

ใบความรู้ที่ 4 เรื่อง ข้อมูลและการเก็บรวบรวมข้อมูล

1. ความหมายของข้อมูล

ข้อมูลเป็นข้อความจริงหรือสิ่งที่บ่งบอกถึงสภาพ สถานการณ์ หรือปรากฏการณ์ใด ปรากฏการณ์หนึ่ง โดยที่ข้อมูลอาจเป็นตัวเลขหรือข้อความก็ได้ เช่น ในปี พ.ศ.2547 คณะรัฐมนตรี กำหนดมาตรการควบคุมการแพร่ระบาดของเชื้อโรคไข้หวัดนก หรือในเดือนกันยายน 2547 น้ำมันเบนซิน 91 จำหน่ายในเขตกรุงเทพและปริมณฑลราคาลิตรละ 20.99 บาท โดยทั่วไป ข้อมูลมักจะอยู่ในรูปตัวเลขซึ่งมีหลาย ๆ จำนวนที่สามารถนำมาเปรียบเทียบขนาดกันได้ เช่น ตัวแต่วันที่ 1 มกราคม ถึง 30 กันยายน 2547 ไทยส่งออกข้าวไปยังประเทศหนึ่งรวม 2.88 ล้านตัน ลดลงจาก 5.00 ล้านตันของการส่งออกในช่วงเวลาเดียวกันของปีก่อนร้อยละ 42.4

ข้อมูลเชิงสถิติเป็นข้อมูลที่สามารถนำมาประมวลผลหรือวิเคราะห์ด้วยกระบวนการหรือวิธีการต่าง ๆ เพื่อตอบคำถามในประเด็นต่าง ๆ ได้

2. ประเภทของข้อมูล

ประเภทของข้อมูลสามารถจำแนกได้จากวิธีการเก็บรวบรวมข้อมูล และจากลักษณะของข้อมูล การจำแนกข้อมูลตามวิธีการเก็บรวบรวมข้อมูล

เมื่อจำแนกประเภทของข้อมูลตามวิธีการเก็บรวบรวมข้อมูลจะแบ่งได้เป็น 2 ประเภทใหญ่ ๆ คือ ข้อมูลปฐมภูมิ (primary data) และข้อมูลทุติยภูมิ (secondary data)

1) **ข้อมูลปฐมภูมิ** คือ ข้อมูลที่ผู้วิจัยจะต้องเก็บรวบรวมจากผู้ให้ข้อมูลหรือแหล่งที่มาของข้อมูลโดยตรง ซึ่งอาจทำได้โดยการสัมภาษณ์ วัด นับ หรือสังเกตจากแหล่งข้อมูลโดยตรง เนื่องจากข้อมูลเหล่านี้ไม่เคยมีผู้ใดเก็บรวบรวมไว้ก่อน

การเก็บรวบรวมข้อมูลประเภทนี้ทำได้ 2 วิธี คือ การสำมะโน (census) และการสำรวจจากกลุ่มตัวอย่าง (sample survey)

(1) การสำมะโน คือ การเก็บรวบรวมข้อมูลจากทุก ๆ หน่วยของประชากร หรือสิ่งที่เราต้องการศึกษา ซึ่งการเก็บข้อมูลในลักษณะนี้ทำให้เสียเวลาและค่าใช้จ่ายในการเก็บรวบรวมข้อมูลมาก การเก็บรวบรวมข้อมูลโดยวิธีนี้จึงไม่ค่อยนิยมใช้ในทางปฏิบัติ ยกเว้นกรณีที่ประชากรมีขนาดเล็กหรือมีขอบเขตไม่กว้างขวางนัก

(2) การสำรวจจากกลุ่มตัวอย่าง คือ การเก็บรวบรวมข้อมูลจากบางหน่วยที่เลือกมาเป็นตัวแทนจากทุก ๆ หน่วยของประชากรหรือสิ่งที่เราต้องการศึกษาเท่านั้น เนื่องจากการเก็บรวบรวมข้อมูลจากทุกหน่วยของประชากร อาจทำให้เสียเวลาและค่าใช้จ่ายโดยไม่จำเป็น เพราะสิ่งที่ต้องการศึกษาอาจจะมีบางกลุ่มที่มีลักษณะที่ต้องการศึกษาอยู่เหมือนกัน ๆ กัน หรือใกล้เคียงกันมาก การ

เลือกตัวอย่างหรือตัวแทนของแต่ละกลุ่มมาทำการศึกษา ก็เป็นการเพียงพอที่จะทำให้สามารถประมาณค่าของสิ่งที่เราต้องการศึกษาทั้งหมดได้ เช่น การสำรวจราคาเฉลี่ยของสินค้าชนิดหนึ่งที่มีขนาดบรรจุใกล้เคียงกันจากร้านค้าปลีกทั่วประเทศ ราคามักจะใกล้เคียงกันด้วย ดังนั้น เราอาจเลือกร้านค้าปลีกเพียงบางร้านมาเป็นตัวแทนจากร้านค้าปลีกทั้งหมดได้ แต่จำนวนร้านค้าปลีกที่เลือกมาเป็นตัวแทนจะมีจำนวนมากหรือน้อยเพียงใดขึ้นอยู่กับความต้องการของผู้เก็บรวบรวมข้อมูลว่าต้องการให้ราคาเฉลี่ยของราคาสินค้าชนิดนั้นที่หาได้จากราคาสินค้าในร้านค้าตัวอย่างที่เลือกขึ้นมาเพื่อเก็บรวบรวมข้อมูลนี้ใกล้เคียงกับค่าที่ควรเป็นจริงซึ่งได้จากการเก็บรวบรวมข้อมูลจากร้านค้าปลีกทุกๆ ร้านมากน้อยเพียงใด ถ้าต้องการให้ได้ผลใกล้เคียงมากก็ควรเลือกตัวอย่างร้านค้าปลีกมาเก็บรวบรวมข้อมูลเป็นจำนวนมาก

2) **ข้อมูลทุติยภูมิ** คือ ข้อมูลที่ผู้ใช้ไม่ต้องเก็บรวบรวมจากผู้ให้ข้อมูลหรือแหล่งที่มาของข้อมูลโดยตรง แต่ได้จากข้อมูลที่มีผู้อื่นเก็บรวบรวมไว้แล้ว ข้อมูลประเภทนี้ ผู้ใช้ไม่ต้องเสียเวลาและค่าใช้จ่ายในการเก็บรวบรวมข้อมูลเอง สามารถนำข้อมูลที่มีผู้อื่นเก็บรวบรวมไว้แล้วมาใช้ได้เลย แต่อย่างไรก็ตามผู้ใช้งานจะต้องระมัดระวังในการนำข้อมูลประเภทนี้มาใช้ให้มาก เนื่องจากมีโอกาสผิดพลาดได้มาก ถ้าผู้เก็บรวบรวมข้อมูลดังกล่าวใช้วิธีเก็บรวบรวมข้อมูลที่ไม่เหมาะสม

แหล่งที่มาของข้อมูลทุติยภูมิที่สำคัญ

(1) รายงานต่าง ๆ ของหน่วยงานราชการและองค์การของรัฐบาล โดยทั่วไป หน่วยงานราชการหรือองค์การของรัฐบาล มักจะมีรายงานแสดงข้อมูลพิมพ์ออกมาเผยแพร่เป็นประจำ ซึ่งอาจเป็นรายงานรายเดือน รายสามเดือน หรือรายปี ข้อมูลที่ได้จากรายงานต่างๆ ของหน่วยงานราชการและองค์การของรัฐบาลนี้อาจถือได้ว่าเป็นที่มาของข้อมูลทุติยภูมิที่สำคัญที่สุด

(2) รายงานและบทความจากหนังสือหรือรายงานจากหน่วยงานเอกชน หน่วยงานของเอกชนบางแห่ง โดยเฉพาะหน่วยงานใหญ่ ๆ จะพิมพ์รายงานเกี่ยวกับผลการดำเนินงานของตนออกเผยแพร่เช่นเดียวกับหน่วยงานของราชการ เช่น รายงานประจำเดือนของธนาคารพาณิชย์ นอกจากนี้หนังสือพิมพ์รายวันหรือสื่ออื่นๆ มักจะมีข้อมูลทุติยภูมิประกอบบทความหรือรายงานด้วย

การจำแนกประเภทของข้อมูลตามลักษณะของข้อมูล

เมื่อจำแนกประเภทของข้อมูลตามลักษณะของข้อมูลจะจำแนกเป็น 2 ลักษณะใหญ่ๆ คือ ข้อมูลเชิงปริมาณ (quantitative data) และข้อมูลเชิงคุณภาพ (qualitative data)

1) **ข้อมูลเชิงปริมาณ** คือ ข้อมูลที่ใช้แทนขนาดหรือปริมาณซึ่งวัดออกมาเป็นจำนวนที่สามารถนำมาใช้เปรียบเทียบกันได้โดยตรง เช่น ปริมาณการผลิตน้ำมันดิบของกลุ่มโอเปกในแต่ละปี อัตราดอกเบี้ยเงินกู้ของธนาคารพาณิชย์ จำนวนสมาชิกโดยเฉลี่ยของครอบครัวไทย

2) **ข้อมูลเชิงคุณภาพ** คือ ข้อมูลที่ไม่สามารถวัดออกมาเป็นจำนวนได้โดยตรง แต่อธิบายลักษณะหรือคุณสมบัติในเชิงคุณภาพได้ เช่น เพศของสมาชิกในครอบครัว สถานภาพสมรสของพนักงานในบริษัทห้างร้าน หรือความคิดเห็นของประชาชน การวิเคราะห์ข้อมูลประเภทนี้ส่วนใหญ่

ทำโดยการนับจำนวน จำแนกตามลักษณะเชิงคุณภาพ เช่น นับจำนวนพนักงานที่เป็นโสด ที่สมรสแล้ว ที่หย่าร้าง และที่เป็นหม้ายว่ามีอย่างละกี่คน ข้อมูลเชิงคุณภาพบางลักษณะสามารถวัดออกมาเป็นลำดับที่หรือตำแหน่งที่ได้ เช่น ความชอบ วัดในรูป ชอบมากที่สุด ชอบมาก ชอบปานกลาง ชอบน้อย ไม่ชอบเลย ความคิดเห็น วัดในรูปเห็นด้วยอย่างยิ่ง เห็นด้วย ไม่มีความเห็น ไม่เห็นด้วย ไม่เห็นด้วยอย่างยิ่ง ประสิทธิภาพในการดำเนินงานวัดในรูป ดีขึ้นมาก ดีขึ้น คงเดิม เลวลง เลวลงมาก หรือวัดในรูป สูง ปานกลาง ต่ำ เป็นต้น การกำหนดลำดับที่หรือตำแหน่งที่ของข้อมูลเชิงคุณภาพนี้เมื่อนำข้อมูลไปวิเคราะห์จะต้องแทนลำดับที่หรือตำแหน่งที่เหล่านี้ด้วยตัวเลข เช่น ให้ตัวเลขที่มีค่ามากใช้แทนลักษณะหรือความรู้สึกที่ดี

ชอบมากที่สุด	หรือ	เห็นด้วยอย่างยิ่ง	แทนด้วย	4
ชอบมาก	หรือ	เห็นด้วย	แทนด้วย	3
ชอบปานกลาง	หรือ	ไม่มีความเห็น	แทนด้วย	2
ชอบน้อย	หรือ	ไม่เห็นด้วย	แทนด้วย	1
ไม่ชอบเลย	หรือ	ไม่เห็นด้วยอย่างยิ่ง	แทนด้วย	0

ในกรณีที่ข้อมูลเชิงคุณภาพใดไม่สามารถวัดออกมาเป็นลำดับที่หรือตำแหน่งที่ได้เช่น กลุ่มนักเรียนของโรงเรียนรัฐบาลกับกลุ่มนักเรียนของโรงเรียนเอกชน หรือกลุ่มพนักงานชายกับกลุ่มพนักงานหญิง หากมีความจำเป็นต้องกำหนดเป็นจำนวนเพื่อใช้ในการวิเคราะห์ด้วยวิธีการทางสถิติอาจใช้ 0 แทนกลุ่มนักเรียนของโรงเรียนรัฐบาลหรือกลุ่มพนักงานชาย และใช้ 1 แทนกลุ่มนักเรียนโรงเรียนเอกชนหรือกลุ่มพนักงานหญิง จำนวนที่ใช้แทนข้อมูลเชิงคุณภาพเหล่านี้ไม่สามารถนำไปตีความหมายในเชิงปริมาณได้ ความหมายของจำนวนที่ใช้แทนข้อมูลเหล่านี้เป็นเพียงสัญลักษณ์ที่ใช้แทน “กลุ่ม” ต่าง ๆ เท่านั้น

3. วิธีเก็บรวบรวมข้อมูล

1) วิธีเก็บรวบรวมข้อมูลทุติยภูมิ ซึ่งส่วนใหญ่มักจะอยู่ในหนังสือ รายงาน บทความ หรือเอกสารต่างๆ ควรดำเนินการดังต่อไปนี้

(1) พิจารณาตัวบุคคลผู้เขียนรายงาน บทความหรือเอกสารเหล่านั้นเสียก่อนว่าเป็นผู้มีความรู้และมีความเชี่ยวชาญในเรื่องที่จะเขียนถึงขั้นพอที่จะเชื่อถือได้หรือไม่ การเขียนอาศัยเหตุผลและหลักวิชาการมากน้อยเพียงใด ข้อมูลที่จะนำมาใช้ซึ่งรวบรวมจากรายงาน บทความหรือเอกสารดังกล่าวควรใช้ข้อมูลของผู้เขียนเก็บรวบรวมมาเองโดยตรง เช่น ข้อมูลที่ได้จากการสำรวจหรือสัมภาษณ์ หากไม่มีความจำเป็นไม่ควรใช้ข้อมูลของผู้เขียนนำมาจากแหล่งข้อมูลอื่นเนื่องจากอาจมีการคลาดเคลื่อนจากข้อมูลที่ควรจะเป็นจริงได้มาก

(2) ถ้าข้อมูลที่ต้องการเก็บรวบรวมสามารถหาได้จากหลาย ๆ แหล่ง ควรเก็บรวบรวมมาจากหลาย ๆ แหล่งเพื่อใช้ในการเปรียบเทียบว่าข้อมูลที่ต้องการมีความผิดพลาดเนื่องจากการลอคคิมพ์ผิดหรือเข้าใจผิดหรือไม่ นอกจากนี้ ผู้เก็บรวบรวมข้อมูลควรจะใช้ความรู้ความชำนาญของตนเองเกี่ยวกับข้อมูลเรื่องนั้น ๆ มาพิจารณาว่าข้อมูลที่จะนำมาใช้นั้นน่าจะเป็นไปได้หรือไม่ เช่น จำนวนประชากรของประเทศไทยในปี พ.ศ.2546 ที่นำเสนออยู่ในรายงานฉบับหนึ่งเป็น 36 ล้านคน จำนวนดังกล่าวน่าจะเป็นไปได้ ที่ถูกต้องควรจะเป็น 63 ล้านคน ความผิดพลาดดังกล่าว อาจเนื่องมาจากการคัดลอกของผู้นำเสนอหรือการพิมพ์ก็ได้ กล่าวคือคัดลอกหรือพิมพ์เลขโดดกลับกัน

(3) พิจารณาจากลักษณะของข้อมูลที่ต้องการเก็บรวบรวมว่าเป็นข้อมูลที่เป็นข้อความจริง ข้อมูลที่ได้จากทะเบียน ข้อมูลที่เป็นความคิดเห็นหรือเจตคติ ข้อมูลประเภทความลับหรือข้อมูลซึ่งผู้ตอบอาจต้องเสียประโยชน์จากการตอบ ถ้าเป็นข้อมูลที่เป็นข้อความจริง ข้อมูลที่ได้จากทะเบียนหรือข้อมูลที่เป็นความคิดเห็นหรือเจตคติ ส่วนใหญ่มักจะมีความถูกต้องเชื่อถือได้สูง แต่ถ้าเป็นข้อมูลประเภทความลับหรือข้อมูลซึ่งผู้ตอบอาจต้องเสียประโยชน์จากการตอบ ส่วนใหญ่มักจะไม่ถูกต้องเชื่อถือได้น้อย

(4) ถ้าข้อมูลที่จะเก็บรวบรวมได้มาจากการสำรวจจากกลุ่มตัวอย่าง หรือต้องผ่านขั้นตอนการวิเคราะห์โดยใช้วิธีการทางสถิติมาก่อน ควรจะต้องตรวจสอบวิธีการที่ใช้ในการเลือกกลุ่มตัวอย่าง ขนาดกลุ่มตัวอย่าง และวิธีการวิเคราะห์ว่าเหมาะสมที่จะใช้หรือไม่

2. วิธีเก็บรวบรวมข้อมูลปฐมภูมิ การเก็บรวบรวมข้อมูลปฐมภูมิซึ่งอาจทำได้โดยการสัมภาษณ์หรือสำรวจสามารถทำได้หลายวิธี แต่วิธีที่นิยมใช้กันทั่ว ๆ ไปมี 5 วิธี คือ

(1) **การสัมภาษณ์** การเก็บรวบรวมข้อมูลโดยการสัมภาษณ์ นิยมใช้กันมากกว่าการเก็บรวบรวมข้อมูลโดยวิธีอื่น ๆ เนื่องจากโอกาสที่จะได้คำตอบกลับคืนมามีมาก นอกจากนี้หากผู้ตอบคำถามไม่เข้าใจคำถามใด ๆ ก็สามารถสอบถามได้จากผู้สัมภาษณ์โดยตรง แต่การเก็บรวบรวมข้อมูลโดยวิธีนี้ ผู้สัมภาษณ์ต้องมีความซื่อสัตย์ ไม่ตอบคำถามแทนผู้ถูกสัมภาษณ์ เพราะจะทำให้ข้อมูลที่รวบรวมได้มีความคลาดเคลื่อนจากที่ควรจะเป็นจริงมาก

(2) **การสอบถามทางไปรษณีย์** การเก็บรวบรวมโดยวิธีนี้สามารถประหยัดค่าใช้จ่ายในการเก็บรวบรวมได้มาก และค่อนข้างแน่ใจได้ว่าผู้ตอบแบบสอบถามทุกคนได้รับแบบสอบถาม ยกเว้นคนที่มีการย้ายที่อยู่เท่านั้น นอกจากนี้ผู้ตอบแบบสอบถามจะได้รับความสะดวกในการตอบคำถาม กล่าวคือจะตอบคำถามเมื่อไรก็ได้ภายในระยะเวลาที่ผู้สำรวจได้กำหนดไว้ คำตอบที่ผู้สำรวจได้รับจะมีความถูกต้องและเชื่อถือได้มาก โดยเฉพาะอย่างยิ่งคำตอบที่จะทำให้ผู้ตอบแบบสอบถามอาจเสียประโยชน์จากการตอบคำถามนั้นๆ เนื่องจากผู้ตอบแบบสอบถามไม่จำเป็นต้องระบุชื่อของตนลงในแบบสอบถามก็ได้ แต่อาจมีจุดอ่อนถ้าผู้ตอบแบบสอบถามไม่เข้าใจปัญหาที่ถาม อาจทำให้คำตอบผิดพลาดได้ อีกประการหนึ่งผู้ถูกถามอาจไม่ได้เป็นผู้ตอบคำถามเองแต่ไปให้ผู้อื่นตอบแทน ข้อมูลที่

รวบรวมได้ก็อาจผิดพลาดได้ เช่นเดียวกัน นอกจากนี้ผู้สำรวจยังไม่สามารถประมาณจำนวนแบบสอบถามที่จะได้รับกลับคืนมาว่าจะมีจำนวนมากน้อยเพียงใด ซึ่งบางครั้งผู้สำรวจได้แบบสอบถามกลับคืนมาไม่เพียงพอที่จะทำการสรุปผลทั้งหมดให้มีความเชื่อถือได้

(3) การสอบถามทางโทรศัพท์ การสอบถามวิธีนี้นิยมใช้น้อยกว่าวิธีอื่น ถึงแม้ว่าการเลือกตัวอย่างผู้ตอบทำได้ง่ายและเสียค่าใช้จ่ายในการเก็บรวบรวมข้อมูลน้อยก็ตาม ทั้งนี้เนื่องจากการเก็บรวบรวมข้อมูลทำได้เฉพาะผู้มีโทรศัพท์เท่านั้น

การสอบถามทางโทรศัพท์โดยทั่ว ๆ ไป มักใช้กับแบบสอบถามที่ไม่ใช้เวลาในการสัมภาษณ์มากนัก และข้อมูลที่ต้องการถามจากผู้ตอบสัมภาษณ์เป็นข้อมูลที่ผู้ตอบสัมภาษณ์สามารถตอบได้ทันทีโดยไม่ต้องไปค้นหาหลักฐานหรือสอบถามจากผู้อื่น การสอบถามทางโทรศัพท์ที่ใช้กันอยู่เสมอ ๆ เช่น การสำรวจความคิดเห็นเกี่ยวกับเหตุการณ์ที่กำลังได้รับความสนใจ

(4) การสังเกต การเก็บรวบรวมข้อมูลโดยการสังเกตมักใช้ประกอบกับการเก็บรวบรวมข้อมูลโดยวิธีอื่นๆ ในกรณีที่ผู้สำรวจไม่แน่ใจว่าข้อมูลที่เก็บรวบรวมได้โดยวิธีนั้นๆ เชื่อถือได้ ซึ่งอาจจะมีสาเหตุมาจากความไม่ร่วมมือของผู้ให้ข้อมูล หรืออาจจะเกิดจากความรู้อันพื้นฐานหรือความรู้เกี่ยวกับเรื่องนั้นๆ ของผู้ตอบไม่ดีพอ เช่น การสอบถามเกี่ยวกับรายได้ของครอบครัวหรือกำไรของบริษัทห้างร้านต่างๆ ข้อมูลดังกล่าวนี้ผู้ตอบไม่ต้องการเปิดเผย นอกจากนี้อาจใช้การสังเกตเมื่อต้องการรวบรวมข้อมูลในเชิงลึก เช่น ครูสังเกตพฤติกรรมของนักเรียนในการทำงานร่วมกัน และการมีทักษะกระบวนการทางคณิตศาสตร์

(5) การทดลอง การเก็บรวบรวมข้อมูลจากการทดลอง ต้องมีการทดลองหรือปฏิบัติเพื่อให้ได้ข้อมูลที่ต้องการ ส่วนใหญ่จะเป็นการเก็บรวบรวมข้อมูลทางวิทยาศาสตร์ เช่น การเปรียบเทียบประสิทธิภาพของยาแก้ปวดหลาย ๆ ชนิด ข้อมูลที่เก็บรวบรวมได้จากการทดลองนี้ จะมีความถูกต้องและเชื่อถือได้มาก ถ้าไม่เกิดความคลาดเคลื่อนจากการวัดหรือการวางแผนการทดลอง

4. ปัญหาในการใช้ข้อมูล

ปัญหาในการใช้ข้อมูลทุกขุม การใช้ข้อมูลทุกขุมมักจะมีปัญหาเกี่ยวกับเรื่องต่าง ๆ ดังนี้

- 1) ความถูกต้องเชื่อถือได้ของข้อมูล
- 2) ความทันสมัยของข้อมูล
- 3) การขาดหายไปของข้อมูลบางรายการ

ปัญหาในการใช้ข้อมูลปฐมภูมิ มักจะมีปัญหาเกี่ยวกับเรื่องต่าง ๆ ดังต่อไปนี้

- 1) ไม่ทราบว่าจะใช้วิธีเลือกตัวอย่างหรือวางแผนการทดลองแบบใดจึงจะเหมาะสม
- 2) ไม่ทราบว่าจะประเมินความถูกต้องเชื่อถือได้ของข้อมูลที่เก็บรวบรวมมาได้อย่างไร
- 3) ไม่ทราบว่าจะวิเคราะห์ข้อมูลอย่างไรในกรณีข้อมูลที่เก็บรวบรวมได้ไม่ครบถ้วน

หรือขาดหายไปมากเนื่องจากไม่ได้รับความร่วมมือจากผู้ให้ข้อมูล