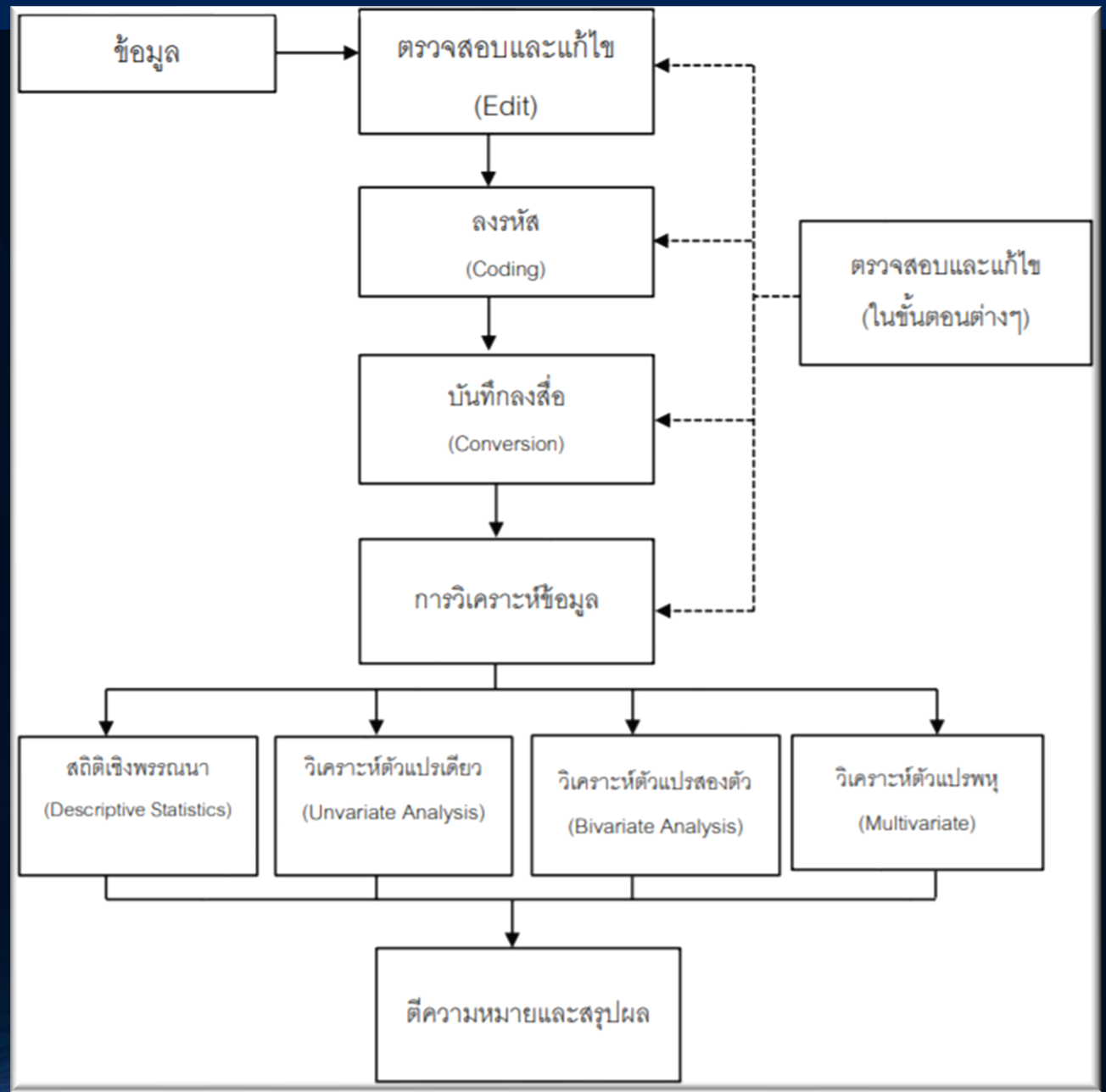
สถิติเบื้องต้นสำหรับการวิจัย

เมื่อผู้วิจัยได้ค้นคว้า เก็บรวบรวมข้อมูลจากแหล่งข้อมูลมาแล้วขั้นตอนต่อไปผู้วิจัยจำเป็นที่จะต้องนำข้อมูลนั้น ๆ มาวิเคราะห์ แปลความหมายและสรุปผลจากข้อมูลที่เก็บรวบรวมมาได้ซึ่งตรงนี้เองที่ “สถิติ” จะเข้ามามีบทบาทในกระบวนการวิจัยเป็นอย่างยิ่ง

สถิติจะช่วยจัดการข้อมูลที่มีอยู่อย่างกระจัดกระจายให้มารวบรวมไว้เป็นหมวดหมู่ โดยใช้วิธีการต่าง ๆ เช่น การแจกแจงความถี่ การหาค่าร้อยละ และการหาค่าความสัมพันธ์ เป็นต้น ตลอดจนช่วยให้ทราบเกี่ยวกับคุณลักษณะต่าง ๆ ของข้อมูล เช่น การวัดแนวโน้มเข้าสู่ส่วนกลาง และการวัดการกระจายของข้อมูล เป็นต้น

ขั้นตอนของการเตรียมข้อมูลและวิเคราะห์



สถิติเบื้องต้นสำหรับการวิจัย

การวิเคราะห์ข้อมูลเป็นขั้นตอนหนึ่งของกระบวนการวิจัย ซึ่งเป็นการนำข้อมูลที่เก็บรวบรวมได้มาทำการแยกแยะเพื่อ**หาข้อสรุป** หรือ**ตอบคำถามตามวัตถุประสงค์การวิจัย**

สำหรับในการเลือกใช้สถิติเพื่อการวิเคราะห์ข้อมูลการวิจัยนั้น **ผู้วิจัยจะเลือกใช้สถิติโดยพิจารณาจากธรรมชาติของข้อมูลและเป้าหมายการวิจัย**

การวัด เป็นการกำหนดตัวเลขให้กับสิ่งที่ต้องการศึกษาภายใต้กฎเกณฑ์ที่แน่นอน เช่น เพศชายแทนด้วยเลข 1 และเพศหญิงแทนด้วยเลข 0 ผู้วิจัยจำเป็นจะต้องทราบคุณลักษณะของข้อมูลที่ถูกวัดเพื่อใช้ในการพิจารณาว่าจะเลือกใช้วิธีการทางสถิติใดจึงจะเหมาะสม ดังนั้นจึงควรทราบว่า**ข้อมูลที่ถูกวัดมานั้นอยู่ในมาตราการวัดระดับใด**

ธรรมชาติของข้อมูล

จากข้อมูลที่ได้เก็บรวบรวมมานั้น ผู้วิจัยจะต้องพิจารณาถึง**ระดับการวัดของตัวแปรและการแจกแจงของข้อมูล**

ค่าที่วัดได้จากตัวแปรสามารถแบ่งออกเป็น 4 ระดับ คือ **มาตรานามบัญญัติ (nominal scale) มาตราเรียงอันดับ (ordinal scale) มาตราอันตรภาค (interval scale) และมาตราอัตราส่วน (ratio scale)**

มาตรานามบัญญัติ (Nominal Scale)

มาตรานามบัญญัติ ใช้สำหรับวัดค่าตัวแปรขั้นพื้นฐานโดยเป็นมาตราวัดที่**กำหนดสัญลักษณ์ขึ้นเพื่อเรียกเรียกแทนชื่อ จำแนก** หรือ **จัดประเภทสิ่งของตามคุณลักษณะ**

มาตรานามบัญญัติจึงเป็นระดับที่ใช้จำแนกความแตกต่างของสิ่งที่ต้องการวัดออกเป็นกลุ่ม ๆ โดยใช้ตัวเลข เพื่อใช้ในการสื่อความหมายและสะดวกในการวัด เช่น จำแนกคนออกเป็น 2 เพศ คือ เพศหญิง-เพศชาย หรือจำแนกคนที่นับถือศาสนาต่าง ๆ ออกเป็น พุทธ อิสลาม คริสต์ และ ฮินดู หรือ ห้องหมายเลข 814 สื่อความหมายว่าเป็นห้อง ๆ หนึ่ง ไม่ได้หมายความว่าตัวเลขห้องที่มีค่ามากกว่าจะสามารถจุคนได้มากกว่าห้องที่มีหมายเลขน้อยกว่า

ในการวิจัย ข้อมูลระดับนี้มักจะใช้ เป็นตัวแปรต้น (Independent Variable)

มาตรานามบัญญัติ

1. เพศ	☐ 1. ชาย	☐ 2. หญิง
2. สถานภาพสมรส	☐ 1. แต่งงาน	☐ 2. โสด
	☐ 3. หย่า	☐ 4 ม่าย
3. อาชีพ	☐ 1. รับราชการ	☐ 2. รับจ้าง
	☐ 3. ธุรกิจส่วนตัว	☐ 4. อื่น ๆ (โปรดระบุ).....

*ตัวเลขเหล่านี้ไม่มีค่าในเชิงปริมาณ เป็นตัวเลขที่ใช้เรียกเท่านั้น

*เป็นการกำหนดตัวเลขในแต่ละช่องเพื่อสะดวกต่อการป้อนข้อมูลทางสถิติ

ไม่สามารถนำหมายเลขที่กำหนดขึ้นเพื่อใช้แทนตัวแปรดังกล่าวมาคิดคำนวณโดยการบวก ลบ คูณ หาร กันได้ ลักษณะที่สามารถทำได้ คือ **การนับเพื่อดูความถี่ (Frequency)** หรือ **การเกิดซ้ำ ๆ** กันของข้อมูลแต่ละประเภทเท่านั้น

มาตรานามบัญญัติ (ต่อ)

ตัวอย่างที่ 1 การสำรวจคุณวุฒิของครูโรงเรียนหนึ่ง พบว่าเป็นครูที่มีวุฒิการศึกษาต่ำกว่าปริญญาตรี 35 คน มีวุฒิปริญญาตรี 29 คน และมีวุฒิสูงกว่าปริญญาตรี 18 คน ดังนั้นวุฒิของครูโรงเรียนแห่งนี้จำแนกออกเป็น 3 ประเภท คือ

วุฒิทางการศึกษา	ความถี่
ต่ำกว่าปริญญาตรี	35
ปริญญาตรี	29
สูงกว่าปริญญาตรี	18

จากการนับคุณวุฒิทางการศึกษา ตัวเลขเหล่านี้เรียกว่า “**ความถี่ที่ได้จากการนับ**” ซึ่งเป็นตัวเลขที่จะนำไปใช้ในการวิเคราะห์ข้อมูลต่อไป ดังนั้นข้อมูลที่อยู่ในมาตรานี้จะเป็นชนิด ความถี่ (Frequency)

มาตราเรียงอันดับ (Ordinal Scale)

การวัดในระดับนี้**เป็นการกำหนดตัวเลขแทนลำดับของข้อมูลจากมากไปหาน้อย** หรือ **จากน้อยไปหามาก** แต่**ไม่ได้บอกถึงปริมาณของแต่ละลำดับว่ามีความแตกต่างกันมากน้อยเพียงใด** ความแตกต่างระหว่างแต่ละหน่วยของการวัดตามมาตรานี้มีค่าไม่เท่ากัน และไม่มีศูนย์แท้ (Non Absolute Zero) **ข้อมูลที่อยู่ในมาตรานี้จะเป็นตัวเลข**

ตัวเลขที่แตกต่างกันแสดงให้เห็นว่าสิ่งนั้นมีคุณสมบัติที่ต่างกัน แต่**ตัวเลขเหล่านั้นไม่สามารถบอกให้ทราบว่า สิ่งต่าง ๆ เหล่านั้นมีคุณสมบัติแตกต่างกันเป็นปริมาณเท่าใด** และ**ตัวเลขที่ต่างกัน 1 หน่วยไม่ได้แทนปริมาณที่เท่ากันเหมือนข้อมูลในมาตราอันตรภาคหรืออัตราส่วน** เช่น ความแตกต่างระหว่างอันดับที่ 1 และอันดับที่ 2 มีค่าไม่เหมือนกับความแตกต่างระหว่างน้ำหนัก 71 กก. และ 70 กก.

ตัวอย่างมาตราเรียงอันดับ

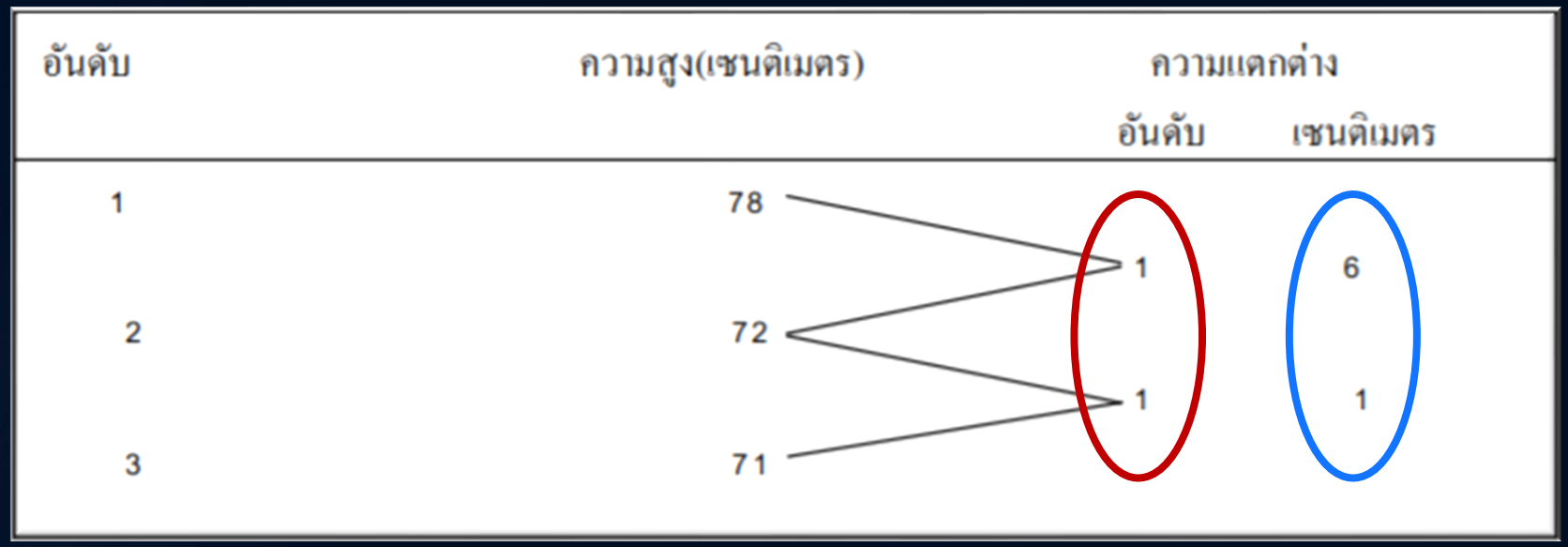
ในการประกวดภาพเขียนด้วยสีน้ำมันของนักเรียน ภาพที่มีความสวยงามมากก็จะแทนด้วยหมายเลข 1 และ ภาพที่มีความสวยงามรองลงมาก็จะกำหนดด้วยหมายเลข 2 และ 3 ตามลำดับ

หรือในการสอบแข่งขันเพื่อชิงทุนรัฐบาลไปศึกษาต่อต่างประเทศ ผู้ที่ได้คะแนนสูงสุดก็จะได้ลำดับที่ 1 ผู้ที่ได้คะแนนรองลงมาก็จะได้ลำดับที่ 2 และ 3 ตามลำดับ

จากตัวอย่างดังกล่าวข้างต้นจะเห็นว่า การวัดในระดับนี้ **ตัวเลขที่กำหนดขึ้นนั้นใช้แทนการบอกลำดับที่หรือตำแหน่งเท่านั้น** **ตัวเลขดังกล่าวไม่สามารถนำมาบวก ลบ คูณ หาร** กันได้

ตัวอย่างมาตราเรียงอันดับ (ต่อ)

การจัดแถวนักเรียนในห้องตามลำดับความสูง คนที่สูงที่สุดจัดเป็นอันดับ 1 สูงรองลงมาจัดเป็นอันดับ 2 จัดเช่นนี้เรื่อยไป จะเห็นได้ว่า อันดับ 1 กับอันดับ 2 ต่างกัน 1 อันดับ เช่นเดียวกับ อันดับ 2 และอันดับ 3 ซึ่งความสูงที่ได้จากการวัดเป็นเซนติเมตรของ 3 คนนี้ คือ



จะเห็นได้ว่าอันดับที่ 1 กับอันดับที่ 2 ต่างกัน 1 หน่วย เช่นเดียวกับระหว่างอันดับที่ 2 กับอันดับที่ 3 แต่ถ้ามองในแง่ของความสูงเป็น ซม. จะเห็นได้ว่า อันดับที่ 1 กับอันดับที่ 2 ต่างกัน 6 ซม. แต่ อันดับที่ 2 กับอันดับที่ 3 ต่างกัน 1 ซม. **จะเห็นว่ามีความต่างกัน 1 อันดับเหมือนกันแต่ไม่ได้แทนปริมาณที่เท่ากัน**

มาตราอันตรภาค (Interval Scale)

การวัดในระดับนี้เป็นการกำหนดตัวเลขที่สามารถบอกระยะห่างได้ว่ามีระยะห่างกันมากน้อยเพียงใด **สามารถนำตัวเลขมาเปรียบเทียบกันได้ว่าว่ามีปริมาณมากน้อยเท่าใด แต่ไม่สามารถบอกได้ว่าเป็นกี่เท่าของกันและกัน**

ตัวอย่างมาตราอันตรภาค

ตัวอย่างที่ 1 ในการสอบวิชาหนึ่งมีคะแนนเต็ม 50 คะแนน ถ้าจิราภา สอบได้ 40 คะแนน อนันต์สอบได้ 30 คะแนน และสมศักดิ์สอบได้ 20 คะแนน จากตัวอย่าง ดังกล่าวทำให้ทราบว่า คะแนนของจิราภาต่างจากคะแนนของอนันต์เท่ากับ 40 - 30 = 10 คะแนน และคะแนนของอนันต์ต่างจากคะแนนของสมศักดิ์เท่ากับ 30 - 20 = 10 คะแนน

ตัวอย่างที่ 2 ในการวัดอุณหภูมิของวันจันทร์เท่ากับ 20 องศาเซลเซียส อุณหภูมิของวันอังคารเท่ากับ 30 องศาเซลเซียส และอุณหภูมิของวันพุธเท่ากับ 40 องศาเซลเซียสจากตัวอย่างดังกล่าวทำให้ทราบว่าอุณหภูมิของวันจันทร์ต่างจากวันอังคารเท่ากับ 30 - 20 = 10 องศาเซลเซียส และอุณหภูมิของวันพุธต่างจากวันอังคารเท่ากับ 10 องศาเซลเซียสเช่นเดียวกัน

ตัวอย่างมาตราอันตรภาค (ต่อ)

จากตัวอย่างดังกล่าว พบว่า**สามารถนำตัวเลขมาบวก ลบ กันได้** **แต่ไม่สามารถนำมาคูณ และ หาร** จากตัวอย่างนี้ ไม่สามารถกล่าวได้ว่า จิราภาเก่งเป็น 2 เท่าของสมศักดิ์ หรือ อุณหภูมิของวันพุธร้อนเป็น 2 เท่าของวันจันทร์ เนื่องจาก**การวัดในระดับนี้ไม่มีจุดศูนย์แท้** (true or absolute zero) **แต่เป็นศูนย์สมมติ** (arbitrary zero) เพราะฉะนั้นการที่เรากล่าวว่า อุณหภูมิเท่ากับ 0 องศาเซลเซียส ไม่ได้หมายความว่า **"ไม่มีอุณหภูมิ"** เพราะ 0 องศาเซลเซียสมีค่าเท่ากับ 32 องศาฟาเรนไฮต์

หรือการที่เรากล่าวว่า นักศึกษาสอบได้ศูนย์คะแนนในวิชาใดวิชาหนึ่ง ไม่ได้หมายความว่านักศึกษาไม่มีความรู้ในวิชานั้นเลย จริง ๆ แล้วเขามีความรู้ แต่เราไม่ได้วัดในสิ่งที่เขารู้

มาตราอัตราส่วน (Ratio Scale)

การวัดในระดับนี้เป็นการวัดที่สมบูรณ์แบบ **มีจุดศูนย์แท้** (Absolute zero) กล่าวคือ **มีการจัดเรียงลำดับของข้อมูลจากน้อยไปหามากอย่างต่อเนื่องโดยมีจุดเริ่มต้นที่ศูนย์** ตัวอย่างเช่น ความเร็ว ระยะทาง น้ำหนัก ส่วนสูง และพื้นที่ เป็นต้น

เป็นระดับที่สามารถกำหนดค่าตัวเลขให้กับสิ่งที่ต้องการวัดโดยมีศูนย์แท้ เช่น น้ำหนักเท่ากับศูนย์แสดงว่าไม่มีน้ำหนัก เป็นต้น

ระดับนี้สามารถนำตัวเลขมาบวก ลบ คูณ หาร หรือหาอัตราส่วนกันได้ คือ สามารถบอกได้ว่าถนนสายหนึ่งยาว 50 กิโลเมตร มีความยาวเป็น 2 เท่าของถนนอีกสายหนึ่งที่ยาวเพียง 25 กิโลเมตร

ตัวอย่างมาตราอัตราส่วน

ตัวอย่าง พ่อสูง 160 เซนติเมตร ลูกสาวสูง 165 เซนติเมตร ตัวเลข 160 และ 165 เซนติเมตร แสดงถึงปริมาณความสูง และแสดงว่าลูกสาวสูงกว่าพ่อ 5 เซนติเมตร ขณะเดียวกันถ้าลูกชายสูง 155 เซนติเมตร แสดงว่าพ่อสูงกว่าลูกชาย 5 เซนติเมตร 5 เซนติเมตรตัวแรกกับ 5 เซนติเมตรตัวหลังห่างกันเป็นระยะเท่ากัน (Equal Interval) และถ้าเรากล่าวว่าสูง 0 เซนติเมตรแปลว่าไม่มีความสูงเลย “0” ตัวนี้เรียกว่า “ศูนย์แท้” (Absolute Zero)

สรุป **มาตราวัดแต่ละลักษณะเหมาะสำหรับสถิติที่ใช้ในการวิเคราะห์ข้อมูลที่แตกต่างกัน ถ้าผู้วิจัยไม่สามารถระบุว่าข้อมูลที่รวบรวมได้อยู่ในมาตราวัดใด ก็อาจจะทำให้ผู้วิจัยเลือกใช้เทคนิคการวิเคราะห์ข้อมูลที่ไม่เหมาะสม ซึ่งจะทำให้การสรุปผลการวิจัยนั้น ๆ ผิดพลาดได้**

สรุปคุณสมบัติของการวัดระดับทั้ง 4 ระดับ

คุณสมบัติ	สเกลนามบัญญัติ	สเกลอันดับ	สเกลอันตรภาค	สเกลอัตราส่วน
1. จำแนกความต่าง	✓	✓	✓	✓
2. เรียงลำดับ		✓	✓	✓
3. การ + , -			✓	✓
4. การ ×,÷				✓
5. มีศูนย์แท้				✓

เป้าหมายการวิจัย

<u>เป้าหมายการวิจัย เป็นการพิจารณาว่า การวิจัยนั้นต้องการตอบคำถามในเรื่องใด</u>
ซึ่งโดยทั่วไปแล้วคำถามการวิจัย แบ่งออกเป็น 3 ลักษณะ คือ

1. คำถามเชิงพรรณนา (Descriptive questions)
2. คำถามเชิงความสัมพันธ์ (Relationship questions)
3. คำถามเชิงความแตกต่าง (Different questions)

ลักษณะของคำถามในการวิจัย

1. คำถามเชิงพรรณนา

รูปแบบของคำถามจะถามว่า "**คืออะไร**" เช่น ผู้วิจัยกำหนดเป้าหมายการวิจัยว่า "ศึกษาความคิดเห็นของนักเรียนระดับมัธยมศึกษาตอนปลายต่อการอนุรักษ์สิ่งแวดล้อม"

โดยความหมายของข้อความดังกล่าว คือ "ความคิดเห็นของนักเรียนระดับมัธยมศึกษาตอนปลายต่อการอนุรักษ์สิ่งแวดล้อมคืออะไร"

หรือ ผู้วิจัยกำหนดเป้าหมายการวิจัยว่า"ศึกษาเจตคติของนักเรียนต่อวิชาวิทยาศาสตร์"

โดยความหมายของข้อความดังกล่าว คือ "เจตคติของนักเรียนต่อวิชาวิทยาศาสตร์คืออะไร"

ลักษณะของคำถามในการวิจัย

2. คำถามเชิงความสัมพันธ์

รูปแบบของคำถามจะถามว่า **"ตัวแปร 2 ตัว หรือมากกว่า 2 ตัว มีความสัมพันธ์กันหรือไม่"** เช่น

ผู้วิจัยกำหนดเป้าหมายการวิจัยว่า

"ผลสัมฤทธิ์ทางการเรียนมีความสัมพันธ์กับความสนใจในการเรียนหรือไม่"

หรือ

"พื้นฐานความรู้เดิม สภาพแวดล้อมทางบ้าน แรงจูงใจใฝ่สัมฤทธิ์ มีความสัมพันธ์กับผลสัมฤทธิ์ทางการเรียนหรือไม่"

ลักษณะของคำถามในการวิจัย

3. คำถามเชิงความแตกต่าง

รูปแบบของคำถามจะถามว่า **"ตัวแปร A แตกต่างจากตัวแปร B หรือไม่"** หรือ **"ตัวแปร X ตัวแปร Y และตัวแปร Z แตกต่างกันหรือไม่"** เช่น ผู้วิจัยกำหนดเป้าหมายการวิจัยว่า

"ผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ของนักเรียนโดยใช้สื่อการสอน A จะแตกต่างจากการใช้สื่อการสอน B หรือไม่"

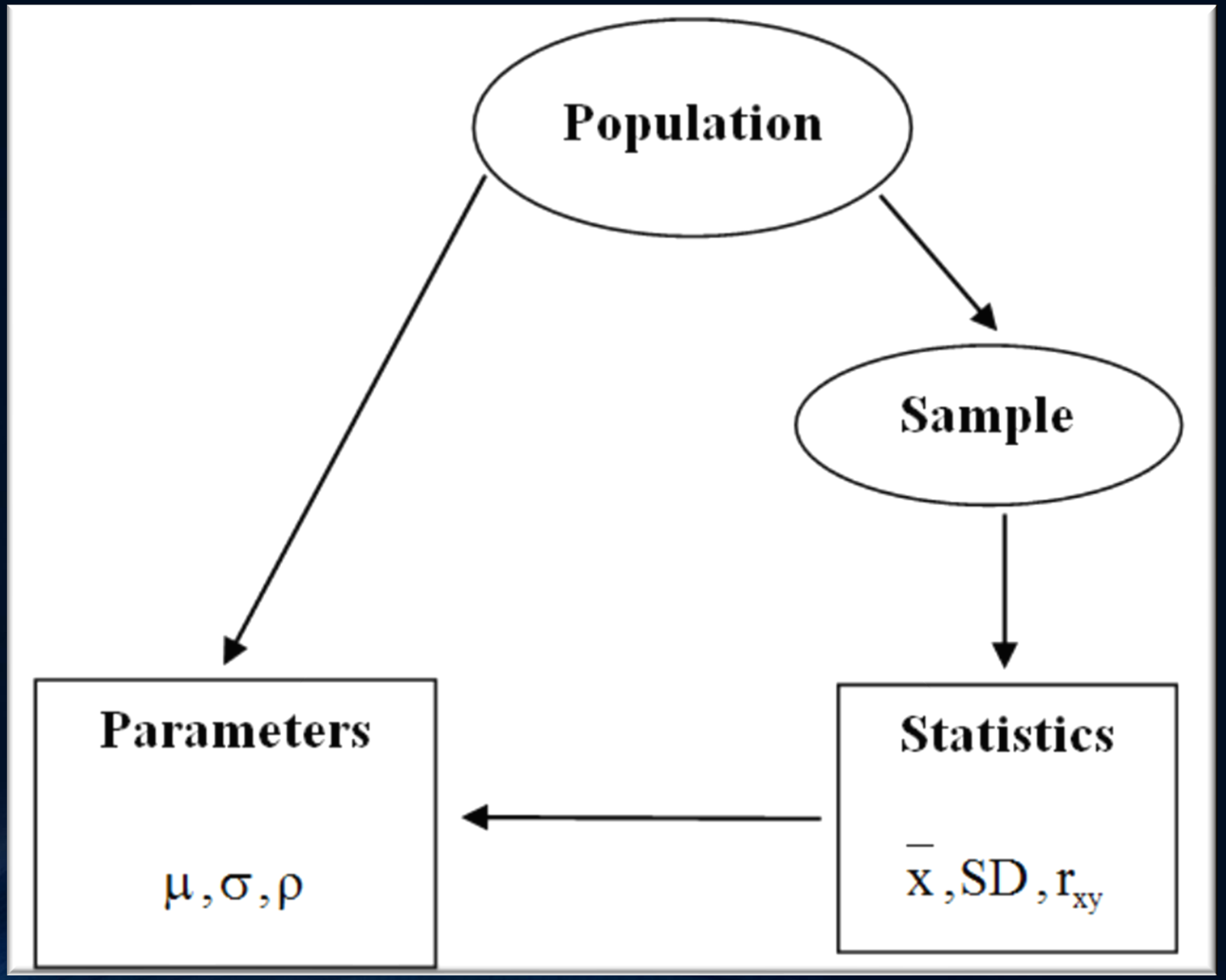
"นักเรียนที่ได้รับการอบรมเลี้ยงดูแบบเข้มงวด แบบปล่อยปละละเลย และแบบประชาธิปไตย จะมีความสนใจในการเรียนต่างกันหรือไม่"

ลักษณะของคำถามในการวิจัย (ต่อ)

"เพื่อเปรียบเทียบความแตกต่างระหว่างสื่อการสอน A และสื่อการสอน B กับผลการเรียนของนักศึกษา"

"เพื่อเปรียบเทียบความรู้ของผู้เข้ารับการฝึกอบรมก่อนและหลังการฝึกอบรมหลักสูตรพัฒนานักวิจัย"

จากคำถามเชิงความแตกต่าง สิ่งที่ผู้วิจัยจะต้องพิจารณาต่อไปคือ **ลักษณะของกลุ่มที่จะนำมาเปรียบเทียบกันเป็นกลุ่มที่อิสระจากกัน (Independent sample) หรือไม่อิสระจากกัน (dependent sample) และจำนวนของกลุ่ม (number of groups) ที่จะนำมาเปรียบเทียบกัน**ด้วย เนื่องจากสถิติที่จะนำมาใช้ในการวิเคราะห์ข้อมูลจะแตกต่างกัน


Population
Sample
Parameters
μ, σ, ρ
Statistics
$\bar{x}, SD, r_{xy}$

สถิติที่ใช้ในการวิเคราะห์ข้อมูล

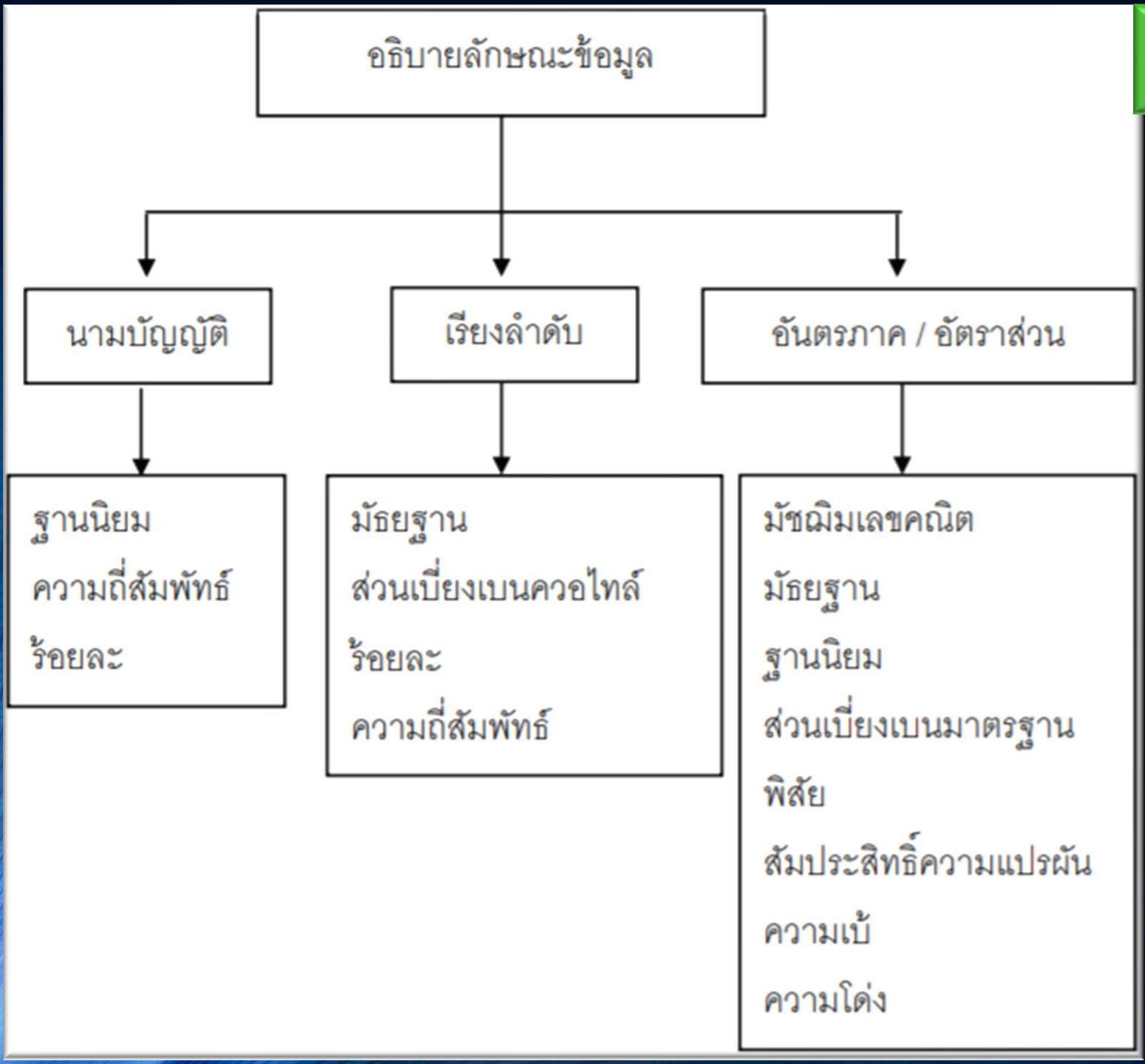
จากคำถามการวิจัยทั้ง 3 ลักษณะดังกล่าว สามารถแบ่งสถิติที่ใช้ในการวิเคราะห์ข้อมูลออกเป็น 3 กลุ่ม คือ

1. **สถิติที่ใช้อธิบายลักษณะของข้อมูล**
2. **สถิติที่ใช้หาความสัมพันธ์ของตัวแปร**
3. **สถิติที่ใช้ในการเปรียบเทียบความแตกต่างของตัวแปร**

สถิติที่ใช้อธิบายลักษณะของข้อมูล

สถิติที่ใช้อธิบายลักษณะของข้อมูล คือ สถิติพรรณนา (Descriptive statistic) ซึ่งประกอบด้วย**การหาค่าร้อยละ การหาความถี่สัมพัทธ์ การวัดค่ากลาง** ซึ่งได้แก่ **มัชฌิมเลขคณิต (ค่าเฉลี่ยเลขคณิต) มัธยฐาน ฐานนิยม** และ **การวัดการกระจาย** ได้แก่ **พิสัย ส่วนเบี่ยงเบนควอไทล์ ส่วนเบี่ยงเบนมาตรฐาน ความแปรปรวน สัมประสิทธิ์ความแปรผัน** ตลอดจน**การวัดความเบ้** และ **ความโด่ง**

สถิติที่ใช้อธิบายลักษณะของข้อมูล



การแจกแจกความถี่

อาชีพ	การแจงนับ	ความถี่
รับราชการ	////	4
ค้าขาย	///// //	7
เกษตรกรรม	///// ///// /	11
รับจ้าง	///// /	6
อื่น ๆ	//	2
รวม	30	30

ค่าร้อยละ

$$\text{นักศึกษาชาย} = \frac{18}{30} \times 100 = 60.0\%$$

$$\text{นักศึกษาหญิง} = \frac{12}{30} \times 100 = 40.0\%$$

การวัดแนวโน้มเข้าสู่ศูนย์กลาง

การหาค่ากลางสำหรับข้อมูลเชิงปริมาณ ประกอบด้วย

1. ค่าเฉลี่ย (Mean) เป็นค่ากลางนิยมใช้กันมากที่สุดใช้ได้กับข้อมูล**มาตราอัตรภาค**และ**อัตราส่วน**

ค่าเฉลี่ยประชากร $\mu = \frac{\sum_{i=1}^{N} x_i}{N}$ โดยที่ N = จำนวนประชากร

ค่าเฉลี่ยตัวอย่าง $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$ โดยที่ n = จำนวนกลุ่มตัวอย่าง

การวัดแนวโน้มเข้าสู่ศูนย์กลาง (ต่อ)

2. ค่ามัธยฐาน (Median) เป็นค่าของข้อมูลที่มีตำแหน่งตรงกลางของข้อมูล เมื่อ**นำข้อมูลมาเรียงลำดับจากค่าน้อยไปมาก** ดังนั้นจะมีข้อมูลครึ่งหนึ่ง (50%) ที่มีค่าน้อยกว่าค่ามัธยฐาน และอีกครึ่งหนึ่งที่มีค่ามากกว่าค่ามัธยฐาน สามารถใช้กับข้อมูล**มาตราอัตรภาค**และ**อัตราส่วน**

ค่ามัธยฐานของ [1 1 3 5 6 9 12] คือ 5

ค่ามัธยฐานของ [1 1 3 5 6 9 12 18] คือ (5+6)/2 = 5.5

การวัดแนวโน้มเข้าสู่ศูนย์กลาง (ต่อ)

3. ค่าฐานนิยม (Mode) ของข้อมูลชุดหนึ่ง คือ **ค่าของข้อมูลที่เกิดขึ้นบ่อยที่สุด** หรือ**มีความถี่สูงสุด** สามารถใช้ได้กับทั้ง 4 มาตรา คือ มาตรานามบัญญัติ (nominal scale) มาตราเรียงอันดับ (ordinal scale) มาตราอันตรภาค (interval scale) และมาตราอัตราส่วน (ratio scale)

ตัวอย่าง ข้อมูลชุดหนึ่งมีจำนวนดังนี้ 5 6 7 8 15 16 17 18

ขอมูลชุดนี้ไมมีฐานนิยม เพราะไม่มีตัวเลขที่ซ้ำกัน

ตัวอย่าง ข้อมูลชุดหนึ่งมีจำนวนดังนี้ 15 14 15 17 18 16

ขอมูลชุดนี้มีฐานนิยม คือ 15

ในการวิจัยไม่นิยมใชฐานนิยม เพราะในการสุ่มกลุมตัวอย่าง 2 ครั้งจากประชากรชุดเดียวกันอาจจะมีฐานนิยมต่างกันได้

การวัดการกระจาย

ค่ากระจายของข้อมูลเชิงปริมาณ เป็นค่าที่แสดงความแตกต่างของข้อมูล **ถ้าข้อมูลชุดใดมีการกระจายมากแสดงว่าค่าข้อมูลชุดนั้นมีความแตกต่างกันมาก**

ค่าการกระจายที่ใช้กันทั่วไป คือ

1. ค่าพิสัย (Range)
2. ค่าแปรปรวน (Variance)
3. ค่าเบี่ยงเบนมาตราฐาน (Standard Deviation)

ค่าพิสัย (Range)

- ผลต่างระหว่าง**ค่าสูงสุดและต่ำสุดของข้อมูล**
- ถ้าพิสัยมีค่าน้อย แสดงว่าข้อมูลกระจุกตัวมาก กระจายตัวน้อย
- ถ้าพิสัยมีค่ามาก แสดงว่าข้อมูลกระจุกตัวกันน้อย กระจายตัวมาก
- เป็นการวัดการกระจายอย่างหยาบ และง่ายที่สุด
- เหมาะสำหรับใช้วัดการกระจายที่มีจำนวนข้อมูลไม่มากนัก และใช้วัดอย่างคร่าว ๆ

คำนวณได้ จากสูตรดังนี้

$$R = H - L$$

เมื่อ R คือ คาพิสัย

H คือ คาสูงสุด

L คือ คาต่ำสุด

ตัวอย่าง คะแนนสอบของนักเรียนมีดังนี้

28 20 19 25 20 18 14 13

พิสัยของคะแนนชุดนี้มีคา 28 - 13 = 15

ค่าเบี่ยงเบนมาตรฐาน (Standard Deviations)

ค่าเบี่ยงเบนมาตรฐาน หรือ ส่วนเบี่ยงเบนมาตรฐาน**เป็นค่าเบี่ยงเบนเฉลี่ยที่แต่ละตัวเบี่ยงออกไปจากค่าเฉลี่ยของคะแนนชุดนั้น** หรือ รากที่สองของค่าเฉลี่ยผลรวมทั้งหมดของค่าเบี่ยงเบนยกกำลังสอง ซึ่งก็คือ **รากที่สองของความแปรปรวนนั่นเอง**

การหาส่วนเบี่ยงเบนมาตรฐาน (Standard Deviations)

การหาส่วนเบี่ยงเบนมาตรฐาน (S.D.) ในกรณีข้อมูลไม่ได้มีการแจกแจงความถี่สามารถหาได้จากสูตร

สูตรที่ 1 $S.D. = \sqrt{\frac{\sum(xi - \bar{x})^2}{(n-1)}}$ หรือ สูตรที่ 2 $S.D. = \sqrt{\frac{n\sum x^2 - (\sum x)^2}{n(n-1)}}$

เมื่อ S.D. คือ ส่วนเบี่ยงเบนมาตรฐาน

x คือ ข้อมูล (ตัวที่ 1, 2, 3, ..., n)

$\bar{x}$ คือ ค่าเฉลี่ยเลขคณิต

n คือ จำนวนข้อมูลทั้งหมด

<u>หมายเหตุ</u> ในกรณีที่ $\bar{x}$ เป็นทศนิยมทำให้เกิดความยุ่งยากในการคำนวณ จึงควรเลือกใช้สูตรที่ 2

ค่าแปรปรวน (Variance)

เป็นค่าที่ใช้วัดการกระจายที่นิยมใช้กัน พิจารณาจากค่าเฉลี่ยของค่าแตกต่างระหว่างค่าข้อมูลแต่ละค่ากับค่าเฉลี่ยยกกำลังสอง **ถ้าค่าแปรปรวนมีค่ามากแสดงว่าข้อมูลชุดนั้นมีการกระจายมาก**

$$S^2 = \frac{\sum(x - \bar{x})^2}{N - 1}$$

$$= (\text{S.D.})^2$$

การคำนวณค่าเบี่ยงเบนมาตราฐาน และ ความแปรปรวน

ตัวอย่าง จงหาส่วนเบี่ยงเบนมาตรฐาน และ ความแปรปรวนของข้อมูลต่อไปนี้ 1, 2, 4, 6, 8, 9

วิธีทำ ใช้สูตร $$S.D. = \sqrt{\frac{n\sum x^2 - (\sum x)^2}{n(n-1)}}$$

หาค่า $\Sigma x^2 = 1^2 + 2^2 + 4^2 + 6^2 + 8^2 + 9^2$

$= 1 + 4 + 16 + 36 + 64 + 91$

$= 212$

หาค่า $\Sigma x = 1 + 2 + 4 + 6 + 8 + 9$

$= 30$

$(\Sigma x)^2 = 30^2$

$= 900$

$n = 6$

การคำนวณค่าเบี่ยงเบนมาตรฐาน และ ความแปรปรวน (ต่อ)

แทนค่าในสูตร $S.D. = \sqrt{\frac{6(212) - 900}{6(6-1)}}$

$= \sqrt{\frac{372}{30}}$

$= 3.52$

จาก $S^2 = (\text{S.D.})^2$

$= (3.52)^2 = 12.4$

ดังนั้น ส่วนเบี่ยงเบนมาตรฐาน และ ความแปรปรวนของข้อมูลต่อไปนี้เท่ากับ 3.52 และ 12.4 ตามลำดับ

ค่าสถิติที่ใช้ได้กับประเภทข้อมูล

สถิติ	ระดับข้อมูล			
	นามบัญญัติ	เรียงอันดับ	อันตรภาค	อัตราส่วน
ความถี่	✓	✓	✓	✓
ร้อยละ	✓	✓	✓	✓
ค่าเฉลี่ย			✓	✓
มัธยฐาน		✓	✓	✓
ฐานนิยม	✓	✓	✓	✓
พิสัย			✓	✓
ส่วนเบี่ยงเบนมาตรฐาน			✓	✓
ความแปรปรวน			✓	✓