

หน่วยที่ 7 การจัดการข้อมูลการวิจัย

การจัดการข้อมูล หมายถึง การนำข้อมูลในแบบสอบถามหรือแบบทดสอบ หรือแบบบันทึกข้อมูลที่รวบรวมมาได้ มาสร้างรหัสเป็นตัวเลขตามความต้องการเพื่อการวิเคราะห์ด้วยมือหรือด้วยคอมพิวเตอร์

ความสำคัญของการจัดการข้อมูล คือ การจัดการข้อมูลที่ดีจะช่วยให้การวิเคราะห์ข้อมูลเกิดความรวดเร็ว ถูกต้อง ครบถ้วน ประหยัดเวลาและงบประมาณ

ลักษณะการจัดการข้อมูล จะมี 2 ลักษณะคือ

1. **การจัดการข้อมูลจำนวนน้อย** จากตัวอย่างที่มีจำนวนไม่มาก วิเคราะห์ด้วยมือหรือเครื่องคำนวณธรรมดา
2. **การจัดการข้อมูลจำนวนมาก** ในแบบสอบถามหรือแบบบันทึกที่ได้มาจากตัวอย่างที่มีจำนวนมากหรือการวิเคราะห์ข้อมูลที่มีการใช้สถิติทดสอบสมมุติฐานหรือตอบตามวัตถุประสงค์ ที่ใช้เก็บข้อมูลสำหรับการวิเคราะห์ด้วยเครื่องคอมพิวเตอร์ สื่อที่ใช้ในการบันทึกข้อมูลได้แก่ คิสเก็ตต์ เทปแม่เหล็ก คิสก์ ซีดี Compact Flash, Handy drive, Removable disk

แนวทางการจัดการข้อมูล มี 2 แนวทางคือ

1. **การจัดการข้อมูลกรณีเก็บข้อมูลครั้งเดียว** เช่น ข้อมูลที่ได้จากการสำรวจครั้งเดียว กลุ่มตัวอย่างจะถูกสัมภาษณ์หรือกรอกข้อมูลเพียงครั้งเดียว
2. **การจัดการข้อมูลกรณีเก็บข้อมูลหลายครั้ง** เช่น ข้อมูลจากการทดลอง มีการเก็บข้อมูลหลายครั้ง หรือก่อนและหลังการทดลอง

หลักการและขั้นตอนการจัดการข้อมูล

หลักการในการจัดการข้อมูล ได้แก่ การนำข้อมูลแต่ละชุดที่ได้มาจากตัวอย่างที่ทำการสำรวจ หรือทดลองนั้นมาตรวจสอบความครบถ้วน ความถูกต้อง ข้อมูลชุดใดที่มีข้อมูลที่มีข้อมูลขาดหายไปจำนวนมากอาจพิจารณาตัดทิ้งไป ทั้งนี้เพื่อการวิเคราะห์เป็นไปอย่างรวดเร็ว

ขั้นตอนการจัดการข้อมูล โดยทั่วไปมีอยู่ 3 ขั้นตอน ด้านการวิเคราะห์ด้วยเครื่องคอมพิวเตอร์คือ การตรวจสอบข้อมูลเบื้องต้น หรือการบรรณาธิการ การสร้างรหัสข้อมูล การบันทึกข้อมูล

1 การตรวจสอบข้อมูลเบื้องต้น ได้แก่ การตรวจสอบ แก้ไข แบบสอบถามหรือแบบทดสอบหรือแบบวัดต่าง ๆ ก่อนที่จะให้รหัส (code) มีอยู่ 3 ประการคือ

1. **ความครบถ้วนสมบูรณ์ (completeness)** ตรวจสอบความเรียบร้อยสมบูรณ์ของคำถามทั้งหมดที่ได้รับมา ข้อใดไม่ตอบต้องมีเหตุผลว่าทำไมไม่ตอบ เช่น ท่านเคยรับประทานอาหารที่โรงพยาบาลแห่งนี้หรือไม่ ถ้าตอบว่าไม่เคย ต้องไม่ตอบข้อต่อไปว่าอาหารที่นี่รสชาติเป็นอย่างไร

2. **ความสอดคล้องของคำตอบ (consistency)** คือการตรวจสอบคำถามและคำตอบที่ได้รับมาต้องมีความสอดคล้องกัน เช่น คำถามเกี่ยวกับความสูง แต่คำตอบที่ได้รับเป็นเรื่องเกี่ยวกับน้ำหนัก

3. **คุณลักษณะของตัวอย่าง (characteristics of sample)** เป็นคุณลักษณะส่วนบุคคลของตัวอย่าง เช่น เพศ อายุ ลำดับที่ของบุคคลในครัวเรือน ควรตรวจสอบคุณลักษณะความถูกต้องของหน่วยแฉ่งนับ เช่น คุณลักษณะของตัวอย่างคือบุตรคนโต แต่ไปเก็บข้อมูลบ้านที่ไม่มีบุตรหรือ บุตรคนเล็ก

2. การสร้างรหัสข้อมูล ข้อมูลที่ได้จากการเก็บรวบรวม หรือเรียกว่าข้อมูลดิบ (raw data) เมื่อผ่านการตรวจสอบในขั้นแรกมาแล้ว จะนำมาให้รหัสตามความเหมาะสมของข้อมูลแต่ละชนิด มีการแยกประเภทของข้อมูล เป็นสิ่งสำคัญต่อการต่อการวิเคราะห์และแปลความหมายข้อมูล

1. การแบ่งแยกประเภทข้อมูลเป็นประเภทหรือพวกหรือกลุ่มนั้น จำแนก มีลักษณะอย่างน้อย 4 ประการ คือ

- ความครอบคลุมของข้อมูล เช่นแบ่งระดับการศึกษา อนุบาล ประถม มัธยม
- การแบ่งแยกจากกันโดยเด็ดขาด เช่น แบ่งกลุ่มอายุนักเรียน 5-9, 10-14, 15-19, 20-24 ปี ไม่ควรที่จะแบ่งปี 5-10 , 10-15
- ขนาดของข้อมูลในการจัดกลุ่ม กลุ่มพวก หรือประเภทที่แบ่งต้องไม่กว้างเกินไป เช่น แบ่งอาชีพของประชาชน
- ความสอดคล้องกับวัตถุประสงค์ของการวิจัย

2. การให้รหัสของข้อมูล เป็นการเปลี่ยนข้อมูลซึ่งอาจอยู่ในรูปของข้อความและตัวเลขต่าง ๆ ผสมกัน โดยเปลี่ยนให้เป็นตัวเลขในแบบฟอร์มเดียวกันทั้งหมดสำหรับนำไปวิเคราะห์ด้วยเครื่องคิดเลขหรือเครื่องคอมพิวเตอร์

1. ประเภทของข้อมูลที่จะลงรหัส โดยทั่วไปมี 3 แบบ

ก. ข้อมูลที่ใช้จำแนก เป็นข้อมูลที่ใช้รหัสเป็นตัวแทนหมายเลขในการกำหนดคุณลักษณะของสิ่งตัวอย่างในการวิจัย ซึ่งรวมถึงพื้นที่ที่ใช้สัมภาษณ์ หรือขึ้นอยู่กับขนาดโครงการวิจัย ช่วยให้ติดตามค้นหาตำแหน่งแหล่งที่ของสิ่งตัวอย่างนั้นได้อย่างรวดเร็ว

ข. ข้อมูลเชิงปริมาณ เป็นข้อมูลที่มีระดับการวัดเป็นช่วงมาตรา หรืออัตราส่วนมาตรา ได้มาจากการ ชั่ง ตวง วัด เช่น น้ำหนัก ปริมาตร อายุ ความสูง ความเร็วของแสง ปริมาณคลอรีนในน้ำดื่ม พลังงานของแสง

ค. ข้อมูลเชิงคุณภาพ เช่นเพศ อาชีพ ศาสนา สี หรืออันดับมาตรา เช่น มาก ปานกลาง น้อย ได้มาจากการนับ หรือการกำหนดน้ำหนักของคะแนน (weighting score)

2. แนวทางการให้รหัสข้อมูล โดยทั่วไปมี 4 แบบใหญ่ คือ

ก. Fact code เป็นรหัสเกี่ยวกับข้อมูลที่เกี่ยวข้องกับข้อเท็จจริง ให้คำตอบแต่ละคำตอบแยกจากกัน จะสะดวกและง่ายในการลงรหัส เช่น เมื่อเจ็บป่วย ไปโรงพยาบาลให้รหัส 1 ไม่ไปโรงพยาบาลให้รหัส 0 ไปคลินิก 1 ไม่ไปคลินิก 0

ข. Field code เป็นรหัสที่ใช้กับข้อมูลเชิงปริมาณ เช่น อายุ น้ำหนัก ส่วนสูง รายได้ ภาษี กำไร นักวิจัยสามารถลงรหัสได้ 2 กรณี คือ ลงจำนวนที่แท้จริง หรือ ลงรหัสเฉพาะเลขจำนวนเต็ม เช่น มีรายได้ 6,400 บาท ลงรหัส 6 หรือ น้ำหนัก 3,500 กรัม ใช้รหัส 0350

ค. Interval code เป็นรหัสที่มีการจัดกลุ่มข้อมูลตามลักษณะที่ต้องการวิเคราะห์ คือ

1) ข้อมูลเชิงปริมาณจะถูกจัดกลุ่มใหม่ เป็นการลดระดับการวัดให้ต่ำลง แล้วจึงให้รหัส เช่นรายได้ของครอบครัว

รายได้ของครอบครัว (บาท/เดือน) ของผู้มีอาชีพ	
รายได้ (บาท/เดือน)	รหัส
	00
1,000 – 4,900	01

2) การให้รหัสข้อมูลที่แสดงคุณลักษณะยังคงใช้ 0 ถึง 9 รหัสที่ใช้จะไม่มี ความหมายในทางคณิตศาสตร์ คือไม่สามารถบวก ลบ คูณหารได้ เพียงแต่แสดงความถี่ ฐานนิยม และใช้วิเคราะห์ทางสถิติที่ใช้แบบกลุ่มเชิงคุณภาพเท่านั้น

ตัวแปร	ข้อมูล	รหัส
ศาสนา	พุทธ	1
	คริสต์	2
	อิสลาม	3
	อื่นๆ ระบุ _____	4
ตำแหน่งที่ปฏิบัติงาน	หัวหน้าตึก	1
	พยาบาลวิชาชีพ	2
	พยาบาลเทคนิค	3
	ผู้ช่วยเหลือคนไข้	4

ง. Pattern code เป็นการให้รายละเอียดในรูปแบบผสมหลาย ๆ คำตอบ ไม่จำเป็นต้องพิจารณาคำตอบจากคำถามเพียงข้อเดียว เช่นในเรื่องเกี่ยวกับความรู้ ทักษะ และ การปฏิบัติ เช่น ตัวอย่าง คำถามเกี่ยวกับการวางแผนครอบครัว 2 คำถามเป็นทัศนคติและการปฏิบัติต้องการจะผสม

คำตอบ	รหัส
เห็นด้วยและกำลังใช้อยู่	01
เห็นด้วยและตั้งใจจะใช้ในอนาคต	02
เห็นด้วยและไม่แน่ใจว่าจะใช้	03
ไม่เห็นด้วยและคงไม่ใช้	04

และรหัสพิเศษสำหรับข้อมูลที่ไม่ครบถ้วน ไม่ทราบ ไม่ตอบ 77 88 99

3. การบันทึกรหัสข้อมูล การลงรหัสข้อมูลอาจทำไปพร้อม ๆ กับการเก็บข้อมูล โดยผู้ตอบแบบสอบถามเองหรือโดยผู้สัมภาษณ์ หรือผู้เก็บข้อมูลหรือทำโดยพนักงานลงรหัสสำนักงาน ทั้งนี้ให้ยึดคู่มือลงรหัสเป็นแม่แบบตลอดไปจนกว่าจะลงรหัสเสร็จ

การสร้างตารางจำลอง

เมื่อนักวิจัยได้ข้อมูลมาแล้ว หลังจากได้มีการจัดการข้อมูลโดยการตรวจสอบข้อมูลเบื้องต้น เพื่อดูความถูกต้องครบถ้วนสมบูรณ์ของข้อมูลทุกชุด สิ่งที่เราควรกระทำต่อไปได้แก่การสร้างตารางจำลอง โดยอิงวัตถุประสงค์และสมมุติฐานของการวิจัยเป็นหลัก

ตารางจำลอง (dummy table) คือ ตารางที่นักวิจัยสร้างขึ้นเสมือนเป็นตารางที่จะนำเสนอจริงในผลการวิจัย เพียงแต่ขาดตัวเลขที่จะกรอกไปเท่านั้น ช่วยเป็นกรอบในการวิเคราะห์ แปรผล ช่วยให้การศึกษามีข้อผิดพลาดได้น้อย ประหยัดค่าใช้จ่ายให้นักวิเคราะห์ด้วยเครื่องคอมพิวเตอร์ การสร้างตารางจำลองที่ดี ทำให้นักวิจัยทราบว่าข้อมูลอะไรขาดหายไป อาจต้องตามไปเก็บข้อมูลเพิ่มเติมถ้าข้อมูลนั้นมีความสำคัญ

องค์ประกอบของตาราง ตารางโดยทั่วไปจะประกอบด้วย

- 1. หมายเลขของตาราง** เป็นการระบุลำดับที่ของตารางว่าเป็นตารางที่เท่าใด โดยทั่วไปมีมากกว่า 1 ตาราง กำหนดให้อยู่ซ้ายมือ
- 2. ชื่อของตาราง** เป็นข้อความที่อยู่ถัดจากหมายเลขของตารางและอยู่แถวเดียวกัน ชื่อตารางจะต้องเป็นข้อความสั้น ๆ แต่ได้ความกระจ่างชัด เพื่อให้ทราบว่าข้อมูลที่แสดงในตารางเป็นเรื่องอะไร ที่ไหนและเมื่อใด
- 3. หัวเรื่องแถว** เป็นส่วนของตารางที่อธิบายถึงข้อมูลที่ปรากฏอยู่ในแนวนอน
- 4. หัวเรื่องสดมภ์** เป็นส่วนของตารางที่อธิบายถึงข้อมูลที่ปรากฏอยู่ในแนวตั้ง
- 5. ตัวเรื่อง** เป็นส่วนของตารางที่เป็นตัวเลข และตัวเลขเหล่านี้จะจำแนกตามรายละเอียดของหัวเรื่องแถวและหัวเรื่องสดมภ์

ชนิดของตารางจำลอง

1. ตาราง 1 ตัวแปร (univariate tabulation or one-dimensional) เป็นตารางอย่างง่ายที่ประกอบไปด้วยตารางที่มีตัวแปรเดียว ในหัวตารางที่เป็นแถวตั้งจะมีเพียงจำนวนและร้อยละ มักจะนำเสนอคุณลักษณะของตัวอย่างไม่ได้แสดงความสัมพันธ์ระหว่างตัวแปร ดังตัวอย่างที่นำเสนอ เพศ (นามมาตรา) และกลุ่มอายุ (เป็นข้อมูลเชิงปริมาณชนิดอัตราส่วนมาตรา ที่นำมาจัดกลุ่มใหม่) ในกรณีข้อมูลเชิงปริมาณ เมื่อนำมาเสนอในรูปแบบของการจัดกลุ่มนักวิจัยต้องนำเสนอค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน รวมทั้งค่าต่ำสุดค่าสูงสุดของตัวแปรนั้นด้วย ดังตัวอย่าง

ตารางที่ 7.4 จำนวนและร้อยละของผู้ตอบแบบสอบถามเรื่องการวางแผนครอบครัว จำแนกตามเพศและอายุ

ตัวแปร	จำนวน	ร้อยละ
เพศ		
ชาย	300	37.5
หญิง	500	62.5
รวม	800	100
กลุ่มอายุ (ปี)		
20 - 29	200	25.5
30 - 49	250	31.3
50 ขึ้นไป	350	43.7
รวม	800	100
ค่าเฉลี่ย	← ส่วนเบี่ยงเบนมาตรฐาน	41.2 ← 12.7
ค่าต่ำสุด-ค่าสูงสุด		20 - 60

2. ตาราง 2 ตัวแปร (bivariate tabulation or two- dimensional) เป็นตารางที่มีตัวแปร 2 ตัว ในหัวตารางที่เป็นแถวอนและแถวตั้งจะเป็นตัวแปรคนละกลุ่ม เช่น ตัวแปรอิสระหรือตัวแปรตาม มักจะเป็นการวิเคราะห์แบบตารางไขว้ การจะจัดตัวแปรตามหรือตัวแปรอิสระเป็นแถวตั้งหรือแถวอนแล้วแต่ความเหมาะสมของหน้ากระดาษ ตัวอย่าง

ตารางที่ 7.6 จำนวนและร้อยละของผู้ป่วยด้วยโรคซาร์สจำแนกตามประวัติการไปต่างประเทศ

ประวัติการไปต่างประเทศ	การเจ็บป่วยด้วยโรคซาร์ส		จำนวน	รวม	ร้อยละ
	ป่วย	ไม่ป่วย			
เคยไป	80	20	100		100.0
ไม่เคยไป	0	100	100		100.0
รวม	80	120	200		100.0

ตามตาราง ที่ 7.6 มี 2 ตัวแปรคือ การเจ็บป่วยด้วยโรคซาร์สและประวัติการไปต่างประเทศ ตัวแปรทั้งสองแยกชั้นย่อยเป็น 2 ชั้น คือ การเจ็บป่วยแยกเป็นป่วยและไม่ป่วย ประวัติการไปต่างประเทศแยกเป็นเคยไปและไม่เคยไป ตามตารางนี้ถือว่าประวัติการไปต่างประเทศเป็นตัวแปรอิสระ การคำนวณร้อยละใช้ประวัติการไปต่างประเทศเป็นฐานในการคำนวณ แต่การแปรผลจะแปลผลตามตัวแปรตามคือ การเจ็บป่วยด้วยโรคซาร์ส

3. ตารางที่มีตัวแปรมากกว่า 2 ตัวแปร

ตารางที่ 7.7 ความสัมพันธ์ระหว่างอาชีพ ระดับการศึกษาและความพึงพอใจต่อการรับบริการในโรงพยาบาลของผู้มารับบริการ

การศึกษา	เกษตรกร		ข้าราชการ		รวม
	พอใจ	ไม่พอใจ	พอใจ	ไม่พอใจ	
ต่ำกว่าปริญญาตรี	80	20	60	40	200
ปริญญาตรีหรือสูงกว่า	20	80	40	60	200
รวม	100	100	100	100	400

ขั้นตอนการสร้างตารางจำลอง

1. การวางแผนในการวิเคราะห์ข้อมูล มีวิธีแยกตัวแปรดังนี้

- แยกข้อมูลหรือตัวแปรจากคำถามแต่ละข้อให้ออกมาเป็นชนิดของตัวแปร แต่ละตัวแปรมีระดับการวัดอย่างไร เช่น นามมาตรา อันดับมาตรา ช่วงมาตราหรืออัตราส่วนมาตรา
- แยกตัวแปรที่บอกคุณลักษณะทั่วไปกลุ่มตัวอย่าง เช่น เพศ อายุ ศาสนา อาชีพ รายได้
- แยกตัวแปรที่จะต้องวิเคราะห์ข้อมูลเพื่อตอบสมมุติฐานและวัตถุประสงค์ของการวิจัย สมมุติฐานหรือวัตถุประสงค์หนึ่งๆ อาจใช้ตัวแปรมากกว่า 2 ตัวขึ้นไป การเขียนตารางจากคำถามเป็นข้อ ๆ ช่วยให้นักวิจัยจัดกลุ่มได้ครบตามวัตถุประสงค์และสมมุติฐาน และสามารถสร้างตารางจำลอง (dummy table) ได้

คำถามข้อที่	ชื่อตัวแปร	ชนิดของตัวแปร			ระดับการวัด			ตอบ สมมติ	ตอบวัตถุ
		ตัวแปร อิสระ	ตัวแปรตาม	ตัวแปร ธรรมดา	นาม มาตรา	อันดับ มาตรา	อัตราส่วน/ ช่วงมาตรา		
1	เพศ	✓ □			✓			1	1
2	อายุ	✓					✓	1	1
3	ศาสนา			✓	✓				
4	น้ำหนัก	✓					✓	2	2
5	SBP		✓				✓	2	2
6	DBP		✓				✓	2	2

หมายเหตุ SBP = ความดันโลหิตตัวบน DBP = ความดันโลหิตตัวล่าง

การสร้างตารางจำลอง เมื่อวางแผนแล้วให้พิจารณาตามตารางข้างบนจะเห็นว่าตัวแปรที่เป็นตัวแปรอิสระ (อายุ เพศ น้ำหนัก) ตัวแปรธรรมดา (ศาสนา) ที่ไม่เกี่ยวกับการตอบวัตถุประสงค์หรือสมมุติฐานแต่อธิบายลักษณะประชากรที่ศึกษา ในกรณีนี้เพศ และอายุถือว่าเป็นลักษณะของประชากรด้วย มีตัวแปรได้แก่ความดันโลหิตตัวบน ความดันโลหิตตัวล่าง

1. ตารางจำลองที่นำเสนอลักษณะของประชากร เป็นการนำเสนอข้อมูลลักษณะของประชากรที่นักวิจัยศึกษา ตารางที่ 7.8 จำนวนร้อยละของผู้ตอบแบบสอบถาม จำแนกตามข้อมูลทั่วไป

ลักษณะของประชากร		จำนวน	ร้อยละ
เพศ	ชาย		
	หญิง		
	รวม		
กลุ่มอายุ (ปี)	20 - 29		
	30 - 49		
	50 ขึ้นไป		
	รวม		
	ค่าเฉลี่ย ← ส่วนเบี่ยงเบนมาตรฐาน		
	ค่าต่ำสุด - ค่าสูงสุด		
น้ำหนัก (กก.)	40 - 49		
	50 - 59		
	60 - 69		
	70 - 79		
	80 ขึ้นไป		
	รวม		
	ค่าเฉลี่ย ← ส่วนเบี่ยงเบนมาตรฐาน		
ค่าต่ำสุด - ค่าสูงสุด			

2. ตารางจำลองที่นำเสนอผลการวิเคราะห์ข้อมูล เป็นการนำเสนอผลการวิเคราะห์ข้อมูลที่จะตอบวัตถุประสงค์หรือสมมุติฐาน มักจะอยู่ตามหลังตารางที่เป็นลักษณะของประชากรที่ศึกษา

ตารางที่ 7.9 ความสัมพันธ์ระหว่างความดันโลหิตกับเพศ

เพศ	ความดันโลหิต		รวม
	ปกติ	สูง	
ชาย			
หญิง			
รวม			
X ² ;df.;p-value			

การจัดการข้อมูลสำหรับการวิเคราะห์ด้วยคอมพิวเตอร์

การใช้คอมพิวเตอร์ในการวิเคราะห์ข้อมูล

องค์ประกอบของระบบคอมพิวเตอร์ ระบบคอมพิวเตอร์ประกอบไปด้วย 5 องค์ประกอบได้แก่

1. **ตัวเครื่องหรือฮาร์ดแวร์ (hard wares)** คือ ส่วนที่เป็นเครื่องคอมพิวเตอร์ทางกายภาพ หรือสามารถสัมผัสได้ แบ่งเป็น 3 ส่วน
- **หน่วยรับข้อมูล (input devices)** ออกแบบเพื่อป้อนคำสั่ง ข้อมูล ข้อมูลจากการกดปุ่มจากแป้นพิมพ์ เมาส์ เครื่องอ่านฟลอปปี ดิสก์ เครื่องอ่านเทปแม่เหล็ก เครื่องสแกน เครื่องอ่านสื่อ อื่นๆ
- **หน่วยประมวลผล (processor units)** เป็นหัวใจของคอมพิวเตอร์ ทำหน้าที่รับข้อมูลจากหน่วยรับข้อมูลแล้วประมวลผลตามคำสั่ง ให้ผลลัพธ์ส่งต่อไป แบ่งเป็น 3 ส่วน คือ หน่วยความจำหลัก หน่วยคำนวณและตรรกะ หน่วยควบคุม ส่วนของหน่วยคำนวณและตรรกะและหน่วยควบคุมเรียกว่า ประมวลผลกลาง (central processing unit) CPU
- **อุปกรณ์ส่งข้อมูล (output devices)** เป็นอุปกรณ์ที่แสดงผลลัพธ์ทำหน้าที่ตรงข้ามกับอุปกรณ์รับข้อมูล เช่น จอภาพ เครื่องพิมพ์ สื่อต่าง ๆ ที่ใช้กับคอมพิวเตอร์
2. **ส่วนของโปรแกรมหรือซอฟต์แวร์ (Softwares)** โปรแกรมหรือชุดคำสั่ง
3. **ข้อมูล (Data)** รายละเอียดความจริง อาจอยู่ในรูป ตัวเลข ตัวหนังสือ หรือรูปภาพ
4. **กระบวนการ (procedures)** หมายถึง ขั้นตอนที่จะบอกว่าคุณต้องทำอะไร อย่างไร สุดท้ายผลลัพธ์ที่ต้องการคืออะไร
5. **บุคลากร (people)** หมายถึง ผู้ปฏิบัติงานตามกระบวนการในกิจกรรมต่าง ๆ

ความสำคัญของคอมพิวเตอร์ในการวิเคราะห์ข้อมูล

คอมพิวเตอร์ช่วยให้ข้อมูลที่มีจำนวนมากและต้องการเวลามากในการวิเคราะห์สำเร็จลงได้อย่างรวดเร็วและแม่นยำ

ขั้นตอนการนำคอมพิวเตอร์มาใช้ในการวิเคราะห์ข้อมูล

1. ก่อนอื่นนักวิจัยควรมีพื้นฐานความรู้เกี่ยวกับเครื่องคอมพิวเตอร์บ้างพอสมควร คอมพิวเตอร์ต้องรองรับโปรแกรมได้
2. โปรแกรมคอมพิวเตอร์สำเร็จรูปบาง โปรแกรมที่ใช้ในการวิเคราะห์ อาจมีความรู้โดยการศึกษาด้วยตนเองหรืออบรมเพิ่มเติม เช่น ระบบดอส (Dos system) โปรแกรมใช้บันทึกข้อมูล เช่น Excel, SYSTAT, Access, EPI INFO, SPSS/FW หรือ WordStar

การบันทึกข้อมูลเพื่อการวิเคราะห์ด้วยคอมพิวเตอร์

เนื่องจากข้อมูลจำนวนมาก และ/หรือต้องใช้สถิติต่าง ๆ ในการวิเคราะห์ข้อมูลเพื่อตอบวัตถุประสงค์หรือสมมุติฐานที่นักวิจัยตั้งไว้ แล้วสรุปผล นักวิจัยจำเป็นต้องใช้คอมพิวเตอร์มาช่วยในการวิเคราะห์ข้อมูล

ความหมายของการบันทึกข้อมูล คือ การนำข้อมูลที่รวบรวมมาได้ และตรวจสอบแล้วว่ามีความถูกต้องครบถ้วน สมบูรณ์ ผ่านการลงรหัสข้อมูลที่สร้างไว้ ในการบันทึกข้อมูลมีขั้นตอนแบ่งเป็น 2 ระยะคือ

1. การบันทึกข้อมูลลงในแบบฟอร์มที่สร้างขึ้นเพื่อกรอกรหัสที่เป็นตัวเลขตามตัวแปรต่าง ๆ
2. การบันทึกข้อมูลลงในคอมพิวเตอร์

ความสำคัญของการบันทึกข้อมูลด้วยคอมพิวเตอร์ คือ อำนวยความสะดวกให้นักวิจัย เพื่อใช้อ่านและวิเคราะห์ แล้วจึงส่งผ่านข้อมูลออกมา ข้อมูลที่บันทึกสามารถแก้ไขข้อมูลได้ แทรกข้อมูลได้

สื่อที่ใช้ในการบันทึกข้อมูลด้วยคอมพิวเตอร์ ได้แก่ การ์ด แผ่นดิสเก็ตต์ เทปแม่เหล็ก ดิสก์หรือจานแม่เหล็ก แผ่นซีดี คอมแพคแฟลช การ์ดมัลติมีเดีย สื่ออื่นๆ เช่น removable disk

โปรแกรมคอมพิวเตอร์ที่ใช้ในการบันทึกข้อมูล

1. Excel เป็นโปรแกรมสำหรับกรอกข้อความ ตัวเลขและรูปภาพลงในช่องตารางที่เรียกว่า Worksheet ใช้ข้อมูลในลักษณะของ database ข้อมูลสามารถนำไปใช้กับโปรแกรมอื่นได้
2. Access เป็นโปรแกรมพัฒนาระบบจัดเก็บข้อมูล ซึ่งมีระบบการค้นหา การคัดเลือก สามารถค้นหาข้อมูลได้ตามเงื่อนไขที่ต้องการโดยไม่เสียเวลาบันทึกข้อมูลซ้ำซ้อน ไม่ต้องเปิดข้อมูลทั้งหมดเหมือน Excel กำหนดเงื่อนไขในการเรียกดูข้อมูลได้
3. WordStar ใช้งานใน Dos

ปัจจุบันมีโปรแกรมที่สร้าง/เตรียมและวิเคราะห์ได้โดยตัวของโปรแกรมสำเร็จรูปนั้น ๆ ที่นิยมใช้ในงานสาธารณสุข เช่น SPSS, EPI INFO (ระบาดวิทยา)

ตัวอย่างการใช้โปรแกรมสำเร็จรูป หน้า 330