

## บทที่ 3

### การเก็บรวบรวมข้อมูล และการนำเสนอข้อมูล

#### 3.1 การเก็บรวบรวมข้อมูล (Collection of data)

วิธีการเก็บรวบรวมข้อมูลมีหลายวิธีด้วยกัน แต่ก่อนอื่นควรจะทราบความหมายของคำว่า ประชากร (Population) พารามิเตอร์ (Parameter) และตัวอย่าง (Sample) เสียก่อน ดังนี้

##### ประชากร (Population)

ประกอบด้วยหน่วย (Units) ต่าง ๆ ทั้งหมดที่สนใจ บางทีเรียกว่า ครอบคลุม (Coverage) ซึ่งคำว่าหน่วยในที่นี้หมายถึงสิ่งหนึ่ง หรือกลุ่มของสิ่งของต่าง ๆ ที่เราอาจจะวัดหรือสังเกตข้อเท็จจริงทางสถิติมาได้ และคำว่าหน่วยนี้บางทีเราก็เรียกว่า หน่วยแจงนับ (Enumeration Units)

ประชากร แบ่งออกเป็น 2 ประเภทด้วยกัน คือ

ก. ประชากรชนิดรู้จบ หรือจำกัด (Finite population) เป็นประชากรที่ประกอบด้วยหน่วยแจงนับที่มีจำนวนรู้จบ หรือจำกัด

ข. ประชากรชนิดไม่รู้จบ หรืออนันต์ (Infinite population) เป็นประชากรที่ประกอบด้วยหน่วยแจงนับที่มีจำนวนไม่รู้จบหรืออนันต์

##### พารามิเตอร์ (Parameter)

หมายถึงตัวคงที่ ซึ่งแสดงคุณลักษณะของประชากรซึ่งพารามิเตอร์จะเป็นฟังก์ชันของค่าของหน่วยต่าง ๆ ทั้งหมดในประชากร เช่น ค่าเฉลี่ยประชากร ( $\mu$ ) สัดส่วนประชากร ( $\pi$ ) เป็นต้น

##### ตัวอย่าง (Sample)

หมายถึงกลุ่มของบรรดาหน่วยที่เลือกได้จากประชากรตามวิธีการสำรวจด้วยตัวอย่าง (Sample Survey) เพื่อที่จะได้ตัวอย่างที่เป็นตัวแทนที่ดีของประชากร การเก็บรวบรวมข้อมูลวิธีที่นิยมใช้กันอยู่มี 3 วิธีด้วยกัน คือ

## 1. การสำมะโน (Census or complete enumeration)

หมายถึงการสำรวจหรือแจงนับหน่วยทุกหน่วยที่อยู่ในประชากรที่สนใจ เช่น การทำสำมะโนประชากร สำมะโนการเกษตร เป็นต้น

โดยทั่วไปแล้ว การทำสำมะโนต้องการเงิน เวลา และกำลังงานมาก ดังนั้น ถ้าไม่จำเป็นจริง ๆ แล้วก็จะไม่ค่อยทำกัน แต่ถ้าต้องการจะได้ข้อเท็จจริงจากทุก ๆ หน่วย แล้วก็จำเป็นต้องใช้วิธีการสำมะโน สำหรับสำมะโนประชากรในประเทศไทยนั้น โดยทั่วไปแล้ว จะทำกับทุก ๆ 5 ปี หรือ 10 ปี

### ข้อดีของการทำสำมะโน

1. ยอดข้อมูลสถิติที่รวบรวมได้ในทุกเรื่องสามารถแสดงออกในเขตบริหาร หรือเขตภูมิศาสตร์ที่เล็กที่สุดได้ ทั้งนี้เพราะมีข้อมูลจากทุกหน่วยแจงนับในคัมรวม เช่น ในสำมะโนเกษตร เราสามารถแสดงตารางข้อมูลสถิติในระดับ จังหวัด อำเภอ ตำบล หรือหมู่บ้านก็สามารถทำได้

2. ข้อมูลสถิติที่ได้นั้นเป็นข้อมูลหลักซึ่งจะนำไปใช้ในการวางแผนเก็บข้อมูลอื่น ๆ ได้อีก

### ข้อเสียของการทำสำมะโน

1. ใช้ทรัพยากร เช่น กำลังคน และงบประมาณมาก จึงไม่สามารถจะทำงานสำมะโนได้ทุกปี ดังนั้นจึงทำให้ข้อมูลมีไม่ครบทุกปี และอาจไม่พอเพียงกับการที่ต้องใช้

2. เสียเวลามาก เนื่องจากมีหน่วยแจงนับมาก จึงทำให้ปริมาณงานมาก ไม่สามารถทำงานให้เสร็จได้ทันท่วงทีได้

3. คุณภาพของข้อมูลที่รวบรวมได้ยังเป็นที่น่าสงสัย เพราะต้องใช้เจ้าหน้าที่ร่วมทำงานสำมะโนเป็นจำนวนมาก ซึ่งยากแก่การควบคุมการบริหารงาน และควบคุมคุณภาพในการแจงนับ

4. ผู้ให้คำตอบในการแจงนับบางรายไม่สามารถให้ข้อมูลที่ถูกต้องแก่พนักงานแจงนับได้ จึงทำให้ยอดข้อมูลสถิติผิดไปจากความเป็นจริง

## 2. การสำรวจด้วยตัวอย่าง (Sample Survey)

หมายถึงการรวบรวมข้อเท็จจริงจากหน่วยตัวอย่าง (Sampling units) ที่เลือกมาเป็นตัวแทนของประชากรที่สนใจ แล้วใช้ข้อเท็จจริงที่ได้ไปกะประมาณค่าพารามิเตอร์ที่เราสนใจ

เหตุผลที่เราต้องใช้การสำรวจด้วยตัวอย่าง คือ

1. ประหยัดเงินในการดำเนินงานในการรวบรวมข้อมูล
2. ทุ่นเวลาในการทำงาน การสำรวจด้วยตัวอย่างจะประหยัดเวลาในการเก็บข้อมูล ประมวลผลข้อมูล และวิเคราะห์ข้อมูล

3. รวบรวมข้อมูลได้กว้างขวาง

4. ความถูกต้องของข้อมูลมีมาก

แต่อย่างไรก็ตาม การสำรวจด้วยตัวอย่างก็มีข้อเสียเหมือนกัน คือ

1. ไม่สามารถประมาณข้อมูลในระดับย่อย ๆ หรือห้องที่เล็ก ๆ ได้ และการวิเคราะห์ข้อมูลในรูปตารางก็ไม่สามารถจำแนกข้อมูลในรายละเอียดได้มากนัก เพราะได้ข้อมูลมาไม่พอเพียงที่จะจำแนก หรือถ้าทำได้ก็อาจจะมีความคลาดเคลื่อนสูงมาก

2. พวกสถิติที่ประมวลได้จากตัวอย่างเป็นเพียงค่าประมาณ (Estimate) ของคุณลักษณะที่เราสนใจ ไม่ใช่ยอดสถิติจริง

วิธีการที่จะเลือกตัวอย่างให้เป็นตัวแทนที่ดีของประชากรที่เราสนใจนั้น มีหลายวิธีด้วยกัน ดังนี้ คือ

1. การสุ่มตัวอย่างแบบง่าย (Simple Random Sampling หรือ SRS.) เป็นการสุ่มตัวอย่างที่ทุก ๆ หน่วยในประชากรมีโอกาสที่จะถูกเลือกเป็นตัวอย่างเท่า ๆ กัน ซึ่งวิธีการเลือกตัวอย่างดังกล่าว กระทำได้ 2 วิธีด้วยกัน คือ

ก. ใช้วิธีจับสลาก มักจะใช้ในกรณีที่ประชากรมีจำนวนไม่มากนัก การจับสลากทำได้ 2 วิธี คือ เลือกโดยไม่มีการแทนที่ (Sampling without replacement) และเลือกโดยมีการแทนที่ (Sampling with replacement) ซึ่งในทางปฏิบัตินิยมใช้แบบแทนที่

ข. ใช้ตารางเลขสุ่ม (Random number tables) ซึ่งเป็นวิธีที่นิยมกันมากในทางปฏิบัติ เพราะใช้ง่ายและสะดวกรวดเร็ว การใช้ตารางเลขสุ่มขั้นแรกต้องใส่หมายเลขกำกับให้กับหน่วยประชากรทั้งหมด ขึ้นต่อมาจึงใช้ตารางเลขสุ่มช่วยในการเลือกหน่วยตัวอย่าง

ตัวอย่างเช่น ถ้าเรามีประชากรทั้งหมด 500 หน่วย

ขั้นแรก ใส่หมายเลขกำกับแก่หน่วยประชากรทั้งหมดดังนี้

หน่วย	$P_1$	$P_2$ .....	$P_{500}$
หมายเลข	001	002 .....	500

**ขั้นที่สอง** กำหนดขนาดตัวอย่างที่เราต้องการสมมติให้เท่ากับ 30 เราเลือกใช้ตารางเลขสุ่มที่มี 3 หลัก มีค่าไม่เกิน 500 แล้วทำการเลือกสุ่มในช่วง 001-500 ขึ้นมาทีละ 3 ตัว โดยไม่เจาะจง เช่น ถ้าได้ 025 หน่วยที่ตกเป็นตัวอย่าง คือ P<sub>025</sub> เลือกต่อไปเรื่อย ๆ จนครบ 30 หน่วยที่เราต้องการ

**ข้อดีและข้อเสียในการสุ่มตัวอย่างแบบง่าย**

**ข้อดี**

เป็นวิธีการที่ง่ายไม่สลับซับซ้อนและเข้าใจได้ง่ายกว่าวิธีอื่น ๆ และควรใช้กับประชากรที่มีลักษณะคล้ายคลึงกันมากที่สุด

**ข้อเสีย**

เมื่อหน่วยตัวอย่างในประชากรมีความแปรปรวนสูง ตัวอย่างที่เลือกได้จะไม่เป็นตัวแทนที่ดีของประชากร

## 2. การสุ่มแบบแบ่งเป็นชั้นภูมิ (Stratified Random Sampling)

เนื่องจากในบางครั้งประชากรประกอบด้วยหน่วยต่าง ๆ ไม่เหมือนกันทางด้านคุณลักษณะที่ต้องการศึกษา ดังนั้นจึงต้องจำแนกประชากรออกเป็นกลุ่มย่อยโดยให้หน่วยต่าง ๆ ที่เหมือนกันอยู่ในกลุ่มเดียวกัน ซึ่งเราเรียกว่า “ชั้นภูมิ” (Strata) จากนั้นจึงทำการสุ่มตัวอย่างมาจากแต่ละชั้นภูมิ โดยใช้วิธีการสุ่มตัวอย่างแบบง่าย (SRS) ตัวอย่างเช่น การสำรวจค่าใช้จ่ายของนักศึกษามหาวิทยาลัยแห่งหนึ่ง เราจำแนกนักศึกษาออกเป็นกลุ่มต่าง ๆ ตามคณะ หลังจากนั้นจึงทำการสุ่มตัวอย่างนักศึกษามาจากแต่ละคณะ โดยวิธี SRS

**ข้อดีและข้อเสียของการสุ่มแบบแบ่งเป็นชั้นภูมิ**

**ข้อดี**

1. มีประสิทธิภาพสูงกว่า แบบ SRS
2. ได้รายละเอียด แยกเป็นรายชั้นภูมิ ทำให้สามารถที่จะศึกษาเปรียบเทียบระหว่างกลุ่มได้

3. การควบคุมด้านบริหารสะดวกกว่าแบบ SRS

**ข้อเสีย**

ต้องจัดเตรียมงานล่วงหน้าเพื่อจัดแบ่งเป็นชั้นภูมิ

### 3. การสุ่มแบบมีระบบ (Systematic Sampling)

เป็นการสุ่มตัวอย่างโดยเริ่มจากจุดเริ่มต้นแบบสุ่ม (Random Start) แล้วจึงเลือกตัวอย่างต่อไปอีกทุก ๆ หน่วยที่ K จากประชากรที่เรียงลำดับไว้ Sampling interval (K) =  $\frac{N}{n}$  เมื่อ N เป็นจำนวนประชากร และ n เป็นจำนวนตัวอย่าง เช่น ต้องการสำรวจรายได้ของครัวเรือนในชุมชนหนึ่ง ที่มีทั้งหมด 50 ครัวเรือน โดยมีแผนที่ของครัวเรือนดังกล่าว และต้องการเลือกตัวอย่างมา 10 ครัวเรือน  $\therefore K = \frac{50}{10} = 5$

วิธีการเลือกตัวอย่างโดยการสุ่มแบบมีระบบนี้ ขั้นแรกเขียนเลขที่กำกับให้กับครัวเรือนทั้ง 50 ครัวเรือน ซึ่งจะได้เป็นเลข 1, 2, 3, ..., 50 จากนั้นก็เลือกจุดเริ่มต้นแบบสุ่ม ๆ สมมติได้เลข 6 หมายความว่า ครัวเรือนแรกที่ถูกเลือกมาเป็นตัวอย่างคือ ครัวเรือนเลขที่ 6 ต่อไปก็คือครัวเรือนเลขที่ (6+5) = 11 และต่อไปก็คือครัวเรือนที่ (11+5) = 16 21, 26, 31, 36, 41, 46 ซึ่งได้เพียง 9 ครัวเรือน จึงต้องขึ้นต้นใหม่จะได้เลขที่ 1 ดังนั้นครัวเรือนทั้ง 10 ที่เลือกมาเป็นตัวอย่างคือครัวเรือนที่ 1, 6, 11, 16, 21, 26, 31, 36, 41 และ 46

#### ข้อดีและข้อเสียของการสุ่มแบบมีระบบ

##### ข้อดี

1. การเลือกตัวอย่างทำได้สะดวกและรวดเร็ว
2. เหมาะสำหรับงานสำรวจที่ต้องการให้หน่วยตัวอย่างกระจายโดยทั่วประชากร
3. เมื่อต้องการสุ่มตัวอย่างจากประชากรที่เรียงกันเป็นแฟ้ม หรือเป็นบัตร

รายการใช้วิธีวัดระยะเอาก็ได้

##### ข้อเสีย

มีความเที่ยงตรงน้อย ถ้าประชากรเป็นวัฏจักรหรือเมื่อประชากรมีลักษณะไม่สม่ำเสมอ

### 4. การสุ่มแบบกลุ่ม (Cluster Sampling)

เป็นการสุ่มตัวอย่างโดยแบ่งประชากรออกเป็นกลุ่มย่อย ๆ เรียกว่า cluster ก่อนแล้วจึงทำการสุ่มตัวอย่างกลุ่มของหน่วยตัวอย่าง (Cluster) โดยใช้วิธีการสุ่มแบบง่าย หรือแบบมีระบบจากนั้น จึงรวบรวมข้อมูลมาจากหน่วยย่อยทุกหน่วยของกลุ่มที่สุ่มได้

ลักษณะของการสุ่มแบบนี้มีลักษณะคล้ายกับการสุ่มแบบชั้นภูมิตรงที่ไม่ซ้ำซ้อนกัน แต่ต่างกันตรงที่ Cluster ประกอบด้วยหน่วยตัวอย่างที่มีลักษณะแตกต่างกัน ส่วนชั้นภูมิประกอบด้วยหน่วยตัวอย่างที่มีลักษณะคล้ายคลึงกัน ซึ่งแต่ละ Cluster จะรวมลักษณะทั้งหลายประชากรไว้ด้วยกัน ตัวอย่างเช่น ต้องการสำรวจรายชื่อของพนักงาน 20 คน จากแผนกต่าง ๆ 5 แผนก แต่ละแผนกมีพนักงาน 10 คน วิธีการก็คือต้องเตรียมบัญชีรายชื่อแผนกต่าง ๆ ทั้ง 5 แผนก แล้วสุ่มมาเพียง 2 แผนก จาก 5 แผนก แล้วเก็บข้อมูลจากพนักงาน 2 แผนกที่สุ่มได้นั้นก็จะได้ข้อมูลจากพนักงาน 20 คน ตามที่ต้องการ

#### **ข้อดีและข้อเสียของการสุ่มแบบกลุ่ม**

##### **ข้อดี**

1. ไม่ต้องทำบัญชีรายชื่อหน่วยตัวอย่างทุกหน่วยในประชากรเหมือนกับวิธี SRS และวิธีสุ่มแบบแบ่งเป็นชั้นภูมิ
2. ทำให้ลดค่าใช้จ่ายในการเตรียมรายชื่อ
3. ประหยัดค่าใช้จ่ายในการเดินทาง

##### **ข้อเสีย**

มีประสิทธิภาพด้อยกว่าแบบ SRS และแบบชั้นภูมิ

#### **5. การสุ่มแบบหลายขั้นตอน (Multistage Sampling)**

เป็นการเลือกตัวอย่างที่ต้องทำตั้งแต่ 2 ขั้นตอนขึ้นไปโดยในขั้นแรก ประชากรจะถูกแบ่งเป็นหน่วยตัวอย่างในขั้นที่หนึ่งก่อนแล้วสุ่มตัวอย่างมาจำนวนหนึ่ง ขั้นต่อไปก็คือการแบ่งตัวอย่างที่สุ่มได้มาเป็นกลุ่มย่อย ๆ อีก แล้วสุ่มมาอีกจำนวนหนึ่งและถ้าอยากแบ่งต่อไปอีกก็สามารถทำได้

#### **ข้อดีและข้อเสียของการสุ่มแบบหลายขั้นตอน**

##### **ข้อดี**

1. ประหยัดค่าใช้จ่ายในการเดินทาง
2. ลดค่าใช้จ่ายในการเตรียมบัญชีรายชื่อ
3. มีประสิทธิภาพและยืดหยุ่นได้มากกว่าการสุ่มแบบขั้นตอนเดียว

##### **ข้อเสีย**

1. เป็นวิธีการที่สลับซับซ้อน เข้าใจยาก
2. ต้องใช้การวางแผนละเอียดมากก่อนการเลือกตัวอย่าง

วิธีการสุ่มตัวอย่างทั้ง 5 วิธี ข้างต้นนั้นเป็นวิธีการสุ่มตัวอย่างแบบสุ่ม (Random Sampling) ซึ่งใช้กฎความน่าจะเป็นมาประยุกต์กับวิธีการเลือก นอกจากนี้ยังมีวิธีการสุ่มตัวอย่างที่นิยมใช้กันอย่างกว้างขวาง คือ การสุ่มตัวอย่างแบบไม่สุ่ม (Nonrandom Sampling) เป็นการเลือกตัวอย่างโดยอาศัย การพิจารณาของผู้ชำนาญความสะดวกสบาย หรือเกณฑ์อื่น ๆ เป็นต้น ซึ่งอาศัยพื้นฐานที่ต่างจากกฎความน่าจะเป็น การเลือกตัวอย่างแบบไม่สุ่มมีดังนี้

ก. การเลือกตัวอย่างแบบโควตา (Quota Sampling) เป็นการเลือกหน่วยตัวอย่างโดยไม่สนใจว่าตัวอย่างที่เลือกมานั้นจะเลือกมาโดยวิธีไหน เพียงแต่ให้มีจำนวนหน่วยครบตามที่กำหนดไว้ในแต่ละโควตาเท่านั้น ตัวอย่างเช่น ต้องการสำรวจขนาดครอบครัว โดยกำหนดว่าจะทำการสอบถาม 20 ครอบครัวที่มีคุณลักษณะตามที่ต้องการ เช่น แต่งงานแล้ว อยู่ด้วยกันมีบุตร เป็นต้น เมื่อพบครอบครัวที่มีลักษณะดังกล่าวก็ทำการเก็บข้อมูลจากครอบครัวนั้น จนกระทั่งได้ครบตามโควตาคือ 20 ครอบครัว

ข. การเลือกตัวอย่างเชิงพินิจพิจารณาหรือแบบมีจุดหมาย (Judgement or Purposive Sampling) เป็นวิธีการเลือกหน่วยตัวอย่างโดยอาศัยการพิจารณาของผู้สุ่มตัวอย่างเองว่า จะเลือกหน่วยไหนมาเป็นตัวอย่าง การเลือกวิธีนี้อาจจะใช้สำหรับทดสอบคำถามหรือศึกษาแนวทาง

ค. การเลือกตัวอย่างแบบใช้ความสะดวก (Convenience Sampling) เป็นวิธีการเลือกหน่วยตัวอย่างโดยอาศัยความสะดวกสบายของผู้สุ่มเอง

หลักในการพิจารณาเลือกวิธีการสุ่มตัวอย่างว่าจะใช้วิธีใดนั้น ผู้สำรวจจะต้องพิจารณาลักษณะของงานสำรวจ ความถูกต้องแม่นยำของค่าประมาณ ค่าใช้จ่ายในการดำเนินงาน ซึ่งมักจะมีอยู่จำกัด และการควบคุมด้านบริหารจึงจำเป็นต้องพิจารณาหลายอย่างพร้อม ๆ กัน แล้วเลือกวิธีการสุ่มตัวอย่างที่ให้ผลตอบแทนที่สูงที่สุด ตามข้อจำกัดที่เรามีอยู่

### 3. การทะเบียน (Registration)

เป็นการเก็บรวบรวมข้อมูลจากแหล่งที่มีการบันทึกข้อมูลไว้แล้ว ข้อมูลจากทะเบียนบางประเภทสมบูรณ์และทันสมัยดี แต่บางประเภทไม่สมบูรณ์และไม่ทันสมัย ซึ่งถ้าการบันทึกลงทะเบียนไม่สมบูรณ์ ผิดพลาดก็จะมีผลกระทบกระเทือนต่อการวิเคราะห์และสรุปผลได้

การเก็บรวบรวมข้อมูล ไม่ว่าจะเป็นวิธีสำมะโน การสำรวจด้วยตัวอย่าง หรือ การทะเบียน ข้อมูลที่เราได้มานั้นส่วนใหญ่จะไดมาจาก

#### ก. การสัมภาษณ์ (Interview)

โดยการส่งพนักงานที่ได้รับการฝึกอบรมแล้วไปสัมภาษณ์ผู้ซึ่งมีข้อมูลที่เรา ต้องการ โดยใช้แบบสอบถาม (Questionnaire)

แบบสอบถามเป็นเครื่องมือที่ใช้ในการบันทึกข้อมูลจากประชากร หรือจาก ตัวอย่างที่เป็นแหล่งข้อมูล รูปแบบของคำถามที่ปรากฏในแบบสอบถามจะมี 2 แบบใหญ่ ๆ ด้วยกันคือ

1. คำถามแบบปลายปิด (Close-ended question) คือคำถามแบบกำหนดรายการ คำตอบตอบทั้งหมดไว้ให้เลือกตอบ

2. คำถามแบบปลายเปิด (open-ended question) คือคำถามแบบที่กำหนด รายการคำตอบเกือบทั้งหมด เหลือเพียงบางส่วนไว้เพื่อให้ผู้ตามเลือกตอบตามคำตอบ ที่กำหนดไว้ให้และตอบเป็นอย่างอื่นที่นอกเหนือจากคำตอบที่กำหนดไว้ให้ก็ได้ หรือ เป็นคำถามที่ไม่ได้กำหนดคำตอบไว้ให้ เพื่อให้ผู้ตอบมีอิสระมากที่สุดในการที่จะแสดง ความคิดเห็นของตนเองในการตอบคำถาม

ต่ออย่างไรก็ตาม ว่าจะสร้างคำถามในรูปแบบ ภาษาและสำนวนที่ใช้ใน คำถามจะต้องเข้าใจง่าย ไม่กำกวมหรือใช้ศัพท์ทางเทคนิคมากเกินไป และคำถามจะ ต้องไม่มากจนเกินไปเพราะจะทำให้ผู้ตอบเบื่อหน่ายที่จะตอบคำถาม นอกจากนี้คำถาม นั้นจะต้องมีลักษณะที่ง่ายต่อการที่จะนำไปวิเคราะห์ทางสถิติต่อไปด้วย

#### ข. การส่งแบบสอบถามให้กรอก

เป็นการส่งแบบสอบถาม หรือข้อความที่ต้องการไปให้ผู้ตอบกรอกข้อมูล ลงในแบบสอบถาม หรือแบบฟอร์มที่ต้องการแล้วให้ผู้กรอกส่งกลับไปทางไปรษณีย์ หรือผู้ส่งแบบสอบถามไปรับเองก็ได้ ซึ่งในกรณีนี้คำถามที่อยู่ในแบบสอบถามหรือใน แบบฟอร์มจะต้องเป็นคำถามง่าย ๆ และคำถามมีไม่มากนัก

### 3.2 การนำเสนอข้อมูล

เป็นการแสดงข้อมูลที่รวบรวมมาแล้วในข้อ 3.1 ซึ่งเป็นข้อมูลที่เรียกว่าข้อมูล ดิบ (Raw data) ให้คนทั่ว ๆ ไปเข้าใจ โดยนำมาจัดระเบียบ และเสนอข้อมูลให้อยู่ในรูป ที่น่าสนใจ เพื่อเตรียมพร้อมที่จะนำเอาข้อมูลเหล่านั้นไปวิเคราะห์



การนำเสนอข้อมูลมีหลายแบบด้วยกันคือ

1. การเสนอในรูปแบบของบทความ
2. การเสนอเป็นตาราง
3. การเสนอเป็นกราฟเส้น
4. การเสนอเป็นกราฟแท่ง
5. การเสนอเป็นแผนภาพวงกลม
6. การเสนอเป็นแบบรูปภาพ
7. การเสนอเป็นแบบฮิสโตแกรม
8. การเสนอเป็นแบบตารางแจกแจงความถี่

ซึ่งการที่จะเสนอแบบใดจึงจะเหมาะสมหรือดีนั้นมีวิธีการพิจารณาดังนี้

1. อ่านเข้าใจง่าย
2. ช่วยให้สามารถเข้าใจความหมายของข้อมูลนั้นได้ดี
3. ใช้ได้เหมาะสมกับข้อมูลแบบต่าง ๆ
4. สะดวกในการวิเคราะห์
5. ช่วยให้เข้าใจผลของการศึกษาได้ถูกต้องละเอียดและมีประสิทธิภาพ

### 3.2.1 การเสนอในรูปแบบของบทความ

การนำเสนอแบบนี้มักจะปรากฏในรายการวิทยุ หรือรายงานต่าง ๆ เป็นบทความสั้น ๆ และมียอดตัวเลขผสมอยู่ด้วยเช่น ในปีพ.ศ. 2524 มีนักศึกษาที่จบจากมหาวิทยาลัยรามคำแหง.....คน เป็นคณะวิทยาศาสตร์.....คน, คณะรัฐศาสตร์.....คน เป็นต้น

### 3.2.2 การเสนอเป็นตาราง

เพื่อให้ดูง่าย และเข้าใจง่าย อาจจัดเป็นรูปตารางเช่น

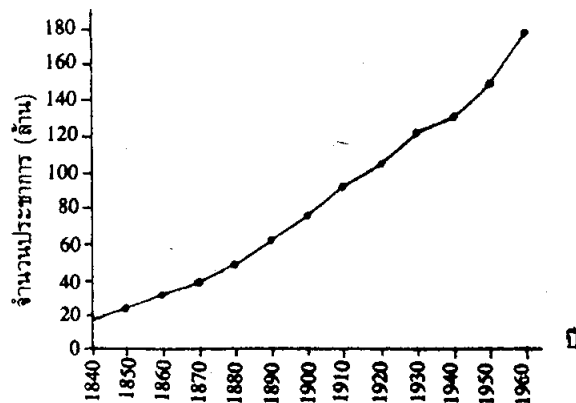
อายุ	จำนวนคนงานที่เป็นผู้ชาย
14-19 ปี	218,000
20-24 ปี	313,000
25-54 ปี	977,000
55 ปีและมากกว่า	255,000
รวม	1,763,000

แหล่งที่มา : United States Department of Labour, January 1965

การนำเสนอแบบตารางนี้ ควรจะมีชื่อตารางและคำอธิบายของแต่ละช่องในตารางอย่างละเอียด ซึ่งถ้ามีหน่วยก็ควรใส่ไว้ด้วย นอกจากนี้ควรใส่แหล่งที่มาของข้อมูลด้วยเพื่อความสะดวกแก่ผู้ที่ต้องการตรวจสอบ

### 3.2.3 การเสนอเป็นกราฟเส้น

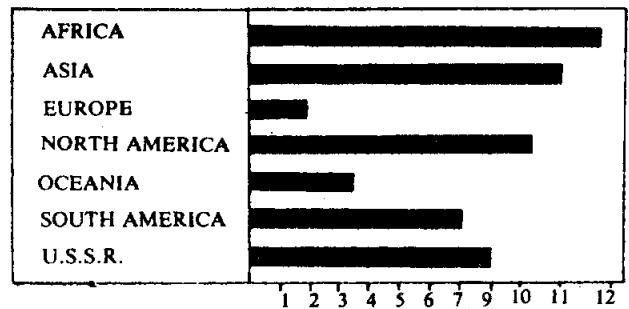
การเสนอโดยวิธีนี้ช่วยให้มองเห็นแนวโน้มของการเพิ่มหรือลดได้ชัดเจน เช่น แนวโน้มของการขาย ค่าใช้จ่ายทางการศึกษา เป็นต้น ดังรูป



กราฟแสดงจำนวนประชากรของประเทศสหรัฐอเมริกา (เป็นล้าน) ในปี 1840, 1850, ..., 1860

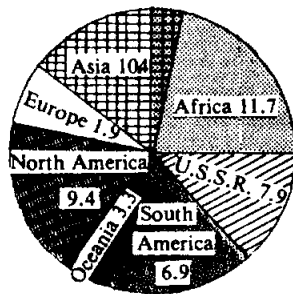
### 3.2.4 การเสนอเป็นกราฟแท่ง








เช่นเดียวกับกราฟเส้นใช้ได้ผลดีเมื่อต้องการให้เห็นแนวโน้มทั่วไปในระยะเวลานึง เช่นกราฟแท่งที่แสดงพื้นที่ของทวีปต่าง ๆ ดังรูป



### 3.2.5 การเสนอเป็นแผนภาพวงกลม หรือเรียกอีกอย่างว่าแผนภาพกึ่ง (Pie Diagram)

แสดงด้วยวงกลมโดยแบ่งออกเป็นส่วนย่อย ๆ ต่าง ๆ กัน การสร้างแผนภาพกึ่ง เราจะแบ่งจากจุดศูนย์กลางของวงกลมโดยเทียบให้ 360° เป็น 100% แล้วคำนวณหาว่าเปอร์เซ็นต์ที่ได้แต่ละเปอร์เซ็นต์จะเป็นมุมเท่าใด แล้วจึงลากแบ่งเป็นส่วนย่อย ๆ และระบายสีแต่ละส่วนให้แตกต่างกัน หรือใช้เครื่องหมายให้แตกต่างกันและเขียนคำอธิบายไว้ข้าง ๆ เช่น แผนภาพวงกลมแสดงพื้นที่ของทวีปต่าง ๆ ในโลก



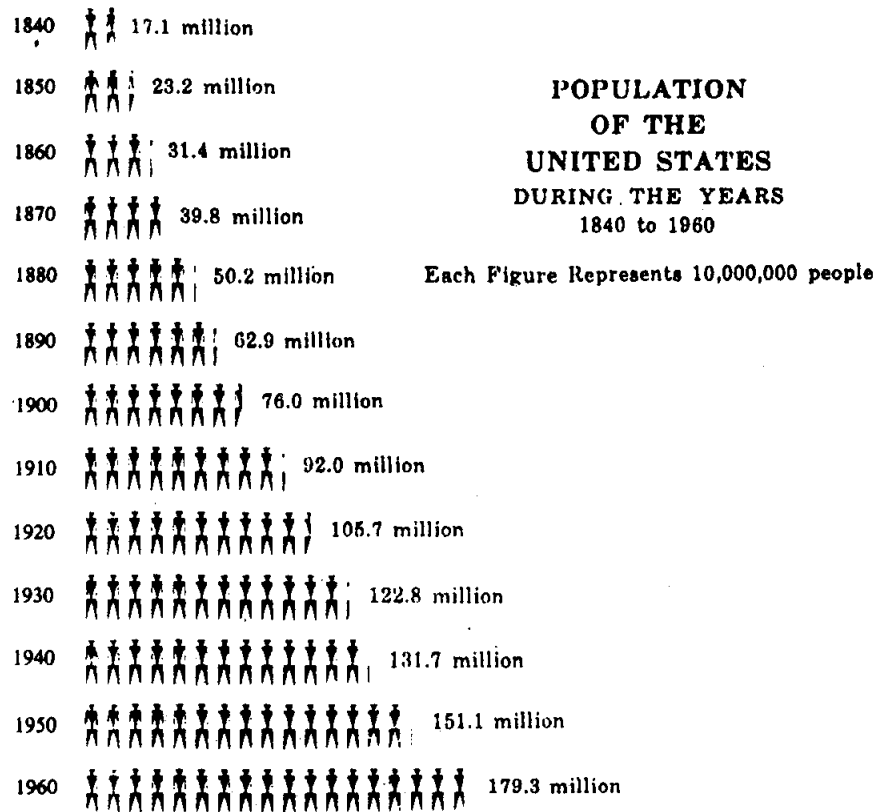
-  = Asia
-  = Africa
-  = u.S.S.R.
-  = South America
-  = Oceania
-  = North America
-  = Europe

แผนภาพวงกลมนี้มักจะใช้เมื่อต้องการเน้นให้เห็นขนาดของส่วนประกอบแต่ละอย่างต่อทั้งหมด

### 3.2.6 การเสนอเป็นแบบรูปภาพ (Pictographs)

เพื่อให้การนำเสนอข้อมูลดูง่าย และเข้าใจง่ายขึ้นอาจใช้รูปภาพแทนได้ ส่วนการที่จะออกแบบภาพแสดงเป็นรูปใดนั้นขึ้นอยู่กับสิ่งที่เกี่ยวข้องในข้อมูลนั้น เช่น ถ้า

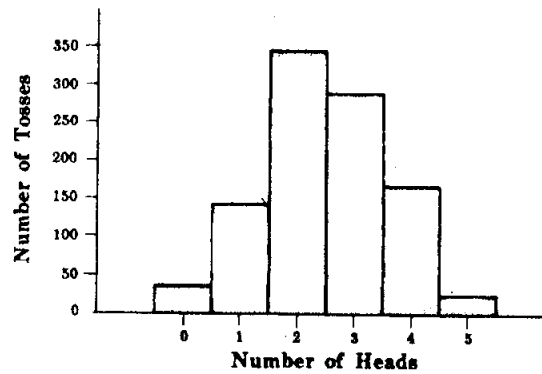
เกี่ยวข้องกับจำนวนคนก็ใช้ภาพเป็นรูปคนถ้าเกี่ยวข้องกับรูปสัตว์ก็ใช้รูปสัตว์เช่น แผนภาพแสดงจำนวนประชากรของประเทศสหรัฐอเมริการะหว่างปี คศ. 1840 ถึง 1960 แต่ละภาพแทนประชากร 10,000,000 คน ดังรูป



### 3.2.7 การเสนอเป็นแบบฮิสโตแกรม (Histogram)

เป็นการจัดแสดงข้อมูลที่ใช้มากที่สุดและใช้กับข้อมูลที่ต่อเนื่องก็ได้ ในกรณีที่ได้จัดข้อมูลเป็นการแจกแจงความถี่ตามช่วงของค่าสังเกตที่ได้จัดขึ้นแล้ว การแสดงแบบฮิสโตแกรม ช่วยให้เข้าใจความหมายได้ง่าย

ลักษณะของฮิสโตแกรมจะเหมือนแผนภาพแท่งแต่แท่งวางเรียงชิดติดกัน เพราะแสดงข้อมูลต่อเนื่อง เช่น



### 3.2.8 การเสนอแบบตารางแจกแจงความถี่

ตารางแจกแจงความถี่เป็นตารางที่แสดงจำนวนข้อมูลหรือข้อสังเกตที่เกิดขึ้นในแต่ละชั้นหรือกลุ่มตารางการแจกแจงความถี่นี้จะสรุปข้อมูลโดยการรวมกลุ่ม ซึ่งตารางนี้จะทำให้เห็นคุณลักษณะทั่ว ๆ ไปของข้อมูลชุดนั้นด้วยการสร้างตารางแจกแจงความถี่มีขั้นตอนดังนี้

1. หาพิสัย (Range) โดยการพิจารณาข้อมูลตัวที่มากที่สุดกับตัวที่น้อยที่สุด แล้วหาพิสัย จากสูตร

$$\text{พิสัย} = \text{ข้อมูลตัวที่มากที่สุด} - \text{ข้อมูลตัวที่น้อยที่สุด}$$

2. หาจำนวนชั้น (class) จากสูตร

$$\text{จำนวนชั้น} = 1 + 3.3 \log_{10} (N)$$

$$N = \text{จำนวนข้อมูลทั้งหมด}$$

3. หาช่วงระหว่างชั้น (class interval) คือหาความกว้างของแต่ละชั้นโดยใช้สูตร

$$\text{class interval} = \frac{\text{พิสัย}}{\text{จำนวนชั้น}}$$

4. หาขีดจำกัดล่าง (Lower class limit) โดยใช้กะเอาเอง คือให้ต่ำกว่าค่าต่ำสุดเล็กน้อย หรือใช้สูตร

$$\text{ขีดจำกัดล่างของชั้นแรก} = \text{ข้อมูลตัวที่น้อยที่สุด} -$$

$$\frac{1}{2} [(\text{จำนวนชั้น} \times \text{ช่วงระหว่างชั้น}) - \text{พิสัย}]$$

$$\text{ขีดจำกัดบนของชั้นล่าง} = \text{ขีดจำกัดล่างของชั้นล่าง} + \text{ช่วงระหว่างชั้น} \\ (\text{class interval})$$

5. หาจำนวนข้อมูล หรือค่าสังเกตที่ตกอยู่ในแต่ละชั้น คือหาความถี่ของแต่ละชั้นนั่นเอง การหาความถี่ใช้รอยขีด (tally) หรือแผ่นคะแนน (Score sheet)

ตัวอย่าง ข้อมูลที่ได้เป็นดังนี้ จงสร้างตารางแจกแจงความถี่

8.25	8.75	1.45	3.60	8.00
8.95	<u>14.95</u>	4.75	6.95	2.75
7.25	6.70	1.50	6.50	10.50
10.50	7.50	3.40	8.75	5.95
9.95	8.25	5.40	8.75	7.95
7.50	11.00	7.00	3.25	14.75
4.75	1.75	1.75	<u>1.40</u>	1.65
7.95	3.50	9.50	7.50	12.50

$$\begin{aligned} 1. \text{พิสัย} &= 14.95 - 1.40 \\ &= 13.55 \end{aligned}$$

$$\begin{aligned} 2. \text{จำนวนชั้น} &= 1 + 3.3 \log_{10} (N) \\ &= 1 + 3.3 \log_{10} (40) \\ &= 6.30 \approx 7 \text{ ชั้น} \end{aligned}$$

$$\therefore \text{จำนวนชั้น} = 7$$

$$3. \text{ช่องระหว่างชั้น} = \frac{\text{พิสัย}}{\text{จำนวนชั้น}} = \frac{13.55}{7}$$

$$4. \text{ขีดจำกัดล่างของชั้นแรก ให้เป็น } 1.00 - 2.99$$

5. พิจารณาจำนวนข้อมูลของแต่ละชั้น ซึ่งจะได้ตารางแจกแจงความถี่ดังนี้

ขีดจำกัด	ค่ากึ่งกลาง ( $x_i$ )	รอยขีด (tally)	ความถี่ $f_i$
1.00 – 2.99	1.995	//////	7
3.00 – 4.99	3.995	//////	6
5.00 – 6.99	5.995	////	5
7.00 – 8.99	7.995	//// // ////	14
9.00 – 10.99	9.995	////	4
11.00 – 12.99	11.995	//	2
13.00 – 14.99	13.995	//	2
รวม			40

แบบฝึกหัด

1. ค่าแรง(รายวัน)ของพนักงาน 50 คนในบริษัทแห่งหนึ่ง ซึ่งได้รับค่าแรงต่าง ๆ กัน ดังนี้

คนที่	ค่าแรง	คนที่	ค่าแรง	คนที่	ค่าแรง	คนที่	ค่าแรง	คนที่	ค่าแรง
1	92	11	51	21	90	31	56	41	59
2	67	12	74	22	96	32	73	42	80
3	92	13	68	23	87	33	79	43	94
4	68	14	65	24	93	34	90	44	80
5	87	15	92	25	79	35	95	45	89
6	74	16	96	26	85	36	93	46	72
7	49	17	77	27	65	37	70	47	62
8	76	18	87	28	74	38	63	48	94
9	86	19	86	29	69	39	88	49	89
10	97	20	90	30	54	40	83	50	73

- ก. จงสร้างตารางแจกแจงความถี่  
ข. จงเขียนฮิสโตแกรมของข้อมูลชุดนี้  
ค. จงเขียนกราฟเส้น ของข้อมูลชุดนี้



2. สมมติว่ามีนักศึกษาที่เรียนวิชาสถิติ 100 คน ทำคะแนนสอบได้คะแนนดัง ต่อไปนี้

77 44 49 33 38 33 76 55 68 39  
 44 59 36 55 47 61 53 32 65 51  
 29 41 32 45 83 58 73 47 40 26  
 59 43 66 44 41 25 39 72 37 55  
 34 47 66 53 55 58 49 45 61 41  
 55 92 83 77 45 62 45 36 78 48  
 54 50 51 66 80 73 57 61 56 50  
 45 82 71 48 46 69 38 72 56 64  
 38 45 51 44 41 68 45 92 43 12  
 37 16 44 57 63 71 40 64 57 51

ก. จงเรียงลำดับข้อมูลจากน้อยไปหามาก

ข. สมมติว่าข้อมูลชุดนี้ต้องการแจกแจงความถี่ใหม่ชั้น 10 ชั้น คือ 10.5-19.5, 19.5-28.5, . . . . , 82.5-91.5 และ 91.5-100.5 จงสร้างตารางแจกแจงความถี่

ค. ให้ใช้ข้อมูลข้างต้นนี้ เขียนฮิสโตแกรมและกราฟเส้น

3. โยนเหรียญ 5 เหรียญ 50 ครั้งและจดจำนวนหัวที่ได้จากการโยนแต่ละครั้ง  
 จงสร้างตารางแจกแจงความถี่และเขียนฮิสโตแกรม

4. โยนลูกเต๋า 2 ลูก 50 ครั้งและจดผลบวกของจุดที่ปรากฏขึ้นมาของลูกเต๋าทิ้ง สอง  
 จงสร้างตารางแจกแจงความถี่และเขียนฮิสโตแกรม

5. ความสูง (นิ้ว) ของคนงานหญิง 50 คนในโรงงานแห่งหนึ่งมีดังนี้

คนที่	ส่วนสูง	คนที่	ส่วนสูง	คนที่	ส่วนสูง	คนที่	ส่วนสูง	คนที่	ส่วนสูง
1	65	11	64	21	64	31	63	41	64
2	63	12	65	22	65	32	65	42	64
3	65	13	64	23	64	33	63	43	63
4	63	14	72	24	71	34	70	44	69
5	69	15	68	25	68	35	67	45	67
6	67	16	66	26	66	36	66	46	66
7	53	17	55	27	56	37	57	47	58
8	58	18	57	28	59	38	59	48	60
9	60	19	60	29	61	39	61	49	61
10	61	20	62	30	62	40	62	50	62

ก. จงสร้างตารางแจกแจงความถี่

ข. จงเขียนฮิสโตแกรมและกราฟเส้น

6. จงสร้างตารางแจกแจงความถี่จากจำนวนอุบัติเหตุ (ต่อวัน) ใน 50 วัน  
ของถนนสายหนึ่งดังนี้

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

7. จากข้อมูลแสดงจำนวนคำที่พิมพ์ผิด (ใน 1 หน้า) ของพนักงานพิมพ์ คนหนึ่งมีดังนี้

9	7	4	3	6	5	8	2	3	6	2	3	0	3	0	2	1	3	1	5
11	7	4	2	3	2	4	7	3	2	1	3	2	1	0	1	2	2	2	3
3	2	5	4	3	6	2	8	2	3	4	1	2	1	6	1	3	2	1	1

จงสร้างตารางแจกแจงความถี่