บทที่ 6

การวิเคราะห์ข้อมูลเชิงปริมาณ (Quantitative Data Analysis)

Excel • SPSS • Data Cleaning • Data Analysis • Data Visualization

รายละเอียดรายวิชา ผู้สอน : ผู้ช่วยศาสตราจารย์ ดร. นัฐพงศ์ ส่งเนียม

ชื่อวิชา: คอมพิวเตอร์สำหรับบัณฑิตศึกษา (Computer for Graduate Studies) **รหัสวิชา:** 4125101

หน่วยกิต: 3(2-2-5)

- 2 ชั่วโมง บรรยาย
- 2 ชั่วโมง ปฏิบัติ
- 5 ชั่วโมง ศึกษาด้วยตนเอง

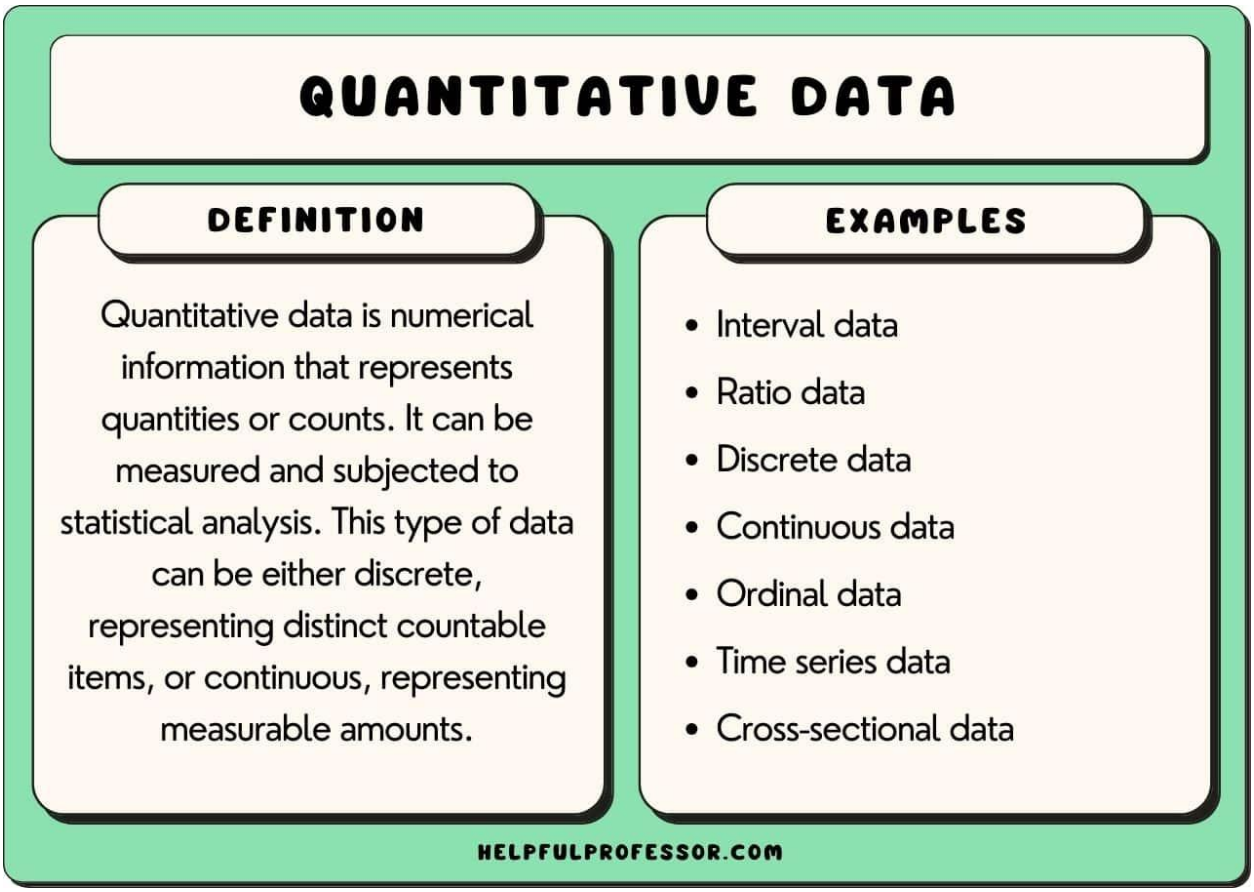
วัตถุประสงค์การเรียนรู้ (Learning Objectives)

เมื่อศึกษาบทนี้แล้ว ผู้เรียนสามารถ

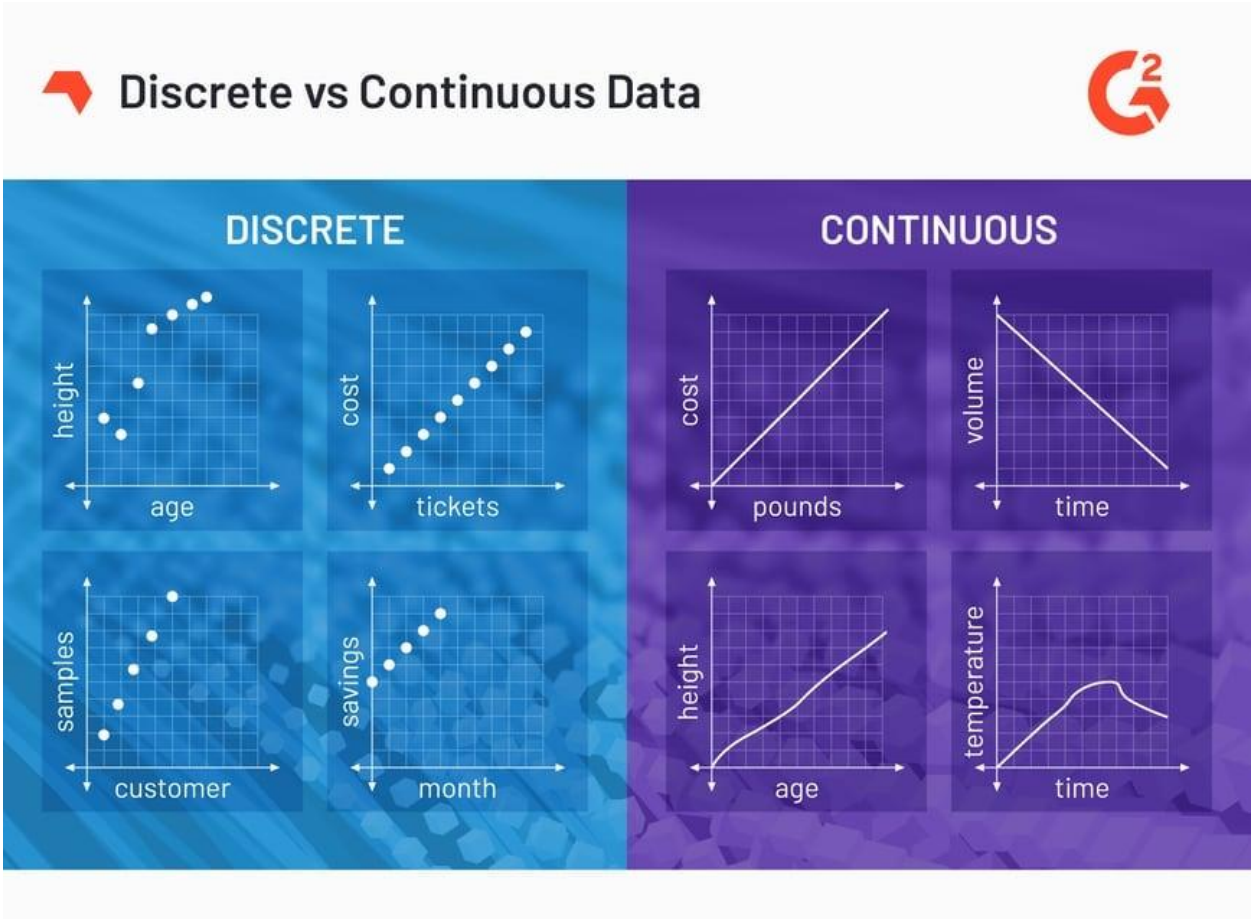
1. อธิบายความหมายและประเภทของข้อมูลเชิงปริมาณได้
2. ใช้ Excel และ SPSS ในการจัดการและวิเคราะห์ข้อมูลเชิงปริมาณ
3. ทำความสะอาดข้อมูล (Data Cleaning) ได้อย่างถูกต้อง
4. วิเคราะห์ข้อมูลเชิงพรรณนาและเชิงอนุมาน
5. เลือกและสร้างกราฟเพื่อแสดงผลข้อมูลได้อย่างเหมาะสม

6.1 ความหมายของข้อมูลเชิงปริมาณ (Quantitative Data)

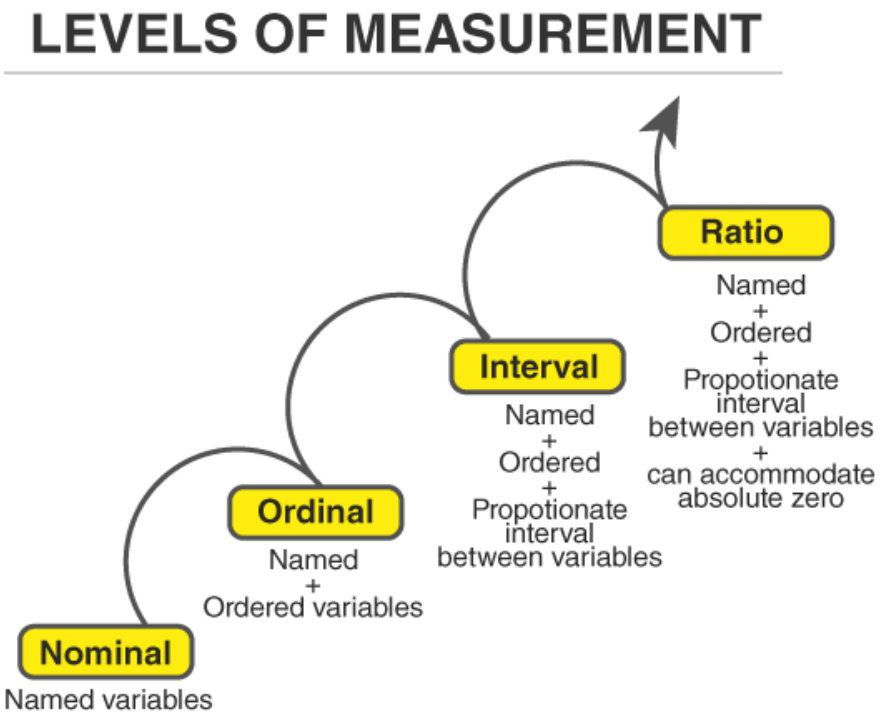
ข้อมูลเชิงปริมาณ (Quantitative Data) คือข้อมูลที่สามารถแสดงค่าในรูป ตัวเลข และสามารถนำไปใช้ในการ คำนวณ วิเคราะห์ และประมวลผลทางสถิติ เพื่ออธิบายลักษณะ แนวโน้ม ความสัมพันธ์ หรือความแตกต่างของข้อมูลได้อย่างเป็นระบบ ข้อมูลประเภทนี้เป็นพื้นฐานสำคัญของการวิเคราะห์ข้อมูล การวิจัยเชิงปริมาณ และการตัดสินใจเชิงวิชาการและเชิงบริหาร



ภาพที่ 6.1 ข้อมูลเชิงปริมาณ



ภาพที่ 6.2 ข้อมูลต่อเนื่องและไม่ต่อเนื่อง



ภาพที่ 6.3 ระดับของการวัด

ตัวอย่างของข้อมูลเชิงปริมาณ ได้แก่ อายุ รายได้ คะแนนสอบ น้ำหนัก ส่วนสูง เวลา อุณหภูมิ ปริมาณยอดขาย หรือจำนวนผู้ใช้บริการ เป็นต้น

ตัวอย่างข้อมูลเชิงปริมาณ (Quantitative Data)

ตารางที่ 6.1 ตัวอย่างข้อมูลเชิงปริมาณหลายประเภท

ลำดับ	อายุ (ปี)	รายได้ต่อเดือน (บาท)	คะแนนสอบ	น้ำหนัก (กก.)	ส่วนสูง (ซม.)	เวลาเข้าเรียน (นาที)	อุณหภูมิ (°C)	ปริมาณยอดขาย (ชิ้น)	จำนวนผู้ใช้บริการ (คน)
1	18	8,500	78	55.2	165	45	32.5	120	35
2	19	9,200	85	60	170	40	31.8	150	42
3	20	10,500	92	62.5	168	50	33.1	180	48
4	21	12,000	67	58	160	55	30.9	95	28
5	22	15,000	88	70.3	175	35	34	210	55

การจำแนกตัวอย่างข้อมูลเชิงปริมาณตามประเภท

1) ข้อมูลไม่ต่อเนื่อง (Discrete Data)

เป็นข้อมูลที่ได้จากการ **นับจำนวน** และมีค่าเป็นจำนวนเต็ม

ตัวอย่างจากตาราง

- คะแนนสอบ
- ปริมาณยอดขาย (ชิ้น)
- จำนวนผู้ใช้บริการ (คน)

📌 *เหมาะสำหรับการวิเคราะห์ความถี่ การเปรียบเทียบ และกราฟแท่ง (Bar Chart)*

2) ข้อมูลต่อเนื่อง (Continuous Data)

เป็นข้อมูลที่ได้จากการ **วัดค่า** และสามารถมีค่าทศนิยมได้

ตัวอย่างจากตาราง

- อายุ (ปี)
- รายได้ต่อเดือน (บาท)
- น้ำหนัก (กิโลกรัม)
- ส่วนสูง (เซนติเมตร)
- เวลา (นาที)
- อุณหภูมิ (องศาเซลเซียส)

📌 *เหมาะสำหรับการวิเคราะห์ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน Histogram และ Boxplot*

ตัวอย่างการนำข้อมูลไปใช้ในการวิเคราะห์

- **อายุ / คะแนนสอบ** ⟶ วิเคราะห์ค่าเฉลี่ยและการกระจาย
- **รายได้** ⟶ วิเคราะห์ Median เพื่อหลีกเลี่ยงผลของ Outliers
- **น้ำหนัก / ส่วนสูง** ⟶ วิเคราะห์ความสัมพันธ์ (Correlation)
- **ยอดขาย / ผู้ใช้บริการ** ⟶ วิเคราะห์แนวโน้ม (Trend Analysis)

ตัวอย่างการเชื่อมโยงกับ Excel และ SPSS

- Excel :
    - AVERAGE, MEDIAN, STDEV.S
    - Histogram / Line Chart / PivotTable
- SPSS :
    - Descriptive Statistics
    - Correlation / Regression

ข้อมูลเชิงปริมาณมีลักษณะเด่นคือ

- สามารถนำไปคำนวณด้วยสูตรทางคณิตศาสตร์ได้
- สามารถวิเคราะห์ด้วยสถิติทั้งเชิงพรรณนาและเชิงอนุมาน
- สามารถนำเสนอในรูปแบบตาราง กราฟ และแผนภูมิได้อย่างชัดเจน

ประเภทของข้อมูลเชิงปริมาณ

ข้อมูลเชิงปริมาณสามารถจำแนกออกเป็น 2 ประเภทหลัก ตามลักษณะของค่าที่วัดได้ ดังนี้

1) ข้อมูลไม่ต่อเนื่อง (Discrete Data)

ข้อมูลไม่ต่อเนื่อง คือข้อมูลที่มีค่าเป็น **จำนวนเต็ม** และไม่สามารถแยกย่อยเป็นค่าทศนิยมได้ โดยค่าของข้อมูลจะเพิ่มขึ้นเป็นขั้น ๆ (stepwise) มักได้มาจากการ **นับจำนวน (Counting)**

ลักษณะสำคัญ

- มีค่าเป็นจำนวนเต็มเท่านั้น
- ไม่สามารถมีค่ากลางระหว่างจำนวนเต็มได้
- มักเกี่ยวข้องกับจำนวนหรือความถี่

ตัวอย่าง

- จำนวนพนักงานในองค์กร
- จำนวนสินค้าที่ขายได้ต่อวัน
- จำนวนครั้งที่นักศึกษาเข้าเรียน
- จำนวนบุตรในครอบครัว

1) ตัวอย่างข้อมูล: จำนวนพนักงานในองค์กร

ตารางที่ 6.2 ตัวอย่างข้อมูล จำนวนพนักงานในองค์กร

แผนก	จำนวนพนักงาน (คน)
บริหาร	12
การเงิน	8
การตลาด	15
ผลิต	25
IT	6

2) ตัวอย่างข้อมูล: จำนวนสินค้าที่ขายได้ต่อวัน

ตารางที่ 6.3 ตัวอย่างข้อมูล จำนวนสินค้าที่ขายได้ต่อวัน

วันที่	จำนวนสินค้าที่ขายได้ (ชิ้น)
1 ม.ค. 2569	120
2 ม.ค. 2569	135
3 ม.ค. 2569	98
4 ม.ค. 2569	150
5 ม.ค. 2569	160

3) ตัวอย่างข้อมูล: จำนวนครั้งที่นักศึกษาเข้าเรียน

ตารางที่ 6.4 ตัวอย่างข้อมูลจำนวนครั้งที่นักศึกษาเข้าเรียน

รหัสนักศึกษา	จำนวนครั้งที่เข้าเรียน (ครั้ง/ภาคเรียน)
65001	28
65002	30
65003	25
65004	27
65005	30

4) ตัวอย่างข้อมูล: จำนวนบุตรในครอบครัว

ตารางที่ 6.5 ตัวอย่างข้อมูลจำนวนบุตรในครอบครัว

ครัวเรือน	จำนวนบุตร (คน)

ครัวเรือนที่ 1	1
ครัวเรือนที่ 2	2
ครัวเรือนที่ 3	0
ครัวเรือนที่ 4	3
ครัวเรือนที่ 5	2

การวิเคราะห์ที่เหมาะสม (สำหรับข้อมูลไม่ต่อเนื่อง)

- ความถี่ (Frequency) หมายถึง จำนวนครั้งที่ข้อมูลแต่ละค่าปรากฏ ใช้เพื่อดูการกระจายของข้อมูลว่าแต่ละค่ามีจำนวนมากน้อยเพียงใด
- ค่าเฉลี่ย (Mean) หมายถึง ค่ากลางของข้อมูล ที่ได้จากการนำข้อมูลทั้งหมดมารวมกันแล้วหารด้วยจำนวนข้อมูล
- มัธยฐาน (Median) หมายถึง ค่าที่อยู่ตรงกลางของข้อมูล เมื่อเรียงข้อมูลจากน้อยไปมาก
- ฐานนิยม (Mode) หมายถึง ค่าที่พบมากที่สุดในชุดข้อมูล อาจมีหนึ่งค่า หลายค่า หรือไม่มีเลย
- กราฟแท่ง (Bar Chart) หมายถึง กราฟที่ใช้แสดงและเปรียบเทียบข้อมูลเชิงปริมาณแบบไม่ต่อเนื่อง โดยความสูงของแท่งแสดงค่าหรือความถี่ของข้อมูล

2) ข้อมูลต่อเนื่อง (Continuous Data)

ข้อมูลต่อเนื่อง คือข้อมูลที่สามารถมีค่าเป็น **จำนวนเต็มหรือทศนิยม** และสามารถวัดค่าได้อย่างต่อเนื่องภายในช่วงหนึ่ง มักได้มาจากการ **วัด (Measurement)**

ลักษณะสำคัญ

- มีค่าทศนิยมได้
- สามารถมีค่าระหว่างตัวเลขสองค่าใด ๆ ได้
- มีความละเอียดของข้อมูลสูง

ตัวอย่าง

- น้ำหนัก (กิโลกรัม)
- ส่วนสูง (เซนติเมตร)
- เวลา (ชั่วโมง นาที วินาที)
- อุณหภูมิ (องศาเซลเซียส)

- รายได้ต่อเดือน

ตัวอย่างข้อมูลเชิงปริมาณแบบต่อเนื่อง (Continuous Data)

ตารางที่ 6.6 ตัวอย่างข้อมูลการวัดค่าทางกายภาพและเศรษฐกิจ

ลำดับ	น้ำหนัก (กก.)	ส่วนสูง (ซม.)	เวลา (ชั่วโมง:นาที:วินาที)	อุณหภูมิ (°C)	รายได้ต่อเดือน (บาท)
1	52.5	160.3	1:15:30	31.5	15,500
2	58	165.8	0:45:20	32	18,200
3	63.7	170.2	2:10:45	33.4	22,000
4	70.2	175.5	1:05:10	30.8	25,600
5	68.9	168	0:55:40	29.9	20,300

การอธิบายตัวอย่างข้อมูล

- น้ำหนัก (กิโลกรัม)
  เป็นข้อมูลที่ได้จากการวัด สามารถมีค่าทศนิยมได้ แสดงถึงมวลของบุคคล
- ส่วนสูง (เซนติเมตร)
  เป็นข้อมูลต่อเนื่องที่แสดงความสูงของบุคคล วัดได้ละเอียดถึงทศนิยม
- เวลา (ชั่วโมง นาที วินาที)
  เป็นข้อมูลต่อเนื่องที่สามารถวัดได้ละเอียดมาก ใช้ในการวิเคราะห์ระยะเวลาหรือประสิทธิภาพ
- อุณหภูมิ (องศาเซลเซียส)
  เป็นข้อมูลต่อเนื่องที่ใช้ในการวิเคราะห์สภาพแวดล้อมหรือสภาพอากาศ
- รายได้ต่อเดือน (บาท)
  เป็นข้อมูลเชิงปริมาณที่แสดงฐานะทางเศรษฐกิจ มักมีค่าทศนิยมเมื่อคิดเป็นค่าเฉลี่ย

การวิเคราะห์ที่เหมาะสม

- ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน
- การกระจายของข้อมูล
- กราฟเส้น (Line Chart), Histogram, Boxplot

ความสำคัญของการจำแนกประเภทข้อมูลเชิงปริมาณ

การจำแนกข้อมูลเชิงปริมาณออกเป็นข้อมูลไม่ต่อเนื่องและข้อมูลต่อเนื่องมีความสำคัญอย่างยิ่ง เนื่องจาก

- ช่วยให้เลือก **วิธีวิเคราะห์ทางสถิติ** ได้อย่างเหมาะสม
- ช่วยให้เลือก **รูปแบบกราฟหรือแผนภูมิ** ได้ถูกต้อง
- ลดความผิดพลาดในการแปลผลข้อมูล

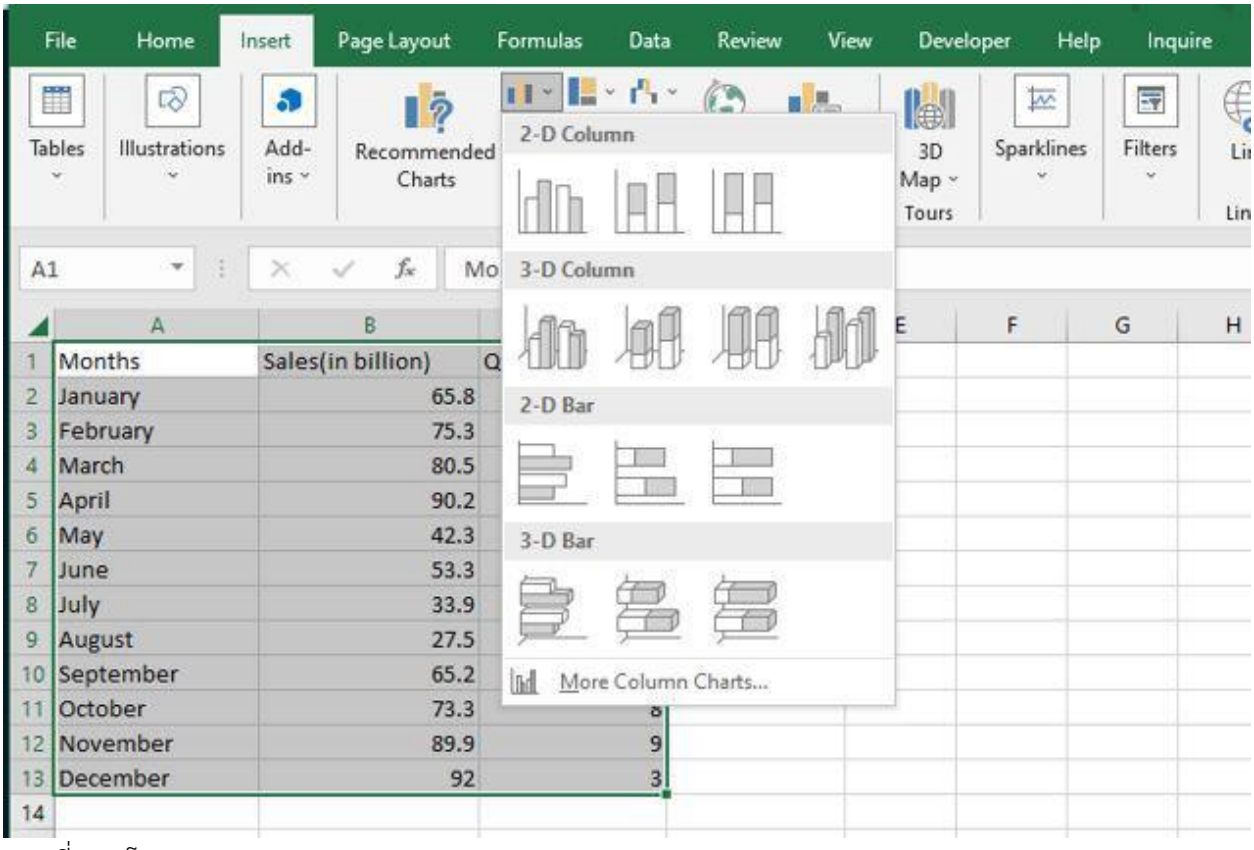
หากเลือกวิธีวิเคราะห์หรือการแสดงผลไม่เหมาะสมกับประเภทข้อมูล อาจทำให้ผลการวิเคราะห์คลาดเคลื่อนหรือสื่อสารข้อมูลผิดพลาดได้

สรุป

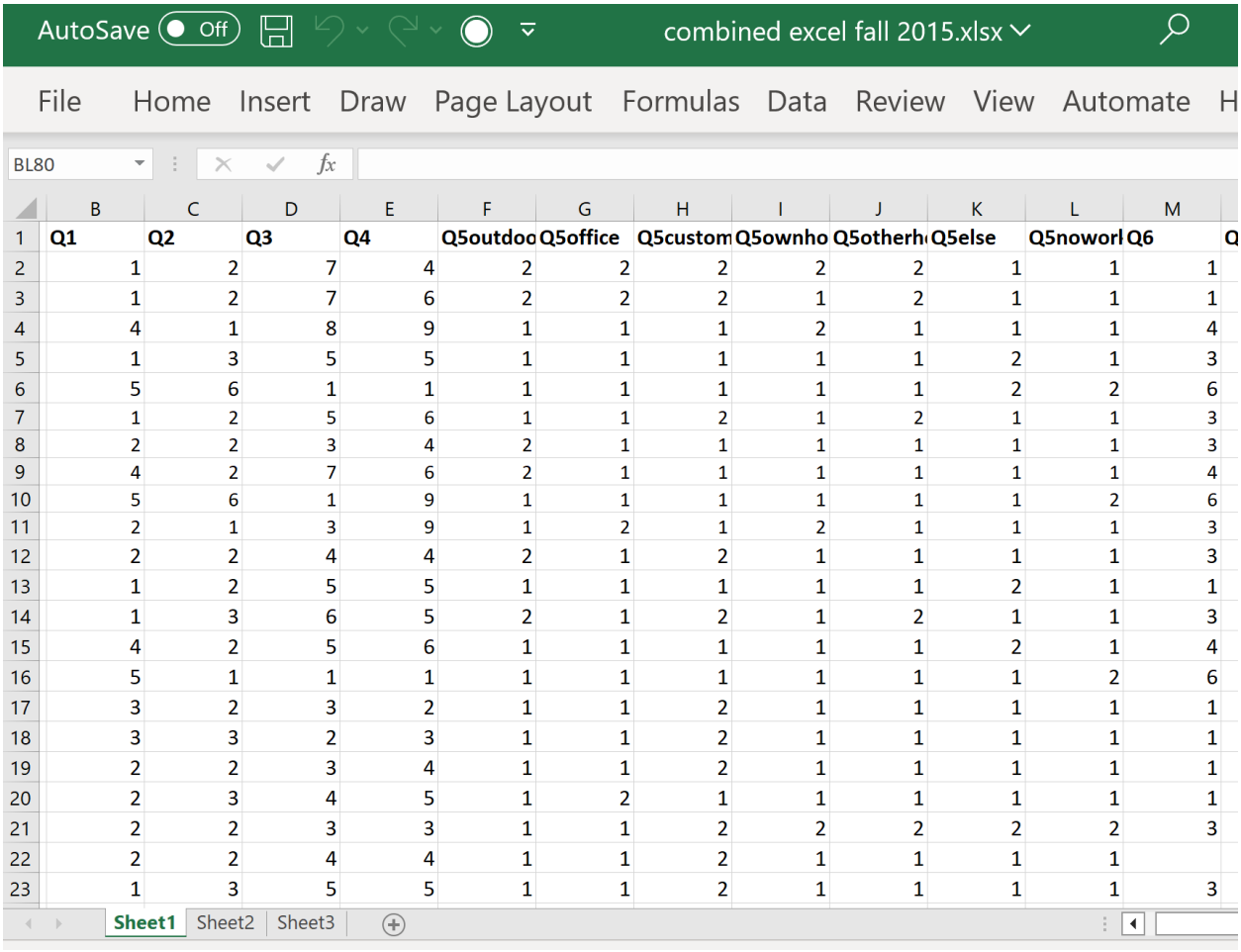
ข้อมูลเชิงปริมาณเป็นข้อมูลพื้นฐานที่มีบทบาทสำคัญในการวิเคราะห์ข้อมูลและการวิจัย โดยสามารถจำแนกออกเป็นข้อมูลไม่ต่อเนื่องและข้อมูลต่อเนื่อง ซึ่งแต่ละประเภทมีลักษณะและวิธีการวิเคราะห์ที่แตกต่างกัน การเข้าใจประเภทของข้อมูลอย่างถูกต้องจะช่วยให้การวิเคราะห์ข้อมูลในขั้นตอนต่อไปมีความถูกต้องและมีประสิทธิภาพมากยิ่งขึ้น

6.2 เครื่องมือที่ใช้ในการจัดการข้อมูลเชิงปริมาณ

6.2.1 Microsoft Excel



ภาพที่ 6.4 โปรแกรม MS-Excel



	B	C	D	E	F	G	H	I	J	K	L	M
1	Q1	Q2	Q3	Q4	Q5outdoo	Q5office	Q5custom	Q5ownho	Q5otherh	Q5else	Q5noworl	Q6
2	1	2	7	4	2	2	2	2	2	1	1	1
3	1	2	7	6	2	2	2	1	2	1	1	1
4	4	1	8	9	1	1	1	2	1	1	1	4
5	1	3	5	5	1	1	1	1	1	2	1	3
6	5	6	1	1	1	1	1	1	1	2	2	6
7	1	2	5	6	1	1	2	1	2	1	1	3
8	2	2	3	4	2	1	1	1	1	1	1	3
9	4	2	7	6	2	1	1	1	1	1	1	4
10	5	6	1	9	1	1	1	1	1	1	2	6
11	2	1	3	9	1	2	1	2	1	1	1	3
12	2	2	4	4	2	1	2	1	1	1	1	3
13	1	2	5	5	1	1	1	1	1	2	1	1
14	1	3	6	5	2	1	2	1	2	1	1	3
15	4	2	5	6	1	1	1	1	1	2	1	4
16	5	1	1	1	1	1	1	1	1	1	2	6
17	3	2	3	2	1	1	2	1	1	1	1	1
18	3	3	2	3	1	1	2	1	1	1	1	1
19	2	2	3	4	1	1	2	1	1	1	1	1
20	2	3	4	5	1	2	1	1	1	1	1	1
21	2	2	3	3	1	1	2	2	2	2	2	3
22	2	2	4	4	1	1	2	1	1	1	1	
23	1	3	5	5	1	1	2	1	1	1	1	3

ภาพที่ 6.5 ตัวอย่างข้อมูลใน MS-Excel

Pivot Table 1

Sales	Sep	Oct	Nov	Total
Apples	**250**	**590**		**840**
John		180		180
Mike		120		120
Pete		290		290
Sally	250			250
Bananas		**430**	**600**	**1030**
John			400	400
Mike			200	200
Pete		180		180
Sally		250		250
Cherries	**580**	**910**		**1490**
John		250		250
Mike	250	330		580
Pete		330		330
Sally	330			330
Oranges		**120**	**720**	**840**
John			120	120
Mike			400	400
Pete		120		120
Sally			200	200
Total	**830**	**2050**	**1320**	**4200**

Pivot Table 2

Month	(All)				
Sales / Reseller	Product / Apples	Bananas	Cherries	Oranges	Total
John	$180	$400	$250	$120	$950
Mike	$120	$200	$580	$400	$1,300
Pete	$290	$180	$330	$120	$920
Sally	$250	$250	$330	$200	$1,030
Total	**$840**	**$1,030**	**$1,490**	**$840**	**$4,200**

Pivot Table 3

Product	(All)			
Sales / Reseller	Month / Sep	Oct	Nov	Total
John		$430	$520	$950
Mike	$250	$450	$600	$1,300
Pete		$920		$920
Sally	$580	$250	$200	$1,030
Total	**$830**	**$2,050**	**$1,320**	**$4,200**

ภาพที่ 6.6 ตาราง Pivot

Microsoft Excel เป็นโปรแกรมสเปรดชีตที่ได้รับความนิยมอย่างแพร่หลายสำหรับการจัดการข้อมูลเชิงปริมาณในระดับพื้นฐานจนถึงระดับกลาง เนื่องจากมีความยืดหยุ่น ใช้งานง่าย และสามารถรองรับข้อมูลจำนวนมากได้ Excel เหมาะสำหรับการเตรียมข้อมูล การคำนวณ การวิเคราะห์ และการนำเสนอข้อมูลก่อนนำไปวิเคราะห์ขั้นสูงด้วยโปรแกรมอื่น เช่น SPSS หรือเครื่องมือด้าน Data Analytics

บทบาทของ Excel ในการจัดการข้อมูลเชิงปริมาณ

Excel สามารถนำมาใช้ในการจัดการข้อมูลเชิงปริมาณได้ในหลายขั้นตอน ได้แก่

1. **การบันทึกและจัดตารางข้อมูล (Data Entry & Structuring)**

   ผู้ใช้สามารถบันทึกข้อมูลในรูปแบบตาราง โดยกำหนดแถวเป็นหน่วยข้อมูล (เช่น บุคคล รายการ หรือช่วงเวลา) และคอลัมน์เป็นตัวแปร (เช่น อายุ คะแนน รายได้) ซึ่งช่วยให้ข้อมูลมีโครงสร้างที่ชัดเจนและพร้อมต่อการวิเคราะห์

2. **การคำนวณค่าทางสถิติพื้นฐาน (Basic Statistical Calculation)**
   Excel มีฟังก์ชันสำเร็จรูปสำหรับคำนวณค่าทางสถิติ เช่น ค่าเฉลี่ย ค่าต่ำสุด ค่าสูงสุด และการนับจำนวนข้อมูล ช่วยลดเวลาและลดความผิดพลาดจากการคำนวณด้วยมือ
3. **การสร้างกราฟและรายงานผล (Visualization & Reporting)**
   Excel สามารถสร้างกราฟและแผนภูมิหลายรูปแบบ เช่น กราฟแท่ง กราฟเส้น และ Histogram เพื่อใช้ในการสื่อสารผลการวิเคราะห์ข้อมูลให้เข้าใจง่ายและชัดเจน

ฟังก์ชันที่สำคัญในการวิเคราะห์ข้อมูลเชิงปริมาณ

1) ฟังก์ชันทางสถิติพื้นฐาน

- **SUM** : ใช้หาผลรวมของข้อมูล
- **AVERAGE** : ใช้คำนวณค่าเฉลี่ยของข้อมูล
- **MIN / MAX** : ใช้หาค่าต่ำสุดและค่าสูงสุดของข้อมูล

ฟังก์ชันกลุ่มนี้เหมาะสำหรับการสรุปลักษณะทั่วไปของข้อมูลในขั้นต้น

2) ฟังก์ชันการนับและเงื่อนไข

- **COUNT** : นับจำนวนเซลล์ที่เป็นตัวเลข
- **COUNTIF** : นับจำนวนข้อมูลตามเงื่อนไขที่กำหนด

ฟังก์ชันเหล่านี้ช่วยในการวิเคราะห์ความถี่และการกระจายของข้อมูลเชิงปริมาณ

3) ฟังก์ชันตรรกะและการค้นหาข้อมูล

- **IF** : ใช้กำหนดเงื่อนไขเพื่อจำแนกหรือจัดกลุ่มข้อมูล
- **VLOOKUP / XLOOKUP** : ใช้ค้นหาข้อมูลจากตารางอ้างอิง

ฟังก์ชันกลุ่มนี้มีประโยชน์ในการจัดการข้อมูลขนาดใหญ่ และการรวมข้อมูลจากหลายแหล่งเข้าด้วยกัน

4) PivotTable

PivotTable เป็นเครื่องมือสำคัญของ Excel สำหรับการวิเคราะห์ข้อมูลเชิงปริมาณแบบสรุปผล โดยไม่จำเป็นต้องเขียนสูตรที่ซับซ้อน

ความสามารถของ PivotTable

- สรุปข้อมูลจำนวนมากให้อยู่ในรูปแบบตาราง
- คำนวณค่าเฉลี่ย ผลรวม และจำนวนข้อมูล

- เปรียบเทียบข้อมูลตามกลุ่ม เช่น เพศ ชั้นปี หรือช่วงเวลา
- ใช้ร่วมกับ PivotChart เพื่อสร้างกราฟอัตโนมัติ

PivotTable ช่วยให้ผู้เรียนสามารถมองเห็นภาพรวมของข้อมูล และเข้าใจความสัมพันธ์ระหว่างตัวแปรได้ง่ายขึ้น

ข้อดีของการใช้ Excel ในงานข้อมูลเชิงปริมาณ

- ใช้งานง่าย เหมาะสำหรับผู้เริ่มต้น
- เห็นโครงสร้างข้อมูลได้ชัดเจน
- รองรับการคำนวณและกราฟพื้นฐานได้ครบถ้วน
- เป็นขั้นตอนเตรียมข้อมูลก่อนการวิเคราะห์เชิงสถิติขั้นสูง

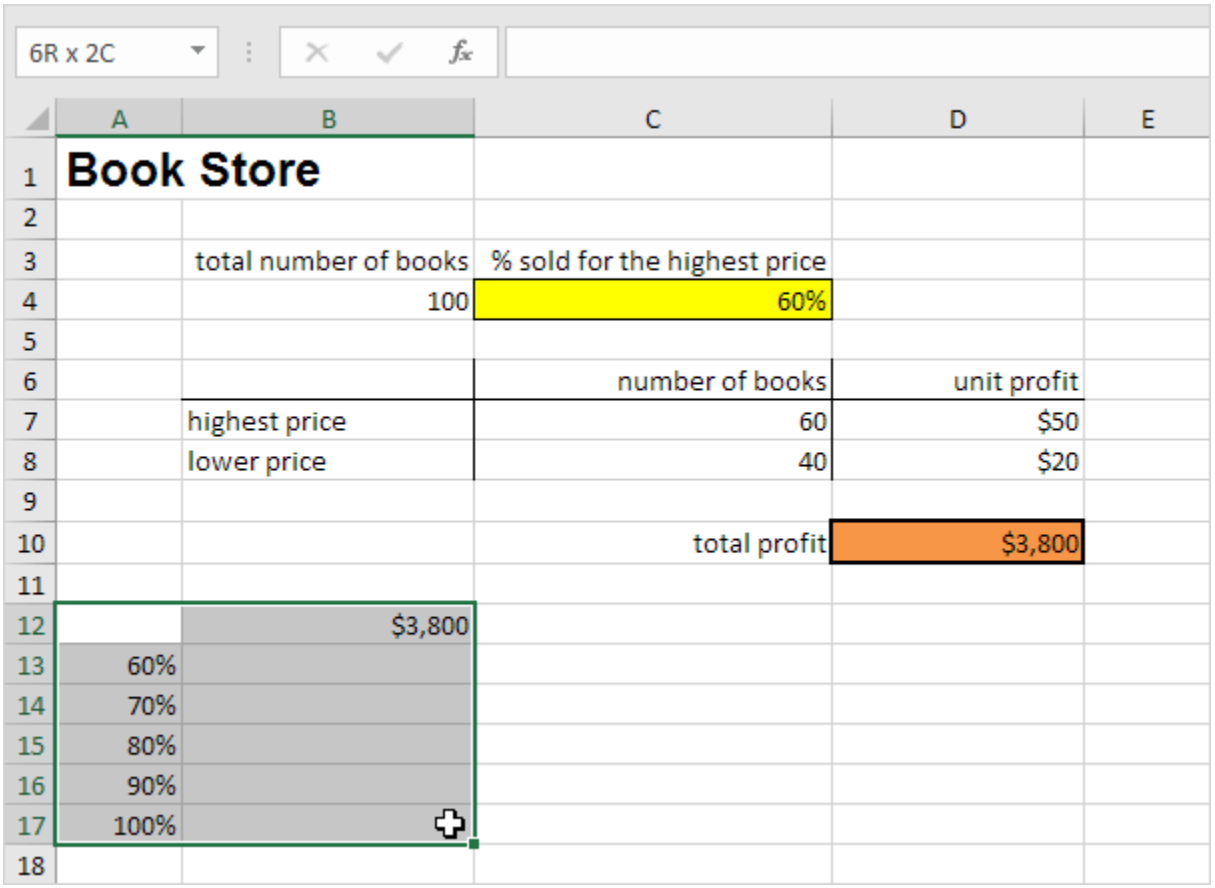
ข้อจำกัดของ Excel

- ไม่เหมาะสำหรับข้อมูลขนาดใหญ่มากหรือการวิเคราะห์ขั้นสูง
- การวิเคราะห์เชิงสถิติซับซ้อนมีข้อจำกัดเมื่อเทียบกับ SPSS

ดังนั้น Excel มักถูกใช้เป็นเครื่องมือในขั้นตอน **Data Preparation และ Exploratory Analysis** ก่อนนำข้อมูลไปวิเคราะห์เชิงลึกด้วยโปรแกรมอื่น

Microsoft Excel เป็นเครื่องมือพื้นฐานที่มีบทบาทสำคัญในการจัดการข้อมูลเชิงปริมาณ ตั้งแต่การบันทึกข้อมูล การคำนวณค่าทางสถิติ การสรุปข้อมูล ไปจนถึงการสร้างกราฟ การใช้ Excel อย่างถูกต้องจะช่วยให้การวิเคราะห์ข้อมูลมีความถูกต้อง เป็นระบบ และพร้อมสำหรับการวิเคราะห์ในระดับที่สูงขึ้น

ตัวอย่างการใช้ Microsoft Excel สำหรับข้อมูลเชิงปริมาณ (เชิงปฏิบัติ)

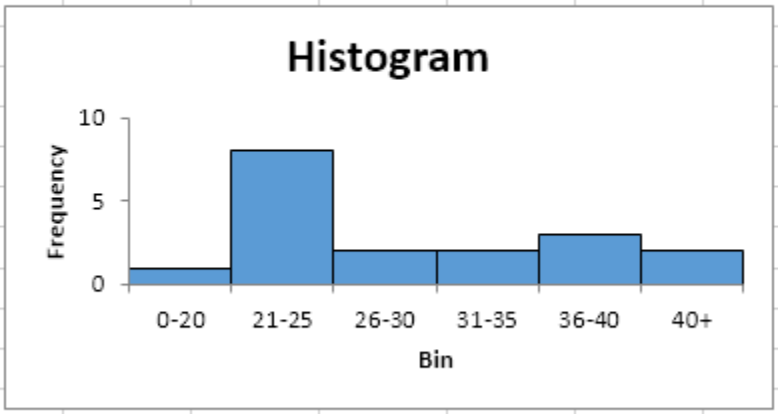


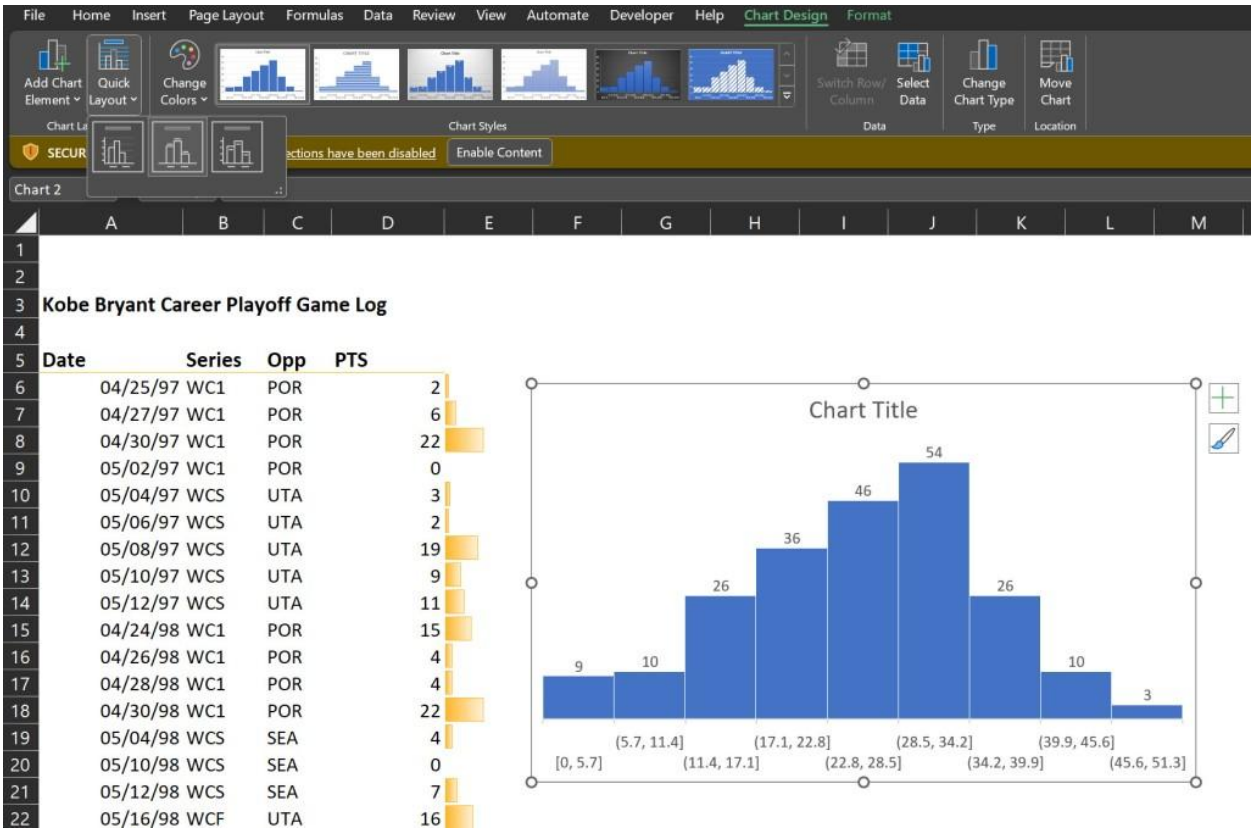
6R x 2C

	A	B	C	D	E
1	**Book Store**				
2					
3		total number of books	% sold for the highest price		
4		100	60%		
5					
6			number of books	unit profit	
7		highest price	60	$50	
8		lower price	40	$20	
9					
10			total profit	$3,800	
11					
12		$3,800			
13	60%				
14	70%				
15	80%				
16	90%				
17	100%				
18					

AVEDEV	average of absolute deviations of data elements from their mean
AVERAGE	average (i.e. mean) of a data set
AVERAGEIF	average of data elements satisfying some criteria (as for SUMIF)
AVERAGEIFS	average of data elements satisfying a set of criteria (as for SUMIFS)
COUNT	number of elements in a data set
DEVSQ	sum of squares of deviations
GEOMEAN	geometric mean of a data set
HARMEAN	harmonic mean of a data set
KURT	kurtosis of a data set
LARGE	kth largest element in a data set
MAX	largest element in a data set
MEDIAN	median of a data set
MIN	smallest element in a data set
MODE	mode of a data set
PERCENTRANK	percentage rank of an element in a data set
PERCENTILE	kth percentile of a data set
QUARTILE	kth quartile of a data set
RANK	rank of an element in a data set
SKEW	skewness of a data set
SMALL	kth smallest element in a data set
STANDARDIZE	standardized value of a data element
STDEV	sample standard deviation of a data set
STDEVP	population standard deviation of a data set
TRIMMEAN	mean of a data set excluding the smallest and largest elements
VAR	sample variance of a data set
VARP	population variance of a data set

Bin	*Frequency*
0-20	1
21-25	8
26-30	2
31-35	2
36-40	3
40+	2





ตัวอย่างที่ 1 : ตารางข้อมูลคะแนนสอบนักศึกษา

ตารางที่ 6.7 ตัวอย่างตารางข้อมูล (Raw Data)

ลำดับ	รหัสนักศึกษา	เพศ	คะแนนสอบ
1	65001	ชาย	78
2	65002	หญิง	85
3	65003	หญิง	92
4	65004	ชาย	67
5	65005	ชาย	74
6	65006	หญิง	88
7	65007	ชาย	90
8	65008	หญิง	81
9	65009	ชาย	69

10	65010	หญิง	95

📌 กำหนดให้คอลัมน์ **คะแนนสอบ** อยู่ที่เซลล์ D2:D11

ตัวอย่างที่ 2 : การคำนวณค่าทางสถิติพื้นฐานด้วย Excel

1) ค่าเฉลี่ย (Mean)

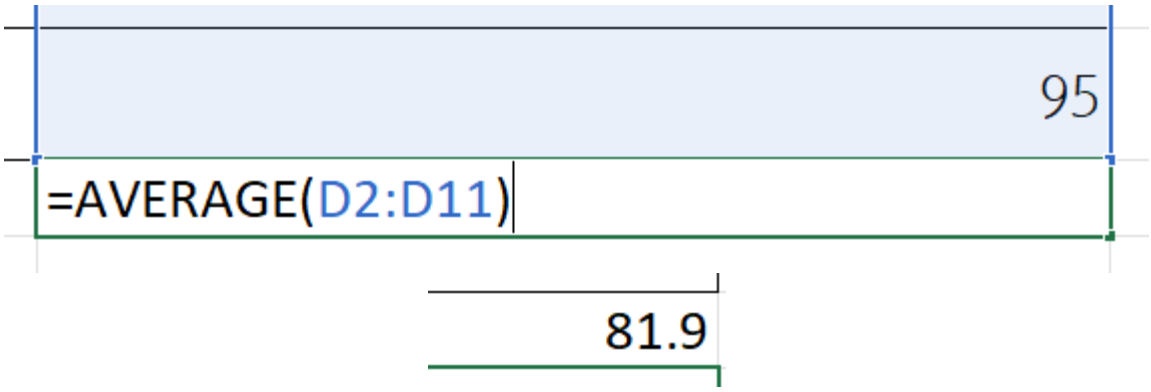
=AVERAGE(D2:D11)

ลำดับ	รหัสนักศึกษา	เพศ	คะแนนสอบ
1	65001	ชาย	78
2	65002	หญิง	85
3	65003	หญิง	92
4	65004	ชาย	67
5	65005	ชาย	74
6	65006	หญิง	88
7	65007	ชาย	90
8	65008	หญิง	81
9	65009	ชาย	69
10	65010	หญิง	95
			=A

ABS
ACCRINT
ACCRINTM
ACOS
ACOSH
ACOT
ACOTH
ADDRESS
AGGREGATE
AMORLINC
AND
ARABIC

Returns the absolute value

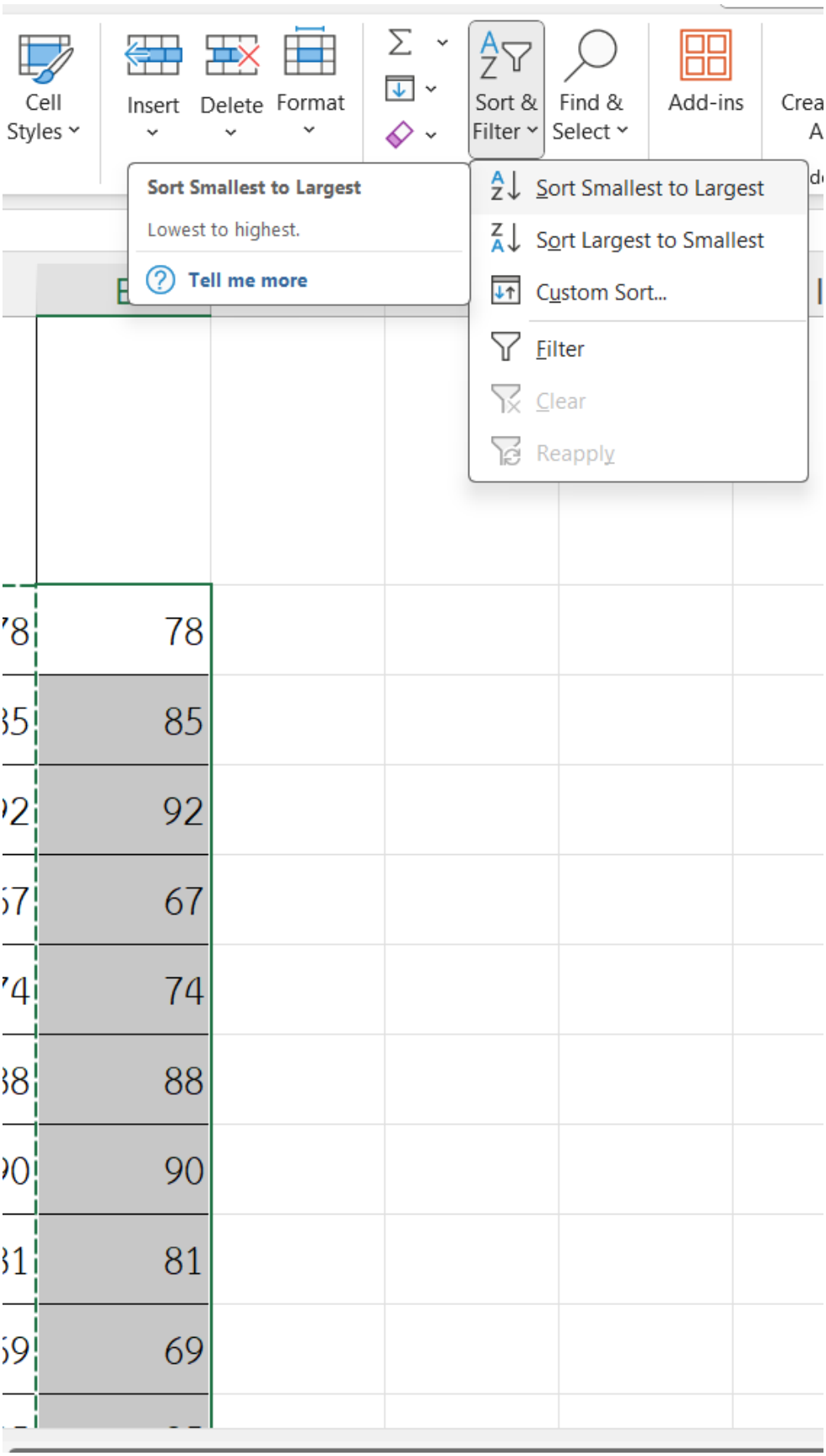
Sheet1



ผลลัพธ์

2) ค่ามัธยฐาน (Median)

=MEDIAN(D2:D11)


Cell Styles
Insert
Delete
Format
Sort & Filter
Find & Select
Add-ins
Sort Smallest to Largest
Lowest to highest.
Tell me more
Sort Smallest to Largest
Sort Largest to Smallest
Custom Sort...
Filter
Clear
Reapply
78
85
92
67
74
88
90
81
69

3) ค่าสูงสุด / ต่ำสุด

=MAX(D2:D11)

=MIN(D2:D11)

4) ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation)

=STDEV.S(D2:D11)

ค่าเฉลี่ย :	81.9
ค่ามัธยฐาน :	83
ค่า MODE :	#N/A
ค่า S.D.	9.734817238

คำอธิบาย:

- STDEV.S ใช้กับข้อมูลตัวอย่าง (Sample)
- หากเป็นข้อมูลประชากร ใช้ STDEV.P

ตัวอย่างที่ 3 : การนับจำนวนข้อมูลและใช้เงื่อนไข

1) นับจำนวนนักศึกษาทั้งหมด

=COUNT(D2:D11)

ค่าเฉลี่ย :	81.9
ค่ามัธยฐาน :	83
ค่า MODE :	#N/A
ค่า S.D.	9.734817238
จำนวน นศ. ทั้งหมด :	10

2) นับจำนวนนักศึกษาที่สอบผ่าน (≥ 80 คะแนน)

=COUNTIF(D2:D11, ">=80")

ค่าเฉลี่ย :	81.9
ค่ามัธยฐาน :	83
ค่า MODE :	#N/A
ค่า S.D.	9.734817238
จำนวน นศ. ทั้งหมด :	10
จำนวน นศ. ที่สอบผ่าน 80:	6
จำนวน นศ. ที่สอบไม่ผ่าน :	4

ตัวอย่างที่ 4 : การใช้ IF เพื่อจัดกลุ่มผลการเรียน

เพิ่มคอลัมน์ **ผลการสอบ** (คอลัมน์ E)

=IF(D2>=80,"ผ่าน","ไม่ผ่าน")

ลำดับ	รหัสนักศึกษา	เพศ	คะแนนสอบ			
1	65001	ชาย	78	67		ไม่ผ่าน
2	65002	หญิง	85	69		ผ่าน
3	65003	หญิง	92	74		ผ่าน
4	65004	ชาย	67	78		ไม่ผ่าน
5	65005	ชาย	74	81	83	ไม่ผ่าน
6	65006	หญิง	88	85		ผ่าน
7	65007	ชาย	90	88		ผ่าน
8	65008	หญิง	81	90		ผ่าน
9	65009	ชาย	69	92		ไม่ผ่าน
10	65010	หญิง	95	95		ผ่าน
		ค่าเฉลี่ย :	81.9			
		ค่ามัธยฐาน :	83			
		ค่า MODE :	#N/A			
		ค่า S.D.	9.734817238			
		จำนวน นศ. ทั้งหมด :	10			

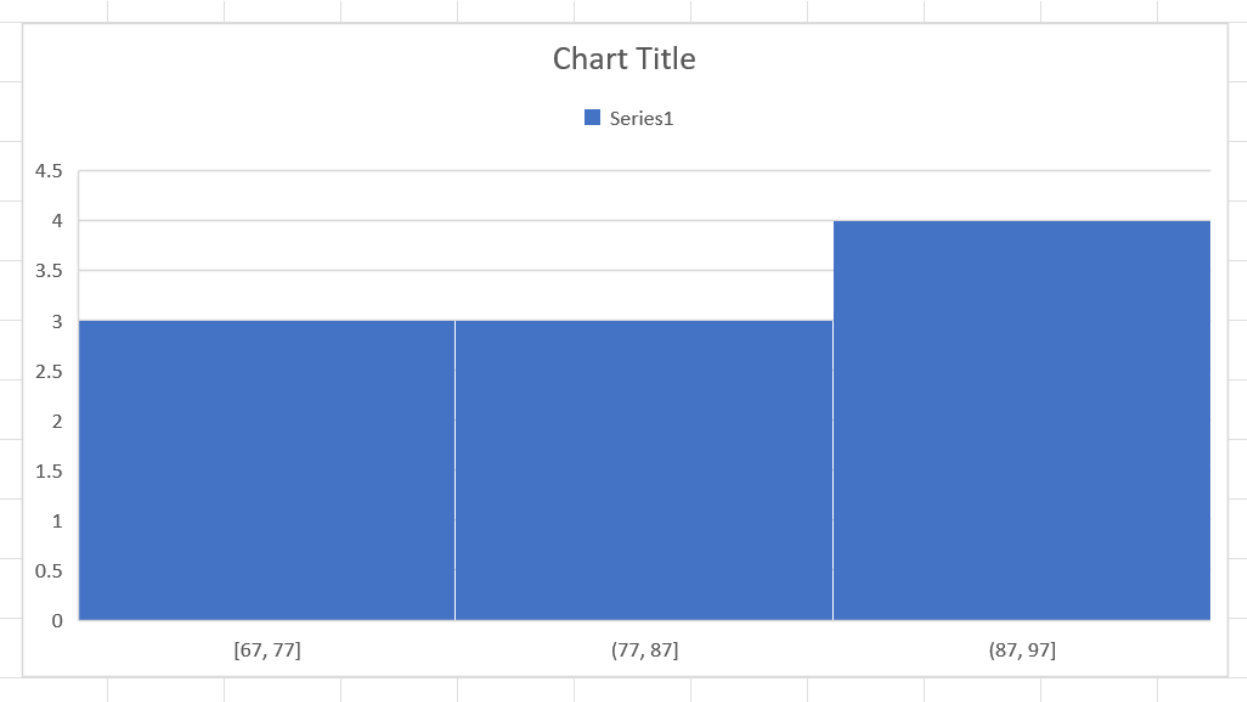
แล้วลากสูตรลงทั้งคอลัมน์

ตัวอย่างที่ 5 : การสร้าง Histogram แสดงการกระจายคะแนน

ขั้นตอน

1. เลือกข้อมูล D2:D11
2. ไปที่ **Insert ⟶ Statistic Chart ⟶ Histogram**
3. ปรับ Bin ให้เหมาะสม (เช่น ช่วงละ 10 คะแนน)

📌 ใช้ Histogram เพื่อวิเคราะห์การกระจายของข้อมูลเชิงปริมาณแบบต่อเนื่อง



ตัวอย่างที่ 6 : การใช้ PivotTable วิเคราะห์ข้อมูล

วัตถุประสงค์

เปรียบเทียบ **ค่าเฉลี่ยคะแนนสอบจำแนกตามเพศ**

ขั้นตอน

1. เลือกตารางข้อมูลทั้งหมด
2. ไปที่ **Insert ⟶ PivotTable**
3. กำหนด

- Rows : เพศ
- Values : คะแนนสอบ ⟶ Average

ตารางที่ 6.8 ผลลัพธ์ตัวอย่าง

เพศ	**คะแนนเฉลี่ย**
ชาย	75.6
หญิง	88.2

สามารถสร้าง **PivotChart** เพื่อแสดงผลในรูปกราฟแท่งได้ทันที

ตัวอย่างที่ 7 : การเตรียมข้อมูลก่อนส่งต่อไป SPSS

Excel สามารถใช้เป็นเครื่องมือ **Data Preparation** โดย

- ตรวจสอบ Missing Values
- ตรวจสอบ Outliers เบื้องต้น
- จัดรูปแบบข้อมูลให้เป็นตัวเลข
- บันทึกไฟล์เป็น .xlsx หรือ .csv

6.2.2 SPSS

SPSS เป็นซอฟต์แวร์สำหรับการวิเคราะห์ข้อมูลเชิงสถิติ เหมาะสำหรับงานวิจัยและการทดสอบสมมติฐาน ความสามารถหลัก

- Descriptive Statistics
- t-test, ANOVA
- Correlation, Regression
- Factor Analysis

6.3 การเตรียมและทำความสะอาดข้อมูล (Data Cleaning)

การทำความสะอาดข้อมูลเป็นขั้นตอนที่สำคัญก่อนการวิเคราะห์ เพื่อให้ผลลัพธ์มีความถูกต้องและน่าเชื่อถือ

ปัญหาที่พบบ่อย

- ข้อมูลสูญหาย (Missing Values)
- ข้อมูลซ้ำซ้อน (Duplicate Data)
- ค่าผิดปกติ (Outliers)
- รูปแบบข้อมูลไม่สอดคล้องกัน

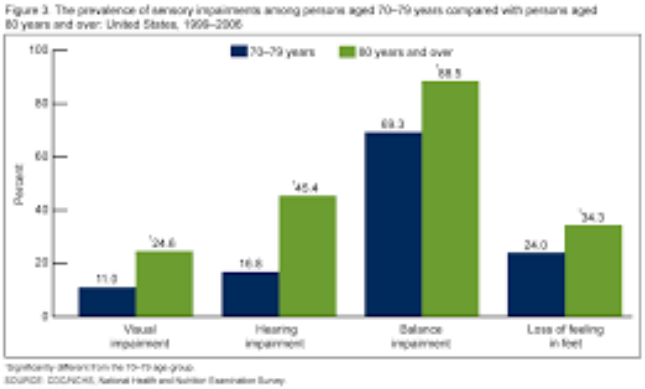
วิธีจัดการข้อมูล

- Excel : Filter, Remove Duplicates, Conditional Formatting
- SPSS : Frequencies, Recode, Boxplot

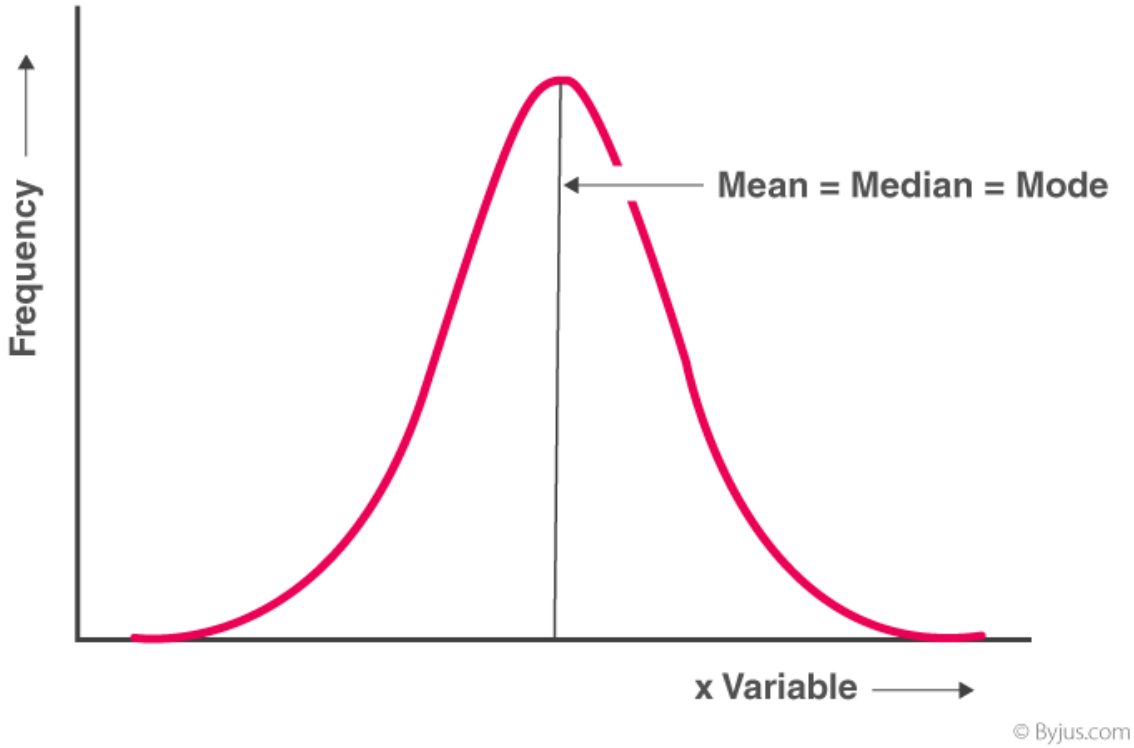
6.4 การวิเคราะห์ข้อมูล (Data Analysis)

การวิเคราะห์ข้อมูลเชิงปริมาณ หมายถึง กระบวนการนำข้อมูลที่อยู่ในรูปของตัวเลขมาจัดการ ประมวลผล และวิเคราะห์ด้วยวิธีการทางสถิติและคณิตศาสตร์ เพื่ออธิบายลักษณะของข้อมูล ค้นหารูปแบบ ความสัมพันธ์ ความแตกต่าง หรือแนวโน้มของข้อมูล และนำผลที่ได้ไปใช้ในการสรุปผล การตัดสินใจ หรือการทดสอบสมมติฐานทางวิชาการอย่างมีเหตุผลและเป็นระบบ

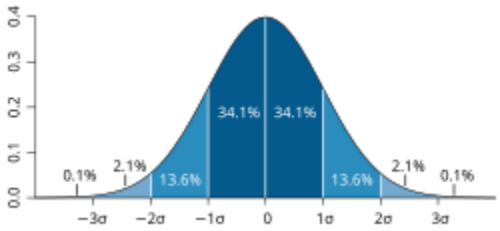
6.4.1 การวิเคราะห์เชิงพรรณนา (Descriptive Statistics)



ภาพที่ 6.6 ...



ภาพที่ 6.7 ...



ภาพที่ 6.8 ...

การวิเคราะห์เชิงพรรณนา (Descriptive Statistics) เป็นการวิเคราะห์ข้อมูลเชิงปริมาณในขั้นพื้นฐาน โดยมีวัตถุประสงค์เพื่อ อธิบายลักษณะทั่วไปของข้อมูล ที่รวบรวมมา เช่น ค่ากลาง การกระจาย และแนวโน้มของข้อมูล โดยไม่มุ่งเน้นการอนุมานหรือทดสอบสมมติฐานไปยังประชากร

การวิเคราะห์เชิงพรรณนาเป็นขั้นตอนสำคัญก่อนการวิเคราะห์เชิงอนุมาน เนื่องจากช่วยให้ผู้วิเคราะห์เข้าใจโครงสร้างของข้อมูล ตรวจสอบความผิดปกติ และเลือกวิธีวิเคราะห์ขั้นต่อไปได้อย่างเหมาะสม

องค์ประกอบหลักของการวิเคราะห์เชิงพรรณนา

การวิเคราะห์เชิงพรรณนาประกอบด้วย 2 ส่วนสำคัญ ได้แก่

1. ค่ากลางของข้อมูล (Measures of Central Tendency)
2. การกระจายของข้อมูล (Measures of Dispersion)

1) ค่ากลางของข้อมูล (Measures of Central Tendency)
ค่ากลางเป็นค่าที่แสดงตำแหน่งศูนย์กลางหรือค่าที่พบได้บ่อยในชุดข้อมูล

1.1 ค่าเฉลี่ย (Mean)

ค่าเฉลี่ย คือผลรวมของข้อมูลทั้งหมดหารด้วยจำนวนข้อมูล เป็นค่ากลางที่ใช้บ่อยที่สุดในการวิเคราะห์ข้อมูลเชิงปริมาณ

สูตร

$$\text{Mean} = \frac{\sum x}{n}$$

ลักษณะเด่น

- ใช้ข้อมูลทุกค่าในการคำนวณ
- เหมาะกับข้อมูลที่กระจายตัวปกติ
- ไวต่อค่าผิดปกติ (Outliers)

ตัวอย่าง

คะแนนสอบ: 60, 70, 80, 90

Mean = (60+70+80+90) / 4 = 75

การคำนวณใน Excel

=AVERAGE(D2:D11)

การคำนวณใน SPSS

Analyze ⟶ Descriptive Statistics ⟶ Descriptives

ตัวอย่างข้อมูลตัวอย่าง (Sample Data) สำหรับงานวิจัยเชิงปริมาณ
หัวข้อวิจัยตัวอย่าง
“ผลสัมฤทธิ์ทางการเรียนของนักศึกษาในรายวิชาสถิติพื้นฐาน”

ตารางที่ 6.9 ข้อมูลตัวอย่างของคะแนนสอบนักศึกษา

ลำดับ	รหัสนักศึกษา	เพศ	คะแนนสอบ (100 คะแนน)
1	65001	ชาย	60
2	65002	หญิง	70
3	65003	หญิง	80
4	65004	ชาย	90
5	65005	ชาย	75
6	65006	หญิง	85
7	65007	ชาย	68
8	65008	หญิง	92
9	65009	ชาย	77
10	65010	หญิง	88

การนำข้อมูลตัวอย่างไปใช้ในการวิเคราะห์

1) การวิเคราะห์เชิงพรรณนา (Descriptive Statistics)

ตัวแปรที่วิเคราะห์: คะแนนสอบ

- ค่าเฉลี่ย (Mean)
- มัธยฐาน (Median)
- ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation)

การคำนวณใน Excel

สมมติคะแนนสอบอยู่ที่ช่วงเซลล์ **D2:D11**

=AVERAGE(D2:D11)

=MEDIAN(D2:D11)

=STDEV.S(D2:D11)

การคำนวณใน SPSS

1. Analyze ⟶ Descriptive Statistics ⟶ Descriptives
2. เลือกตัวแปร “คะแนนสอบ”
3. เลือก Mean, Std. Deviation, Minimum, Maximum

ตัวอย่างการเขียนผลการวิเคราะห์ (เชิงงานวิจัย)

จากการวิเคราะห์คะแนนสอบของนักศึกษา จำนวน 10 คน พบว่าคะแนนสอบมีค่าเฉลี่ยเท่ากับ 78.5 คะแนน ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 10.2 คะแนน แสดงให้เห็นว่านักศึกษามีผลสัมฤทธิ์ทางการเรียนอยู่ในระดับปานกลางถึงดี และมีการกระจายของคะแนนในระดับปานกลาง

การต่อยอดเป็นงานวิจัยขั้นถัดไป

จากข้อมูลชุดเดียวกัน สามารถตั้งโจทย์เพิ่มได้ เช่น

- เปรียบเทียบคะแนนสอบจำแนกตามเพศ ⟶ **Independent t-test**
- วิเคราะห์ความสัมพันธ์ระหว่างเวลาเรียนกับคะแนนสอบ ⟶ **Correlation**
- พยากรณ์คะแนนสอบจากจำนวนชั่วโมงเรียน ⟶ **Regression**

สรุปเชิงการสอน

ชุดข้อมูลตัวอย่างนี้

- เป็นข้อมูลเชิงปริมาณจริง
- มีรายละเอียดพอสำหรับการฝึกวิเคราะห์
- ใช้ได้ทั้ง Excel และ SPSS
- เหมาะสำหรับงานวิจัยเชิงพรรณนาและเชิงอนุมานเบื้องต้น

1.2 มัธยฐาน (Median)

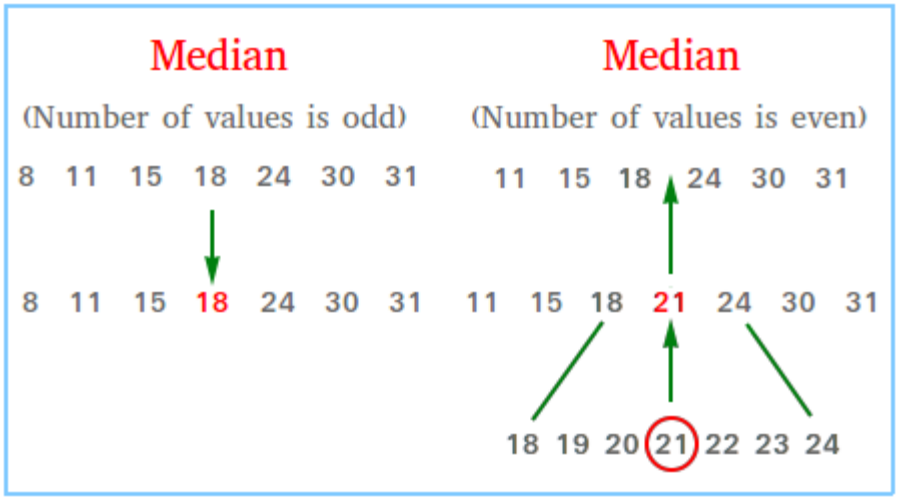
1, 3, 3, **6**, 7, 8, 9

Median = **6**

1, 2, 3, **4**, **5**, 6, 8, 9

Median = (4 + 5) ÷ 2

= **4.5**



ความหมาย

มัธยฐาน (Median) คือ ค่าที่อยู่ตรงกลางของชุดข้อมูล เมื่อทำการเรียงข้อมูลจากค่าน้อยไปหาค่ามาก โดยมัธยฐานจะแบ่งข้อมูลออกเป็นสองส่วนที่มีจำนวนข้อมูลเท่ากัน

อธิบายรายละเอียดจากตัวอย่าง

ชุดข้อมูลตัวอย่าง

คะแนนสอบของนักศึกษา

60, 70, 80, 90

ขั้นตอนที่ 1 : เรียงลำดับข้อมูล

เรียงข้อมูลจากน้อยไปมาก (ในตัวอย่างนี้เรียงไว้แล้ว)

60 ⟶ 70 ⟶ 80 ⟶ 90

ขั้นตอนที่ 2 : ตรวจสอบจำนวนข้อมูล

จำนวนข้อมูลทั้งหมด = 4 ค่า

เป็นจำนวน คู่

📌 หลักการสำคัญ

- ถ้าจำนวนข้อมูลเป็น คี่ ⟶ มัธยฐานคือค่าตรงกลาง
- ถ้าจำนวนข้อมูลเป็น คู่ ⟶ มัธยฐานคือค่าเฉลี่ยของ “สองค่ากลาง”

ขั้นตอนที่ 3 : ระบุค่ากลาง

ข้อมูล 4 ค่า จะมีค่ากลาง 2 ค่า คือ

- ค่าที่ 2 = 70
- ค่าที่ 3 = 80

ขั้นตอนที่ 4 : คำนวณมัธยฐาน

$$\text{Median} = \frac{70 + 80}{2} = 75$$

📌 ดังนั้น มัธยฐานของชุดข้อมูลนี้เท่ากับ 75 คะแนน

การตีความหมายของผลลัพธ์

มัธยฐาน = 75 หมายความว่า

- คะแนนสอบครึ่งหนึ่งของนักศึกษามีค่าน้อยกว่าหรือเท่ากับ 75
- อีกครึ่งหนึ่งมีค่ามากกว่าหรือเท่ากับ 75

มัธยฐานจึงเป็นค่าที่สะท้อน “ตำแหน่งกึ่งกลางของข้อมูล” ได้ดี โดยไม่ถูกดึงค่าไปทางใดทางหนึ่ง

เหตุผลที่มัธยฐานเหมาะกับข้อมูลบางประเภท

จากตัวอย่าง หากมีคะแนนที่ผิดปกติ (Outlier) เช่น

60, 70, 80, 150

- ค่าเฉลี่ย (Mean) จะสูงขึ้นมาก
- แต่มัธยฐานยังคงเป็นค่ากลางที่สะท้อนข้อมูลส่วนใหญ่ได้ดีกว่า

📌 จึงนิยมใช้มัธยฐานกับ

- ข้อมูลรายได้
- คะแนนสอบที่มีค่าผิดปกติ
- ข้อมูลที่มีการกระจายไม่สมมาตร (Skewed Data)

การคำนวณมัธยฐานด้วยโปรแกรม

การคำนวณใน Excel

หากข้อมูลอยู่ในช่วงเซลล์ D2:D11

=MEDIAN(D2:D11)

Excel จะจัดการเรียงข้อมูลและคำนวณค่ากลางให้อัตโนมัติ

การคำนวณใน SPSS

1. Analyze ⟶ Descriptive Statistics ⟶ Frequencies
2. เลือกตัวแปรคะแนนสอบ
3. คลิก Statistics ⟶ เลือก Median
4. ดูผลลัพธ์ในตาราง Output

จากตัวอย่างนี้แสดงให้เห็นว่า

- มัธยฐานเป็นค่ากลางที่คำนวณจากตำแหน่งของข้อมูล
- เหมาะกับข้อมูลที่มี Outliers หรือการกระจายไม่สมมาตร
- สามารถคำนวณได้ง่ายทั้งด้วยมือ Excel และ SPSS

1.3 ฐานนิยม (Mode)

ฐานนิยม คือค่าที่ปรากฏบ่อยที่สุดในชุดข้อมูล

ลักษณะเด่น

- ใช้ได้กับข้อมูลเชิงปริมาณและเชิงคุณภาพ
- เหมาะกับการวิเคราะห์ข้อมูลเชิงหมวดหมู่
- อาจไม่มีหรือมีหลายค่าได้

ตัวอย่าง

ข้อมูล: 70, 80, 80, 90 ⟶ Mode = 80

การคำนวณใน Excel

=MODE.SNGL(D2:D11)

การคำนวณใน SPSS

Analyze ⟶ Descriptive Statistics ⟶ Frequencies

2) การกระจายของข้อมูล (Measures of Dispersion)

การวัดการกระจายช่วยอธิบายว่าข้อมูลกระจายตัวมากน้อยเพียงใดรอบค่ากลาง

2.1 ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation)

ส่วนเบี่ยงเบนมาตรฐานเป็นค่าที่บอกระดับการกระจายของข้อมูลจากค่าเฉลี่ย

ลักษณะเด่น

- ค่ายิ่งสูง แสดงว่าข้อมูลกระจายมาก
- ค่ายิ่งต่ำ แสดงว่าข้อมูลเกาะกลุ่ม
- ใช้ร่วมกับค่าเฉลี่ยเพื่ออธิบายข้อมูลได้ชัดเจน

การตีความ

- SD ต่ำ ⟶ คะแนนใกล้เคียงกัน
- SD สูง ⟶ คะแนนแตกต่างกันมาก

การคำนวณใน Excel

=STDEV.S(D2:D11)

การคำนวณใน SPSS

Analyze ⟶ Descriptive Statistics ⟶ Descriptives

การใช้ Descriptive Statistics ร่วมกับกราฟ

การวิเคราะห์เชิงพรรณนามักใช้ร่วมกับการแสดงผลข้อมูล เช่น

- Histogram ⟶ การกระจายของข้อมูล
- Boxplot ⟶ ค่ากลางและ Outliers
- Bar Chart ⟶ เปรียบเทียบค่ากลางระหว่างกลุ่ม

ตัวอย่างการแปลผลเชิงพรรณนา

จากการวิเคราะห์คะแนนสอบของนักศึกษา พบว่าคะแนนเฉลี่ยเท่ากับ 78.5 คะแนน ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 6.2 คะแนน แสดงว่าคะแนนสอบของนักศึกษาส่วนใหญ่มีค่าใกล้เคียงกับค่าเฉลี่ย และไม่มีการกระจายตัวมากนัก

สรุปท้ายหัวข้อ 6.4.1

การวิเคราะห์เชิงพรรณนาเป็นขั้นตอนพื้นฐานที่สำคัญในการวิเคราะห์ข้อมูลเชิงปริมาณ ช่วยให้ผู้วิเคราะห์เข้าใจลักษณะทั่วไปของข้อมูล ทั้งในด้านค่ากลางและการกระจาย ซึ่งเป็นพื้นฐานสำคัญก่อนการวิเคราะห์เชิงอนุมานและการตัดสินใจทางสถิติในขั้นต่อไป

6.4.2 การวิเคราะห์เชิงอนุมาน (Inferential Statistics)

ใช้ในการสรุปผลจากกลุ่มตัวอย่างไปยังประชากร

- t-test
- ANOVA
- Correlation
- Regression

6.4.2 การวิเคราะห์เชิงอนุมาน (Inferential Statistics)

การวิเคราะห์เชิงอนุมาน (Inferential Statistics) คือการวิเคราะห์ข้อมูลเชิงปริมาณที่มีวัตถุประสงค์เพื่อ **สรุปผลจากกลุ่มตัวอย่าง (Sample) ไปอ้างอิงถึงประชากร (Population)** โดยอาศัยหลักความน่าจะเป็นและสถิติ เพื่อช่วยในการตัดสินใจและการทดสอบสมมติฐานทางวิชาการ

การวิเคราะห์เชิงอนุมานมักใช้เมื่อต้องการทราบว่า

- ความแตกต่างที่พบในข้อมูล **มีนัยสำคัญทางสถิติหรือไม่**
- ตัวแปรต่าง ๆ **มีความสัมพันธ์กันหรือไม่**
- ตัวแปรหนึ่ง **สามารถใช้พยากรณ์อีกตัวแปรหนึ่งได้หรือไม่**

Comparison of MEANS	Degrees of Freedom	Application	Assumptions	Test Statistic
One Sample Z-Test	Not Applicable	Testing the difference of a sample mean, x-bar, with a known population mean, μ (fixed mean, historical mean, or targeted mean)	Normal distribution Known population σ.	$Z = \frac{\overline{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$
One Sample t-test	n-1	Testing the difference of one sample mean, x-bar, with a known population mean, μ (fixed mean, historical mean, or targeted mean)	Normal distribution Population standard deviation, σ, is unknown.	$t = \frac{\overline{x} - \mu}{\frac{s}{\sqrt{n}}}$
Two Sample t-test	$n_1 + n_2 - 2$	Testing difference of two sample means when population variances unknown but considered equal	Normal Distribution Requires standard pooled deviation calculation, s_p	$t = \frac{\overline{x}_1 - \overline{x}_2}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$
Paired t-test	n - 1	Testing two sample means when their respective population standard deviations are unknown but considered equal. Data recorded in pairs and each pair has a difference, d.	Normal Distribution Two dependent samples Always two-tailed test S_d = standard deviation of the differences of all samples	$t = \frac{\overline{d}\sqrt{n}}{s_d}$
One-Way ANOVA	$n_1 - 1$ & $n_2 - 1$	Testing the difference of three or more population means	Normal Distribution s_1^2 and s_2^2 represent sample variances	$F = \frac{(s_1)^2}{(s_2)^2}$

ภาพที่ 6.

			Self-esteem scale				Self-handicapping scale			
		df	Sum of squares	Mean square	F	p	Sum of squares	Mean square	F	p
Level of education	BG	3	9.34	1.781	0.204	.89	188.54	62.845	0.850	.47
	WG	78	681.55	8.738			5769.72	73.971		
	Total	81	686.89				5958.26			
Referee category	BG	3	37.20	12.399	1.489	.22	378.92	126.307	1.766	.16
	WG	78	649.69	8.329			5579.34	71.530		
	Total	81	686.89				5958.26			
Experience (yrs)	BG	4	33.42	8.356	0.985	.42	667.44	166.859	2.428	.05
	WG	77	653.47	8.487			5290.82	68.712		
	Total	81	686.89				5958.26			
Profession	BG	5	7.50	1.499	0.168	.97	396.61	79.921	1.093	.37
	WG	76	679.39	8.939			5558.65	73.140		
	Total	81	686.89				5958.26			
Level of monthly income	BG	3	20.82	6.939	0.813	.49	124.94	41.647	0.557	.65
	WG	78	666.07	8.539			5833.32	74.786		
	Total	81	686.89				5958.26			
Place of residence	BG	2	77.06	38.527	4.951	.01*	65.80	21.934	0.290	.83
	WG	79	609.83	7.719			5892.46	75.544		
	Total	81	686.89				5958.26			
Proficiency in English	BG	4	35.35	8.837	1.044	.39	227.34	56.834	0.764	.55
	WG	77	651.54	8.462			5730.92	74.428		
	Total	81	686.89				5958.26			

* p<.01; BG: between groups; WG: within groups

ภาพที่ 6.

หลักการพื้นฐานของการวิเคราะห์เชิงอนุมาน

1. ใช้ข้อมูลจาก **กลุ่มตัวอย่าง** แทนประชากรทั้งหมด
2. มีการกำหนด **สมมติฐาน (Hypothesis)**
    - สมมติฐานศูนย์ (H_0)
    - สมมติฐานทางเลือก (H_1)
3. ใช้ระดับนัยสำคัญทางสถิติ (เช่น α = 0.05)

4. ตัดสินใจจากค่า **p-value (Sig.)**

📌 หากค่า Sig. < 0.05 ⟶ ปฏิเสธสมมติฐานศูนย์

📌 หากค่า Sig. ≥ 0.05 ⟶ ไม่ปฏิเสธสมมติฐานศูนย์

เทคนิคการวิเคราะห์เชิงอนุมานที่สำคัญ

1) t-test

t-test ใช้สำหรับทดสอบว่า **ค่าเฉลี่ยของข้อมูล 2 กลุ่มแตกต่างกันอย่างมีนัยสำคัญหรือไม่**

ลักษณะสำคัญ

- ใช้กับข้อมูลเชิงปริมาณ
- เปรียบเทียบค่าเฉลี่ย
- เหมาะกับกรณีมีกลุ่มตัวอย่างขนาดไม่ใหญ่มาก

ประเภทของ t-test

- Independent Samples t-test : เปรียบเทียบ 2 กลุ่มอิสระ
- Paired Samples t-test : เปรียบเทียบข้อมูลก่อน–หลัง
- One-Sample t-test : เปรียบเทียบกับค่าเกณฑ์

ตัวอย่าง

- คะแนนสอบก่อนเรียนและหลังเรียน
- รายได้ของพนักงานชายและหญิง

2) ANOVA (Analysis of Variance)

ANOVA ใช้สำหรับทดสอบว่า **ค่าเฉลี่ยของข้อมูลตั้งแต่ 3 กลุ่มขึ้นไปแตกต่างกันหรือไม่**

ลักษณะสำคัญ

- เป็นการขยายแนวคิดจาก t-test
- ใช้เปรียบเทียบหลายกลุ่มพร้อมกัน
- ลดความผิดพลาดจากการทดสอบหลายครั้ง

ตัวอย่าง

- เปรียบเทียบคะแนนสอบของนักศึกษา 3 ห้องเรียน
- เปรียบเทียบผลผลิตจากปุ๋ย 4 สูตร

📌 หากผล ANOVA มีนัยสำคัญ ต้องทำ **Post Hoc Test** เพื่อดูว่ากลุ่มใดแตกต่างจากกลุ่มใด

3) Correlation (สหสัมพันธ์)

Correlation ใช้ในการวิเคราะห์ว่า **ตัวแปรสองตัวมีความสัมพันธ์กันหรือไม่** และมีทิศทางอย่างไร

ลักษณะสำคัญ

- ค่าวัดอยู่ระหว่าง -1 ถึง +1
- ค่าบวก ⟶ ความสัมพันธ์ทิศทางเดียวกัน
- ค่าลบ ⟶ ความสัมพันธ์ทิศทางตรงข้าม
- ค่าใกล้ 0 ⟶ ความสัมพันธ์ต่ำหรือไม่มี

ตัวอย่าง

- ส่วนสูงกับน้ำหนัก
- เวลาอ่านหนังสือกับคะแนนสอบ

📌 Correlation ไม่ได้บอกเหตุ–ผล เพียงบอกความสัมพันธ์เท่านั้น

4) Regression (การถดถอย)

Regression ใช้ในการวิเคราะห์ว่า **ตัวแปรอิสระสามารถอธิบายหรือพยากรณ์ตัวแปรตามได้มากน้อยเพียงใด**

ลักษณะสำคัญ

- ใช้เพื่อการพยากรณ์ (Prediction)
- วิเคราะห์ความสัมพันธ์เชิงเหตุผล
- มีทั้ง Simple Regression และ Multiple Regression

ตัวอย่าง

- พยากรณ์ยอดขายจากงบโฆษณา
- พยากรณ์คะแนนสอบจากชั่วโมงเรียน

ผลลัพธ์ที่ได้ เช่น

- สมการพยากรณ์
- ค่า R^2 (ความสามารถในการอธิบายข้อมูล)

ชุดตัวอย่างโจทย์วิจัย + สถิติที่เหมาะสม

1) ก่อน–หลังเรียน (ผลของการสอน/นวัตกรรม)

โจทย์วิจัย: “คะแนนหลังเรียนของนักศึกษาหลังใช้สื่อการสอนออนไลน์สูงกว่าก่อนเรียนหรือไม่”

ตัวแปร: คะแนนก่อนเรียน / คะแนนหลังเรียน (คนเดิม)

สถิติที่เหมาะสม: Paired Samples t-test

เหตุผลสั้น ๆ: เปรียบเทียบค่าเฉลี่ย “2 ครั้ง” ของกลุ่มเดิม

ตัวอย่างโจทย์ + ชุดข้อมูลตัวอย่าง (ก่อน–หลังเรียน) สำหรับงานวิจัย “คะแนนหลังเรียนสูงกว่าก่อนเรียนหรือไม่” ที่เหมาะกับ Paired Samples t-test

1) โจทย์วิจัย (ตัวอย่างเขียนเป็นทางการ)

ชื่อเรื่อง (ตัวอย่าง): ผลของสื่อการสอนออนไลน์ต่อผลสัมฤทธิ์ทางการเรียนของนักศึกษา

คำถามวิจัย: หลังใช้สื่อการสอนออนไลน์ นักศึกษามีคะแนนหลังเรียนสูงกว่าก่อนเรียนหรือไม่

สมมติฐาน

- H0: ค่าเฉลี่ยคะแนนหลังเรียน ไม่สูงกว่า ก่อนเรียน (หรือ μหลัง − μก่อน ≤ 0)
- H1: ค่าเฉลี่ยคะแนนหลังเรียน สูงกว่า ก่อนเรียน (μหลัง − μก่อน > 0) ⟵ (ทดสอบด้านเดียว)

ตัวแปร

- คะแนนก่อนเรียน (Pretest): 0–100
- คะแนนหลังเรียน (Posttest): 0–100 (คนเดิม)

2) ตัวอย่างข้อมูล (สมมติ) 20 คน

โครงสร้างข้อมูล: 1 แถว = นักศึกษา 1 คน มีคะแนนก่อนและหลัง

ตารางที่ 6.10 ตัวอย่างข้อมูลคะแนนก่อน-หลังเรียนของนักศึกษาจำนวน 20 คน

รหัส	ก่อนเรียน	หลังเรียน	ผลต่าง (หลัง-ก่อน)
S01	52	64	12
S02	61	70	9
S03	45	55	10
S04	70	78	8
S05	58	66	8
S06	63	72	9

S07	49	60	11
S08	55	65	10
S09	62	74	12
S10	47	57	10
S11	66	75	9
S12	59	68	9
S13	53	63	10
S14	71	80	9
S15	46	56	10
S16	60	71	11
S17	57	67	10
S18	64	76	12
S19	50	61	11
S20	68	79	11

หมายเหตุ: ข้อมูลนี้ถูกทำให้เห็นแนวโน้ม “คะแนนหลังสูงขึ้น” ชัดเจน เพื่อเป็นตัวอย่างสำหรับ paired t-test

3) รูปแบบโจทย์ในแบบฝึกวิเคราะห์ (ใส่ในรายงาน/แบบฝึกได้เลย)
กำหนดให้: ตารางคะแนน Pretest/Posttest ของนักศึกษา 20 คนหลังใช้สื่อการสอนออนไลน์
ให้ทำ:

1. ตั้งสมมติฐาน H0/H1 ที่ระดับนัยสำคัญ 0.05
2. วิเคราะห์ด้วย Paired Samples t-test (one-tailed)
3. สรุปผลว่าคะแนนหลังเรียนสูงกว่าก่อนเรียนหรือไม่
4. (เสริม) รายงานค่าเฉลี่ยก่อน–หลัง, ค่าเฉลี่ยผลต่าง, t, df, p-value และข้อสรุป

1) โครงสร้างข้อมูล

ข้อมูลต้องเป็น “คนเดิม 2 ครั้ง” เช่น 1 แถว = นักศึกษา 1 คน

คอลัมน์อย่างน้อย:

- Pretest (ก่อนเรียน)
- Posttest (หลังเรียน)

2) ตั้งสมมติฐาน (ตามโจทย์ “หลังเรียนสูงกว่า” = ด้านเดียว)

กำหนดระดับนัยสำคัญ α = 0.05

- H0: μหลัง − μก่อน ≤ 0 (หลังไม่สูงกว่า)
- H1: μหลัง − μก่อน > 0 (หลังสูงกว่า) ☑ one-tailed

3) คำนวณ “ผลต่าง” ก่อน (หัวใจของ Paired t-test)

สร้างคอลัมน์ผลต่าง

$$d = \text{Post} - \text{Pre}$$

จากนั้นหา:

- $\bar{d}$= ค่าเฉลี่ยของผลต่าง
- s_d= ส่วนเบี่ยงเบนมาตรฐานของผลต่าง
- n= จำนวนคู่ข้อมูล
- $SE = \frac{s_d}{\sqrt{n}}$

แล้วคำนวณค่าสถิติ t:

$$t = \frac{\bar{d}}{\frac{s_d}{\sqrt{n}}}$$

และ df = n − 1

4) ตัดสินผล

- ถ้า p-value (one-tailed) < 0.05 ⟶ ปฏิเสธ H0 ⟶ หลังเรียนสูงกว่าก่อนเรียน “อย่างมีนัยสำคัญ”
- ถ้า p ≥ 0.05 ⟶ ไม่ปฏิเสธ H0

5) ตัวอย่าง “คำนวณจริง” จากข้อมูล 20 คนของคุณ

ผมคำนวณจากชุดข้อมูลที่คุณส่งมาได้ผลดังนี้:

- จำนวนตัวอย่าง $n = 20$
- ค่าเฉลี่ยผลต่าง $\bar{d} = 10.05$คะแนน
- SD ของผลต่าง $s_d \approx 1.234$
- $t \approx 36.41$
- $df = 19$
- p(one-tailed) $\approx 2.4 \times 10^{-19}$(เล็กมาก)

สรุป: p < 0.05 ⇒ คะแนนหลังเรียน สูงกว่าก่อนเรียนอย่างมีนัยสำคัญทางสถิติ ที่ระดับ .05

หมายเหตุ: ค่า t สูงมากเพราะข้อมูลตัวอย่างถูกสร้างให้ “เพิ่มขึ้นสม่ำเสมอ” เพื่อสาธิต

6) วิธีทำในโปรแกรม (เลือกใช้ได้)

A) SPSS

1. ใส่ข้อมูล 2 คอลัมน์: pre, post
2. Analyze ⟶ Compare Means ⟶ Paired-Samples T Test
3. จับคู่ pre กับ post ⟶ OK
4. อ่านผล:
    - “Paired Samples Test” ดูค่า Sig. (2-tailed)
    - ถ้าโจทย์เป็น “ด้านเดียว” ให้เอา p(two-tailed)/2 (ในกรณีทิศทางถูกต้องคือ post > pre)

B) Excel

1. ทำคอลัมน์ผลต่าง d = post - pre
2. หา AVERAGE(d), STDEV.S(d), COUNT(d)
3. คำนวณ t = mean_d / (sd_d / SQRT(n))
4. p-value (ด้านเดียว) ใช้ฟังก์ชัน:

- =T.DIST.RT(t, df) (Right-tail = ด้านเดียว แบบ post>pre)

7) ตัวอย่างประโยคสรุปใส่รายงาน (พร้อมใช้)

“ผลการทดสอบ Paired Samples t-test พบว่าคะแนนหลังเรียนสูงกว่าคะแนนก่อนเรียนอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 (t(19)=36.41, p<.001)”

2) เปรียบเทียบ 2 กลุ่มอิสระ (ชาย–หญิง / กลุ่มทดลอง–ควบคุม)

โจทย์วิจัย: “รายได้เฉลี่ยต่อเดือนของพนักงานชายและหญิงแตกต่างกันหรือไม่”

ตัวแปร: เพศ (2 กลุ่ม) ⟶ รายได้ต่อเดือน

สถิติที่เหมาะสม: Independent Samples t-test

เหตุผล: เปรียบเทียบค่าเฉลี่ยของ “2 กลุ่มอิสระ”

ตัวอย่างโจทย์ + ชุดข้อมูลตัวอย่าง สำหรับงานวิจัย “รายได้เฉลี่ยของพนักงานชายและหญิงแตกต่างกันหรือไม่” ที่เหมาะกับ Independent Samples t-test

1) โจทย์วิจัย (ตัวอย่างเขียนเป็นทางการ)

ชื่อเรื่อง (ตัวอย่าง): การเปรียบเทียบรายได้เฉลี่ยต่อเดือนของพนักงานชายและหญิงในองค์กร X

คำถามวิจัย: รายได้เฉลี่ยต่อเดือนของพนักงานชายและหญิงแตกต่างกันหรือไม่

สมมติฐาน (ทดสอบสองด้าน)

- H0: รายได้เฉลี่ยชาย = รายได้เฉลี่ยหญิง (μชาย = μหญิง)
- H1: รายได้เฉลี่ยชาย ≠ รายได้เฉลี่ยหญิง (μชาย ≠ μหญิง)

ตัวแปร

- ตัวแปรอิสระ: เพศ (ชาย/หญิง) = 2 กลุ่มอิสระ
- ตัวแปรตาม: รายได้ต่อเดือน (บาท)

2) ตัวอย่างข้อมูล (สมมติ) 2 กลุ่มอิสระ

โครงสร้างข้อมูล: 1 แถว = พนักงาน 1 คน

กลุ่มชาย (n=12)

ตารางที่ 6.11 ข้อมูลรายได้แยกตามเพศชาย

รหัส	**เพศ**	**รายได้/เดือน (บาท)**
M01	ชาย	34,000
M02	ชาย	30,500
M03	ชาย	36,200
M04	ชาย	28,700
M05	ชาย	33,100

M06	ชาย	31,800
M07	ชาย	29,400
M08	ชาย	37,500
M09	ชาย	32,600
M10	ชาย	35,000
M11	ชาย	30,200
M12	ชาย	34,800

กลุ่มหญิง (n=12)

ตารางที่ 6.11 ข้อมูลรายได้แยกตามเพศหญิง

รหัส	**เพศ**	**รายได้/เดือน (บาท)**
F01	หญิง	29,800
F02	หญิง	27,900
F03	หญิง	31,200
F04	หญิง	26,500
F05	หญิง	28,600
F06	หญิง	30,100
F07	หญิง	27,300
F08	หญิง	32,000
F09	หญิง	28,900
F10	หญิง	29,500
F11	หญิง	26,900
F12	หญิง	30,700

ชุดข้อมูลนี้ทำให้เห็นแนวโน้มรายได้ “ชายสูงกว่าเล็กน้อย” เพื่อให้เหมาะกับตัวอย่างการทดสอบ t-test

3) รูปแบบโจทย์แบบฝึกวิเคราะห์ (ใส่รายงานได้เลย)

กำหนดให้: ข้อมูลรายได้ต่อเดือนของพนักงานชาย 12 คน และหญิง 12 คน

ให้ทำ:

1. ตั้งสมมติฐาน H0/H1 ที่ระดับนัยสำคัญ 0.05
2. วิเคราะห์ด้วย Independent Samples t-test (two-tailed)
3. ตรวจสอบเงื่อนไข: ความเป็นอิสระของกลุ่ม, การกระจายใกล้ปกติ (หรือ sample พอ), และความแปรปรวนเท่ากัน (Levene's test)
4. สรุปผลว่ารายได้เฉลี่ยแตกต่างกันหรือไม่

ได้เลยครับ — วิธีทำ Independent Samples t-test (เปรียบเทียบรายได้ชาย–หญิง) แบบทำตามได้จริง มี 2 ทาง: ทำ "ด้วยมือ (แนวคิด+สูตร)" และทำ "ด้วยโปรแกรม (SPSS/Excel)"

A) ขั้นตอนวิธีทำ (แนวคิดมาตรฐานในงานวิจัย)

1) ระบุโจทย์/ตัวแปร

- กลุ่ม 1: ชาย
- กลุ่ม 2: หญิง
- ตัวแปรตาม: รายได้ต่อเดือน (ต่อเนื่อง)

2) ตั้งสมมติฐาน

- H0: μชาย = μหญิง
- H1: μชาย ≠ μหญิง (สองด้าน)

  กำหนดระดับนัยสำคัญ เช่น α = 0.05

3) ตรวจสอบเงื่อนไขก่อนใช้ t-test (สำคัญในรายงาน)

1. กลุ่มอิสระ: คนละคนกัน (ชายกับหญิงไม่ซ้ำคน) ☑
2. การกระจายใกล้ปกติ (แนะนำตรวจ):
   - n น้อย: ใช้ Shapiro–Wilk ในแต่ละกลุ่ม
   - n มาก: ดู histogram/QQ plot ก็พอ
3. ความแปรปรวนเท่ากัน? ใช้ Levene's test
   - ถ้า p(Levene) > 0.05 ⟶ ใช้ t-test แบบ Equal variances assumed
   - ถ้า p(Levene) ≤ 0.05 ⟶ ใช้แบบ Welch's t-test (Equal variances not assumed)

4) คำนวณค่าเฉลี่ย/SD ของแต่ละกลุ่ม

รายงานอย่างน้อย:

- Mean ชาย, SD ชาย
- Mean หญิง, SD หญิง

5) ทำ Independent t-test แล้วดูผล

ผลหลักที่ใช้ตัดสิน:

- ค่า t, df, p-value
- ถ้า $p < 0.05$ ⟶ ปฏิเสธ H0 ⟶ รายได้เฉลี่ย “แตกต่างกัน”
- ถ้า $p \geq 0.05$ ⟶ ไม่ปฏิเสธ H0 ⟶ “ไม่แตกต่างกันอย่างมีนัยสำคัญ”

6) รายงานขนาดอิทธิพล (แนะนำ)

- ใช้ Cohen’s d

  ช่วยบอกว่า “ต่างมากน้อยแค่ไหน” ไม่ใช่ดู p อย่างเดียว

B) ทำใน SPSS (ง่ายสุดสำหรับรายงาน)

1. ใส่ข้อมูล 2 คอลัมน์:
    - gender (ชาย=1, หญิง=2)
    - income
2. ไปที่ Analyze ⟶ Compare Means ⟶ Independent-Samples T Test
3. ย้าย income ไปช่อง Test Variable(s)
4. ย้าย gender ไปช่อง Grouping Variable
5. กด Define Groups… ใส่ 1 และ 2
6. กด OK

อ่านผล SPSS

- ตาราง Group Statistics: ดู Mean/SD ของแต่ละกลุ่ม
- ตาราง Independent Samples Test
    - ดู Levene’s test ก่อน
    - แล้วเลือกอ่านแถวที่ถูกต้อง:
        - *Equal variances assumed* หรือ *not assumed*
    - ดูค่า Sig. (2-tailed) = p-value

C) ทำใน Excel (ทำได้ถ้ามี Data Analysis ToolPak)

เตรียมข้อมูล

- คอลัมน์ A: รายได้ชาย
- คอลัมน์ B: รายได้หญิง

วิธีทำ

1. เปิด ToolPak: File ⟶ Options ⟶ Add-ins ⟶ Excel Add-ins ⟶ ติ๊ก Analysis ToolPak
2. ไปที่ Data ⟶ Data Analysis ⟶ t-Test: Two-Sample Assuming Equal Variances
    - ถ้าสงสัยความแปรปรวนไม่เท่ากัน ให้ใช้ t-Test: Two-Sample Assuming Unequal Variances
3. กำหนดช่วงข้อมูลชาย/หญิง และ α=0.05
4. อ่านค่า P(T<=t) two-tail (สองด้าน)

D) สูตร (กรณีอธิบายในรายงาน/ทำด้วยมือ)

1. Pooled t-test (เมื่อทั้งสองกลุ่มมีแปรปรวนเท่ากัน)

ใช้เมื่อ: ต้องการเปรียบเทียบค่าเฉลี่ยของสองกลุ่ม และทั้งสองกลุ่มมีความแปรปรวนใกล้เคียงกัน

สูตร:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

โดยที่

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

- $\bar{x}_1, \bar{x}_2$: ค่าเฉลี่ยของแต่ละกลุ่ม
- s_1, s_2: ส่วนเบี่ยงเบนมาตรฐานของแต่ละกลุ่ม
- n_1, n_2: ขนาดกลุ่ม
- s_p: ส่วนเบี่ยงเบนรวม (pooled)
- $df = n_1 + n_2 - 2$: ดีกรีอิสระ

ใช้เมื่อ:

- กลุ่มมีขนาดใกล้เคียงกัน
- แปรปรวนของสองกลุ่มไม่แตกต่างกันมาก (ตรวจสอบด้วย Levene's Test ได้)

2. Welch’s t-test (เมื่อแปรปรวนไม่เท่ากัน)

ใช้เมื่อ: ต้องการเปรียบเทียบค่าเฉลี่ยของสองกลุ่ม แต่ แปรปรวนแตกต่างกัน หรือ ขนาดกลุ่มต่างกัน

สูตร:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

ดีกรีอิสระ (df):

ใช้สูตร Welch–Satterthwaite ซึ่งซับซ้อนกว่า และโดยมากปล่อยให้โปรแกรมคำนวณ

ใช้เมื่อ:

- กลุ่มมีขนาดต่างกันมาก
- แปรปรวนของกลุ่มแตกต่างกัน (unequal variances)

3. Cohen’s d (ค่าขนาดอิทธิพล Effect Size)

ใช้วัดว่า ความแตกต่างระหว่างกลุ่มมีขนาด "มากน้อยแค่ไหน" โดยไม่ขึ้นกับขนาดกลุ่ม

สูตร:

$$d = \frac{\bar{x}_1 - \bar{x}_2}{s_p}$$

ตีความ:

- 0.2 = ขนาดเล็ก
- 0.5 = ขนาดกลาง
- 0.8 = ขนาดใหญ่

ใช้ประกอบกับค่า t-test เพื่อให้รู้ว่า "ผลลัพธ์ที่มีนัยสำคัญนั้น แท้จริงแล้วมีน้ำหนักมากแค่ไหน"

E) ตัวอย่าง “ประโยคสรุปผล” ใส่รายงาน (ปรับตัวเลขตามผลจริง)

- กรณีมีนัยสำคัญ:
  “ผลการทดสอบ Independent t-test พบว่ารายได้เฉลี่ยของพนักงานชายแตกต่างจากพนักงานหญิงอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 (t(df)=..., p=...).”
- กรณีไม่มีนัยสำคัญ:
  “ไม่พบความแตกต่างของรายได้เฉลี่ยระหว่างพนักงานชายและหญิงอย่างมีนัยสำคัญทางสถิติ (t(df)=..., p=...).”

1) รูปแบบการทดสอบที่ใช้

- สถิติ: Independent Samples t-test
- ลักษณะการทดสอบ: สองด้าน (Two-tailed)
- ระดับนัยสำคัญ: α = 0.05

2) สมมติฐาน (ต้องเขียนแบบนี้)

- H0: รายได้เฉลี่ยต่อเดือนของพนักงานชายและหญิง ไม่แตกต่างกัน
  μชาย=μหญิง
- H1: รายได้เฉลี่ยต่อเดือนของพนักงานชายและหญิง แตกต่างกัน
  μชาย≠μหญิง

***จุดสำคัญของ “สองด้าน” คือ ไม่กำหนดทิศทาง แค่ดูว่า “ต่างกันหรือไม่” ไม่สนว่าใครมากกว่าใคร

3) ขั้นตอนการวิเคราะห์ (เขียนในบทที่ 3 ได้เลย)
ขั้นที่ 1 ตรวจสอบเงื่อนไข

1. กลุ่มเป็นอิสระจากกัน (ชาย ≠ หญิง) ☑
2. ตัวแปรเป็นข้อมูลเชิงปริมาณ (รายได้) ☑
3. ตรวจสอบความแปรปรวนด้วย Levene’s test
   - p > .05 ⟶ ใช้ Equal variances assumed
   - p ≤ .05 ⟶ ใช้ Equal variances not assumed (Welch)

ขั้นที่ 2 คำนวณสถิติเชิงพรรณนา

รายงาน:

- ค่าเฉลี่ย (Mean)
- ส่วนเบี่ยงเบนมาตรฐาน (SD)
  แยก ชาย / หญิง

ขั้นที่ 3 วิเคราะห์ Independent Samples t-test

ดูค่า:

- t
- df
- Sig. (2-tailed) ⟵ ★ สำคัญที่สุด

4) เกณฑ์การตัดสินผล (จำสูตรนี้ได้เลย)

- ถ้า Sig. (2-tailed) < 0.05
  ⟶ ✖ ปฏิเสธ H0
  ⟶ รายได้เฉลี่ย แตกต่างกันอย่างมีนัยสำคัญ
- ถ้า Sig. (2-tailed) ≥ 0.05
  ⟶ ☑ ไม่ปฏิเสธ H0
  ⟶ รายได้เฉลี่ย ไม่แตกต่างกัน

5) ตัวอย่างการเขียน “สรุปผลการวิจัย” (เลือกใช้ตามผล)

☑ กรณี *แตกต่างกัน*

“ผลการทดสอบ Independent Samples t-test แบบสองด้าน พบว่ารายได้เฉลี่ยต่อเดือนของพนักงานชายและหญิงแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

(t(df)=..., p < .05)”

✖ กรณี *ไม่แตกต่างกัน*

“ผลการทดสอบ Independent Samples t-test แบบสองด้าน พบว่ารายได้เฉลี่ยต่อเดือนของพนักงานชายและหญิงไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05
(t(df)=..., p > .05)”

6) สรุปสั้น ๆ สำหรับจำสอบ / ทำรายงาน

- คำว่า แตกต่างกันหรือไม่ ⟶ ใช้ Two-tailed
- ดูค่า Sig. (2-tailed) เท่านั้น
- อย่าเขียนว่า “ชายมากกว่าหญิง” ถ้าโจทย์ไม่ได้กำหนดทิศทาง

3) เปรียบเทียบมากกว่า 2 กลุ่ม (หลายห้อง/หลายสูตร/หลายระดับ)

โจทย์วิจัย: “คะแนนเฉลี่ยวิชาคณิตศาสตร์ของนักศึกษา 3 ห้องเรียนแตกต่างกันหรือไม่”
ตัวแปร: ห้องเรียน (3 กลุ่ม) ⟶ คะแนน
สถิติที่เหมาะสม: One-way ANOVA
เหตุผล: เปรียบเทียบค่าเฉลี่ยมากกว่า 2 กลุ่ม
ถ้าพบแตกต่าง (Sig. < .05) ให้ทำ Post Hoc (เช่น Tukey) เพื่อดูว่ากลุ่มใดต่างจากกลุ่มใด

1) โจทย์วิจัย (ตัวอย่างเขียนเป็นทางการ)
ชื่อเรื่อง (ตัวอย่าง): การเปรียบเทียบผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ของนักศึกษา 3 ห้องเรียน
คำถามวิจัย: คะแนนเฉลี่ยวิชาคณิตศาสตร์ของนักศึกษา 3 ห้องเรียนแตกต่างกันหรือไม่
สมมติฐาน (Two-sided ตามธรรมชาติของ ANOVA)

- H0: ค่าเฉลี่ยคะแนนคณิตฯ ของทั้ง 3 ห้อง เท่ากัน

  $\mu_1=\mu_2=\mu_3$

- H1: มีอย่างน้อย 1 ห้องที่ค่าเฉลี่ย แตกต่าง จากห้องอื่น

  อย่างน้อยหนึ่ง μ ต่างจากที่เหลือ

ตัวแปร

- ตัวแปรอิสระ: ห้องเรียน (3 กลุ่ม: ห้อง A, B, C)
- ตัวแปรตาม: คะแนนวิชาคณิตศาสตร์ (0–100)

2) ตัวอย่างข้อมูล (สมมติ) 3 ห้องเรียน (ห้องละ 10 คน)

โครงสร้างข้อมูล: 1 แถว = นักศึกษา 1 คน

ตารางที่ 6.12 ข้อมูลคะแนนของนักเรียนแยกตามห้องเรียน

รหัส	**ห้องเรียน**	**คะแนนคณิตฯ**
A01	A	62
A02	A	68
A03	A	65
A04	A	70
A05	A	66
A06	A	64
A07	A	69
A08	A	63
A09	A	67
A10	A	71
B01	B	72
B02	B	75
B03	B	78
B04	B	74
B05	B	77
B06	B	73
B07	B	79
B08	B	76
B09	B	80
B10	B	74

C01	C	58
C02	C	60
C03	C	55
C04	C	62
C05	C	57
C06	C	59
C07	C	61
C08	C	56
C09	C	63
C10	C	54

หมายเหตุ: ชุดนี้ตั้งใจให้เห็นแนวโน้มว่า B สูงกว่า A และ C เพื่อให้มีโอกาส Sig. < .05 และต่อด้วย Tukey ได้

3) รูปแบบโจทย์ในแบบฝึกวิเคราะห์ (ใส่รายงาน/แบบฝึกได้เลย)
กำหนดให้: คะแนนวิชาคณิตศาสตร์ของนักศึกษา 3 ห้องเรียน (A, B, C) ห้องละ 10 คน
ให้ทำ:

1. ตั้งสมมติฐาน H0/H1 ที่ระดับนัยสำคัญ 0.05
2. วิเคราะห์ด้วย One-way ANOVA
3. ถ้า Sig. < .05 ให้ทำ Post Hoc (Tukey HSD)
4. สรุปว่าห้องใด “แตกต่างจากห้องใด” พร้อมรายงานค่าเฉลี่ยแต่ละห้อง

ถ้าคุณต้องการ ผมทำต่อให้ครบ “วิธีทำ” แบบเดียวกับ 2 ข้อก่อนหน้าได้เลย เช่น

- ทำตารางค่าเฉลี่ย/SD รายห้อง
- คำนวณ/สรุปผล ANOVA (F, df, p)
- สรุปผล Tukey ว่าคู่ไหนต่างกัน
  บอกได้เลยว่าคุณอยากทำใน SPSS / Excel / หรือสรุปขั้นตอนแบบรายงานบทที่ 3-4 ครับ

4) เปรียบเทียบหลายปัจจัย (2 ปัจจัยขึ้นไป)

โจทย์วิจัย: “วิธีสอน (บรรยาย/โครงงาน) และเพศ มีผลต่อคะแนนสอบหรือไม่ และมีปฏิสัมพันธ์ร่วมกันหรือไม่”

ตัวแปร: วิธีสอน (2 กลุ่ม) + เพศ (2 กลุ่ม) ⟶ คะแนน

สถิติที่เหมาะสม: Two-way ANOVA

เหตุผล: วิเคราะห์ผลของ “2 ปัจจัย” และ interaction

4) เปรียบเทียบหลายปัจจัย (2 ปัจจัยขึ้นไป)

ตัวอย่างโจทย์วิจัย + ชุดข้อมูลตัวอย่าง (2×2) สำหรับ Two-way ANOVA และ วิธีทำแบบเป็นขั้นตอน (พร้อมแนวทางสรุปผล/รายงาน)

1) โจทย์วิจัย (ตัวอย่างเขียนเป็นทางการ)

ชื่อเรื่อง (ตัวอย่าง): ผลของวิธีสอนและเพศต่อคะแนนสอบ และการมีปฏิสัมพันธ์ระหว่างปัจจัย

คำถามวิจัย:

1. วิธีสอน (บรรยาย/โครงงาน) มีผลต่อคะแนนสอบหรือไม่
2. เพศ (ชาย/หญิง) มีผลต่อคะแนนสอบหรือไม่
3. วิธีสอนและเพศมี ปฏิสัมพันธ์ (Interaction) ต่อคะแนนสอบหรือไม่

ตัวแปร

- ตัวแปรอิสระ (ปัจจัยที่ 1): วิธีสอน 2 ระดับ = บรรยาย / โครงงาน
- ตัวแปรอิสระ (ปัจจัยที่ 2): เพศ 2 ระดับ = ชาย / หญิง
- ตัวแปรตาม: คะแนนสอบ (0–100)

2) สมมติฐาน (ของ Two-way ANOVA จะมี 3 ชุด)

ให้กำหนด α = 0.05

(A) ผลหลักของ “วิธีสอน” (Main effect A)

- H0A: μบรรยาย = μโครงงาน
- H1A: อย่างน้อยหนึ่งค่าเฉลี่ยต่างกัน

(B) ผลหลักของ “เพศ” (Main effect B)

- H0B: μชาย = μหญิง
- H1B: อย่างน้อยหนึ่งค่าเฉลี่ยต่างกัน

(C) ปฏิสัมพันธ์ “วิธีสอน×เพศ” (Interaction A×B)

- H0AB: ไม่มีปฏิสัมพันธ์ (ผลของวิธีสอนไม่ขึ้นกับเพศ)
- H1AB: มีปฏิสัมพันธ์ (ผลของวิธีสอนขึ้นกับเพศ)

---

3) ชุดข้อมูลตัวอย่าง (สมมติ) แบบ 2×2 (กลุ่มละ 8 คน = รวม 32 คน)
ตั้งใจทำให้เห็น “มี Interaction เล็กน้อย” เพื่อให้ใช้สอนการแปลผลได้

ตารางที่ 6.13 ข้อมูลผลคะแนนของนักเรียนแยกตามเพศ

รหัส	วิธีสอน	เพศ	คะแนน
L_M01	บรรยาย	ชาย	62
L_M02	บรรยาย	ชาย	64
L_M03	บรรยาย	ชาย	66
L_M04	บรรยาย	ชาย	63
L_M05	บรรยาย	ชาย	65
L_M06	บรรยาย	ชาย	61
L_M07	บรรยาย	ชาย	67
L_M08	บรรยาย	ชาย	64
L_F01	บรรยาย	หญิง	70
L_F02	บรรยาย	หญิง	72
L_F03	บรรยาย	หญิง	69
L_F04	บรรยาย	หญิง	71
L_F05	บรรยาย	หญิง	73
L_F06	บรรยาย	หญิง	70
L_F07	บรรยาย	หญิง	74
L_F08	บรรยาย	หญิง	71
P_M01	โครงงาน	ชาย	78
P_M02	โครงงาน	ชาย	80

P_M03	โครงงาน	ชาย	77
P_M04	โครงงาน	ชาย	79
P_M05	โครงงาน	ชาย	81
P_M06	โครงงาน	ชาย	78
P_M07	โครงงาน	ชาย	82
P_M08	โครงงาน	ชาย	79
P_F01	โครงงาน	หญิง	75
P_F02	โครงงาน	หญิง	76
P_F03	โครงงาน	หญิง	74
P_F04	โครงงาน	หญิง	77
P_F05	โครงงาน	หญิง	76
P_F06	โครงงาน	หญิง	75
P_F07	โครงงาน	หญิง	78
P_F08	โครงงาน	หญิง	74

ภาพรวมแนวโน้ม (จากการตั้งข้อมูล):

- ผู้ชายได้คะแนน “สูงขึ้นมาก” เมื่อเรียนแบบโครงงาน (บรรยาย⟶โครงงาน เพิ่มเยอะ)
- ผู้หญิงได้คะแนน “สูงขึ้นไม่มาก/บางทีลดเล็กน้อย” เมื่อเป็นโครงงาน

  ⇒ จึงมีโอกาสเกิด Interaction

4) วิธีทำ Two-way ANOVA (ขั้นตอนมาตรฐานในรายงาน)

ขั้นที่ 1 ตรวจเงื่อนไข (Assumptions)

1. กลุ่มอิสระ (คนละคน) ☑
2. คะแนนเป็นเชิงปริมาณ ☑
3. ความเป็นปกติของเศษเหลือ (Residual) — ดู Q-Q plot หรือ Shapiro-Wilk ของ residual
4. ความแปรปรวนเท่ากัน (Homogeneity) — Levene’s test

ถ้าเงื่อนไขแปรปรวนไม่เท่ากันมาก อาจพิจารณาแปลงข้อมูล/ใช้วิธี robust แต่ในงานทั่วไป Two-way ANOVA ยังใช้ได้ถ้าสมดุลและไม่ผิดหนัก

ขั้นที่ 2 สถิติเชิงพรรณนา

ทำตาราง Mean, SD แยก 4 กลุ่ม:

- บรรยาย-ชาย, บรรยาย-หญิง, โครงงาน-ชาย, โครงงาน-หญิง
  และทำกราฟเส้น Interaction plot (ช่วยอ่านง่ายมาก)

ขั้นที่ 3 รัน Two-way ANOVA

ผลจะได้ p-value 3 บรรทัดหลัก:

- วิธีสอน (A)
- เพศ (B)
- วิธีสอน×เพศ (A×B)

ขั้นที่ 4 แปลผล (กฎสำคัญ)

- ถ้า Interaction มีนัยสำคัญ ($p < .05$):
  ให้ “โฟกัสที่ Interaction ก่อน” แล้วทำ Simple effects (เช่น เปรียบเทียบวิธีสอนภายในเพศชาย และภายในเพศหญิง แยกกัน)
- ถ้า Interaction ไม่มีนัยสำคัญ ($p \geq .05$):
  แปลผล Main effects ของวิธีสอน และเพศ ได้ตรง ๆ

ขั้นที่ 5 (แนะนำ) รายงานขนาดอิทธิพล

รายงาน Partial Eta Squared (ηp^2) สำหรับ A, B, A×B

5) ทำใน SPSS (วิธีทำที่ใช้จริงในงานวิจัย)

1. ใส่ข้อมูล 3 คอลัมน์:
   - method (บรรยาย=1, โครงงาน=2)
   - gender (ชาย=1, หญิง=2)
   - score
2. ไปที่ Analyze ⟶ General Linear Model ⟶ Univariate...
3. ย้าย score ไปที่ Dependent Variable
4. ย้าย method และ gender ไปที่ Fixed Factor(s)
5. กด Model... เลือก Full factorial (เพื่อให้มี interaction)
6. กด Options...

- ติ๊ก Descriptive statistics
- ติ๊ก Homogeneity tests (Levene)
- ติ๊ก Estimates of effect size (ηp^2)
- ใน Display Means for: เลือก method*gender และกด Compare (ถ้าต้องการ)

7. กด Plots... (แนะนำ)
    - Horizontal axis: method
    - Separate lines: gender
    - Add ⟶ Continue
8. OK

อ่านผล SPSS

- ตาราง Levene's test: ดูความแปรปรวนเท่ากันไหม
- ตาราง Tests of Between-Subjects Effects:
    - ดู Sig. ของ method, gender, method*gender
- กราฟ interaction plot: ดูเส้น "ตัดกัน/ไม่ขนาน" ⟶ มีแนวโน้ม interaction

6) ถ้าพบ Interaction แล้ว "ทำอะไรต่อ" (คำตอบที่อาจารย์ชอบ)

ทำ Simple Effects เช่น

- ภายในเพศชาย: เปรียบเทียบบรรยาย vs โครงงาน (t-test/EMMeans compare)
- ภายในเพศหญิง: เปรียบเทียบบรรยาย vs โครงงาน

  หรืออีกมุม:
- ภายในบรรยาย: ชาย vs หญิง
- ภายในโครงงาน: ชาย vs หญิง

ใน SPSS ทำได้ด้วย:

- EMMEANS / Estimated Marginal Means ใน Options (เลือก method*gender แล้ว Compare)

  หรือ
- แยกไฟล์ตามเพศแล้วรัน t-test แยก (กรณีสอน/ทำรายงานพื้นฐาน)

7) ตัวอย่างประโยคสรุปผล (เอาไปเติมตัวเลขจริง)

กรณี Interaction มีนัยสำคัญ

“ผลการวิเคราะห์ Two-way ANOVA พบว่ามีปฏิสัมพันธ์ระหว่างวิธีสอนและเพศต่อคะแนนสอบอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ($F(1, ...)=..., p<.05, \eta p^2=...$). จึงทำการวิเคราะห์ผลอย่างง่าย (simple effects) พบว่า....”
กรณี Interaction ไม่มีนัยสำคัญ
“ไม่พบปฏิสัมพันธ์ระหว่างวิธีสอนและเพศ (p>.05) แต่พบว่าวิธีสอนมีผลต่อคะแนนสอบอย่างมีนัยสำคัญ (p<.05) / เพศมีผล (p<.05) ....”

5) หาความสัมพันธ์ของ 2 ตัวแปรเชิงปริมาณ

โจทย์วิจัย: “เวลาอ่านหนังสือ (ชั่วโมง/สัปดาห์) มีความสัมพันธ์กับคะแนนสอบหรือไม่”
ตัวแปร: เวลาอ่านหนังสือ, คะแนนสอบ
สถิติที่เหมาะสม: Correlation (Pearson’s r)
เหตุผล: ทั้งสองเป็นตัวแปรเชิงปริมาณ ต้องการดูความสัมพันธ์
ถ้าข้อมูลไม่เป็นปกติ/เป็นอันดับ ใช้ Spearman’s rho

1) โจทย์วิจัย (ตัวอย่างเขียนเป็นทางการ)
ชื่อเรื่อง (ตัวอย่าง): ความสัมพันธ์ระหว่างเวลาอ่านหนังสือกับคะแนนสอบของนักศึกษา
คำถามวิจัย: เวลาอ่านหนังสือ (ชั่วโมง/สัปดาห์) มีความสัมพันธ์กับคะแนนสอบหรือไม่
สมมติฐาน (แบบสองด้าน ใช้บ่อยสุด)

- H0: เวลาอ่านหนังสือ ไม่มีความสัมพันธ์ กับคะแนนสอบ ($\rho = 0$)
- H1: เวลาอ่านหนังสือ มีความสัมพันธ์ กับคะแนนสอบ ($\rho \ne 0$)

ถ้าโจทย์ต้องการ “ยิ่งอ่านมากยิ่งได้คะแนนสูง” ก็ทำด้านเดียวได้ แต่โดยมาตรฐาน “มีความสัมพันธ์หรือไม่” = สองด้าน
ตัวแปร

- เวลาอ่านหนังสือ (ชั่วโมง/สัปดาห์) = เชิงปริมาณ
- คะแนนสอบ (0–100) = เชิงปริมาณ

2) ชุดข้อมูลตัวอย่าง (สมมติ) 20 คน
1 แถว = นักศึกษา 1 คน

ตารางที่ 6.14 ข้อมูลเวลาและคะแนนของนักศึกษา

รหัส	เวลาอ่าน (ชม./สัปดาห์)	คะแนนสอบ
S01	2	52
S02	3	55
S03	4	57
S04	5	60
S05	6	63
S06	7	65
S07	8	68
S08	9	71
S09	10	73
S10	11	76
S11	12	78
S12	13	80
S13	14	82
S14	15	85
S15	16	86
S16	17	88
S17	18	90
S18	19	92
S19	20	93
S20	21	95

หมายเหตุ: ชุดนี้ตั้งใจให้เห็นแนวโน้มความสัมพันธ์เชิงบวกชัดเจน เพื่อสาธิต Pearson r

3) วิธีทำ (Pearson's r) แบบทีละขั้น

ขั้นที่ 1 ตรวจสอบว่าใช้ Pearson ได้ไหม

ใช้ Pearson's r เมื่อโดยทั่วไป:

- ทั้ง 2 ตัวแปรเป็นเชิงปริมาณ (☑)
- ความสัมพันธ์ "ค่อนข้างเป็นเส้นตรง" (ดู Scatter plot)
- การกระจายไม่เพี้ยนหนัก/มี outlier รุนแรง

ถ้าไม่ผ่าน (ไม่ปกติ, มี outlier, เป็นข้อมูลอันดับ/ลิเคิร์ต) ⟶ ใช้ Spearman's rho

ขั้นที่ 2 ทำกราฟกระจาย (สำคัญมาก)

วาด Scatter plot: แกน X = เวลาอ่าน, แกน Y = คะแนน

- ถ้าเป็นแนวเส้นขึ้น/ลงชัด ⟶ ใช้ Pearson ได้ดี
- ถ้าเป็นโค้ง/กระจายแปลก ⟶ พิจารณา Spearman หรือโมเดลอื่น

ขั้นที่ 3 คำนวณค่าสหสัมพันธ์ r

แนวคิด: r อยู่ระหว่าง -1 ถึง +1

- $r > 0$ = ความสัมพันธ์เชิงบวก
- $r < 0$ = ความสัมพันธ์เชิงลบ
- r ใกล้ 0 = แทบไม่สัมพันธ์

(ปกติทำด้วยโปรแกรม/Excel/SPSS จะได้ค่า r และ p-value เลย)

ขั้นที่ 4 ทดสอบนัยสำคัญ (p-value)

โปรแกรมจะให้ Sig. (2-tailed)

- ถ้า $p < .05$ ⟶ มีความสัมพันธ์อย่างมีนัยสำคัญ
- ถ้า $p \geq .05$ ⟶ ยังสรุปไม่ได้ว่ามีความสัมพันธ์

ขั้นที่ 5 สรุปผล "ทั้งทิศทาง + ความแรง + นัยสำคัญ"

เกณฑ์ความแรงแบบจำง่าย (คร่าว ๆ):

- $|r|$ = 0.00–0.19 น้อยมาก

- |r| = 0.20–0.39 น้อย
- |r| = 0.40–0.59 ปานกลาง
- |r| = 0.60–0.79 สูง
- |r| = 0.80–1.00 สูงมาก

4) ทำใน SPSS (เร็วสุด)

1. ใส่ข้อมูล 2 คอลัมน์: study_hours, exam_score
2. Analyze ⟶ Correlate ⟶ Bivariate…
3. เลือก 2 ตัวแปร ⟶ ติ๊ก Pearson และเลือก Two-tailed ⟶ OK
4. อ่านผล:

- Pearson Correlation = r
- Sig. (2-tailed) = p-value
- N = จำนวนตัวอย่าง

5) ทำใน Excel

ได้ค่า r

- =CORREL(A2:A21, B2:B21) (A=เวลาอ่าน, B=คะแนน)

ถ้าต้องการ p-value (ทางเลือก)

Excel ไม่มี p ของ correlation ให้ตรง ๆ แบบง่ายเท่า SPSS แต่ทำได้ด้วยสูตรทดสอบ t:

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}, \quad df = n-2$$

แล้วหา p(two-tailed) ด้วย:

- =T.DIST.2T(ABS(t), n-2)

6) เมื่อไหร่ควรใช้ Spearman’s rho

ใช้ Spearman เมื่อ:

- ข้อมูลไม่เป็นปกติ/มี outlier มาก
- ความสัมพันธ์ไม่เป็นเส้นตรงแต่ “เพิ่ม/ลดตามลำดับ” (monotonic)

- ตัวแปรเป็น อันดับ (ranking) หรือ Likert ที่ถือเป็นอันดับ

ใน SPSS: Analyze ⟶ Correlate ⟶ Bivariate… ⟶ ติ๊ก Spearman

7) ตัวอย่างประโยคสรุปใส่รายงาน (เติมตัวเลข r, p จากโปรแกรม)

“ผลการวิเคราะห์สหสัมพันธ์เพียร์สันพบว่าเวลาอ่านหนังสือมีความสัมพันธ์เชิงบวกกับคะแนนสอบอย่างมีนัยสำคัญทางสถิติ (r = …, p < .05)”

หรือถ้าไม่เป็นนัยสำคัญ:

“ไม่พบความสัมพันธ์อย่างมีนัยสำคัญระหว่างเวลาอ่านหนังสือกับคะแนนสอบ (r = …, p > .05)”

6) พยากรณ์ด้วยตัวแปรอิสระ 1 ตัว

โจทย์วิจัย: “งบโฆษณาต่อเดือนสามารถพยากรณ์ยอดขายได้หรือไม่”

ตัวแปร: งบโฆษณา ⟶ ยอดขาย

สถิติที่เหมาะสม: Simple Linear Regression

เหตุผล: ต้องการ “พยากรณ์” และสร้างสมการเส้นตรง

7) พยากรณ์ด้วยหลายตัวแปรอิสระ

โจทย์วิจัย: “ยอดขายสามารถพยากรณ์จากงบโฆษณา ราคา และจำนวนพนักงานขายได้หรือไม่”

ตัวแปร: งบโฆษณา + ราคา + จำนวนพนักงานขาย ⟶ ยอดขาย

สถิติที่เหมาะสม: Multiple Regression

เหตุผล: มีตัวแปรอิสระหลายตัว ต้องการดูอิทธิพลร่วมและพยากรณ์

8) ความแตกต่างของสัดส่วน/อัตรา (ตัวแปรกลุ่ม + ผลลัพธ์เป็นผ่าน/ไม่ผ่าน)

โจทย์วิจัย: “สัดส่วนการสอบผ่านของนักศึกษาหลังใช้สื่อใหม่แตกต่างจากวิธีเดิมหรือไม่”

ตัวแปร: วิธีสอน (เดิม/ใหม่) ⟶ ผ่าน/ไม่ผ่าน

สถิติที่เหมาะสม: Chi-square test of independence

เหตุผล: ตัวแปรเป็น “กลุ่ม/จัดประเภท” และเปรียบเทียบ “สัดส่วน”

(อันนี้เป็นตัวอย่างเสริม นอก 4 ตัวหลัก แต่ใช้จริงบ่อยในงานวิจัย)

ตารางที่ 6.15 สรุปเลือกสถิติแบบเร็ว (Decision Guide)

เป้าหมาย	ตัวแปรตาม (Y)	กลุ่ม/ตัวแปรอิสระ (X)	สถิติที่เหมาะสม
เปรียบเทียบ 2 ครั้ง (คนเดิม)	เชิงปริมาณ	ก่อน–หลัง	Paired t-test
เปรียบเทียบ 2 กลุ่ม (คนละกลุ่ม)	เชิงปริมาณ	2 กลุ่มอิสระ	Independent t-test
เปรียบเทียบ ≥ 3 กลุ่ม	เชิงปริมาณ	≥ 3 กลุ่ม	One-way ANOVA
ดูผล 2 ปัจจัย	เชิงปริมาณ	2 ปัจจัย	Two-way ANOVA
ดูความสัมพันธ์	เชิงปริมาณ	เชิงปริมาณ	Correlation
พยากรณ์	เชิงปริมาณ	1 ตัว/หลายตัว	Regression

การใช้งานใน Excel และ SPSS (ภาพรวม)

- **Excel**
    - t-test / Correlation ผ่าน Data Analysis ToolPak
    - Regression เบื้องต้น
- **SPSS**
    - Analyze ⟶ Compare Means ⟶ t-test / ANOVA
    - Analyze ⟶ Correlate ⟶ Bivariate
    - Analyze ⟶ Regression ⟶ Linear

สรุป

การวิเคราะห์เชิงอนุมานเป็นหัวใจสำคัญของการวิจัยเชิงปริมาณและการตัดสินใจเชิงข้อมูล โดยช่วยให้สามารถสรุปผลจากกลุ่มตัวอย่างไปยังประชากรได้อย่างมีหลักการ เทคนิคสำคัญ ได้แก่ t-test, ANOVA, Correlation และ Regression ซึ่งแต่ละวิธีมีวัตถุประสงค์และเงื่อนไขการใช้งานที่แตกต่างกัน การเลือกใช้วิธีที่เหมาะสมจะช่วยให้ผลการวิเคราะห์มีความถูกต้องและน่าเชื่อถือ

6.5 การแสดงผลข้อมูล (Data Visualization)

การนำเสนอข้อมูลด้วยกราฟช่วยให้เข้าใจข้อมูลได้ง่ายขึ้น

ประเภทกราฟ

- Bar Chart : เปรียบเทียบข้อมูล
- Line Chart : แสดงแนวโน้ม
- Histogram : การกระจายข้อมูล
- Pie Chart : สัดส่วน
- Boxplot : ตรวจสอบ Outliers

6.6 กระบวนการวิเคราะห์ข้อมูลเชิงปริมาณ

1. เก็บรวบรวมข้อมูล
2. ทำความสะอาดข้อมูล
3. วิเคราะห์ข้อมูล
4. แสดงผลข้อมูล
5. ตีความและสรุปผล

แบบฝึกหัดท้ายบทที่ 6

แบบฝึกหัดที่ 1 : ความเข้าใจพื้นฐาน

จงตอบคำถามต่อไปนี้

1. ข้อมูลอายุของนักศึกษาเป็นข้อมูลประเภทใด
2. ความแตกต่างระหว่างข้อมูลไม่ต่อเนื่องและข้อมูลต่อเนื่องคืออะไร
3. เพราะเหตุใดการทำ Data Cleaning จึงมีความสำคัญ

แบบฝึกหัดที่ 2 : การใช้ Excel

กำหนดข้อมูลคะแนนสอบของนักศึกษา 10 คน

1. คำนวณค่าเฉลี่ย มัธยฐาน และส่วนเบี่ยงเบนมาตรฐาน
2. สร้างกราฟ Histogram แสดงการกระจายของคะแนน

แบบฝึกหัดที่ 3 : การใช้ SPSS

กำหนดข้อมูลคะแนนก่อนเรียนและหลังเรียน

1. วิเคราะห์ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน
2. ทดสอบว่าคะแนนหลังเรียนสูงกว่าก่อนเรียนหรือไม่ (t-test)

แบบฝึกหัดที่ 4 : การเลือก Visualization

ให้นักศึกษาเลือกกราฟที่เหมาะสมสำหรับข้อมูลต่อไปนี้ พร้อมอธิบายเหตุผล

1. ยอดขายรายเดือน
2. สัดส่วนค่าใช้จ่าย
3. การกระจายของรายได้พนักงาน