บทที่ 9
การวิเคราะห์ข้อมูล

ข้อมูลจะมีความหมายเด่นชัดก็ต่อเมื่อนักวิจัย ได้ทำการวิเคราะห์และตีความข้อมูล การวิเคราะห์และตีความข้อมูลจึงเป็นกระบวนการที่เกี่ยวข้องสัมพันธ์กันอย่างใกล้ชิดนักวิจัยมีข้อมูลจำนวนมากที่จะต้องใช้สถิติในการสรุปข้อมูล ซึ่งนักวิจัยจะต้องทราบว่าจะใช้สถิติอะไรในการวิเคราะห์ข้อมูลของตน จึงจะได้คำตอบได้แน่นอนและเที่ยงตรง ลำพังการวิเคราะห์ข้อมูลในตัวของมันเองไม่อาจให้คำตอบปัญหาที่วิจัยได้ จะต้องมีการตีความด้วย จุดประสงค์ของการวิเคราะห์ข้อมูล คือ เพื่อสรุปผลของการศึกษาในทำนองที่สามารถให้คำตอบกับปัญหาที่วิจัยได้ ในการตีความข้อมูลนั้น นักวิจัยจะต้องใช้ความรู้ ความคิด ตลอดจนสติปัญญาของเขาในการวิเคราะห์และตีความข้อมูล เมื่อแบบของการวิจัยแตกต่างกันไปก็จะทำให้การวิเคราะห์และตีความข้อมูลแตกต่างกันไปด้วย

การวิเคราะห์ข้อมูล คือ การวิเคราะห์ค่าของตัวแปรที่ได้มาจากการรวบรวมข้อมูลเพื่อที่จะทราบลักษณะของตัวแปรนั้น ๆ แล้วนำเอาผลจากการวิเคราะห์นั้นมาตีความเพื่อหาคำตอบให้กับปัญหาที่วิจัยนั่นเอง

ในการวิเคราะห์ข้อมูลเชิงสถิติ เป็นการอธิบายประชากรที่ศึกษา นักวิจัยต้องอธิบายสรุปข้อมูลที่รวบรวมมาได้จากตัวอย่าง ในขั้นนี้ต้องอาศัยการเข้าตารางตัวเลขด้วย นอกจากนั้นต้องคาดคะเนความเชื่อถือได้ของการวางกฎทั่ว ๆ ไปจากข้อมูลด้วย วิธีการทางสถิติจะช่วยอธิบายสรุป และคาดคะเนความเชื่อถือได้ เทคนิคทางสถิติที่นำมาใช้ในการวิเคราะห์ข้อมูลคือ สถิติในเชิงพรรณนา (Descriptive Statistics) ใช้บรรยายกลุ่ม เช่น ค่าเฉลี่ย ความเบี่ยงเบน มาตรฐาน ร้อยละ สัดส่วน และสถิติในเชิงวิเคราะห์ (Inference Statistics) ใช้สำหรับอนุมานไปสู่มวลประชากร การใช้สถิติต้องให้เหมาะสม บรรลุเป้าหมายของการวิจัยเป็นสำคัญ

สถิติ : ความหมายและบทบาทในการวิจัย

สถิติมีความหมายสองประการ คือ (1) สถิติเป็นตัวเลขหรือค่าของสิ่งใดสิ่งหนึ่ง ซึ่งบันทึกไว้ ถ้าเราพูดถึงสถิติคนว่างงานในปี พ.ศ. 2525 เราก็มักจะนึกถึงตัวเลขแสดงจำนวนคนว่างงานในปีนั้น หรือขนาดครอบครัวโดยเฉลี่ยในประเทศไทย สถิติหรือการจดบันทึกตัวเลขเป็นที่รู้จักกันทั่วไป และใช้กันมานานแล้ว (2) สถิติ หมายถึง เทคนิคที่ใช้วิเคราะห์

ข้อมูลเพื่อสรุปผล เช่น การใช้ร้อยละ ค่าเฉลี่ย พิสัย ส่วนเบี่ยงเบนมาตรฐาน เป็นต้น

ปัจจุบันนี้สถิติมีความหมายกว้างขวางมาก ในทัศนะของ Williams "สถิติ" เป็นสาขาของคณิตศาสตร์ประยุกต์ ซึ่งมีระเบียบวิธีการต่าง ๆ ที่ช่วยให้นักวิจัยสามารถอธิบาย (describing) และสรุปเหตุผล (reasoning) ได้จากสิ่งต่าง ๆ ที่สังเกต (observations) (Williams : 1968 : 4) สถิติมีความสำคัญยิ่งสำหรับงานค้นคว้าเป็นเครื่องมือซึ่งนักวิจัยทั้งแขนงวิทยาศาสตร์ และสังคมศาสตร์ ใช้ในการวางแผนทดลอง การเก็บตัวเลข การประมาณค่า และการสรุปผลงานทดลอง สถิติเป็นวิทยาการที่เกี่ยวข้องกับหน้าที่ต่อไปนี้ (1) การวางแผนงานทดลอง และสำรวจ (2) การเก็บตัวเลขและประมวลข้อมูล (3) การตรวจพิสูจน์สมมติฐาน (4) การคาดลักษณะของประชากร (5) การศึกษาความสัมพันธ์ระหว่างตัวแปร

ทุก ๆ ครั้งที่เราศึกษาเอกสารงานวิจัยต่าง ๆ จะพบว่าเกือบทุกเรื่อง ๆ ที่ผ่านสายตาเราต้องมีความเกี่ยวข้องกับการใช้หลักวิชาสถิติเข้าวิเคราะห์ตัวเลขเพื่อสรุปผลดังนั้นวิชาสถิติจึงมีความสำคัญต่อการวิจัยมาก สถิติเป็นเครื่องมือของนักวิจัยที่ช่วยให้นักวิจัยสามารถสรุปคำตอบให้แก่ปัญหา นักวิจัยสนใจได้กับทั้งช่วยให้นักวิจัยสามารถตรวจสอบพิสูจน์ความถูกต้องของคำตอบให้ผู้อื่นยอมรับว่าเป็นข้อเท็จจริงได้โดยวิธีการทางวิทยาศาสตร์

วิธีวิเคราะห์ข้อมูล (Data Analysis) เป็นงานขั้นที่ 3 ในงาน 4 ขั้น ซึ่งถือเป็นงานสถิติ งาน 4 ขั้น นั่นคือ (1) การรวบรวมข้อมูล (Data Collection) (2) การประมวลผลข้อมูล (Data Processing) (3) การวิเคราะห์ข้อมูล (Data Analysis) (4) การนำเสนอข้อมูล (Data Presentation)

วัตถุประสงค์ที่สำคัญในการวิเคราะห์ข้อมูล ก็เพื่อจะให้ข้อมูลที่เก็บรวบรวมมาและได้รับการประมวลผลสำเร็จแล้วอยู่ในรูปที่มีความหมายเด่นชัดยิ่งขึ้น และสามารถนำไปใช้เป็นพื้นฐานในการอธิบายข้อสรุปเกี่ยวกับเรื่องหรือสถานการณ์ต่าง ๆ ที่สามารถอธิบายจากข้อมูลซึ่งเก็บรวบรวมมาได้ ย่อข้อมูลที่เก็บมาได้ในรูปที่อ่านเข้าใจง่ายโดยไม่ต้องกลับไปดูข้อมูลเดิมอีก งานวิเคราะห์ข้อมูลอาจจะทำไปพร้อมกับการประมวลข้อมูลก็ได้โดยเฉพาะการวิเคราะห์ข้อมูลจำนวนมากต้องอาศัยเครื่องคอมพิวเตอร์เข้าช่วย อาจใช้โปรแกรมสำเร็จรูป เช่น SAS (Statistical Analysis System) ประกอบด้วยโปรแกรมสำเร็จรูปใช้ในการวิเคราะห์ข้อมูลเชิงสถิติในเรื่องของ Regression Analysis, Analysis of Variance, Factor Analysis หรือ SPSS (Statistical Package For the Social Sciences) ใช้ในการคำนวณ Frequency Distribution, Percent, Multiple Correlation and Regression, Factor Analysis, Scalogram Analysis ฯลฯ ซึ่งเหมาะสำหรับใช้ในการวิเคราะห์ข้อมูลทางด้านสังคมศาสตร์

ก่อนที่นักวิจัยจะวิเคราะห์ข้อมูล จะต้องตรวจสอบข้อมูลว่ามีลักษณะเป็นตัวแทนของประชากรหรือไม่ แหล่งที่มาของข้อมูลจะต้องเชื่อถือได้ และปราศจากอคติสามารถวัดหรือทดสอบซ้ำได้ และข้อมูลมีความสำคัญต่อทฤษฎีและสามารถนำไปประยุกต์ใช้ได้ ข้อมูลที่เป็น

ตัวเลขควรตรวจสอบด้วยความระมัดระวังก่อนคำนวณ เพื่อป้องกันข้อผิดพลาดในการอนุมานหรือสรุปเกี่ยวกับประชากรที่ศึกษา

ขั้นตอนที่สำคัญของการวิเคราะห์ข้อมูล มีดังต่อไปนี้

1. การแยกประเภทข้อมูล
2. การลงรหัส
3. การเข้าตาราง
4. การอ่านและวิเคราะห์ข้อมูลจากตาราง
5. การพิจารณาและตีความหมายของข้อมูล

การแยกประเภทข้อมูล

เมื่อนักวิจัยรวบรวมข้อมูลมาได้แล้ว นักวิจัยจะต้องจัดแยกประเภทของข้อมูล เพื่อสะดวกในการวิเคราะห์ จัดเข้าตาราง หาความสัมพันธ์ระหว่างกัน การจัดแยกประเภทนั้นจะต้องเป็นไปตามปัญหาและวัตถุประสงค์ในการวิจัยถ้าหากไม่ทำตามความต้องการของปัญหาในวิจัยก็ไม่อาจให้คำตอบในการวิจัยได้ในการแยกประเภทของข้อมูลนั้นเราจะต้องวางแผนไว้ก่อนเก็บข้อมูลซึ่งจะอยู่ในแบบของการวิจัยว่าจะใช้แบบใด เมื่อเก็บข้อมูลมาแล้วจึงนำมาแยกประเภทตามแผนที่วางไว้ ซึ่งมักจะทำเป็นตารางเปล่า (Dummy Table) ไว้ก่อน ข้อมูลที่นำมาวิเคราะห์ต้องจัดเข้าประเภทหนึ่งเท่านั้น (Mutually Exclusive) การที่จะจัดได้สะดวกขึ้นอยู่กับคำจำกัดความในเชิงปฏิบัติ (Operation Definition) โดยให้ความหมายของตัวแปรต่าง ๆ ต้องชัดเจนและไม่คลุมเครือมิฉะนั้นแต่ละรายอาจจัดเข้าช่องได้เกินกว่า 1 ช่องก็ได้

การลงรหัส

การลงรหัส คือ การเปลี่ยนข้อมูลดิบเป็นสัญลักษณ์ ซึ่งปกติก็เป็นตัวเลขทำให้สามารถจัดเข้าตารางและนับจำนวนได้ เช่น รหัส 1 เพศชาย 2 เพศหญิง 1 และ 2 เป็นรหัสของการวิจัยมักจะเป็นตัวเลขเกือบทั้งหมดการลงรหัสนี้ไม่ใช่ทำโดยอัตโนมัติต้องอาศัยการตัดสินใจของผู้ลงรหัสเองในการสร้างแบบสอบถามก็จะต้องกำหนดคำตอบให้เพียงพอ ถ้ามีคำตอบไม่เพียงพอในบางครั้งอาจไม่ตรงกับวัตถุประสงค์ของการวิจัยเพราะฉะนั้นผู้วิจัยจะต้องตั้งคำถามให้ชัดเจน เพื่อจะทำให้ได้คำตอบที่เพียงพอในการลงรหัสโดยเชื่อถือได้

การแจกแจงนับและการเข้าตาราง

การเข้าตาราง คือ การนับเพื่อทราบจำนวนของรายการต่าง ๆ ที่แยกประเภทได้ วิธีการนี้ใช้แบบสอบถาม ทั้งที่ซับซ้อนและไม่ซับซ้อน ใช้กับแบบสอบถามที่ใช้รหัสวิธีจัดข้อมูล

ตารางนี้จะช่วยให้นักวิจัยในการแยกหมวดหมู่ และสะดวกในการแจกแจงความถี่และสามารถเช็คได้หากมีข้อผิดพลาด การเข้าตารางแยกประเภท (Cross Tabulation) หมายถึงการเข้าตารางรายการต่าง ๆ โดยแยกเป็น 2 ประเภทหรือเกินกว่า 2 ประเภท เช่น เข้าตารางโดยมีทั้งการศึกษาสูง และรายได้ต่ำ การเข้าตารางแยกประเภท ถือเป็นขั้นตอนที่สำคัญในการค้นหาหรือทดสอบความสัมพันธ์ระหว่างตัวแปรต่าง ๆ

การเข้าตารางอาจทำโดยใช้มือหรือใช้เครื่องจักรก็ได้ แต่ละอย่างมีทั้งข้อได้เปรียบและเสียเปรียบ โดยทั่วไปการเข้าตารางโดยใช้มือประหยัดกว่าและใช้เวลาน้อยกว่าถ้ามีจำนวนที่ศึกษาไม่มากนัก และเข้าตารางแยกประเภทไม่มาก แต่ถ้าจำนวนมากรายที่ศึกษาและตารางแยกประเภทมากขึ้น การเข้าตารางโดยใช้เครื่องจักรประหยัดกว่า แต่ถ้าคำนึงถึงค่าใช้จ่ายในการใช้เครื่องจักรเป็นรายชั่วโมง นับว่าแพงมาก การใช้เครื่องจักรจึงควรใช้กับข้อมูลจำนวนมาก และมีการวิเคราะห์ทางสถิติซับซ้อน

การอ่านและวิเคราะห์ข้อมูลจากตาราง

การอ่านและการวิเคราะห์ข้อมูลไม่มีหลักแน่นอนว่าจะต้องวิเคราะห์อย่างนั้นอย่างนี้ ขึ้นอยู่กับตัวของนักวิจัยเอง เพราะฉะนั้นในการวิจัย นักวิจัยต้องเลือกวิธีการวิจัยที่เหมาะสมกับปัญหาที่ศึกษา นักวิจัยปกติต้องเลือกแบบของการวิจัย วิธีรวบรวมข้อมูล วิธีวัดและแบบในการวิเคราะห์ส่วนประกอบเหล่านี้ต้องสอดคล้องกัน เข้ากันได้ สิ่งสำคัญก็คือแบบของการวิจัย วิธีรวบรวมข้อมูลการวัด และการวิเคราะห์ทางสถิติ ต้องเหมาะสมกับปัญหาในการวิจัย ในการวิเคราะห์เรื่องที่ซับซ้อน การตีความปรากฏการณ์หลายอย่าง ทั้งนี้ขึ้นอยู่กับปัญหาและคำตอบที่ต้องการ

การพิจารณาและตีความหมายของข้อมูล

การตีความหมายของข้อมูล ได้แก่การแปลตัวเลขจากข้อมูลที่ได้ออกมา เพื่อให้ผู้อ่านเข้าใจว่าตัวเลขในเรื่องนั้นออกมาอย่างนั้นหมายความว่าอย่างไร การตีความข้อมูลทางสถิติ ส่วนใหญ่จะเป็นข้อความที่มีลักษณะน่าจะเป็นไปได้ เราไม่สามารถจะยึดถือการตีความทางสถิติอย่างเดียว ในการตัดสิน หรือการยอมรับว่าสมมติฐานที่ตั้งไว้เป็นความจริง จะต้องทำการวิจัยซ้ำและพยายามเชื่อมโยงข้อมูลที่ได้กับการศึกษาของผู้อื่นที่ได้ทำไว้

สถิติเชิงพรรณนา

ในการวิเคราะห์ข้อมูลเชิงพรรณนา ใช้เพื่ออธิบายลักษณะและสรุปข้อมูลปกติใช้อธิบายข้อมูลในลักษณะดังต่อไปนี้

1. อธิบายว่าอะไรเป็นลักษณะแบบอย่างของกลุ่ม เช่น ต้องการรู้ว่าจำนวนบุคคลที่ไปลงคะแนนเสียงเลือกตั้งสมาชิกสภาผู้แทนราษฎรถัวเฉลี่ยเท่าไร สถิติที่นำมาใช้คือ การหาแนวโน้มเข้าหาศูนย์ (Central Tendency) ได้แก่ การใช้มัชฌิมเลขคณิต (Arithmatic Mean) ถัวเฉลี่ย (Average) เป็นต้น วิธีนี้เป็นการวิเคราะห์ในเชิงสถิติแบบง่าย ๆ

2. ชี้ว่าแต่ละบุคคลในกลุ่มผันแปรไปเพียงใด วิธีวัดที่ใช้กันบ่อย คือการวัดการกระจาย (Dispersion) พิสัย (Range) ส่วนเบี่ยงเบนเฉลี่ย (Average Deviation) ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) เป็นต้น

3. แสดงว่าแต่ละบุคคลจะกระจายไปตามตัวแปรที่วัดอย่างไร โดยแสดงให้เห็นถึงการกระจายเป็นเส้นโค้งแบบใด

สถิติเชิงวิเคราะห์

คือการวิเคราะห์ข้อมูลโดยอาศัยวิธีการทางสถิติเป็นเครื่องชี้นำ โดยมีวัตถุประสงค์ดังต่อไปนี้

1. แสดงความสัมพันธ์ระหว่างตัวแปรต่าง ๆ เช่น ความสัมพันธ์ระหว่างผู้ไปลงคะแนนเสียงเลือกตั้งกับเพศ รายได้และอายุ เป็นต้น เราต้องการแสดงว่าการผันแปรในลักษณะหนึ่งจะเกี่ยวข้องสัมพันธ์กับอีกลักษณะหนึ่งหรือไม่ วิธีแสดงความสัมพันธ์กันมีหลายวิธี เช่น สหสัมพันธ์ (Correlation) ระหว่างตัวแปรและความสัมพันธ์ในเชิงเหตุผล (Causal Relationship) เป็นต้น สำหรับการวัดสหสัมพันธ์วัดได้หลายอย่าง เช่น The product moment coefficient of correlation (r) the rank order coefficient of correlation (p) distance measure (d) การวัดสหสัมพันธ์ดังกล่าวไม่ว่าจะมีวิธีการวัดแตกต่างกันอย่างไร แต่ก็มีวัตถุประสงค์อย่างเดียวกันเป็นการวัดชุดของตัวแปรคู่หนึ่งว่ามีความสัมพันธ์ร่วมกันสำหรับ Coefficient of Correlation ปกติจะมีค่าผันแปรระหว่าง - 1.00 + 1.00 หรือ 0 ถ้าค่าเท่ากับ -1.00 แสดงว่าความสัมพันธ์ในเชิงนิเสธโดยสมบูรณ์ ถ้าค่าเท่ากับ +1.00 แสดงว่ามีค่าในเชิงปฏิฐานโดยสมบูรณ์ ส่วนค่า 0 แสดงว่าไม่มีความสัมพันธ์กันเลย

2. อธิบายความแตกต่างระหว่าง 2 กลุ่ม หรือเกินกว่า 2 กลุ่ม ปกติการเปรียบเทียบการวัดการผันแปรของแต่ละกลุ่มและความสัมพันธ์ระหว่างตัวแปรของทั้ง 2 กลุ่มก็ได้ (อำนวย วิทย์. 2519 : 306 – 307)

เทคนิคที่สำคัญ ๆ ในการวิเคราะห์ข้อมูลเบื้องต้น

เทคนิคการวิเคราะห์ข้อมูลมีอยู่หลายอย่าง และแต่ละอย่างมีความเหมาะสมที่จะใช้ไม่

เหมือนกัน แล้วแต่ว่าข้อมูลที่เขารวบรวมมาเป็นข้อมูลเกี่ยวกับอะไร และแล้วแต่ว่ากำลังศึกษาเรื่องอะไร ข้อมูลแต่ละอย่างใช้ระเบียบวิธีการวิเคราะห์แตกต่างกันออกไป ดังนั้นนักวิจัยจะต้องทำความเข้าใจกับเทคนิคแต่ละอย่างและเลือกใช้ให้เหมาะสมกับลักษณะของข้อมูลให้เหมาะสมกับเรื่องราวที่กำลังวิเคราะห์ หรือพยายามหาข้อสรุปจากข้อมูล

เทคนิคแรก คือ **วิธีการจำแนกข้อมูล** (Data Classification) เนื่องจากข้อมูลที่รวบรวมมาอาจไม่เป็นหมวดหมู่ นักวิจัยจะต้องจัดแบ่งข้อมูลออกเป็นพวก ๆ หรือเป็นกลุ่ม ๆ อย่างมีระบบ ทั้งนี้โดยอาศัยการพิจารณาลักษณะที่เหมือนกันบางอย่างโดยทั่ว ๆ ไป ข้อมูลที่นักวิจัยรวบรวมมาได้นั้นพอจะจำแนกได้ 2 ลักษณะ คือ

1. จำแนกตามปริมาณ (Quantitative Classification) หมายถึง การจำแนกข้อมูลตามจำนวนหรือปริมาณซึ่งวัดออกเป็นตัวเลขได้ชัดเจนเช่น จำแนกครัวเรือน โดยระดับรายได้ จำแนกครัวเรือนเกษตรโดยเนื้อที่ถือครอง เป็นต้น

2. จำแนกตามคุณภาพ (Qualitative Classification) ซึ่งหมายถึงการจำแนกข้อมูลตามลักษณะที่ข้อมูลมีร่วมกันบางอย่างซึ่งลักษณะดังกล่าวไม่อาจวัดเป็นปริมาณหรือตัวเลขได้โดยตรง เช่น การจำแนกประชากรตามกลุ่มอาชีพ การจำแนกครูตามวุฒิ เป็นต้น

หลักเกณฑ์ที่สำคัญในการวิเคราะห์ข้อมูลโดยวิธีการจำแนกข้อมูล ไม่ว่าจะเป็นการจำแนกตามปริมาณหรือคุณภาพ คือ

1. ต้องให้มีความแนบนัยในคำจำกัดความ และมาตรฐาน การจำแนก เช่น อาชีพ การศึกษา หมวดอายุ ขนาด ประเภท ฯลฯ เป็นต้น

2. พยายามอย่าให้มีกลุ่มที่ไม่สามารถจำแนกได้อย่างชัดเจน ซึ่งมักเรียกว่า "อื่น ๆ" มีความถี่ (Frequency) สูงนัก จนข้อมูลขาดความหมายที่เด่นชัดไป

3. ถ้าเป็นการจำแนกตามปริมาณ ก็จะศึกษาและทำความเข้าใจหลักการใช้อันตรภาคชั้น (Class Interval) ให้ถูกต้องและสะดวกกับการคำนวณต่าง ๆ เชิงสถิติด้วย

การวิเคราะห์ข้อมูลโดยการจำแนกข้อมูล อาจเป็นการจำแนกตามลักษณะหนึ่งลักษณะใดอย่างเดียว เรียกกันว่า One - way Classification เช่น จำแนกประชากรตามหมวดอายุ จำแนกโรงงานอุตสาหกรรมตามขนาด จำแนกการใช้เนื้อที่การเกษตรตามพืช หรืออาจจะเป็นการจำแนกตามลักษณะ 2 อย่างพร้อม ๆ กัน เช่น จำแนกข้าราชการตามเพศและชั้น จำแนกโรงงานอุตสาหกรรมตามประเภทและขนาด จำแนกครัวเรือนตามรายได้และอาชีพ ซึ่งเรียกว่า Two - way Classification หรืออาจจะเป็นการจำแนกตามลักษณะ 3 อย่างพร้อม ๆ กัน เช่น การวิเคราะห์ข้อมูลเกี่ยวกับครูในจังหวัด อาจจำแนกข้อมูลออกเป็น 3 ลักษณะ ตาม เพศ คุณวุฒิ ประสบการณ์ในการสอน (นิยม : 2520 : 29-31)

1. ตารางจำแนกข้อมูลด้านเดียว

ตัวอย่างที่ 1

ตารางที่ 1 แสดงอัตราร้อยละของหัวหน้าครัวเรือนจำแนกตามเพศ

เพศ	จำนวน	ร้อยละ
ชาย	224	89.6
หญิง	26	10.4
รวม	250	100.0

2. ตารางจำแนกข้อมูล 2 ด้าน

ตัวอย่างที่ 2

ตารางที่ 2 แสดงความสัมพันธ์ระหว่างระดับการศึกษากับผลการปฏิบัติงาน

ระดับการศึกษา	ผลการปฏิบัติงาน			
	สูง	ปานกลาง	ต่ำ	รวม
ประถมปีที่ 4 และต่ำกว่า	22.5	34.3	28.4	85.2
	(23)	(35)	(29)	(87)
สูงกว่าประถมปีที่ 4	10.8	3.0	1.0	14.8
	(11)	(3)	(1)	(15)
รวม	33.3	37.3	29.4	100.0
	(34)	(38)	(30)	(102)

3. ตารางจำแนกข้อมูล 3 ด้าน

ตารางที่ 3 อัตราอุบัติเหตุในการขับรถของผู้ชายและผู้หญิง แยกตามระยะทางที่ขับ

เคยมีอุบัติเหตุ	ผู้ชาย		ผู้หญิง	
	ขับเกินกว่า 1000 ไมล์ %	ขับ 1000 ไมล์ หรือต่ำกว่า %	ขับเกินกว่า 1000 ไมล์ %	ขับ 1000 ไมล์ หรือน้อยกว่า %
อย่างน้อยหนึ่งครั้ง	52	25	52	24
ไม่เคยมีอุบัติเหตุ	48	75	48	75
รวม	100 (5,010)	100 (2,070)	100 (1,915)	100 (5,035)

เทคนิคการวิเคราะห์ข้อมูลอีกชนิดหนึ่งคือการสรุปย่อข้อมูล (Condensation and Summarizaton of Data) การวิเคราะห์ข้อมูลโดยการสรุปย่อที่สำคัญและใช้กันอย่างแพร่หลายมากที่สุดมีอยู่ 2–3 อย่าง คือ

ก. การทำการแจกแจงความถี่ (Frequency Distribution) คือวิธีที่จะช่วยให้เก็บหรือเข้าใจข้อมูลต่าง ๆ ที่ให้ได้ดีนั้น โดยการนำเอาข้อมูลเหล่านั้นมาเรียงลำดับเข้าให้เป็นพวก โดยอาศัยขนาดและจัดเป็นพวกแบบเลือกช่วงห่างของหมู่ หรืออันตรภาคชั้น (Size of Class - Interval) การเลือกใช้ขนาดของอันตรภาคชั้น และจำนวนของอันตรภาคชั้น โดยทั่ว ๆ ไปไม่มีกฎเกณฑ์ที่ตายตัวว่าควรจะเป็นเท่าไร แต่จะต้องพิจารณาจากจำนวนและค่าตัวเลขของข้อมูลที่วิเคราะห์ โดยอาศัยขั้นตอนที่สำคัญ 3 ขั้น คือ

1. หาค่าสูงสุด (Maximum Value) และค่าต่ำสุด (Minimum Value) ของข้อมูลที่มีอยู่ทั้งหมด
2. หาค่าพิสัย (Range) ของข้อมูลซึ่งก็คือผลต่างระหว่างค่าสูงสุดและค่าต่ำสุดของข้อมูล

$$\text{Range} = X_{max} - X_{min}$$

3. แบ่งพิสัยหรือ Range ออกเป็นจำนวนชั้น (Class) ที่ต้องการซึ่งโดยทั่วไปมักจะใช้ระหว่าง 5 – 20 ชั้น แล้วแต่ว่าเราต้องการละเอียดในการวิเคราะห์ การแจกแจงความถี่มากน้อยแค่ไหนเพียงไร และขึ้นอยู่กับจำนวนข้อมูลที่มีอยู่อีกอย่างหนึ่ง

4. กำหนดขนาดหรือช่องหรือความกว้างของอันตรภาค (Size of Class Interval) ให้มีความต่อเนื่องกันจากน้อยไปหามาก โดยใช้หลักเกณฑ์ว่าขนาดของอันตรภาคชั้น (Size of Class Interval) ไม่ควรจะกว้างหรือแคบเกินไป เพราะถ้ากว้างเกินไปจะทำให้มองไม่เห็นภาพการแจกแจงที่ชัดเจนนัก และในทางตรงกันข้ามการใช้ขนาดของอันตรภาคชั้นที่เล็กหรือช่วงแคบเกินไปก็อาจจะเป็นผล ทำให้เกิดความไม่สม่ำเสมอหรือความไม่เป็นตามปกติธรรมชาติของการแจกแจงของลักษณะที่เราสนใจไปได้เช่นกัน

ข. การสรุปข้อมูลโดยวิธีการหาค่าทางสถิติที่เป็นค่าวัดลักษณะสำคัญ ๆ ของข้อมูลในบางกรณีการวิเคราะห์ข้อมูลเพื่อหาข้อสรุปเกี่ยวกับเรื่องใดเรื่องหนึ่ง อาจทำได้ง่าย ๆ โดยการคำนวณหาอัตราส่วนระหว่างยอดรวมของข้อมูล 2 ชุด ที่มีความสัมพันธ์กัน เช่น การวิเคราะห์ข้อมูลเกี่ยวกับสภาพการศึกษา อาจหาอัตราส่วนระหว่างครูต่อนักเรียนที่เป็นอยู่เปรียบเทียบกัน การวิเคราะห์ข้อมูลเกี่ยวกับสภาพของบริการทางอนามัย หรือสาธารณสุขของท้องถิ่น อาจหาอัตราส่วนระหว่างจำนวนบุคลากรทางการแพทย์ต่อคนไข้ที่มาใช้บริการในช่วงเวลาหนึ่ง ๆ เช่น ต่อวัน ต่อเดือน ต่อปี เป็นต้น

ร้อยละ (Percentage) เป็นระเบียบวิธีทางสถิติที่ใช้ในการเปรียบเทียบอย่างหนึ่ง ในบางครั้งเราไม่อาจวัดพฤติกรรม ทัศนคติ หรือคุณลักษณะบางอย่างได้โดยตรงแต่อาจคิดเป็นร้อยละได้โดยการประมาณจากตัวอย่างหรือข้อมูลที่มีร้อยละหรือเปอร์เซ็นต์ เป็นตัวเลขที่ใช้กันมากในทางสถิติ เพื่อความสะดวกรวดเร็วในการเปรียบเทียบข้อมูล อย่างไรก็ดีการใช้ร้อยละในสรุปข้อมูลหรือเรื่องราวบางอย่างจะต้องทำด้วยความระมัดระวังพอสมควร เพราะถ้าหากใช้ไม่ระวังก็จะทำให้การตีความหมายผิดไปด้วย สาเหตุของความผิดพลาดต่าง ๆ มีอยู่หลายประการด้วยกัน คือ อาจจะเกิดความสับสนในการใช้เลขฐานคำนวณหาร้อยละจากตัวเลขจำนวนน้อยเกินไป ใช้จุดทศนิยมผิด คิดเลขผิดพลาด

วิธีการคำนวณหาร้อยละหรือเปอร์เซ็นต์เพื่อวิเคราะห์สรุปสถานการณ์ต่าง ๆ จากข้อมูลที่มีอยู่ โดยใช้การเปรียบเทียบตัวเลขจำนวนหนึ่งหรือหลายจำนวนกับตัวเลขอีกจำนวนหนึ่ง ตัวเลขที่ใช้เปรียบเทียบด้วยนั้นเรียกว่า ฐาน (Base) เราสามารถคำนวณหาเปอร์เซ็นต์ได้โดยใช้ตัวเลขจำนวนที่เราต้องการจะเปรียบเทียบกับฐานตั้งแล้วหารด้วยฐาน

ในการใช้ร้อยละหรือเปอร์เซ็นต์เรามักจะใช้สรุปสถานการณ์จากข้อมูลจำนวนมาก ๆ เช่น การวิเคราะห์อัตราการเกิดและอัตราการตาย การวิเคราะห์อัตราการว่างงานในเมืองและชนบท การวิเคราะห์การหยั่งเสียงประชามติ ฯลฯ

ค่าเฉลี่ย (Average)

ค่าเฉลี่ย (Average or Mean) ในทางสถิติถือเป็นค่าที่วัดสภาพปานกลาง หรือแนวโน้มเข้าสู่ส่วนกลาง (Measure of Cental Tendency) ของข้อมูลเป็นตัวแทนของข้อมูลต่าง ๆ ซึ่งอาจจะเป็นจำนวนเลข จำนวนคะแนน จำนวนคน จำนวนอายุ ที่แสดงให้เห็นถึงค่ากลาง ๆ ของจำนวนเลขทั้งหมด หรือจำนวนคะแนนทั้งหมดเพราะในวงการสถิตินั้น บางครั้งเราอาจจะสรุปหรือย่อข้อมูลจำนวนมาก ๆ โดยการหาตัวแทนของข้อมูลที่เหมาะสมเพียงบางตัวขึ้นมา ก็พอจะทำให้ผู้สนใจในตัวเลข หรือข้อมูลนั้น ๆ ได้ข้อสรุปโดยไม่ต้องเอาข้อมูลทั้งหมดไปดู. และการรู้ค่าเฉลี่ยนี้จะทำให้ทำการศึกษาสามารถเทียบค่าต่ำของจำนวน ตัวเลข หรือคะแนนในแต่ละหน่วยได้ โดยถือค่ากลางหรือค่าเฉลี่ยเป็นมาตรการในการเปรียบเทียบ

ในทางสถิติ เรามีวิธีการหาค่าเฉลี่ยเพื่อใช้สรุปย่อข้อมูลได้หลายวิธี ที่สำคัญและใช้ได้มาก ในการวิเคราะห์ค่าเฉลี่ยข้อมูลทั่ว ๆ ไป ได้แก่

ก. มัชฌิมเลขคณิต (Arithmetic Mean)

ข. มัธยฐาน (Median)

ส่วนค่าเฉลี่ยอื่น ๆ ที่ใช้กันอยู่บ้างในบางกรณี คือ

ค. มัชฌิมเรขาคณิต (Geometric Mean)

ง. มัชฌิมฮาร์โมนิค (Harmonic Mean)

จ. ฐานนิยม (Mode)

ก. มัชฌิมเลขคณิต (Arithmetic Mean) ในการศึกษาลักษณะใดในทางสังคมศาสตร์หรือแขนงวิชาอื่น ๆ ข้อมูลที่ได้ประกอบค่าสังเกตต่าง ๆ กัน จำนวนมากหรือน้อยขึ้นอยู่กับการทดลอง ท่ามกลางความแตกต่างของค่าเหล่านี้ จะมีค่าหนึ่งที่แสดงให้เห็นว่าส่วนมากค่าทั้งหมดจะตกอยู่ใกล้เคียงกับค่านี้ ดังนั้นจึงเป็นเครื่องชี้ให้เห็นศูนย์กลางของลักษณะนั้น

มัชฌิมเลขคณิต (Arithmetic Mean) ซึ่งคำนวณได้จากผลรวมของค่าทั้งหมด หารด้วยค่าในวิชาสถิติ ซึ่งเขียนสูตรได้ดังนี้

สูตร $$X = \frac{X_1 + X_{2+} \ldots\ldots\ldots\ldots + Xn}{N}$$

ซึ่งมักเขียนสั้น ๆ ด้วย

$$\overline{X} = \sum_{i=1}^{N} X_i / N$$

ในกรณีที่มีข้อมูลจำนวนมาก ๆ และได้มีการแจกแจงความถี่ของข้อมูลออกเป็นชั้นต่าง ๆ หรือสูตรที่ใช้คำนวณค่าเฉลี่ยของข้อมูลมัชฌิมเลขคณิต คือ

$$X = \frac{f_1 x_1 + f_2 x_2 + \ldots\ldots\ldots\ldots + f_k x_k}{f_1 \quad f_2 \ldots\ldots\ldots\ldots + f_k}$$

ซึ่งมักเขียนสั้น ๆ ด้วย

$$X = \frac{\sum_{i=i}^{k} f_i x_i}{\sum_{i=1}^{h} f_i}$$

ในเมื่อ f_1 , f_2 ..... f_k คือความถี่ของค่า x_1 , x_2 , ... x_n ตามลำดับ

ข. มัธยฐาน (Median) หมายถึง ค่าที่เมื่อเรียงค่าต่าง ๆ ตามลำดับจากต่ำไปหาสูงแล้ว ค่านั้นจะตกอยู่ตรงกลาง คือจะมีค่าต่ำกว่า และสูงกว่าอย่างละครึ่ง หรือกล่าวอีกนัยหนึ่งก็คือค่าซึ่งแบ่งข้อมูลออกเป็น 2 พวก เท่า ๆ กัน โดยพวกหนึ่งมีค่าต่ำกว่านี้ และอีกพวกหนึ่งมีค่าสูงกว่านี้ ถ้าหากมีตัวเลขเป็นจำนวนคู่ Median เท่ากับเฉลี่ยของค่า ซึ่งอยู่ตรงกลางสองค่า

ตัวอย่าง ข้อมูลเกี่ยวกับอายุการทำงานของข้าราชการ 14 คน เป็นดังนี้

1, 2, 3, 5, 4, 6, 7, 10, 8, 3, 6, 14, 4, 5

ถ้าวิเคราะห์อายุการทำงานโดยเฉลี่ยโดยมัธยฐาน อาจทำได้โดยเรียงค่าของข้อมูลจากน้อยไปหามาก คือ

1, 2, 3, 3, 4, 4, 5, 5, 6, 6, 7, 8, 10, 14

และค่าที่แบ่งข้อมูลทั้ง 14 ตัวเป็น 2 ส่วนเท่า ๆ กัน ก็ คือ "5" ซึ่งเป็นค่าที่ 7 ของข้อมูลเมื่อเรียงจากน้อยไปหามาก

ฉะนั้น ถ้าหากอายุการทำงานโดยเฉลี่ยของข้าราชการทั้ง 14 คน โดยมัธยฐานจะได้ 5 ปี

แต่ถ้าใช้มัชฌิมเลขคณิตเป็นเฉลี่ยจะได้

$$\frac{1 + 2 + 3 + 3 + \ldots\ldots\ldots\ldots + 10 + 14}{14} \text{ ปี}$$

5.56 ปี

ซึ่งพออนุโลมได้ว่าไม่แตกต่างกันมาก

แต่ถ้าเป็นข้อมูลซึ่งมีค่าบางค่าที่สูงหรือต่ำผิดปกติ ค่าเฉลี่ยโดยมัธยฐานจะมีความเป็นตัวแทนได้ดีกว่าค่าเฉลี่ยโดยมัชฌิมเลขคณิต

ค. ฐานนิยม (Mode) เป็นค่าแสดงความโน้มเข้าหาศูนย์อีกอย่างหนึ่ง อันหมายถึงค่าซึ่งเกิดบ่อยที่สุด เช่น มีข้อมูล 1, 2, 3, 3, 4, 4, 4, 5, 5, 6, ในตัวแทนนี้ฐานนิยมเท่ากับ 4

ในการเลือกใช้ค่าเฉลี่ยเพื่อสรุปย่อข้อมูลนั้นจำเป็นที่จะต้องทราบข้อดีและข้อเสียของค่าเฉลี่ยเหล่านี้เสียก่อน ซึ่งมีอยู่หลายประการด้วยกัน

มัชฌิมเลขคณิต (Arithmetic Mean)

ข้อได้เปรียบ

1. คำนวณได้ง่ายสำหรับข้อมูลทั่ว ๆ ไป
2. เข้าใจง่าย
3. ค่าของมัชฌิมเลขคณิตเป็นค่าที่แน่นอน
4. เหมาะสมสำหรับเข้าสมการคำนวณในขั้นต่อไป

ข้อเสียเปรียบ

1. ต้องทราบค่าของข้อมูลทุกค่าจึงจะคำน่วณได้
2. ค่าของมัชฌิมเลขคณิตจะตรงกับค่าที่เป็นจริงของข้อมูล เพียงไม่กี่รายการหรือไม่ตรงกันเลย

ประโยชน์

1. เหมาะสำหรับจัดค่าความเป็นกลางสำหรับข้อมูลที่มีค่าไม่แตกต่างกันมากนัก

มัธยฐาน (Median)

ข้อได้เปรียบ

1. คำนวณง่าย
2. เข้าใจง่าย
3. ไม่กระทบกระเทือนมากนัก เมื่อมีค่าที่สูงหรือต่ำผิดปกติในข้อมูล
4. ไม่ต้องใช้ค่าของข้อมูลทุกค่า

ข้อเสียเปรียบ

1. ไม่สามารถหาค่าของมัธยฐานได้จากข้อมูลที่มิได้จัดเรียงตามลำดับ
2. ไม่สะดวกในการใช้คำนวณสถิติอื่น ๆ เช่นการรวมค่าเฉลี่ยโดยมัธยฐานต้องหาใหม่หมด

ฐานนิยม (Mode)

ข้อได้เปรียบ

1. เข้าใจง่าย
2. สามารถหาค่าเฉลี่ยของฐานนิยมได้จากข้อมูลที่ได้จัดเรียงลำดับแล้ว
3. เป็นส่วนเฉลี่ยที่อาศัยเสียงข้างมาก

ข้อเสียเปรียบ

1. คำนวณได้ยาก
2. ไม่เหมาะเข้าสมการคำนวณในขั้นต่อไป
3. ค่าของฐานนิยมเปลี่ยนไปตามขนาดของอันตรภาคชั้น

การวิเคราะห์การกระจายหรือความแปรปรวนในข้อมูล (Measure of Dispersion of Variation of Data)

การวิเคราะห์ข้อมูลเพื่อสรุปย่อข้อมูล โดยการใช้ค่าเฉลี่ย (Averge) ในกรณีนี้ที่ข้อมูลมีเป็นจำนวนมากเพียงอย่างเดียวย่อมไม่เพียงพอ เพราะว่าข้อมูลที่มีจำนวนมาก ๆ จะมีความแปรปรวนในค่าตัวเลขอยู่เป็นอันมาก ถ้าหากเราใช้ค่าเฉลี่ยเพียงอย่างเดียวไม่เพียงพอที่จะอธิบายให้เห็นสภาพความแปรปรวนต่าง ๆ ภายในชุดของข้อมูล ดังนั้นในการวิเคราะห์ข้อมูลเพื่อสรุปย่อข้อมูล จึงจำเป็นต้องคำนวณค่าที่ใช้วัดความแปรปรวน (Variation) หรือกระจาย (Dispersion) ของข้อมูลชุดนั้น ๆ กำกับไปกับค่าเฉลี่ยด้วยค่าทางสถิติที่ใช้วัดความแปรปรวนของข้อมูลที่นิยมใช้กันมีอยู่ 2 ตัว คือ พิสัย (Range) และส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation)

พิสัย (Range)

พิสัย (Range) เป็นค่าที่ใช้วัดการกระจายหรือความแปรปรวนในค่าตัวเลขของข้อมูลที่ง่ายที่สุดและคำจำกัดความของพิสัยของข้อมูลชุดหนึ่งชุดใด ก็คือ

ถ้าให้ $x_1, x_2, x_3 \ldots\ldots\ldots, x_n$ แทนค่าตัวเลขของข้อมูลชุดหนึ่งชุดใด

พิสัย $= x_{max} - x_{min}$

ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation)

ส่วนเบี่ยงเบนมาตรฐานเป็นเครื่องมือสำหรับใช้วัดการกระจายหรือความแปรปรวนของข้อมูลได้ดี และนิยมใช้กันมากในงานวิจัย วิธีการหาส่วนเบี่ยงเบนมาตรฐานทำได้ดังต่อไปนี้

ถ้าให้ $x_1, x_2, x_3 \ldots\ldots\ldots x_n$ เป็นข้อมูลชุดหนึ่ง

X = แทนมัชฌิมเลขคณิตของข้อมูลนี้

S = แทนส่วนเบี่ยงเบนมาตรฐาน

ให้ i = $1, 2, 3, \ldots\ldots\ldots\ldots n$

$$S^2 = \frac{(X_1 - \bar{X}^2) + (X_2 - \bar{X}^2) + .. = (X_n - \bar{X})^2}{N}$$

$$\therefore S = \sqrt{\frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \ldots\ldots + (X_n - \bar{X})^2}{N}}$$

$$S = \sqrt{\frac{\Sigma (\bar{X}_i - \bar{X})^2}{N}}$$

เทคนิคในการวิเคราะห์ข้อมูลที่สูงขึ้นไป ได้แก่ การวิเคราะห์ข้อมูลโดยการวิเคราะห์ความสัมพันธ์เชิงฟังก์ชันของข้อมูล (Analysis of Functional Relation) ที่ใช้กันมาก คือ

1. Regression Analysis
2. Correlation Analysis

Regression Analysis คือการวิเคราะห์ความสัมพันธ์เชิงฟังก์ชันระหว่างข้อมูลหรือตัวแปร 2 ตัว หรือมากกว่า 2 ตัว ซึ่งอาจจะมีความสัมพันธ์กันเชิงฟังก์ชันหรือความสัมพันธ์เชิงเหตุผลกัน ส่วน Correlation Analysis เป็นการแจกแจงความถี่ร่วมกันระหว่างตัวแปร การทำ Correlation Analysis ไม่ได้สรุปว่าตัวแปรตัวใดเป็นเหตุ ตัวใดเป็นผล เพียงแต่หาว่าตัวแปรทั้งหลายมีความสัมพันธ์กันหรือไม่เท่านั้น

สถิติที่ใช้ในการทดสอบความสัมพันธ์ระหว่างตัวแปร มีมากมายหลายชนิดซึ่งแต่ละชนิดก็เหมาะสมที่จะใช้กับข้อมูลและลักษณะการกระจายของข้อมูลบางประเภท แต่ที่นิยมใช้สำหรับการวิจัยทางสังคมศาสตร์ ได้แก่

การทดสอบไคสแควร์ (Chi - Square Test) ใช้ได้กับข้อมูลแทบทุกชนิด และสามารถสอบคุณสมบัติหลายประการของข้อมูลในคราวเดียวกัน สามารถใช้ทดสอบความเป็นอิสระ (Test of Independence) คือดูว่าความสัมพันธ์ระหว่างตัวแปร 2 ตัว หรือดูความแตกต่างระหว่างตัวแปร 2 ตัวว่าต่างกันหรือเหมือนกันเพียงใด

สูตรทั่วไปของไคสแควร์ (Chi - Square)

$$X^2 = \frac{\Sigma (O - E)^2}{E}$$

O = ค่าของข้อมูลที่รวบรวมได้ (Observe Value)

E = ค่าของข้อมูลที่คาดว่าจะเป็น (Expect Value) ซึ่งคำนวณได้จากผลคูณของแต่ละ r แถว (row) และแต่ละ $E = \frac{r.k}{N}$ k สดมภ์ (Column) หารด้วยจำนวนทั้งหมด

$$E = \frac{r \times k}{N}$$

ในการแปลความหมายค่าของไคสแควร์ เพื่อทราบว่าผลต่างจะมีนัยสำคัญหรือไม่ ต้องอาศัยตารางไคสแควร์ที่ระดับความมีนัยสำคัญต่าง ๆ และในการดูจากตารางนั้นต้องหาชั้นแห่งความเป็นอิสระ (Degree of Freedom) จากสูตร

$$df = (r - i)(k - i)$$

ข้อจำกัดในการใช้ไคสแควร์ คือในแต่ละ Cell จะต้องมีค่าไม่น้อยกว่า 5 และค่า N ไม่ต่ำกว่า 30 ไคสแควร์จะไม่สามารถบอกได้ว่าความสัมพันธ์นั้นเป็นไปในทิศทางใด และมีความสัมพันธ์อยู่ระดับใด นอกจากนี้ไคสแควร์ยังใช้ในการทดสอบการเป็นตัวแทนประชากรของกลุ่มตัวอย่างที่ใช้ในการวิเคราะห์ (Test of Statistical Significance)

แกมมา (Gamma) ใช้ในการหาค่าความสัมพันธ์ระหว่างตัวแปร 2 ตัวเช่นเดียวกัน แต่สามารถบอกทิศทางของความสัมพันธ์ได้ว่าเป็นไปในทางบวกหรือลบ (Positive or Nagative Association) ใช้สำหรับตัวแปรที่มีระดับในการวัดตั้งแต่ Ordinal scale ขึ้นไป

สูตร Gamma = $\frac{CP - IP}{CP - IP}$

เมื่อ CP = Consistant pair

IP = Inconsistant Pair

ค่าของ Gamma ที่คำนวณได้

.70 หรือมากกว่าถือว่ามีความสัมพันธ์สูงมาก (Strong Positive Association)

.50 – .69 ถือว่ามีความสัมพันธ์ค่อนข้างสูง (Substantial Positive Association)

.30 – .49 ถือว่ามีความสัมพันธ์ปานกลาง (Moderate Positive Association)

.10 – .29 ถือว่าความสัมพันธ์อย่างต่ำ (Low Positive Association)

.01 – .09 ถือว่าแทบจะไม่มีความสัมพันธ์
(Negligible Positive Association)

.00 ถือว่าไม่มีความสัมพันธ์
(No Association)

แต่ถ้าค่าออกมาเป็น – (ลบ) ก็ถือว่ามีความสัมพันธ์ด้านลบ (Negative Association)

สรุป

เทคนิคทางสถิติมีประโยชน์ในการที่จะรวบรวมรายละเอียดของข้อมูล โดยอาศัยสรุปรายละเอียดนั้น ๆ จากข้อมูลจำนวนมากมายที่มีอยู่ และคาดคะเนความเชื่อถือได้สถิติจำแนกออกได้เป็น 2 ประเภท คือ สถิติเชิงพรรณนา และสถิติเชิงวิเคราะห์ สถิติเชิงพรรณนาใช้บรรยายลักษณะของข้อมูล ส่วนสถิติเชิงวิเคราะห์ใช้อนุมานและสรุปข้อมูลที่ได้มาจากตัวอย่างช่วยให้นักวิจัยคาดคะเนความเชื่อถือได้และสามารถวางกฎทั่ว ๆ ไปจากข้อมูลด้วย รวมทั้งการพิสูจน์สมมติฐานหรือการหาคำตอบให้กับคำถามในการวิจัย นอกจากนี้สถิติยังช่วยให้นักวิจัยสามารถคาดการณ์เกี่ยวกับปรากฏการณ์ที่น่าจะเกิดขึ้นในอนาคตอีกด้วย