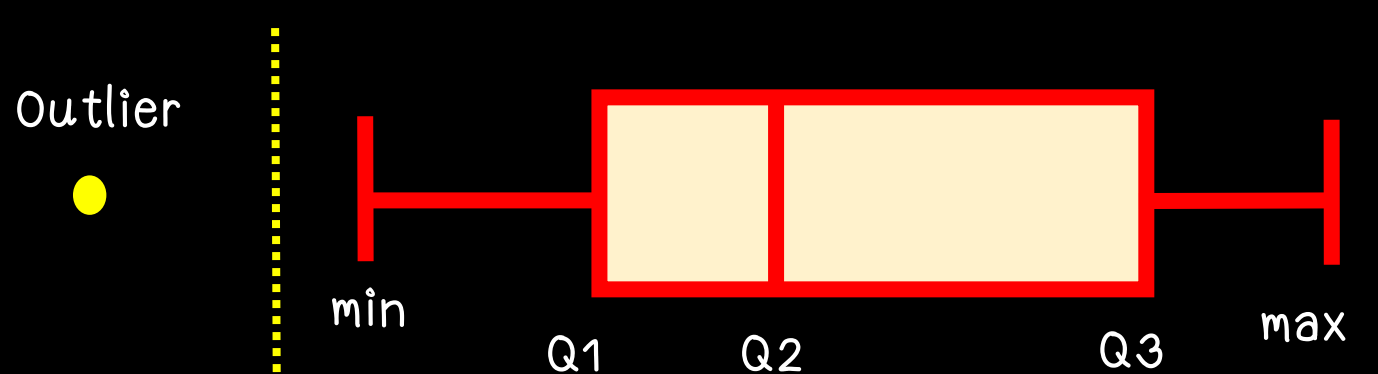


# STATISTICS

## แผนภาพกล่อง (BOX PLOT)

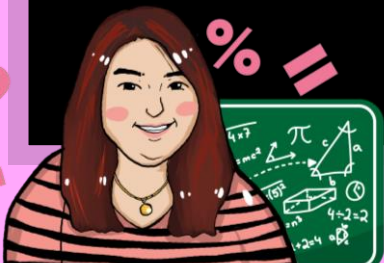


แผนภาพกล่อง (Box plot) เป็น กราฟที่มีลักษณะดังนี้



กล่องที่ประกอบด้วยค่า 3 ค่า คือ ควอไทล์แรก (Q1) ค่ากลาง (Median หรือ Q2) ควอไทล์สาม (Q3)  
มีเส้นที่มีลักษณะคล้ายหนวดแมว แสดงค่าน้อยที่สุดและค่ามากที่สุดที่อยู่ภายในรั้ว  
ค่าที่อยู่นอกรั้วจะถือว่าเป็นค่าที่อยู่นอกกลุ่ม (Outlier)

\*\*\* ดูการกระจายของข้อมูล และหาค่าที่อยู่นอกกลุ่ม (OUTLIER) \*\*\*

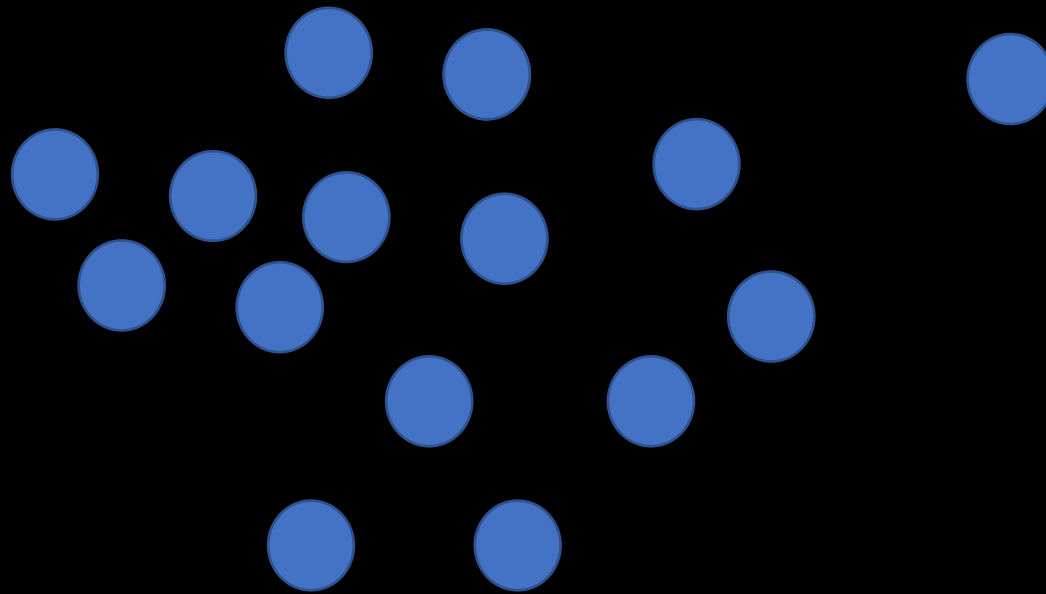


# STATISTICS

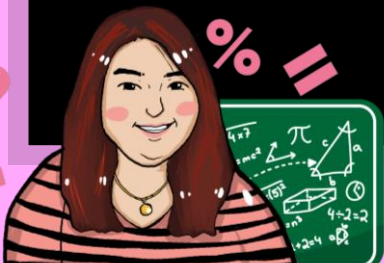
## แผนภาพกล่อง (BOX PLOT)



ข้อมูลชุดหนึ่งมี 15 ค่า

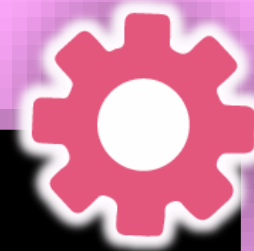


**\*\*อยากรู้ว่าข้อมูลชุดนี้มีการกระจายอย่างไร และมีข้อมูลที่อยู่นอกกลุ่ม (OUTLIER) หรือไม่\*\***

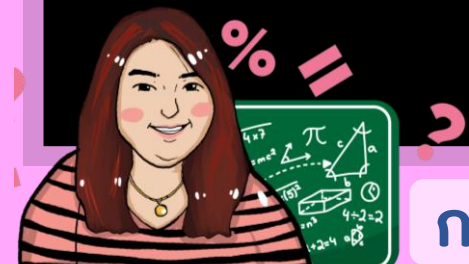
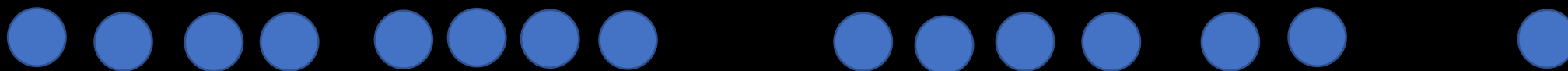


# STATISTICS

## แผนภาพกล่อง (BOX PLOT)

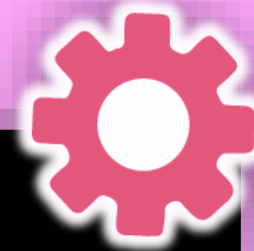


เรียงข้อมูลจากน้อยไปมาก



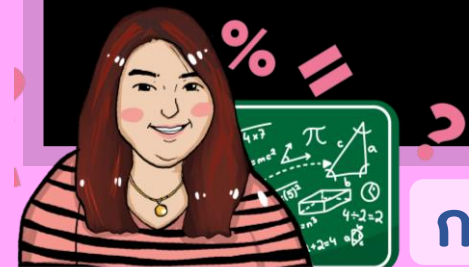
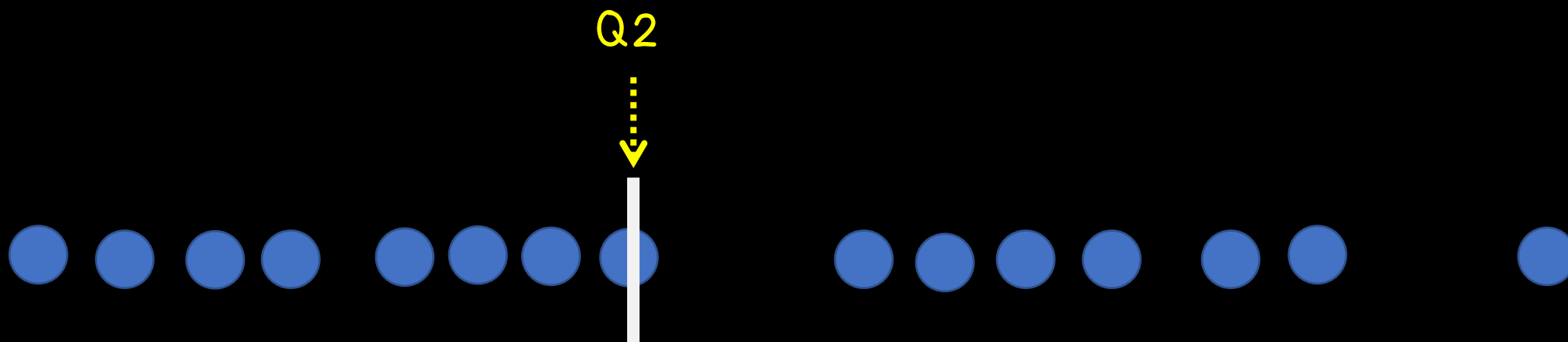
กานต์ ศุภจิตกุล

วิศวกรรมอุตสาหการ คณะวิศวกรรมศาสตร์ ม.นเรศวร



### ค่ากลาง (Median) หรือ Q2

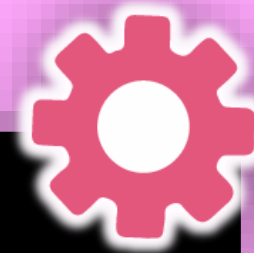
เป็นค่าที่มีตำแหน่งอยู่ตรงกลางของข้อมูลทั้งหมด  
จะแบ่งข้อมูลเป็น 2 กลุ่ม เท่า ๆ กัน





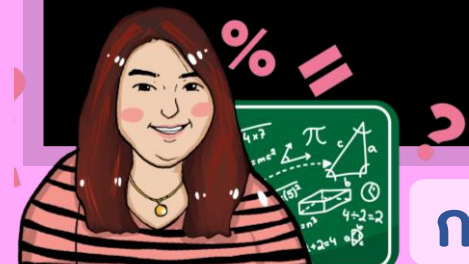
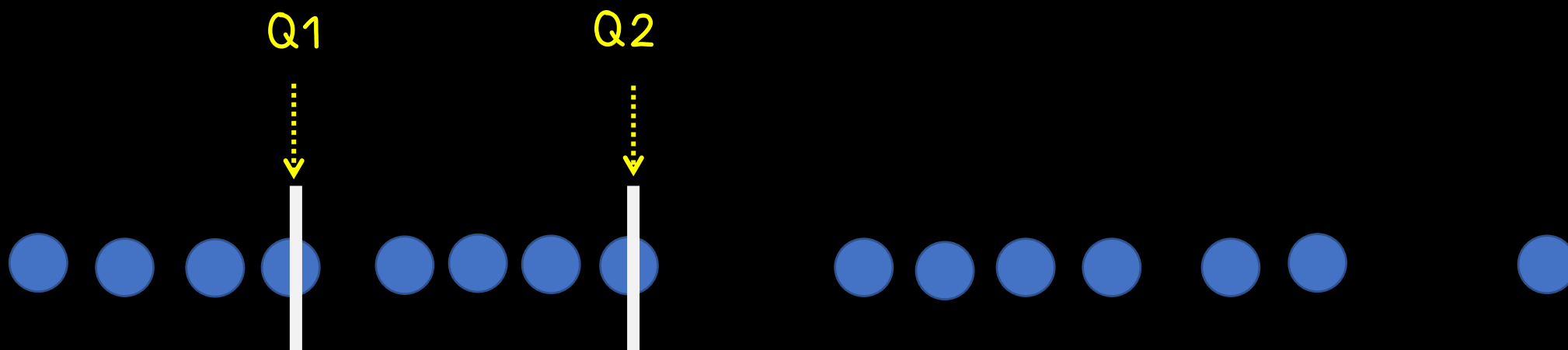
# STATISTICS

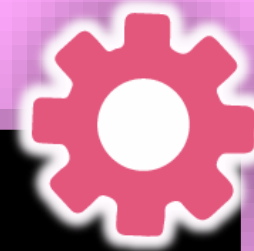
## แผนภาพกล่อง (BOX PLOT)



First quartile (Q1)

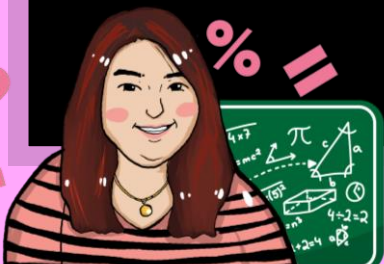
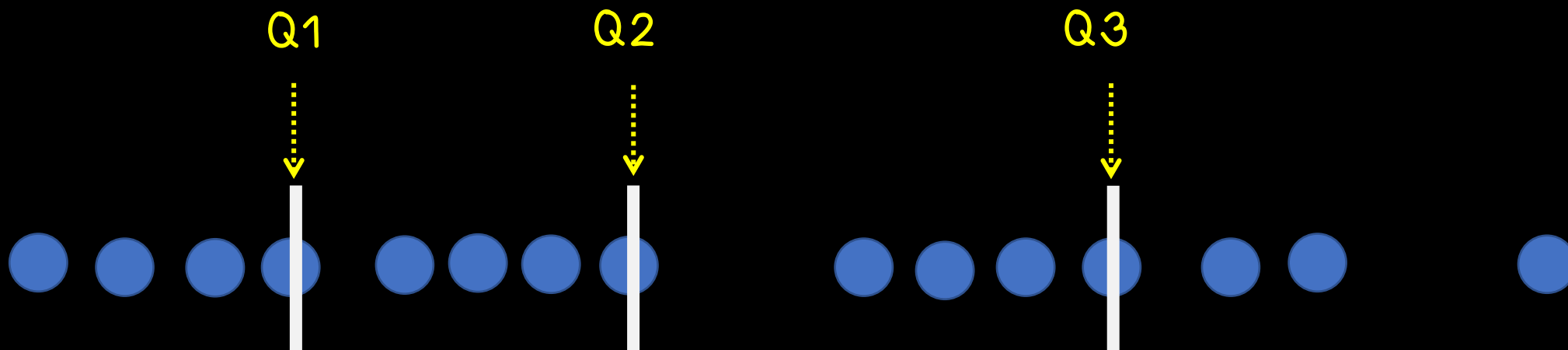
เป็นค่าที่มีตำแหน่งอยู่ตรงกลาง  
ของข้อมูลกลุ่มล่าง





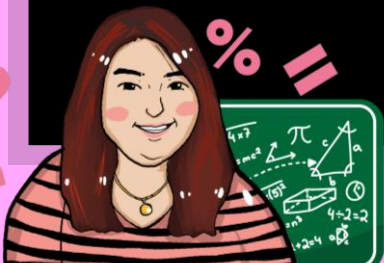
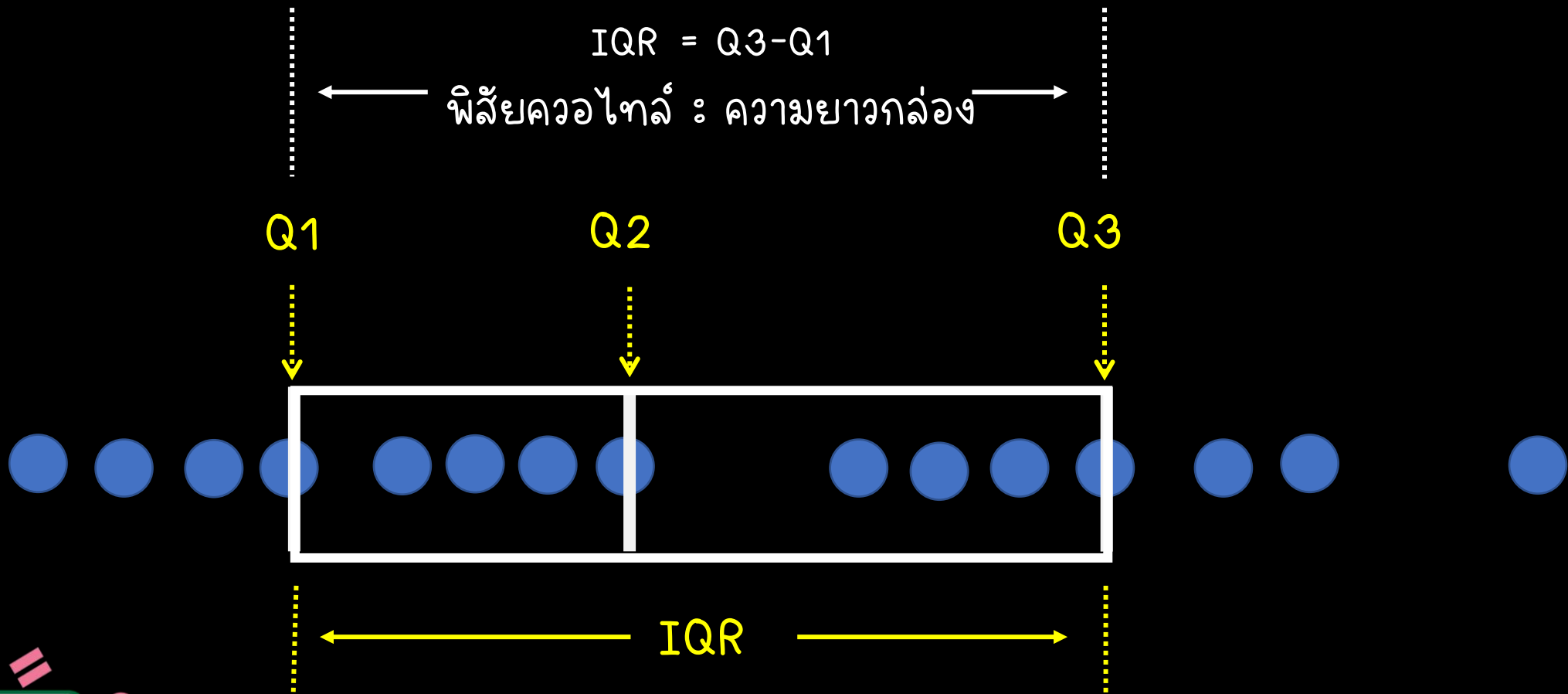
Third quartile (Q3)

เป็นค่าที่มีตำแหน่งอยู่ตรงกลาง  
ของข้อมูลกลุ่มบน



# STATISTICS

## แผนภาพกล่อง (BOX PLOT)



# STATISTICS

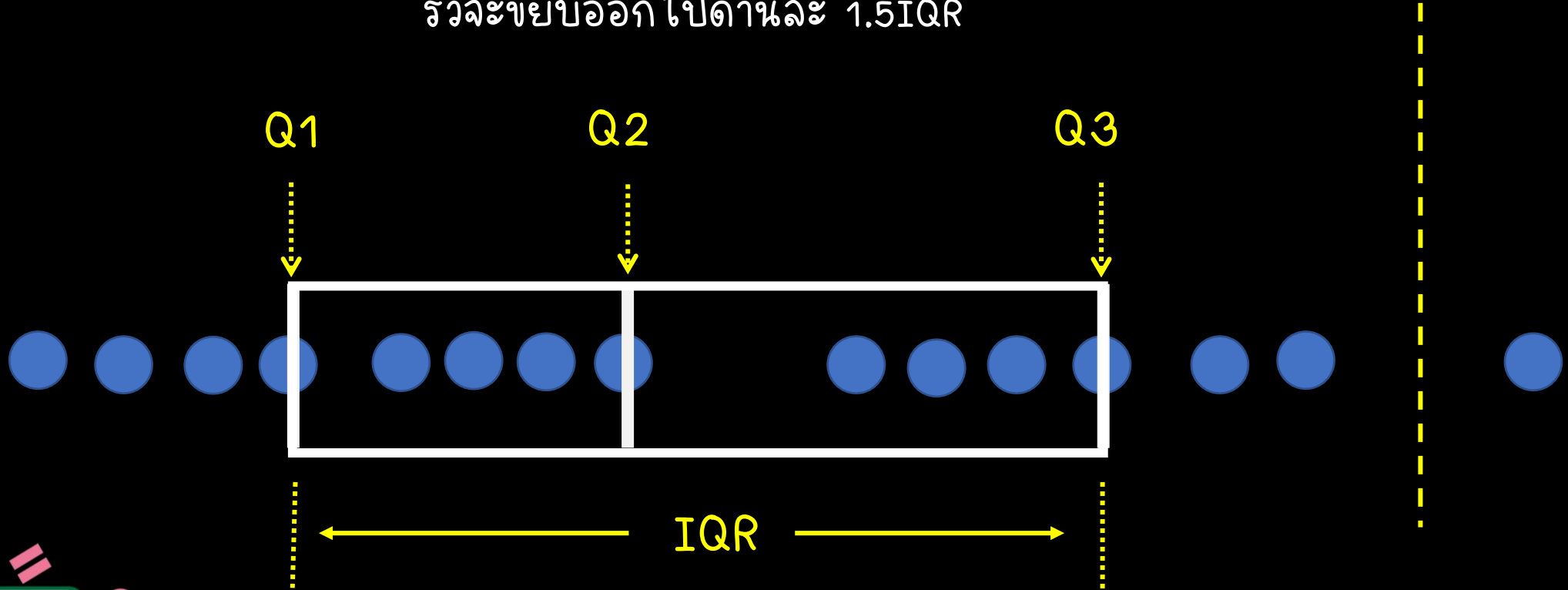
## แผนภาพกล่อง (BOX PLOT)



Inner fence  
=  $Q1 - 1.5IQR$

รั้วชั้นใน (Inner fence)  
เป็นขอบเขตที่กั้นกลุ่มข้อมูลที่ใช้ได้  
รั้วจะขยับออกไปด้านละ  $1.5IQR$

Inner fence  
=  $Q3 + 1.5IQR$



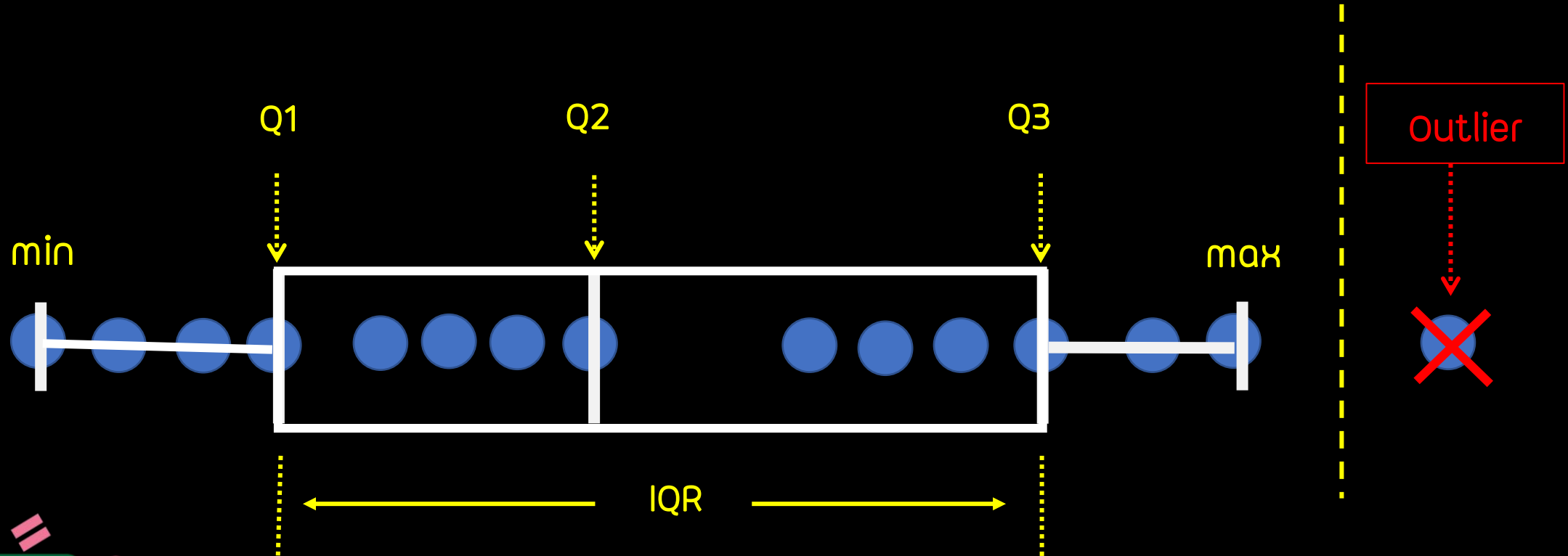
# STATISTICS

## แผนภาพกล่อง (BOX PLOT)

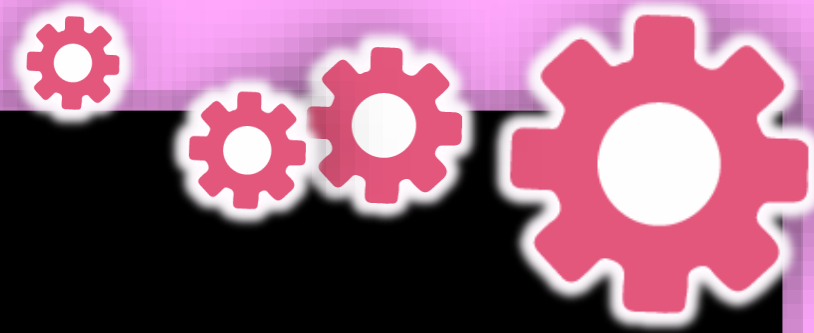


Inner fence  
=  $Q1 - 1.5IQR$

Inner fence  
=  $Q3 + 1.5IQR$



## ขั้นตอนการสร้าง BOXPLOT



1. เรียงข้อมูลจากน้อยไปมาก

2. ตำแหน่งของควอร์ไทล์ที่หนึ่ง      ตำแหน่ง  $Q1 = \frac{1}{4} [N+1]$

3. ตำแหน่งของควอร์ไทล์ที่สอง      ตำแหน่ง  $Q2 = \frac{2}{4} [N+1]$

4. ตำแหน่งของควอร์ไทล์ที่สาม      ตำแหน่ง  $Q3 = \frac{3}{4} [N+1]$

5. ค่าพิสัยควอร์ไทล์ (Inter-Quartile Range : IQR)       $IQR = Q3 - Q1$

6. ขอบรั้วชั้นใน (inner fence)       $= Q1 - 1.5IQR$  และ  $Q3 + 1.5IQR$

7. ขอบรั้วชั้นนอก (Outer fence)       $= Q1 - 3IQR$  และ  $Q3 + 3IQR$

ตำแหน่ง

# STATISTICS

## แผนภาพกล่อง (BOX PLOT)



Inner fence

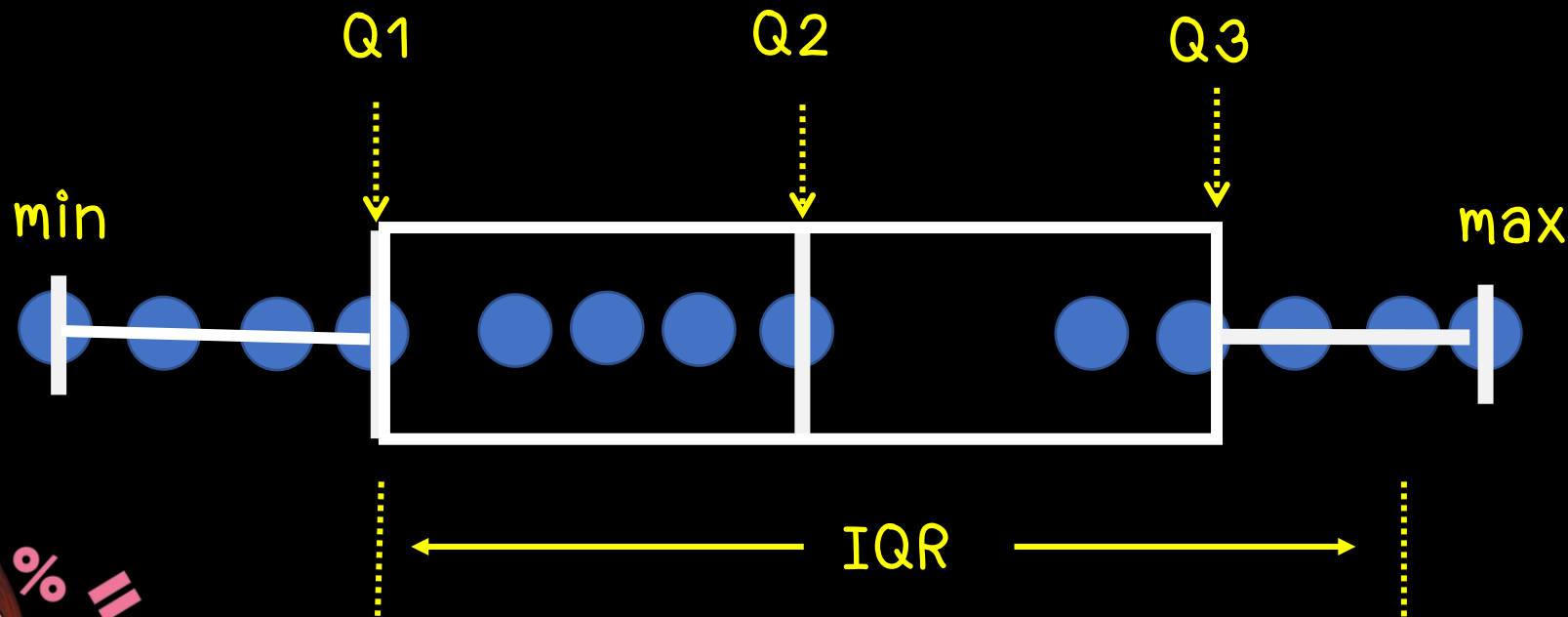
$$= Q1 - 1.5IQR$$

ตกระหว่างรั้ว เป็น **Outlier** ให้พิจารณาว่าจะตัดทิ้งหรือไม่

ตกนอกรั้ว ถือว่าเป็น **Extreme outlier** ให้ตัดทิ้ง

$$\text{Outer fence} = Q3 + 3IQR$$

$$\text{Inner fence} = Q3 + 1.5IQR$$



**EXTREME**  
**Outlier**





## ตัวอย่าง

1      2  
27    32    40    43    45    51    59    62

ตำแหน่ง  $Q1 = 2.25$

$$Q1 = 32 + (0.25 \times 8) = 34$$

40-32  
↓



ตัวอย่างที่ 1 ข้อมูลน้ำหนักสินค้า 20 48 16 10 19 17 13 75 25 27

1. เรียงข้อมูลจากน้อยไปมาก      **1**   **2**   **3**   **4**   **5**   **6**   **7**   **8**   **9**   **10**  
10   13   16   17   19   20   25   27   48   75

2. ตำแหน่ง  $Q_1 = (1/4)[10+1] = 2.75$       ดังนั้น  $Q_1 = 13 + (0.75 \times [16-13]) = 15.25$

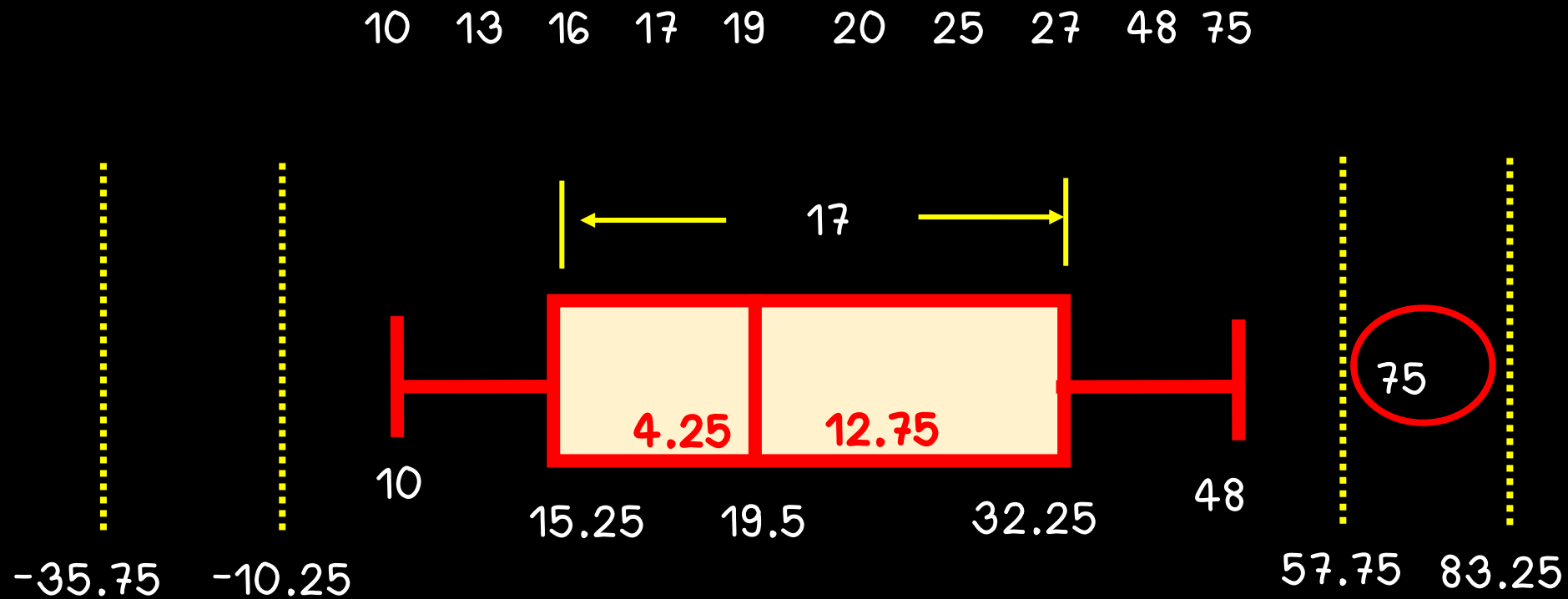
3. ตำแหน่ง  $Q_2 = (2/4)[10+1] = 5.50$       ดังนั้น  $Q_2 = 19.5$

4. ตำแหน่ง  $Q_3 = (3/4)[10+1] = 8.25$       ดังนั้น  $Q_3 = 27 + (0.25 \times [48-27]) = 32.25$

5. ค่า IQR       $Q_3 - Q_1 = 32.25 - 15.25 = 17$

6. Inner fence      IF =  $15.25 - 1.5(17) = -10.25$  และ  $32.25 + 1.5(17) = 57.75$

7. Outer fence      OF =  $15.25 - 3(17) = -35.75$  และ  $32.25 + 3(17) = 83.25$

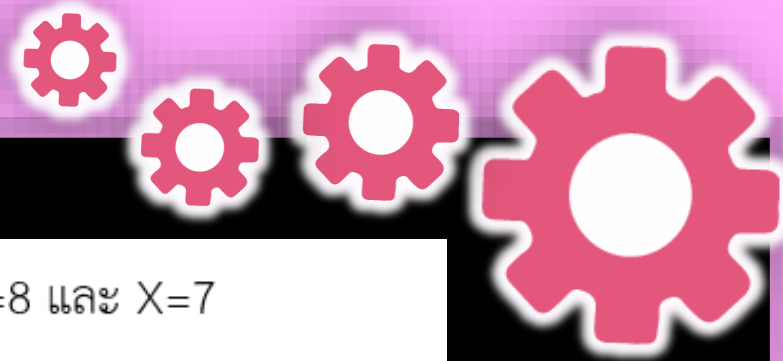


ข้อมูลมีลักษณะ **เบ้ขวา**

Min = 10    Max = 48

พบ Outlier คือค่า 75 ให้พิจารณาว่าจะตัดทิ้งหรือไม่ตัดทิ้ง ขึ้นอยู่กับความเหมาะสม

# QUIZ 1



ให้แทนค่า Z Y X ด้วยเลข 3 ตัวท้ายของรหัสสินค้า เช่น 65361287 จะได้ Z=2 , Y=8 และ X=7

วิศวกรได้สุ่มเก็บค่าปริมาณสารปนเปื้อนของผลิตภัณฑ์ชนิดหนึ่ง แสดงข้อมูลดังตาราง

78	81	72	90	2Y	75	10X	3X
7X	85	79	12Z	77	69	83	7X

จงวิเคราะห์ข้อมูลด้วย Boxplot (แสดงวิธีทำอย่างละเอียด) ให้ใช้ทศนิยม 2 ตำแหน่ง แล้วหาคำตอบต่อไปนี้

- 1) ค่าต่ำสุด =
  - 2) ค่าสูงสุด =
  - 3) ลักษณะกราฟ =
  - 4) Outlier =
  - 5) Extreme Outlier =
- ควรทำอย่างไร
- ควรทำอย่างไร

กรณีทำไม่เสร็จในคาบ ถ้าหมดเวลาให้คืน QUIZ ที่อาจารย์ แล้วไปแก้วันจันทร์ 17.00-19.00 น.  
กรณีไม่ส่งถือว่าได้ 0 คะแนน