

# Chapter 1

Introduction to statistic

คำศัพท์ทางสถิติต่างๆ ที่ควรรู้

**ข้อมูล (Data)** คือ ผลลัพธ์ที่ได้มาจากการเก็บรวบรวมข้อเท็จจริงจากสิ่ง ที่เราสนใจ โดยข้อมูลอาจอยู่ในรูปของตัวเลขหรือข้อความที่ได้มาจากการนับการวัด การสังเกต หรือการบันทึก เป็นต้น

**การเก็บรวบรวมข้อมูล** คือ การให้ได้มาซึ่งข้อมูลที่เราสนใจ เป็นขั้นตอน ที่สำคัญมากสำหรับงานวิจัย เพราะถ้าไม่มีวิธีการเก็บข้อมูลที่ดี ข้อมูลที่ได้มาก็อาจจะไม่ถูกต้อง ในทางสถิติจะนิยมใช้เทคนิคในการสุ่มตัวอย่างเข้ามาใช้ในการรวบรวมข้อมูล

**การนำเสนอข้อมูล** คือ การนำข้อมูลมาจัดใหม่ให้อยู่ในรูปต่างๆ เช่นตารางแจกแจงความถี่ กราฟ ฯลฯ เพื่อให้ง่ายในการทำความเข้าใจและการตีความหมาย

**การวิเคราะห์ข้อมูล** คือ การเอาข้อมูลที่เก็บรวบรวมมาได้มาเข้าสู่ขั้นตอนการวิเคราะห์ เพื่อให้ได้ผลของการวิจัยออกมา ซึ่งส่วนใหญ่ในขั้นตอนนี้จะออกมาในรูปของตัวเลข

**การตีความหมายและสรุปผลข้อมูล** คือ การแปลความหมายและสรุปผล การวิเคราะห์โดยใช้ทฤษฎีทางสถิติ และนำผลที่ได้ไปใช้ประโยชน์

**ประชากร (Population)** คือ ข้อมูลทั้งหมดที่เรากำลังทำการศึกษา หรือ กลุ่มของสิ่งของทั้งหมดที่เราต้องการศึกษา

**กลุ่มตัวอย่าง (Sample)** คือ ส่วนหนึ่งของประชากรซึ่งถูกเลือกมาศึกษา ได้มาจากการสุ่มตัวอย่าง

**ค่าพารามิเตอร์ (Parameter)** คือ ตัวเลขที่คำนวณได้จากประชากร เพื่อแสดงลักษณะบางอย่างของประชากร นิยมแสดงค่าพารามิเตอร์ด้วยอักษรกรีก

**ค่าสถิติ (Statistics)** คือ ตัวเลขที่คำนวณได้จากกลุ่มตัวอย่าง เพื่อแสดง ลักษณะบางอย่างของตัวอย่าง โดยจะใช้ค่าสถิติเป็นตัวประมาณ

ค่าพารามิเตอร์

นิยมแสดงค่าสถิติด้วยอักษรโรมัน

สถิติโดยทั่วไปมี 2 ความหมาย คือ

1. ตัวเลขหรือข้อเท็จจริงของเรื่องราวต่างๆ ที่เราสนใจ แล้วเราไปเก็บรวบรวมข้อมูลบางส่วนมาจากข้อมูลจำนวนมากๆ นั้นด้วยวิธีการต่างๆ เช่นการสำรวจ หรืออาจเป็นค่าของสิ่งใดสิ่งหนึ่งที่ได้มีการบันทึกไว้อยู่แล้วก็ได้เช่น สถิติการส่งออกข้าวของประเทศไทย สถิติการเกิดอุบัติเหตุช่วงวันหยุดสงกรานต์ ฯลฯ

2. วิชาที่เกี่ยวข้องกับกลวิธีที่ใช้ในการศึกษาข้อมูล ได้แก่ การเก็บรวบรวม ข้อมูล การนำเสนอ การวิเคราะห์ข้อมูล การตีความหมายและหาข้อสรุป โดยเฉพาะการหาข้อสรุปของข้อมูลปริมาณมากๆ โดยอาศัยข้อมูลตัวอย่างที่เลือกมาเอามาหาข้อสรุปแล้วเอาไปอธิบายข้อมูลปริมาณมากๆ อีกทีหนึ่ง

สถิติที่เป็นวิชาแยกแบ่งออกเป็น 2 ประเภท คือ

สถิติเชิงบรรยาย (Descriptive Statistics)

สถิติเชิงอนุมาน (Inferential Statistics)

# งานวิจัยคืออะไร

งานวิจัยคือกระบวนการแสวงหาความจริงหรือองค์ความรู้ใหม่ แต่ในกระบวนการเก็บข้อมูลนั้น นักวิจัยมักต้องเผชิญกับ "ความแปรปรวน" (Variability) ของข้อมูล ไม่ว่าจะเป็นผลการทดลองในห้องปฏิบัติการ หรือแบบสอบถามทางสังคมศาสตร์ สถิติ (Statistics) จึงเข้ามามีบทบาทไม่ใช่เพียงแค่เครื่องมือคำนวณตัวเลข แต่เป็น "ตรรกะ" ที่ช่วยเปลี่ยนข้อมูลดิบให้กลายเป็นสารสนเทศที่มีความหมาย

## 1. เป็นระบบและตรวจสอบได้ (Systematic & Verifiable)

- มีขั้นตอนชัดเจนตามระเบียบวิธีวิจัย (Methodology)
- คนอื่นสามารถทำซ้ำ (Reproducible) แล้วได้ผลลัพธ์ไปในทิศทางเดียวกัน

## 2. มีความถูกต้องและแม่นยำ (Validity & Reliability)

- **Validity:** เครื่องมือวัดได้ตรงประเด็น (เช่น ตาชั่งต้องวัดน้ำหนัก ไม่ใช่วัดส่วนสูง)
- **Reliability:** วัดกี่ครั้งก็ได้ค่าคงที่ (น่าเชื่อถือ ไม่รวน)

## 3. ปราศจากอคติ (Objective / Unbiased)

- ใช้ข้อมูลและสถิติตัดสินความจริง ไม่ใช่ความรู้สึกส่วนตัว (Subjectivity)
- แผลผลอย่างตรงไปตรงมา ไม่บิดเบือนข้อมูลเพื่อให้ตรงกับสมมติฐาน

## 4. ริเริ่มสร้างสรรค์ (Originality / Novelty)

- เป็นการค้นพบองค์ความรู้ใหม่ หรือต่อยอดแก้ปัญหาเดิมในมุมมองใหม่
- ไม่คัดลอกผลงานผู้อื่น (No Plagiarism)

## 5. มีประโยชน์และทันสมัย (Applicable & Up-to-date)

- ตอบโจทย์ปัญหาในปัจจุบัน (Research Gap)
- ผลลัพธ์นำไปประยุกต์ใช้ได้ (Practical Implication)

## สถิติในงานวิจัย

**1. สถิติช่วยข้อมูล (Descriptive Statistics)** ข้อมูลดิบมักมีจำนวนมากและกระจัดกระจาย สถิติช่วยสรุปภาพรวมของข้อมูลเหล่านั้นให้เข้าใจง่ายขึ้น ผ่านค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน หรือกราฟ ทำให้เรามองเห็นแนวโน้มและลักษณะสำคัญของกลุ่มตัวอย่างได้อย่างรวดเร็ว

**2. สถิติช่วยขยายผลสู่ภาพใหญ่ (Inferential Statistics)** ในทางปฏิบัติ เราไม่สามารถเก็บข้อมูลจากประชากรทั้งหมด (Population) ได้ สถิติจึงเป็นเครื่องมือที่ช่วยให้เราเก็บข้อมูลจาก "กลุ่มตัวอย่าง" (Sample) เพียงเล็กน้อย แล้วนำผลลัพธ์นั้นไป "อ้างอิง" หรือทำนายผลของประชากรทั้งหมดได้อย่างแม่นยำตามหลักความน่าจะเป็น

**3. สถิติช่วยตัดสินใจและทดสอบ (Hypothesis Testing)** ความรู้สึกส่วนตัวอาจทำให้เราเข้าใจสมมติฐานของตัวเอง แต่สถิติจะให้คำตอบที่เป็นกลางผ่านการทดสอบสมมติฐาน เช่น การใช้ค่า P-value เพื่อยืนยันว่าผลลัพธ์ที่ได้นั้นเกิดขึ้นจริงจากการทดลอง หรือเกิดขึ้นเพียงเพราะความบังเอิญ

**บทสรุป** หากปราศจากสถิติ งานวิจัยจะเป็นเพียงแค่การรวบรวมตัวเลขที่ไร้ทิศทาง สถิติจึงเป็นกุญแจสำคัญที่สร้างความน่าเชื่อถือ (Reliability) ความเที่ยงตรง (Validity) และ ปราศจากอคติ (Objective / Unbiased) ให้กับงานวิจัย เพื่อให้ผลลัพธ์นั้นนำไปใช้ประโยชน์ได้จริง

## Descriptive Statistics (สถิติเชิงพรรณนา)

หัวข้อรอง: การสรุปและบรรยายลักษณะ "สิ่งที่เห็น" ในข้อมูล

### 1. หัวใจสำคัญ (Concept)

- คือการใช้ตัวเลขสรุปภาพรวมของข้อมูล "เฉพาะกลุ่มที่เก็บมาเท่านั้น"
- เป้าหมาย: ย่อข้อมูลจำนวนมากให้เหลือเพียงค่าที่เข้าใจง่าย (ไม่ทำนายข้ามไปหากกลุ่มอื่น)

### 2. เครื่องมือหลัก (Key Measures)

- **หาตัวแทนข้อมูล (Central Tendency):** ค่าที่บอกว่า "คนส่วนใหญ่" อยู่ตรงไหน
  - *Mean ( $\bar{X}$ ):* ค่าเฉลี่ย,
  - *Median:* ค่ามัธยฐาน,
  - *Mode:* ฐานนิยม
- **ดูความแตกต่าง (Dispersion):** ค่าที่บอกว่าข้อมูล "เกาะกลุ่ม" หรือ "กระจัดกระจาย"
  - *SD (ส่วนเบี่ยงเบนมาตรฐาน),*
  - *Range (พิสัย)*

### 3. การนำเสนอ (Visualization)

- แปลงตัวเลขเป็นภาพ: Histogram, Bar Chart, Pie Chart เพื่อให้เห็นแนวโน้มได้ทันที

ค่าสถิติ (Statistic)	สัญลักษณ์
ค่าเฉลี่ย (Mean)	$\bar{x}$
มัธยฐาน (Median)	Me
ฐานนิยม (Mode)	Mo
ค่าเฉลี่ยเรขาคณิต (G.M.)	G.M.
ค่าเฉลี่ยฮาร์โมนิก (H.M.)	H.M.
พิสัย (Range)	R
ส่วนเบี่ยงเบนมาตรฐาน(Sample SD)	s หรือ S.D.
ส่วนเบี่ยงเบนมาตรฐาน(Population SD)	$\sigma$
ความแปรปรวน(Sample Variance)	s <sup>2</sup>
สัมประสิทธิ์ความแปรผัน(C.V.)	C.V.
ควอร์ไทล์ (Quartile)	
เปอร์เซ็นต์ไทล์(Percentile)	P
เดไซล์(Decile)	D
ความเบ้ (Skewness)	
ความโด่ง (Kurtosis)	

## ข้อที่ 1: การหาค่ากลาง (Basic Central Tendency)

โจทย์: คะแนนสอบย่อยของนักเรียนกลุ่มหนึ่งจำนวน 9 คน เป็นดังนี้:

7, 5, 9, 5, 8, 6, 9, 5, 10

จงหาค่าเฉลี่ย (Mean), มัธยฐาน (Median), และฐานนิยม (Mode)

## ข้อที่ 2: การวัดการกระจาย (Dispersion & SD)

โจทย์: ในการผลิตน็อต โรงงานสุ่มตัวอย่างมา 5 ตัว วัดขนาดได้ดังนี้ (หน่วย mm):

10, 12, 14, 16, 18

จงหาส่วนเบี่ยงเบนมาตรฐาน (Sample S.D.)

### ข้อที่ 3: ผลกระทบของ Outlier (Mean vs Median)

โจทย์: เงินเดือนพนักงานในแผนกหนึ่งจำนวน 5 คน คือ:

15,000 | 18,000 | 20,000 | 22,000 | 200,000 (ผู้จัดการ)

จงหาค่าเฉลี่ยและมัธยฐาน และตอบว่าค่าใดเป็นตัวแทนที่ดีกว่า?

#### ข้อที่ 4: เกรดเฉลี่ย (Weighted Mean)

โจทย์: นาย ก. ลงเรียน 3 วิชา ได้เกรดดังนี้

- วิชา A (3 หน่วยกิต): เกรด 4
- วิชา B (2 หน่วยกิต): เกรด 2
- วิชา C (1 หน่วยกิต): เกรด 3

จงหาเกรดเฉลี่ย (GPA)

### ข้อที่ 5: การหาตำแหน่ง (Percentile)

โจทย์: คะแนนสอบของนักเรียน 10 คน เรียงจากน้อยไปมากแล้วดังนี้:

40, 45, 50, 55, 60, 65, 70, 75, 80, 90

จงหาว่าคะแนน 60 คะแนน ตรงกับเปอร์เซ็นต์ไทล์ที่เท่าไร? (ขณะคนอยู่ที่ %)

## ข้อที่ 6: ผลกระทบของ Outlier ในข้อมูล Viral Load

สถานการณ์: ในการตรวจวัดปริมาณไวรัส (Viral Load) ในผู้ป่วย 7 คน ค่าที่ได้มักมีการกระจายตัวสูงและอาจมีผู้ป่วยที่เป็น Super Spreader ข้อมูล (Copies/mL):

1,500 | 2,000 | 1,800 | 2,200 | 1,900 | 1,600 | 250,000 (Super Spreader)

คำถาม:

1. จงหาค่า Mean และ Median
2. ถ้าคุณต้องทำรายงานสรุปสถานการณ์ "ผู้ป่วยทั่วไป" ในกลุ่มนี้ คุณควรเลือกใช้ค่าใดเป็นตัวแทน และเพราะเหตุใด?
3. จงระบุว่าการแจกแจงข้อมูล (Distribution) มีลักษณะเบ้ (Skewness) ไปทางไหน?

## ข้อที่ 7: การคำนวณจากข้อมูลกลุ่ม (Grouped Data) - ขนาดสปอร์เห็ด

สถานการณ์: นักกีฬาวិทยา (Mycologist) วัดขนาดสปอร์ของเห็ดสายพันธุ์ใหม่จำนวน 50 สปอร์ แต่ไม่ได้บันทึกค่าดิบ บันทึกมาเป็น "ช่วงความกว้าง" (Frequency Table) ดังนี้:

จงประมาณค่าเฉลี่ย ( $\bar{x}$ ) ของขนาดสปอร์จากตารางนี้

ขนาดสปอร์ ( $\mu\text{m}$ )	จำนวนสปอร์ (f)	จุดกึ่งกลางชั้น (xi)
4 - 6	5	
7 - 9	15	
10 - 12	20	
13 - 15	10	
รวม		

ค่าสถิติ (Statistic)	สัญลักษณ์	Excel Function	
ค่าเฉลี่ย (Mean)	$\bar{x}$	AVERAGE(ช่วงข้อมูล)	
มัธยฐาน (Median)	Me	MEDIAN(ช่วงข้อมูล)	
ฐานนิยม (Mode)	Mo	MODE.SNGL(ช่วงข้อมูล)	
ค่าเฉลี่ยเรขาคณิต (G.M.)	G.M.	GEOMEAN(ช่วงข้อมูล)	
ค่าเฉลี่ยฮาร์โมนิก (H.M.)	H.M.	HARMEAN(ช่วงข้อมูล)	
พิสัย (Range)	R	ไม่มีฟังก์ชันตรงตัว	ต้องใช้สูตรผสม: =MAX(range) - MIN(range)
ส่วนเบี่ยงเบนมาตรฐาน(Sample SD)	s หรือ S.D.	STDEV.S(ช่วงข้อมูล)	ใช้บ่อยที่สุดในงานวิจัย(คำนวณแบบ n-1)
ส่วนเบี่ยงเบนมาตรฐาน(Population SD)	$\sigma$	STDEV.P(ช่วงข้อมูล)	ใช้เมื่อเก็บข้อมูลได้ครบทุกคน(คำนวณแบบ n)
ความแปรปรวน(Sample Variance)	s <sup>2</sup>	VAR.S(ช่วงข้อมูล)	ใช้คู่กับ STDEV.S
ควอร์ไทล์ (Quartile)		QUARTILE.EXC(range, k)	แบ่งข้อมูลเป็น 4 ส่วน k ใส่เลข 1, 2, หรือ 3
เปอร์เซ็นต์ไทล์(Percentile)	P	PERCENTILE.EXC(ช่วงข้อมูล, k)	แบ่งข้อมูลเป็น 4 ส่วน k ใส่เลข 1-100
เดไซล์(Decile)	D	ใช้ฟังก์ชันเดียวกับ Percentile	แบ่งข้อมูลเป็น 4 ส่วน k ใส่เลข 1-100
ความเบ้ (Skewness)		SKEW(ช่วงข้อมูล)	ค่า + = เบ้ขวา (หางยาวไปทางขวา)ค่า - = เบ้ซ้าย (หางยาวไปทางซ้าย)ค่า 0 = สมมาตร (ระฆังคว่ำสวยงาม)
ความโด่ง (Kurtosis)		KURT(ช่วงข้อมูล)	ค่า + = โด่งมาก (ข้อมูลกระจุกตัวสูง)ค่า - = แบนราบ (ข้อมูลกระจายตัวมาก)

## Inferential Statistics (สถิติเชิงอนุมาน)

หัวข้อรอง: การขยายผลและทำนาย "สิ่งที่มองไม่เห็น" ในประชากร

### 1. หัวใจสำคัญ (Concept)

- คือการนำข้อมูลจาก "กลุ่มตัวอย่างเล็กๆ" (Sample) ไปใช้อ้างอิงหรือทำนายผลของ "ประชากรทั้งหมด" (Population)
- เป้าหมาย: หาข้อสรุปในภาพใหญ่ โดยยอมรับความคลาดเคลื่อนได้ในระดับที่กำหนด (Confidence Level)

### 2. เครื่องมือหลัก (Key Methods)

- การประมาณค่า (Estimation): การคาดเดาค่าจริงของประชากรว่าจะอยู่ที่เท่าไร

• เช่น *Confidence Interval* (ช่วงความเชื่อมั่น)

- การทดสอบสมมติฐาน (Hypothesis Testing): การพิสูจน์ข้อสงสัยทางสถิติ

• เช่น *T-test, ANOVA, Chi-Square* เพื่อดูว่าความแตกต่างที่เห็นนั้น "มีนัยสำคัญ" หรือแค่บังเอิญ

### 3. การนำไปใช้ (Application)

- ใช้หาความสัมพันธ์ (Correlation), เปรียบเทียบผลการทดลอง, หรือพยากรณ์อนาคต (Regression)

สัมประสิทธิ์ความแปรผัน(C.V.)	C.V.	ไม่มีฟังก์ชันตรงตัว	สูตร: $SD / Mean \times 100$
------------------------------	------	---------------------	------------------------------

### ข้อที่ 1: การเปรียบเทียบความแม่นยำ (Precision) ของชุดข้อมูลต่างหน่วย

สถานการณ์: นักวิจัยต้องการเปรียบเทียบ "ความนิ่ง" (Consistency) ของวิธีการวัดโปรตีน 2 วิธี

- วิธี A (Bradford Assay): วัดค่าการดูดกลืนแสง (OD) ได้ค่าเฉลี่ยต่ำ
- วิธี B (Bicinchoninic Acid - BCA): ให้ค่าสัญญาณสูงกวามาก

ข้อมูล (n=5 การทดลองซ้ำ):

- Method A (OD Units): 0.5, 0.6, 0.5, 0.7, 0.5
- Method B (RFU Units): 500, 600, 550, 520, 580

คำถาม:

1. จงคำนวณค่าเฉลี่ย ( $\bar{x}$ ) และส่วนเบี่ยงเบนมาตรฐาน (S) ของทั้งสองวิธี
2. วิธีไหนมีความแม่นยำ (Precision) สูงกว่ากัน?

ข้อที่ 1: การเปรียบเทียบความแม่นยำ (Precision) ของชุดข้อมูลต่างหน่วย

สถานการณ์: นักวิจัยต้องการเปรียบเทียบ "ความนิ่ง" (Consistency) ของวิธีการวัดโปรตีน 2 วิธี

- วิธี A (Bradford Assay): วัดค่าการดูดกลืนแสง (OD) ได้ค่าเฉลี่ยต่ำ
- วิธี B (Bicinchoninic Acid - BCA): ให้ค่าสัญญาณสูงกว่ามาก

ข้อมูล (n=5 การทดลองซ้ำ):

- Method A (OD Units): 0.5, 0.6, 0.5, 0.7, 0.5
- Method B (RFU Units): 500, 600, 550, 520, 580

คำถาม:

1. จงคำนวณค่าเฉลี่ย ( $\bar{x}$ ) และส่วนเบี่ยงเบนมาตรฐาน (S) ของทั้งสองวิธี
2. วิธีไหนมีความแม่นยำ (Precision) สูงกว่ากัน?

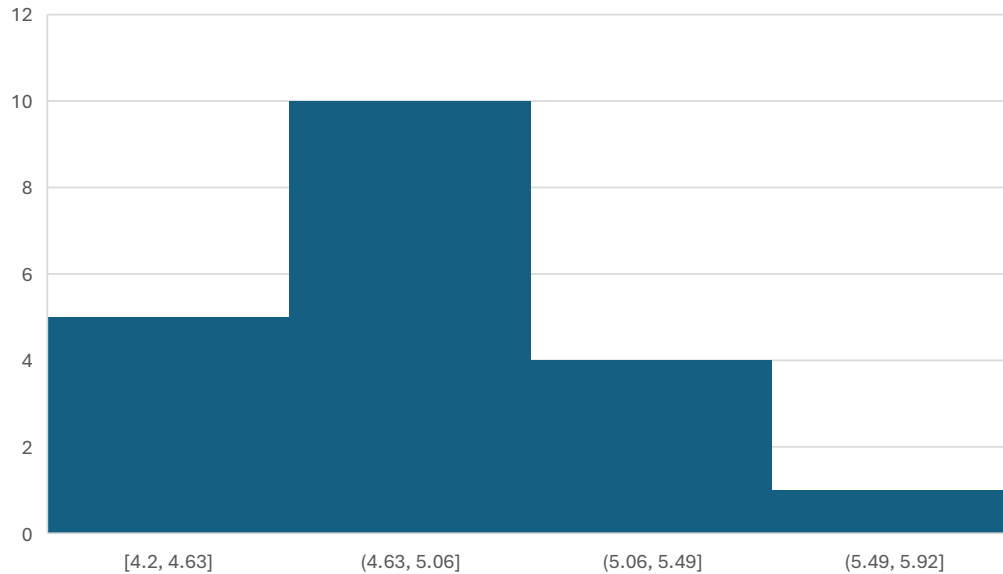
## การนำเสนอ (Visualization)

• แปลงตัวเลขเป็นภาพ: Histogram, Bar Chart, Pie Chart

เพื่อให้เห็นแนวโน้มได้ทันที

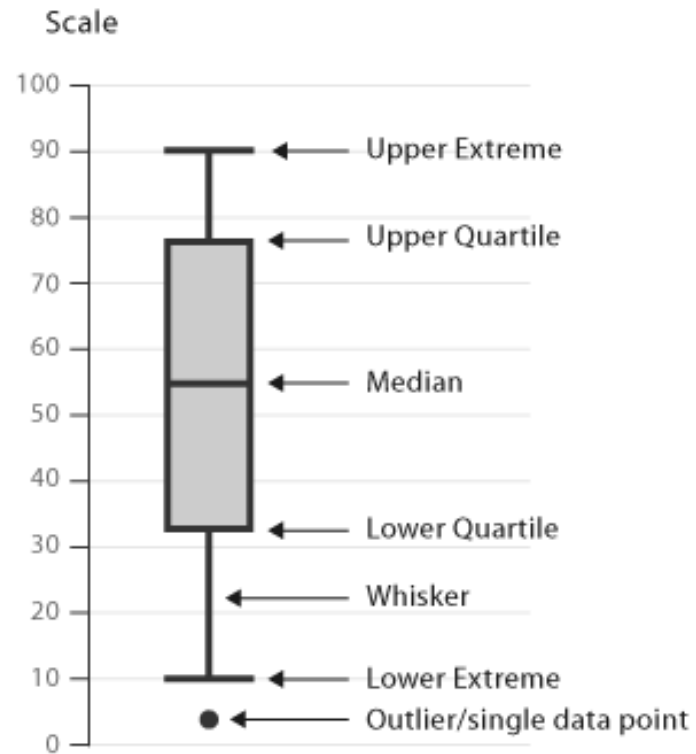
### โจทย์ข้อที่ 1: ขนาดของสปอร์เห็ด (Spore Size Distribution)

- สถานการณ์: นักวิจัยวัดความยาวของสปอร์เห็ดชนิดหนึ่งจำนวน 20 ตัวอย่าง เพื่อดูความแปรปรวนของขนาด
- กราฟที่แนะนำ: Histogram (เพื่อดูว่าขนาดส่วนใหญ่อยู่ที่ช่วงไหน และกราฟเป็นระฆังคว่ำหรือไม่)
- ข้อมูล (หน่วย ): 4.2, 4.5, 5.1, 4.8, 5.0, 4.9, 5.2, 4.3, 4.7, 5.0, 5.5, 4.8, 4.6, 5.1, 4.9, 5.3, 4.7, 5.0, 4.4, 4.8



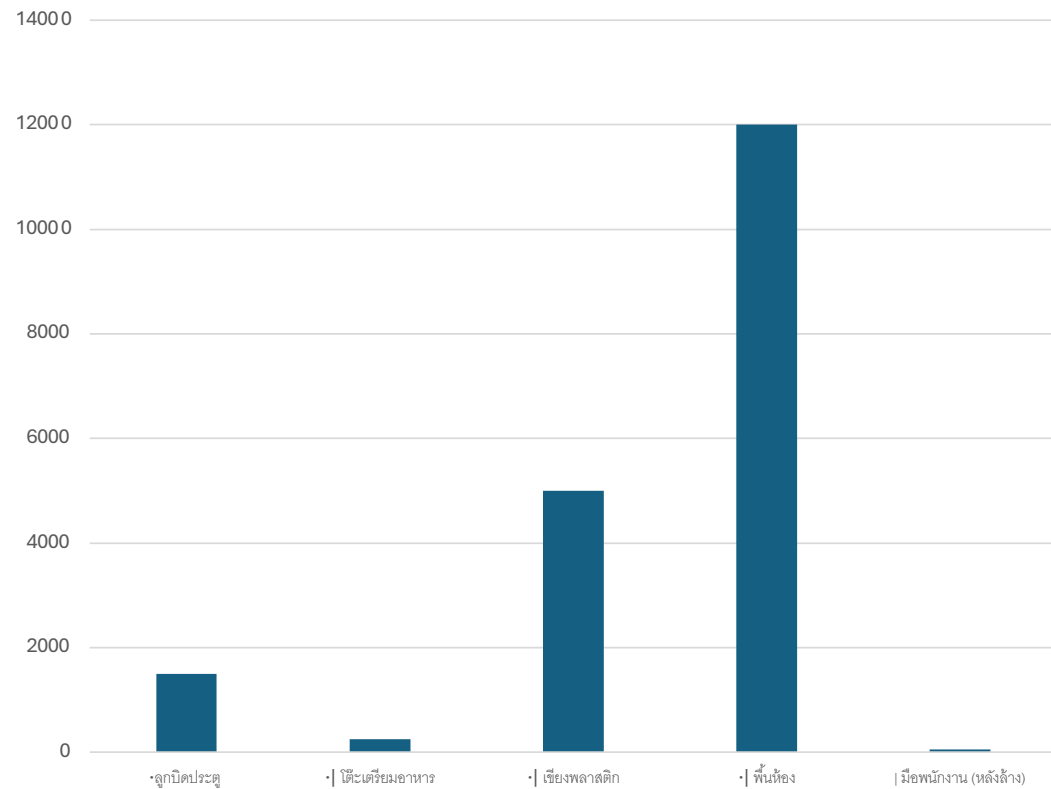
## โจทย์ข้อที่ 2: เปรียบเทียบผลผลิต DNA จากชุดสกัด 2 ยี่ห้อ

- สถานการณ์: เปรียบเทียบความเข้มข้น DNA ที่สกัดได้จากชุด Kit A และ Kit B
- กราฟที่แนะนำ: Box Plot (เพื่อดู Median และการกระจายตัว รวมถึง Outlier)
- ข้อมูล (ng/):
  - Kit A: 25, 28, 30, 22, 120 (Outlier), 26, 29
  - Kit B: 45, 48, 50, 42, 46, 49, 47



### โจทย์ข้อที่ 3: ปริมาณเชื้อแบคทีเรียบนพื้นผิวต่างๆ (Food Safety)

- สถานการณ์: ตรวจสอบเชื้อ (Swab Test) ในจุดต่างๆ ของโรงงานผลิตอาหาร
- กราฟที่แนะนำ: Bar Chart (แนวตั้งหรือแนวนอน)
- ข้อมูล:
  - | พื้นผิว (Surface) | จำนวนเชื้อ (CFU/sq.in) | | :--- | :--- |
  - | ลูกบิดประตู | 1,500 |
  - | โต๊ะเตรียมอาหาร | 250 |
  - | เชียงพลาสติก | 5,000 |
  - | พื้นห้อง | 12,000 |
  - | มือพนักงาน (หลังล้าง) | 50 |

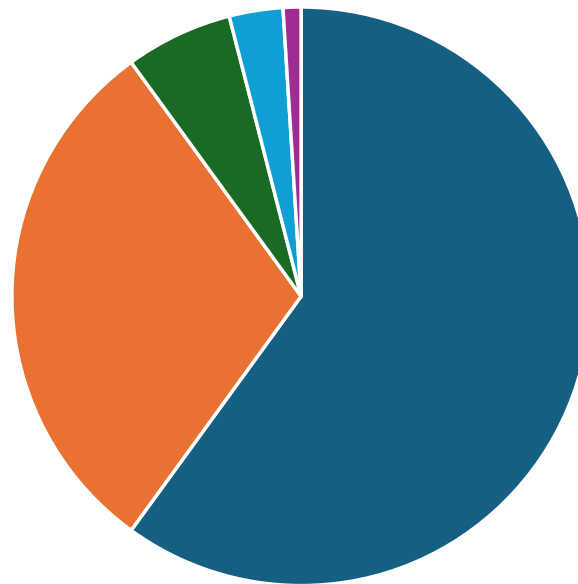


#### โจทย์ข้อที่ 4: ประสิทธิภาพยาปฏิชีวนะ (Antibiotic Susceptibility)

- สถานการณ์: วัดขนาดวงใส (Zone of Inhibition) ของยา 4 ชนิดต่อเชื้อ *E. coli*
- กราฟที่แนะนำ: Column Chart (พร้อม Error Bar ถ้ามีค่า SD)
- ข้อมูล: | ยาปฏิชีวนะ | ขนาดวงใสเฉลี่ย (mm) | | :--- | :--- |
- | Penicillin | 0 (ดื้อยา) |
- | Amoxicillin | 15 |
- | Gentamicin | 22 |
- | Tetracycline | 18 |

### โจทย์ข้อที่ 5: ชนิดของเม็ดเลือดขาว (Differential White Blood Cell Count)

- สถานการณ์: ผลตรวจเลือดของผู้ป่วย คำนวณส่วนเม็ดเลือดขาว 5 ชนิด
- กราฟที่แนะนำ: Pie Chart หรือ Donut Chart
- ข้อมูล: | ชนิดเซลล์ | จำนวน (%) | | :--- | :--- |
- | Neutrophil | 60% |
- | Lymphocyte | 30% |
- | Monocyte | 6% |
- | Eosinophil | 3% |
- | Basophil | 1% |



■ -| Neutrophil   ■ -| Lymphocyte   ■ -| Monocyte   ■ -| Eosinophil   ■ -| Basophil

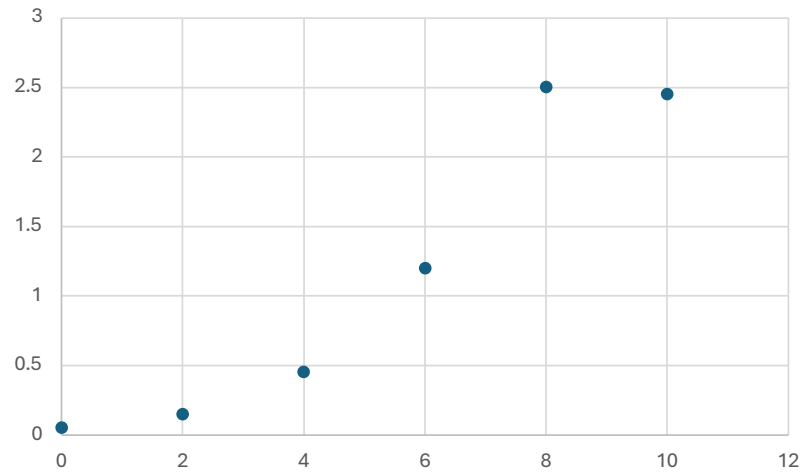
### โจทย์ข้อที่ 6: องค์ประกอบขยะมูลฝอยทางชีวภาพ

- สถานการณ์: แยกประเภทขยะในบ่อหมักปุ๋ย
- กราฟที่แนะนำ: Pie Chart (ระวังอย่าให้มี Category เยอะเกินไป)
- ข้อมูล: เศษอาหาร (45%), เศษใบไม้ (30%), เปลือกผลไม้ (15%), กระดูกสัตว์ (10%)

### โจทย์ข้อที่ 7: กราฟการเจริญเติบโตของแบคทีเรีย (Bacterial Growth Curve)

- สถานการณ์: วัดค่าความขุ่น (OD600) ของเชื้อ *S. aureus* ทุกชั่วโมง
- กราฟที่แนะนำ: Line Chart (Smooth line)
- ข้อมูล:

เวลา (ชั่วโมง)	ค่า OD600
0	0.05
2	0.15
4	0.45
6	1.20
8	2.50
10	2.45



โจทย์ข้อที่ 9: สถานการณ์: นักวิทยาศาสตร์การกีฬาต้องการศึกษาความฟิตของร่างกายระหว่าง "นักกีฬามืออาชีพ" (Athlete) และ "คนทั่วไปที่ไม่ออกกำลังกาย" (Non-Athlete) โดยให้ทั้งคู่ปั่นจักรยานด้วยความเร็วคงที่ เป็นเวลา 15 นาที แล้วหยุดพักทันที (Cool down) จากนั้นทำการวัดอัตราการเต้นของหัวใจ (BPM) ทุกๆ 5 นาที ตั้งแต่เริ่มจนจบการทดลอง

เวลา (นาที)	อัตราการเต้นของหัวใจ (Athlete)	อัตราการเต้นของหัวใจ (Non-Athlete)
0 (เริ่ม)	60	75
5	95	110
10	120	145
15 (หยุดปั่น)	135	170
20 (พัก)	90	150
25 (พัก)	70	130
30 (จบ)	62	110

1. จงสร้าง Line Chart เพื่อแสดงแนวโน้มอัตราการเต้นของหัวใจของทั้ง 2 กลุ่ม

• แกน X: เวลา (นาที)

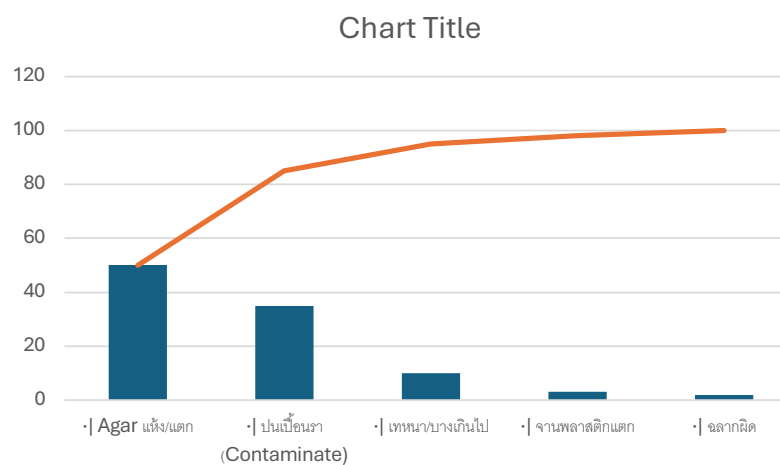
• แกน Y: อัตราการเต้นของหัวใจ (ครั้ง/นาที)

• เส้นกราฟ: ต้องมี 2 เส้น (แยกสี) พร้อมคำอธิบาย (Legend) ว่าเส้นไหนคือใคร

2. การวิเคราะห์: จากกราฟ จงตอบคำถามว่ากลุ่มใดที่มี "อัตราการฟื้นตัว" (Recovery Rate) ดีกว่ากัน? (ดูจากความชันของกราฟช่วงนาทีที่ 15 ถึง 30)

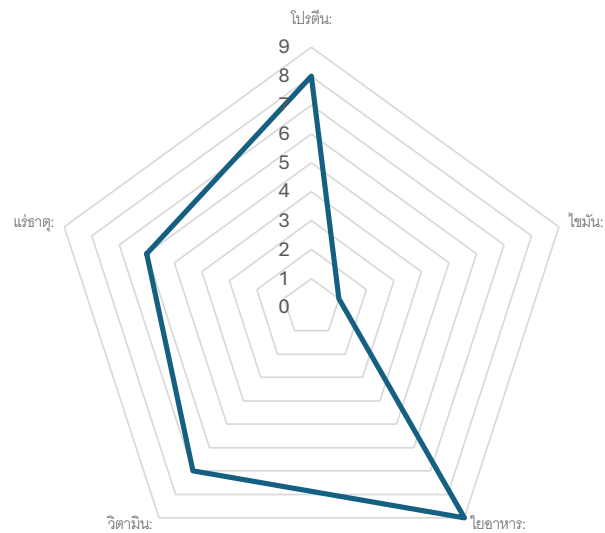
### โจทย์ข้อที่ 9: สาเหตุความเสียหายของงานเพาะเชื้อ (Pareto Chart - HACCP)

- สถานการณ์: ตรวจสอบสาเหตุที่งาน Agar ใช้งานไม่ได้ เพื่อแก้ปัญหาที่สาเหตุหลัก
- กราฟที่แนะนำ: Pareto Chart (กราฟแท่งผสมเส้นสะสม)
- ข้อมูล: | สาเหตุความเสียหาย | จำนวน (งาน) | | :--- | | :--- |
- | Agar แท่ง/แตก | 50 |
- | ปนเปื้อนรา (Contaminate) | 35 |
- | เทหนา/บางเกินไป | 10 |
- | งานพลาสติกแตก | 3 |
- | ฉลากผิด | 2 |



## โจทย์ข้อที่ 10: คุณค่าทางโภชนาการของเห็ด (Nutritional Profile)

- สถานการณ์: เปรียบเทียบสารอาหาร 5 ด้านของเห็ดนางฟ้า
- กราฟที่แนะนำ: Radar Chart (Spider Web)
- ข้อมูล (คะแนนเต็ม 10):
  - โปรตีน: 8
  - ไขมัน: 1
  - โยอาหาร: 9
  - วิตามิน: 7
  - แร่ธาตุ: 6



## การเก็บรวบรวมข้อมูล

- ทดสอบ (Test)
- การสอบถามและสำรวจ (Survey)
- การสังเกต (observation)
- การสัมภาษณ์ (Interview)
- การสนทนากลุ่ม (Focus group Discussion)
- การวิเคราะห์เอกสาร (Documentary analysis)

# ลำดับขั้นตอนการรวบรวมข้อมูล

- กำหนดข้อมูลและตัวชี้วัด
- กำหนดแหล่งข้อมูล
- กำหนดวิธีการเก็บรวบรวมข้อมูล
- การสร้างและหาคุณภาพเครื่องมือที่ใช้ในการเก็บข้อมูล
- การเก็บข้อมูลกับกลุ่มผู้ให้ข้อมูล

## ลักษณะของข้อมูลที่ดี

- Accuracy and validity
- Reliability
- Completeness
- Legibility
- Modernity
- Accessibility

# สิ่งพึงปฏิบัติและต้องคำนึงถึงในการเก็บรวบรวมข้อมูล

- เคารพ ให้เกียรติ และรบกวนผู้ให้ข้อมูลน้อยที่สุดเท่าที่จะเป็นไปได้
- หลีกเลี่ยงการหลอกลวงผู้เข้าร่วมวิจัย/ผู้ให้ข้อมูล
- คำนึงถึงศักยภาพของผู้ให้ข้อมูลแต่ละคนและไม่หาประโยชน์จากผู้ให้ข้อมูล
- ไม่ใช่ประโยชน์จากผู้ให้ข้อมูล โดยพอเก็บรวบรวมข้อมูลได้ก็หนีหายไป
- หลีกเลี่ยงการเก็บรวบรวมข้อมูลที่เป็นอันตรายหรือไม่เป็นทางการ
- สร้างความไว้วางใจ และบอกเล่าเรื่องราวที่จะไปศึกษาหรือเก็บข้อมูลให้กับ ผู้ให้ข้อมูลได้รับทราบ
- บอกวัตถุประสงค์ของการเข้ามาศึกษาและวิธีการที่จะนำข้อมูลไปใช้
- หลีกเลี่ยงคำถามที่จะทำให้ผู้ให้ข้อมูลไม่ประทับใจ และหลีกเลี่ยงการเปิดเผย ข้อมูลสำคัญของผู้ให้ข้อมูลกับผู้อื่น
- ให้รางวัลหรือสิ่งตอบแทนกับผู้ให้ข้อมูล
- คำถามจะต้องอยู่ในขอบเขตที่กำหนดเอาไว้

# ประเภทของข้อมูล

- แบ่งตามแหล่งที่มา
  - Primary data
  - Secondary data
- แบ่งตามแหล่งข้อมูล
  - Quantitative data
  - Qualitative data
- แบ่งตามมาตรวัดข้อมูล
  - Norminal scale data
  - Ordinal scale data
  - Interval scale data
  - Ratio scale data

## คำถามท้ายบท

1. จงระบุว่าข้อมูลต่อไปนี้เป็นข้อมูลที่อยู่ในมาตรวัดใด
  - 1.1 ผลสัมฤทธิ์ทางการเรียน
  - 1.2 ปริมาณน้ำฝน
  - 1.3 ความพึงพอใจ
  - 1.4 ลำดับที่ของการสอบ
  - 1.5 รายได้
  - 1.6 บ้านเลขที่
  - 1.7 อาชีพ
  - 1.8 เวลามาเรียน
  - 1.9 ระดับการศึกษา
  - 1.10 ศาสนา
2. จงอธิบายลักษณะของข้อมูลที่ดี พร้อมยกตัวอย่างประกอบ
3. การเก็บรวบรวมข้อมูลควรคำนึงถึงสิ่งใดบ้าง เพราะเหตุใด
4. จงยกตัวอย่างปัญหาวิจัยที่สนใจมาสัก 1 เรื่อง แล้วระบุวิธีการเก็บรวบรวม ข้อมูล ที่จะทำได้มาซึ่งคำตอบของปัญหาวิจัยนั้นได้
5. สิ่งที่ยังปฏิบัติและต้องคำนึงถึงในการเก็บรวบรวมข้อมูลมีอะไรบ้างเพราะเหตุใดต้องพึงปฏิบัติและคำนึงถึงประเด็นเหล่านั้นด้วย