

ความรู้เบื้องต้นในการเพิ่มทักษะด้านข้อมูล

ปัจจุบันหลายหน่วยงานได้ผลักดันให้มีการรวบรวมข้อมูลในรูปแบบดิจิทัล มีการทำโครงการ Big Data มีการเปิดเผยข้อมูลสาธารณะกันมากมายที่สามารถนำไปใช้ในการวิเคราะห์ได้ แต่ยังมีหลายองค์กรที่ถึงแม้จะมีข้อมูลจำนวนมากหลากหลายที่สามารถนำมาวิเคราะห์ได้ แต่การทำงานภายในองค์กรยังเป็นรูปแบบเดิมๆ ที่ไม่ได้ใช้ข้อมูลมาประกอบในการตัดสินใจ บางองค์กรทำโครงการ Big Data แต่ยังไม่พบว่า ผลงานที่ออกมาจากโครงการกลับถูกนำไปใช้เพียงเล็กน้อยและมีเพียงคนบางกลุ่มที่ได้ใช้งาน

บิกดาต้า (Big Data) คือ คำนิยามของข้อมูลที่มีขนาดใหญ่ทุกชนิดที่อยู่ในองค์กรของเราไม่ว่าจะเป็น ข้อมูลบริษัท ข้อมูลลูกค้า Suppliers พฤติกรรมผู้บริโภค Transaction ไฟล์เอกสารต่างๆที่เกี่ยวข้องทั้งหมด รวมไปถึงจนถึง รูปภาพ URLs ลิงค์ต่างๆที่คุณเก็บไว้ ฯลฯ ที่มีปริมาณมากจนกระทั่งซอฟต์แวร์ปกติทั่วไปไม่สามารถรองรับการเก็บข้อมูลหรือประมวลผลได้อย่างเต็มประสิทธิภาพ ดังนั้น Big data จึงเป็นโครงสร้างพื้นฐานที่รองรับการวิเคราะห์ ซึ่งการวิเคราะห์นั้นใช้คณิตศาสตร์ และการวิเคราะห์เรียกอีกอย่างว่า ศาสตร์ข้อมูล

Data Analytics เป็น Business Intelligence ซึ่งเป็นศาสตร์ของการใช้ข้อมูลต่างๆ จำนวนมาก ที่เรียกว่า Big data มาวิเคราะห์ร่วมกันและแสดงผลเพื่อช่วยในด้านธุรกิจหรือตามวัตถุประสงค์ อื่นๆ ที่ต้องการ ซึ่งเริ่มต้นจากความสามารถของเราในการนำข้อมูลเหล่านั้นมาทำให้อยู่ในรูปแบบที่พร้อมจะประมวลผลได้โดยเทคโนโลยีหรือชุดคำสั่งและแบบจำลองที่สร้างขึ้นเพื่อนำข้อมูลที่ได้ผ่านการวิเคราะห์แล้วมาใช้ หรือแปลความหมายโดยบุคคลที่ได้รับการฝึกอบรมการใช้เทคโนโลยีเหล่านั้น รูปแบบของการวิเคราะห์ข้อมูลสามารถแบ่งได้ดังนี้

ปัจจัยสำคัญในการปรับเปลี่ยนไปสู่องค์กรที่ขับเคลื่อนด้วยข้อมูล (Data-driven Organization) คือ วัฒนธรรมการใช้ข้อมูลขององค์กร (Data Culture) และทักษะของบุคลากรในองค์กร โดยเฉพาะผู้นำองค์กร และผู้บริหารซึ่งจำเป็นต้องใช้ข้อมูลเป็นเครื่องมือในการตัดสินใจ ในการวางกลยุทธ์ แต่ยังไม่พบว่า ทักษะทางด้านข้อมูล (Data Literacy) ไม่ได้ถูกสร้างเป็นค่านิยมหรือเป็นวัฒนธรรมในการใช้ข้อมูลมาประกอบการตัดสินใจในการทำงาน เพราะการตัดสินใจหรือการดำเนินงานในองค์กรส่วนใหญ่หรือบ่อยครั้งที่ใช้ความรู้สึกจากประสบการณ์ ความชอบ หรือความคิดในการตัดสินใจมากกว่าการใช้ข้อมูล

“ทักษะด้านข้อมูล” ในความเข้าใจของคนส่วนใหญ่อาจจะคิดว่าเป็นเรื่องของนักวิเคราะห์ข้อมูล นักวิทยาศาสตร์ข้อมูล หรือต้องเป็นผู้เชี่ยวชาญด้านคณิตศาสตร์และสถิติเพียงเท่านั้น แต่ในความเป็นจริงแล้วการทำงานในปัจจุบันบุคลากรในองค์กรจำเป็นต้องมีทักษะด้านข้อมูล แต่ก็ไม่ได้หมายความว่า บุคลากรทุกคนจะต้องเป็นผู้เชี่ยวชาญในด้านเหล่านั้น หรือจะต้องสามารถสร้างข้อมูล หรือต้องทำ Dashboard หรือพยากรณ์เรื่องต่างๆ จากข้อมูลได้

ดังนั้น จึงปฏิเสธไม่ได้ว่าความรู้เบื้องต้นเกี่ยวกับสถิติจึงเป็นองค์ประกอบสำคัญที่จะส่งเสริมการใช้ “ทักษะด้านข้อมูล” การวิเคราะห์ข้อมูลทางสถิติเป็นการประมวลผลข้อมูลที่มีปริมาณมาก โดยการนำคอมพิวเตอร์มาช่วยในการวิเคราะห์ข้อมูล จะช่วยลดขั้นตอน ประหยัดเวลาและค่าใช้จ่าย มีความสะดวก

มากกว่าที่จะคำนวณด้วยเครื่องคิดเลขทั่ว ๆ ไป และในการใช้คอมพิวเตอร์ ส่วนใหญ่การวิเคราะห์ข้อมูลจึงนิยมที่จะใช้โปรแกรมสำเร็จรูป (Package Program) มากกว่าที่จะเขียนโปรแกรมขึ้นมาเอง โปรแกรมสำเร็จรูปที่นำมาใช้วิเคราะห์ข้อมูลจะเป็นโปรแกรมสำเร็จรูปทางสถิติโดยตรง แต่ในปัจจุบันมีโปรแกรมสำเร็จรูปที่ทำงานด้านอื่น ๆ ที่สามารถนำมาวิเคราะห์ข้อมูลทางสถิติได้ เช่น โปรแกรม EXCEL ซึ่งเป็นโปรแกรมประเภทกระดาด (Spread Sheet) แต่เนื่องจากโปรแกรมประเภทนี้ถูกสร้างขึ้นมาสำหรับงานบางอย่างที่ไม่ใช่สถิติโดยตรง ดังนั้นการวิเคราะห์ข้อมูลด้วยวิธีการทางสถิติบางอย่าง โปรแกรมประเภทนี้อาจทำไม่ได้ หรือให้รายละเอียดน้อยกว่า ดังนั้นการวิเคราะห์ข้อมูลส่วนใหญ่จึงนิยมใช้โปรแกรมสำเร็จรูปทางสถิติมากกว่า ส่วนโปรแกรม EXCEL จะมีความสามารถในการนำเสนอข้อมูลด้วยกราฟหรือแผนภูมิ

1.1 ความหมายของสถิติ

คำว่าสถิติ (Statistics) เป็นคำที่แปลงมาจากศัพท์บัญญัติในภาษาเยอรมันว่า Statistik เป็นคำที่มีรากศัพท์มาจากคำว่า "State" ซึ่งหมายถึง ข้อมูล หรือข่าวสารซึ่งจะเป็นประโยชน์ต่อการบริหารงานของรัฐหรือประเทศในด้านต่างๆ ต่อมาในปลายศตวรรษที่ 19 ความหมายของสถิติได้เปลี่ยนไป หมายถึง ข้อมูลที่เป็นตัวเลขเกี่ยวกับการบริหารงานของรัฐ เช่น การสำรวจสำมะโนครัว เพื่อทราบจำนวนและความมั่งคั่งของพลเมือง จนในศตวรรษที่ 20 "สถิติ" หมายถึง ข้อมูลในเรื่องต่างๆซึ่งได้มีการเก็บรวบรวมแล้วนำมาหาความหมาย เช่น ข้อมูลทางด้านธุรกิจ ข้อมูลสภาพภูมิอากาศ ข้อมูลการเดินทาง เป็นต้น

โดยทั่วไปแล้ว คำว่า สถิติมีความหมายได้ 2 นัย คือ สถิติที่เป็นตัวเลขและสถิติที่เป็นศาสตร์ในความหมายที่แตกต่างกันมีดังนี้

1. สถิติที่เป็นตัวเลข หมายถึง สถิติที่ได้มาจากการเก็บรวบรวมข้อมูลด้วยวิธีการใดๆจากข้อมูลจำนวนมาก เช่น สถิติข้อมูลการดำเนินงานทำของบัณฑิตในรอบ 10 ปี
2. สถิติที่เป็นศาสตร์ หมายถึง วิชาที่ว่าด้วยการเก็บรวบรวมข้อมูล การวิเคราะห์ข้อมูล การนำเสนอข้อมูลชุดใดชุดหนึ่ง และความรู้เกี่ยวกับบ่อนุมาณคุณลักษณะของประชากรเป้าหมายจากข้อมูลบางส่วนที่สุ่มมา สถิติ คือ ตัวเลขต่าง ๆ ที่ได้มีการรวบรวมขึ้นเพื่อบอกข้อเท็จจริง เกี่ยวกับข้อมูลต่างๆที่สามารถแสดงออกมาเป็นตัวเลขได้

1.2 ความสำคัญของสถิติ

สถิติมีบทบาทที่สำคัญดังต่อไปนี้

1. ใช้ในการวิเคราะห์ข้อมูล ได้แก่
 - (1) สถิติพื้นฐานใช้อธิบายคุณลักษณะต่างๆ เช่น ความถี่ ร้อยละ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน เป็นต้น
 - (2) สถิติที่หาค่าความสัมพันธ์ระหว่างตัวแปร เช่น สหสัมพันธ์อย่างง่าย
2. ใช้ตรวจสอบคุณภาพของเครื่องมือก่อนนำไปใช้จริงคุณภาพที่จำเป็นต้องตรวจสอบ เช่น ความเที่ยงตรง ความเชื่อมั่น ค่าอำนาจจำแนกความยากง่าย เป็นต้น

3. ใช้กำหนดขนาดของกลุ่มตัวอย่างในการวิเคราะห์หรือการวิจัยที่ไม่สามารถใช้ประชากรทั้งหมดจึงจำเป็นต้องกำหนดขนาด

1.3 สถิติมี 2 ประเภท คือ

1. สถิติพรรณนา (Descriptive statistics) เป็นสถิติที่ใช้ในการสรุปข้อมูลที่ได้มาจากกลุ่มตัวอย่าง โดยไม่มีการอ้างอิงไปยังประชากร แต่เป็นการบรรยายลักษณะข้อมูลเท่านั้น เช่น การแจกแจงความถี่ (frequency distribution) การวัดแนวโน้มเข้าสู่ส่วนกลาง การวัดการกระจายของข้อมูล ฯลฯ การนำเสนอผลการวิเคราะห์ข้อมูลพรรณนาจะอยู่ในรูปตาราง (table) และแผนภูมิ (Chart) ชนิดต่างๆ

2. สถิติอ้างอิง (Inferential statistics) หรือสถิติอนุมาน เป็นสถิติที่ใช้เพื่อนำผลสรุปที่คำนวณได้จากการสุ่มตัวอย่าง ไปอธิบายหรือสรุปลักษณะของประชากรทั้งหมด วิธีที่ใช้ในการสรุปอ้างอิงไปยังกลุ่มประชากรนั้น คือ การประมาณค่า (Estimation) และการทดสอบสมมติฐาน (Hypothesis testing) การวิเคราะห์ความแปรปรวน (analysis of variance) การวิเคราะห์ความถดถอยและสหสัมพันธ์ (regression and correlation analysis)

สถิติอ้างอิง จำแนกเป็น 2 ชนิดคือ

แบบอ้างอิงพารามิเตอร์ (Parametric statistics) (ทดสอบสมมติฐานโดยใช้สถิติ t-test, z-test, ANOVA, regression analysis ตัวแปรที่ต้องการวัดเป็น interval scale กลุ่มตัวอย่างจะต้องมีการแจกแจงเป็นโค้งปกติ กลุ่มประชากรจะต้องมีความแปรปรวนเท่ากัน)

แบบไม่อ้างอิงพารามิเตอร์ (Nonparametric statistics) (ใช้สถิติ chi-square, median test, sign test กลุ่มตัวอย่างเป็น free distribution เป็นกลุ่มตัวอย่างขนาดเล็ก ไม่ทราบลักษณะการแจกแจงของประชากรที่สนใจจะศึกษา)

พารามิเตอร์ หมายถึง ค่าที่ใช้อธิบายคุณลักษณะประชากร (Population) เช่น ค่าเฉลี่ยของประชากร (population mean)

ค่าสถิติ หมายถึง ค่าที่ได้จากตัวอย่าง (sample) เช่น ค่าเฉลี่ยของตัวอย่าง (sample mean) ลักษณะข้อมูลที่ใช้ในงานวิจัย

– **ข้อมูลเชิงปริมาณ** แบ่งเป็นข้อมูลต่อเนื่อง (continuous data) คือค่าที่มีจุดทศนิยมได้ และข้อมูลไม่ต่อเนื่อง (discrete data) คือค่าที่เป็นจำนวนเต็มหรือจำนวนนับ

– **ข้อมูลเชิงคุณภาพ** เป็นข้อมูลที่แสดงถึงสถานภาพ คุณลักษณะ หรือคุณสมบัติ เช่น เพศ ตำแหน่ง หรือจำแนกตัวแปรตามระดับการวัด ได้แก่

– **นามบัญญัติ (nominal scale)** จำแนกความแตกต่างของสิ่งที่ต้องการวัดออกเป็นกลุ่ม เช่น 1=ชาย 2=หญิง ตัวเลขไม่สามารถนำมาบวก ลบ คูณ หาร ได้

- **เรียงอันดับ (ordinal scale)** ใช้สำหรับจัดอันดับที่หรือตำแหน่งของสิ่งที่ต้องการวัด เช่น ระดับการศึกษา ผลการเรียน ความเก่ง ตัวเลขอันดับที่แตกต่างกันไม่สามารถบ่งบอกถึงปริมาณความแตกต่างได้ เช่น ไม่สามารถบอกได้ว่าเก่งกว่ากันเท่าไร ตัวเลขสามารถนำมาบวกหรือลบกันได้

- **อันดับภาค หรือระดับช่วง (interval scale)** กำหนดค่าตัวเลขโดยมีช่วงห่างระหว่างตัวเลขเท่าๆ กัน สามารถนำตัวเลขมาเปรียบเทียบกันได้ว่ามีปริมาณมากน้อยเท่าใด แต่ไม่สามารถบอกได้ว่าเป็นกี่เท่าของกันและกัน เพราะไม่มีศูนย์ที่แท้จริง เช่น คะแนนสอบ อุณหภูมิ (อุณหภูมิ 0 องศา ไม่ได้หมายความว่า จะไม่มีความร้อน) ตัวเลขสามารถนำมาบวก ลบ คูณ หรือหารกันได้

- **อัตราส่วน (ratio scale)** สามารถกำหนดค่าตัวเลขให้กับสิ่งที่ต้องการวัด มีศูนย์แท้ เช่น น้ำหนัก ความสูง อายุ สามารถนำตัวเลขมาบวก ลบ คูณ หาร หรือหาอัตราส่วนกันได้ (เช่น ถนน 50 กิโลเมตร ยาวเป็น 2 เท่าของถนน 25 กิโลเมตร) หรือจำแนกตามหน้าที่ ได้แก่ ตัวแปรอิสระ (ตัวแปรต้น) และตัวแปรตาม ซึ่งเป็นตัวแปรที่ต้องการศึกษา นอกจากนั้นอาจมีตัวแปรที่ไม่ได้ต้องการศึกษาแต่ต้องควบคุม เช่น ตัวแปรภายนอก (ตัวแปรเกิน หรือตัวแปรแทรกซ้อน) และตัวแปรเชื่อมโยง (ตัวแปรสอดแทรก)

1.4 ข้อมูลและข่าวสาร

1.4.1 ความหมายและองค์ประกอบของข้อมูล

ข้อมูล หมายถึง ข่าวสารหรือข้อเท็จจริงที่เกิดขึ้น ซึ่งอาจเกี่ยวกับคนหรือสิ่งของ ข้อมูลอาจจะอยู่ในรูปของตัวเลขหรือข้อความซึ่งได้จากการนับ การวัด การสังเกต หรือการบันทึก ข้อมูลที่ได้รวบรวมไว้เป็นชุดมีการปรับปรุงแก้ไขให้ตรงกับข้อเท็จจริงที่เป็นอยู่ในปัจจุบัน ซึ่งจะเรียกชุดนั้นว่าฐานข้อมูล (Data Base) ซึ่งฐานข้อมูลนี้จะใช้เป็นหลักฐานอ้างอิงเพื่อค้นหาความจริงได้

ข้อมูลในระบบฐานข้อมูลหนึ่งๆจะมีส่วนประกอบย่อยหลายๆส่วน ซึ่งแบ่งออกเป็นระดับจากระดับย่อยที่สุดดังนี้

- 1.อักขระ (Character) หมายถึง ตัวอักษร ตัวเลข หรือเครื่องหมายตัวใดตัวหนึ่ง
- 2.เขตข้อมูล (Field) หมายถึง กลุ่มของอักขระหรือสัญลักษณ์ที่รวมกันเพื่อแทนความหมายของตัวเลข 1 จำนวน หรือข้อความ 1 ข้อความ
- 3.ระเบียน (Record) หมายถึง กลุ่มของเขตข้อมูลที่มีความหมายเกี่ยวข้องกัน บันทึกอยู่ด้วยกันเพื่ออ้างอิงเรื่องใดเรื่องหนึ่ง
- 4.แฟ้มข้อมูล (File) หมายถึง กลุ่มของระเบียนที่เกี่ยวข้องกับเรื่องเดียวกัน

1.4.2 การแบ่งลักษณะของข้อมูล

การแบ่งลักษณะของข้อมูล แบ่งได้เป็น 2 ประเภท ดังนี้

1.ข้อมูลเชิงคุณภาพ (Qualitative Data) คือ ข้อมูลที่มีค่าไม่ต่อเนื่องและไม่สามารถนำไปคำนวณทางคณิตศาสตร์ได้ เป็นลักษณะของข้อมูลที่บรรยายลักษณะของสิ่งที่ศึกษา จึงไม่อยู่ในรูปของตัวเลข ไม่สามารถบอกได้ว่ามีค่ามากน้อยเท่าไร แต่สามารถกำหนดตัวเลขขึ้นแทนลักษณะของสิ่งที่ศึกษาได้ นั่นคือข้อมูลในมาตรานามบัญญัติ (Nominal Scale) และมาตราเรียงลำดับ (Ordinal Scale) เช่น การแบ่งกลุ่ม

ข้อมูลนักศึกษาของวิทยาลัยอาชีวศึกษาร้อยเอ็ด ได้แก่ ระดับประกาศนียบัตรวิชาชีพ (ปวช.) และระดับประกาศนียบัตรวิชาชีพชั้นสูง (ปวส.) และให้ 1 แทนระดับประกาศนียบัตรวิชาชีพ (ปวช.) และให้ 2 แทนระดับประกาศนียบัตรวิชาชีพชั้นสูง (ปวส.) หรือตัวอย่างการแบ่งกลุ่มคณะสีของนักศึกษาวินิจฉัยอาชีวศึกษาร้อยเอ็ด ได้แก่ คณะสีแดงทับทิม คณะสีเหลืองอำพัน คณะสีน้ำเงินไพริน และคณะสีเขียวมรกต โดยให้นักศึกษาที่อยู่คณะสีแดงทับทิมแทนด้วยหมายเลข 1 คณะสีเหลืองอำพันแทนด้วยหมายเลข 2 คณะสีน้ำเงินไพรินแทนด้วยหมายเลข 3 และคณะสีเขียวมรกตแทนด้วยหมายเลข 4 เป็นต้น

2. ข้อมูลเชิงปริมาณ (Quantitative Data) คือ ข้อมูลที่มีค่าต่อเนื่องกันและสามารถนำไปคำนวณทางคณิตศาสตร์ได้ สามารถเปรียบเทียบความแตกต่างได้ เป็นข้อมูลตัวเลขที่บอกค่าความมากและน้อยเปรียบเทียบในรูปของอัตราส่วนได้ ซึ่งก็คือ ข้อมูลในมาตราอันดับ (Interval Scale) และมาตราอัตราส่วน (Ratio Scale) เช่น ข้อมูลอายุ น้ำหนัก ความสูง พื้นที่ เป็นต้น

1.4.3 แหล่งข้อมูล

ข้อมูลที่เกี่ยวข้องมาได้สามารถจำแนกตามแหล่งที่มาได้ 2 ประเภท คือ

1. ข้อมูลปฐมภูมิ (Primary Data) เป็นข้อมูลที่ได้จากแหล่งข้อมูลโดยตรงหรือที่เรียกว่าแหล่งปฐมภูมิ (Primary Data) ซึ่งเป็นแหล่งที่เกิดของข้อมูล ข้อมูลปฐมภูมิที่ได้มานี้ถือว่าเป็นข้อมูลที่เชื่อถือได้มากที่สุดเพราะได้จากแหล่งข้อมูลโดยตรงและยังไม่มีมีการเปลี่ยนแปลงรูป เปลี่ยนความหมาย เช่น ข้อมูลที่ได้จากแบบสอบถาม แบบสัมภาษณ์ แบบสังเกต แบบทดสอบ ฯลฯ

2. ข้อมูลทุติยภูมิ (Secondary Data) เป็นข้อมูลที่ได้จากแหล่งที่มีผู้เก็บข้อมูลมาแล้ว หรือที่เรียกว่าแหล่งทุติยภูมิ (Secondary Data) ข้อมูลที่ได้จากแหล่งทุติยภูมินี้อาจมีการเปลี่ยนแปลงรูปหรือเปลี่ยนความหมายซึ่งอาจมีความคลาดเคลื่อนจากความเป็นจริงได้ ในบางครั้งผู้วิจัยไม่สามารถนำข้อมูลจากแหล่งปฐมภูมิได้จึงจำเป็นต้องใช้ข้อมูลจากแหล่งทุติยภูมิ เช่น ข้อมูลในอดีต ผู้วิเคราะห์ข้อมูลจำเป็นต้องนำข้อมูลที่ได้จากการจัดบันทึกไว้แล้วมาทำการวิเคราะห์ เป็นต้น

1.5 มาตรการจัด

ในการเตรียมข้อมูลเพื่อการวิเคราะห์ จะต้องเข้าใจข้อมูลที่เก็บมาได้ ซึ่งข้อมูลที่เก็บรวบรวมมาได้นั้นอาจมีลักษณะต่างกันเพื่อการเลือกใช้สถิติที่เหมาะสมในการประมวลผล ซึ่งสามารถจัดระดับของข้อมูลได้ 4 ระดับ ตามวิธีการวัดค่าดังต่อไปนี้

1.5.1 มาตรนามบัญญัติ (Nominal Scale)

มาตรนามบัญญัติ (Nominal Scale) เป็นระดับของข้อมูลที่ได้จากวัดแบบง่ายที่สุดคือ เป็นการแบ่งแยกประชากรที่จะศึกษาออกเป็นกลุ่มหรือเป็นพวก โดยแต่ละกลุ่มและแต่ละพวกมีความเท่าเทียมกัน เช่น

แบ่งประชากร โดยใช้เพศเป็นตัวแบ่ง คือ ชายและหญิง

แบ่งประชากร โดยใช้ภาคเป็นตัวแบ่ง คือ ภาคเหนือ ภาคใต้ ภาคกลาง ภาค

ตะวันออกเฉียงเหนือ และภาคตะวันออก

จะเห็นว่าแต่ละกลุ่มแยกออกจากกันและกันแสดงถึงความแตกต่างของประชากร ในการนำไปใช้อาจกำหนดตัวเลขหรือสัญลักษณ์แทนกลุ่ม เช่น ถ้าเป็นเพศชาย กำหนดให้เป็น M เพศหญิง กำหนดให้เป็น F หรือการกำหนดเบอร์ให้กลับนักฟุตบอล ผู้รักษาประตูเป็น 1 กองหน้าเป็น 2, 3, ฯลฯ ตัวเลขหรือสัญลักษณ์เหล่านี้เป็นเพียงชื่อที่แทนกลุ่มเท่านั้น ไม่สามารถนำมาใช้ในการคำนวณทางคณิตศาสตร์ได้

1.5.2 มาตราเรียงลำดับ (Ordinal Scale)

มาตราเรียงลำดับ (Ordinal Scale) เป็นระดับของข้อมูลที่กำหนดรายละเอียดของการวัดเพิ่มขึ้นจากระดับนามบัญญัติ กล่าวคือ นอกจากจะแบ่งแยกข้อมูลออกเป็นกลุ่มแล้ว ยังสามารถหาระดับความแตกต่างระหว่างกลุ่มได้ด้วย ซึ่งระบบการวัดแบบนี้ใช้หลักของความมากกว่า ความน้อยกว่า เช่น การแบ่งประชากรโดยใช้ความคิดเห็น ซึ่งอาจมีระดับต่างๆ เช่น เห็นด้วยอย่างยิ่ง เห็นด้วย ไม่มีความเห็น ไม่เห็นด้วย ไม่เห็นด้วยอย่างยิ่ง การแบ่งสินค้าโดยใช้คุณภาพของสินค้า คือ ดี พอใช้ ต้องปรับปรุง

จากตัวอย่างดังกล่าว สามารถบอกความแตกต่างของแต่ละกลุ่มข้อมูลได้โดยจัดอันดับของข้อมูล แต่ไม่สามารถกำหนดปริมาณความน้อยกว่าหรือมากกว่าออกมาเป็นตัวเลขที่แน่นอนได้ เช่น สามารถบอกได้ว่าสินค้าที่มีคุณภาพนั้นย่อมจะดีกว่าสินค้าพอใช้ แต่ไม่สามารถบอกได้ว่าดีกว่าเป็นตัวเลขเท่าไร การใช้ตัวเลขหรือสัญลักษณ์ใดๆ กำหนดอันดับความมากกว่าหรือน้อยกว่า จะไม่มีผลต่อข้อมูล เช่น การให้ร้อยตรีติดดาว 1 ดวง ร้อยโทติดดาว 2 ดวง และร้อยเอกติดดาว 3 ดวง อาจกำหนดใหม่ให้ร้อยเอกติดดาว 1 ดวง ร้อยโทติดดาว 2 ดวง และร้อยตรีติดดาว 3 ดวง โดยที่อันดับก่อน และหลังย่อมไม่เปลี่ยนแปลง ตัวเลขที่แทนข้อมูลระดับนี้ยังไม่สามารถนำมาคำนวณทางคณิตศาสตร์ได้

1.5.3 มาตราอันตรภาค (Interval Scale)

มาตราอันตรภาค (Interval Scale) เป็นระดับของข้อมูลที่มีคุณสมบัติที่เพิ่มเติมจากการวัดระดับเรียงระดับ กล่าวคือ สามารถกำหนดปริมาณของความแตกต่างระหว่างอันดับได้ เพราะการวัดแบบนี้หน่วยของการวัดมีลักษณะคงที่ซึ่งเป็นมาตรฐานในการกำหนดค่าเป็นตัวเลข เช่น หน่วยวัดปริมาตรในระบบเมตริก สามารถบอกได้ว่า 1,000 มิลลิลิตร เท่ากับ 1 ลิตร ดังตารางที่ 1.1

ตารางที่ 1.1 การเปรียบเทียบหน่วยวัดปริมาตรในระบบเมตริก

1,000 ลูกบาศก์มิลลิเมตร	เท่ากับ	1 ลูกบาศก์เซนติเมตร
1,000,000 ลูกบาศก์เซนติเมตร	เท่ากับ	1 ลูกบาศก์เมตร
10 มิลลิลิตร	เท่ากับ	1 เซนติลิตร
10 เดซิลิตร	เท่ากับ	1 ลิตร
1,000 มิลลิลิตร	เท่ากับ	1 ลิตร
10 ลิตร	เท่ากับ	1 เดคาลิตร
1 ลูกบาศก์เซนติเมตร	เท่ากับ	1 มิลลิลิตร
1,000 ลูกบาศก์เซนติเมตร	เท่ากับ	1 ลิตร
1,000 ลิตร	เท่ากับ	1 ลูกบาศก์เมตร

1.5.4 มาตรฐานอัตราส่วน (Ratio Scale)

มาตรฐานส่วน Ratio Scale เป็นระดับของข้อมูลที่ดีกว่ามีความสมบูรณ์ที่สุดและเป็นการวัดระดับสูงสุดมีจุดเริ่มต้นเป็นธรรมชาติ คือ มีศูนย์แท้ที่หมายความถึงการไม่มีค่า เช่น น้ำหนัก ส่วนสูง อายุ ฯลฯ ข้อมูลบางชนิดไม่สามารถวัดได้ถึงระดับนี้ เช่น ข้อมูลทางพฤติกรรม ทักษะสติ ข้อมูลในระดับนี้สามารถนำไปคำนวณทางคณิตศาสตร์ได้

1.6 ระเบียบวิธีทางสถิติ

ระเบียบวิธีทางสถิติ เป็นการศึกษาถึงขั้นตอนวิธีการดำเนินการทางสถิติ ซึ่งประกอบด้วยขั้นตอนต่างๆ ดังนี้

1.6.1 การเก็บรวบรวมข้อมูล (Collection of Data)

การเก็บรวบรวมข้อมูล หมายถึง การเก็บรวบรวมข้อความหรือตัวเลขที่ต้องการจากประชากรที่มีคุณสมบัติสอดคล้องตามความต้องการ เช่น จะศึกษาจำนวนชาวเขาที่มีอยู่ในจังหวัดเชียงใหม่ก็ต้องรวบรวมตัวเลขจำนวนชาวเขาที่มีอยู่ในจังหวัดเชียงใหม่ เป็นต้น การเก็บรวบรวมข้อมูล จำแนกได้เป็น 4 วิธีการที่สำคัญ ดังนี้

1. การเก็บรวบรวมจากทะเบียนประวัติ เป็นการเก็บรวบรวมข้อมูลจากแหล่งที่มีข้อมูลเหล่านี้อยู่ เช่น จากกองทะเบียนกรมตำรวจจากทะเบียนประวัติคนไข้ของโรงพยาบาลต่างๆ

2. การเก็บรวบรวมโดยวิธีการสำรวจ เป็นการเก็บรวบรวมข้อมูลโดยใช้แบบสำรวจ ซึ่งผู้ทำการสำรวจได้จัดเตรียมไว้ล่วงหน้า โดยสามารถทำได้ 4 วิธี คือ

(1) การสำรวจโดยการสัมภาษณ์ ผู้สำรวจต้องออกไปสัมภาษณ์ผู้เกี่ยวข้องกับข้อมูล

(2) การสำรวจโดยการสร้างแบบสอบถาม ผู้สำรวจเป็นผู้ส่งแบบสอบถาม การสำรวจโดยการสร้างแบบสอบถาม ผู้สำรวจเป็นผู้ส่งแบบสอบถามไปยังผู้ที่เกี่ยวข้องกับข้อมูล เพื่อตอบคำถามแล้วส่งกลับคืนมายังผู้ทำการสำรวจ

(3) การเก็บรวบรวมข้อมูลโดยใช้วิธีการทดลอง เป็นการเก็บรวบรวมข้อมูลที่เป็นผลจากการทดลอง เช่น การทดลองทางวิทยาศาสตร์ในโรงเรียนการสอบสมรรถภาพของนักเรียนในด้านพลศึกษา เช่น การวิ่ง การกระโดด การทุ่มน้ำหนัก เป็นต้น

(4) การเก็บรวบรวมข้อมูลโดยวิธีการสังเกต การเก็บรวบรวมข้อมูลโดยวิธีนี้ ผู้ทำการสำรวจต้องทำการบันทึกข้อมูลที่สนใจจากการสังเกต เช่น จดบันทึกรถยนต์ที่ผ่านสี่แยกปทุมวัน เป็นต้น

1.6.2 การนำเสนอข้อมูล (Presentation of Data)

การนำเสนอข้อมูล คือ การนำข้อมูลที่รวบรวมมาจัดเป็นหมวดหมู่ให้เรียบร้อย ง่ายต่อการเข้าใจและสะดวกในการวิเคราะห์ และตีความหมาย วิธีการนำเสนอข้อมูลมี 2 วิธี ดังนี้

1 การนำเสนอข้อมูลอย่างไม่เป็นแบบแผน Informal Presentation คือ การนำเสนอข้อมูลไม่มีกฎเกณฑ์หรือแบบอย่างแน่นอนที่ใช้กันมี 2 วิธี คือ

(1) การนำเสนอในรูปแบบข้อความ Text Presentation เป็นการนำเสนอข้อมูลแบบบรรยายรูปสั้นๆ ซึ่งเหมาะสำหรับข้อมูลที่ไม่มากจนเกินไป เช่น มหาวิทยาลัยอัสสัมชัญ มีจำนวนผู้สำเร็จการศึกษา ประจำปี

การศึกษา 2563 จำนวนทั้งสิ้น 3,203 คน จำแนกเป็น ปริญญาตรี ปริญญาโท และปริญญาเอก จำนวน 3,101 คน 75 คน และ 27 คน

(2) การนำเสนอในรูปแบบข้อความกึ่งตาราง Semi-Presentation เป็นการนำเสนอข้อมูลโดยแยกตัวเลข และบทความ เพื่อให้มองเห็นการเปรียบเทียบได้ชัดเจน เช่น จำนวนผู้สำเร็จการศึกษาประจำปี 2563 มีดังนี้

คณะ/วิทยาลัย	ระดับการศึกษา			รวม
	ปริญญาตรี	ปริญญาโท	ปริญญาเอก	
1. คณะเกษตรศาสตร์	216	5	-	221
2. คณะวิทยาศาสตร์	384	20	16	420
3. คณะวิศวกรรมศาสตร์	541	4	10	555
4. คณะศิลปประยุกต์และสถาปัตยกรรมศาสตร์	62	-	-	62
5. คณะเภสัชศาสตร์	112	5	1	118
6. วิทยาลัยแพทยศาสตร์และการสาธารณสุข	161	1	-	162
7. คณะพยาบาลศาสตร์	65	-	-	65
8. คณะศิลปศาสตร์	414	22	-	436
9. คณะบริหารศาสตร์	649	18	-	667
10. คณะนิติศาสตร์	234	-	-	234
11. คณะรัฐศาสตร์	263	-	-	263
ผลรวมทั้งหมด	3,101	75	27	3,203

2. การนำเสนออย่างเป็นแบบแผน Formal Presentation คือ การเสนอข้อมูลตามกฎเกณฑ์หรือแบบอย่างที่กำหนดไว้ ได้แก่ การนำเสนอในรูปแบบต่างๆ ดังนี้

(1) การนำเสนอข้อมูลในรูปตาราง เป็นการนำเสนอข้อมูลโดยจัดข้อมูลให้อยู่ในรูปที่อ่านความหมายได้ทั้งแนวนอน แถว หรือ Row และแนวตั้ง สดมภ์ หรือ Column ทั้งนี้เพื่อให้ข้อมูลอยู่ในรูปที่เข้าใจง่ายและสะดวกต่อการวิเคราะห์

องค์ประกอบของตารางสถิติโดยทั่วไป มีดังนี้

(ก) หมายเลขตารางที่แสดงลำดับตารางในกรณีที่มีตารางสถิติมากกว่าหนึ่งตาราง

(ข) ชื่อเรื่อง อยู่แถวเดียวกับหมายเลขตาราง ชื่อเรื่องต้องสั้นกะทัดรัดและได้ใจความสมบูรณ์

(ค) หมายเหตุคำนำ เป็นข้อความที่อยู่ใต้ชื่อเรื่อง ซึ่งเป็นข้อความที่ทำให้เข้าใจตารางยิ่งขึ้น

(ง) ต้นขั้ว ประกอบด้วย หัวขั้วและตัวขั้ว หัวขั้วจะอธิบายถึงตัวขั้วต่างๆ และตัวขั้วแต่ละอันจะอธิบายถึงข้อมูลที่ปรากฏในแต่ละแถวตามแนวนอน

(จ) หัวเรื่อง ประกอบด้วย หัวสดมภ์ จะมีหัวสดมภ์อันเดียวหรือหลายอันก็ได้และภายในหัวสดมภ์แต่ละอันอาจแบ่งให้ย่อยลงไปอีกก็ได้หัวเรื่องจะอธิบายเกี่ยวกับข้อมูลที่ปรากฏอยู่ในแต่ละสดมภ์ตามแนวตั้ง

(ฉ) ตัวเรื่อง ประกอบด้วย ข้อมูลที่เป็นตัวเลข

(ช) หมายเหตุล่าง เป็นคำอธิบายข้อความบางตอนในตารางให้ชัดเจนขึ้น

(ข) เหตุแห่งที่มา เป็นเหตุที่บอกให้ทราบว่าข้อมูลที่มาั้นได้มาจากไหน ซึ่งจะช่วยให้ผู้อ่านสามารถค้นคว้าเพิ่มเติมได้

(2) การนำเสนอข้อมูลโดยใช้แผนภูมิและแผนภาพต่างๆ

แผนภูมิแท่ง (Bar Chart) คือแผนภูมิที่ประกอบด้วยสี่เหลี่ยมผืนผ้า ซึ่งอาจอยู่ในแนวตั้งก็ได้ โดยเรียกสี่เหลี่ยมผืนผ้าว่าแท่ง ความสูง และความกว้างของแต่ละแท่งจะต้องได้สัดส่วนกับขนาดของข้อมูล และความกว้างของทุกๆแท่งจะต้องเท่ากันทั้งหมด

แผนภูมิแท่งแบ่งออกเป็น 6 ชนิดคือ

(ก) แผนภูมิแท่งเชิงเดี่ยว คือ แผนภูมิแท่งที่แสดงลักษณะของข้อมูลเพียงชุดเดียว

(ข) แผนภูมิแท่งเชิงซ้อน คือ แผนภูมิที่แสดงการเปรียบเทียบให้เห็นถึงลักษณะของข้อมูลตั้งแต่ชุดขึ้นไป

(ค) แผนภูมิแท่งส่วนประกอบ คือแผนภูมิแท่งที่ใช้แสดงรายละเอียดส่วนย่อยของข้อมูลที่นำเสนอ

(ง) แผนภูมิแท่งบวก - ลบ คือ แผนภูมิแท่งที่ใช้กับข้อมูลซึ่งมีค่าบวกและค่าลบ

(จ) แผนภูมิแท่งซ้อนกัน คือ แผนภูมิแท่งเชิงซ้อน เพียงแต่ให้เห็นแท่งสี่เหลี่ยมแต่ละแท่งซ้อนเหลื่อมกัน ทั้งนี้เพื่อประหยัดเนื้อที่ในการนำเสนอด้วย

(ฉ) แผนภูมิแท่งพีระมิด คือ แผนภูมิแท่งเรียงซ้อนกันเป็นรูปสามเหลี่ยม

1.5.3 การวิเคราะห์ข้อมูล (Analysis of Data)

การวิเคราะห์ข้อมูล (Analysis of Data) เมื่อข้อมูลถูกนำเสนอแล้วโดยวิธีใดวิธีหนึ่ง ซึ่งทำให้เกิดความสะดวกในการอ่านและการวิเคราะห์แล้ว หลังจากขั้นตอนของการวางแผนจะศึกษาอะไรก็จะทำการวิเคราะห์ในสิ่งที่ต้องการศึกษานั้นในขั้นตอนตอนนี้

1.5.4 การแปลความหมายของข้อมูล (Interpretation of Data)

การแปลความหมายของข้อมูล (Interpretation of Data) เป็นขั้นตอนของการแปลความหมาย ผลการวิเคราะห์ เพื่อให้บุคคลทั่วไปเข้าใจและเพื่อให้ได้ข้อยุติที่อาจเป็นประโยชน์ยิ่งขึ้นไป

