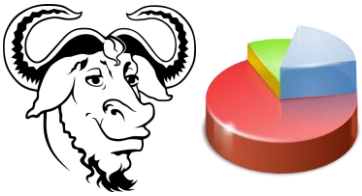
เอกสารคำสอน

วิชา 26234160 การใช้โปรแกรมคอมพิวเตอร์ทางสถิติในด้านสังคมศาสตร์

The Application of Statistical Package in Social Sciences

เรวัต แสงสุริยงค์



GNU **PSPP**

Open Source Statistical Software: OSSS

ภาควิชาสังคมวิทยา คณะมนุษยศาสตร์และสังคมศาสตร์

มหาวิทยาลัยบูรพา

ปีการศึกษา 2563

คำนำ

เอกสารคำสอนวิชา การใช้โปรแกรมคอมพิวเตอร์ทางสถิติในด้านสังคมศาสตร์ (The Application of Statistical Package in Social Sciences) พัฒนาและใช้สอนนิสิตภาควิชาสังคมวิทยาทั้งสาขาวิชาการพัฒนาชุมชนและสาขาวิชาการจัดการบริการสังคมในรหัสวิชาที่แตกต่างกันไปตามการปรับปรุงหลักสูตรจนถึงปัจจุบัน คือ รหัสวิชา 26234160 ตามหลักสูตรศิลปศาสตรบัณฑิต สาขาวิชาการจัดการบริการสังคม หลักสูตรปรับปรุง พ.ศ. 2560 ที่ให้นิสิตเรียนในภาคเรียนที่ 1 ปีการศึกษา 2563 เป็นต้นไป

จุดมุ่งหมายของวิชานี้ คือ การปูพื้นฐานความรู้เกี่ยวกับการใช้โปรแกรมคอมพิวเตอร์ทางสถิติในการวิเคราะห์ข้อมูลเชิงปริมาณด้านสังคมศาสตร์ เนื้อหาจึงประกอบด้วยความรู้เกี่ยวกับการวิจัยด้านสังคมศาสตร์ การวิจัยเชิงปริมาณ สถิติสำหรับการวิจัย การเตรียมข้อมูลเข้าโปรแกรม การติดตั้งและใช้โปรแกรมเบื้องต้น การจัดการข้อมูล การวิเคราะห์เครื่องมือการวิจัย การตรวจสอบข้อมูล การวิเคราะห์ข้อมูลด้วยสถิติเชิงพรรณนาและสถิติเชิงอ้างอิง (การทดสอบแบบอิงค่าพารามิเตอร์และไม่อิงค่าพารามิเตอร์) โดยเน้นการวิเคราะห์ตัวแปรไม่เกิน 2 ตัวแปร และในแต่ละบทยังมีแบบฝึกหัดท้ายบทให้ทบทวนและศึกษาเพิ่มเติมด้วย

เนื้อหาในเอกสารคำสอนใช้โปรแกรม PSPP ที่เป็นโปรแกรมทางสถิติแบบรหัสเปิด (Open Source Statistical Software: OSSS) ภายใต้สัญญาอนุญาตให้ใช้ซอฟต์แวร์แบบสาธารณะทั่วไป (GNU General Public License: GNU-GPL) ที่สามารถหาดาวน์โหลดได้จากอินเทอร์เน็ตและติดตั้งใช้งานได้ด้วยตนเอง เป็นเครื่องมือในการวิเคราะห์และทดสอบข้อมูล

ผู้เขียนหวังว่าผู้เรียนจะได้ความรู้เบื้องต้นเกี่ยวกับการวิเคราะห์ข้อมูลเชิงปริมาณ นำเอาความรู้ที่ได้จากการเรียนไปใช้เป็นเครื่องมือวิเคราะห์ข้อมูลเพื่อจัดทำรายงานระหว่างเรียนอยู่ในมหาวิทยาลัยและนำไปใช้ประโยชน์หลังเรียนจบไปแล้ว

ผู้ช่วยศาสตราจารย์ ดร. เรวัต แสงสุริยงค์

กุมภาพันธ์ 2564

ภาควิชาสังคมวิทยา คณะมนุษยศาสตรและสังคมศาสตร์
มหาวิทยาลัยบูรพา

แผนการสอน

1. รหัสและชื่อรายวิชา

26234160 การใช้โปรแกรมคอมพิวเตอร์ทางสถิติในด้านสังคมศาสตร์

2. จำนวนหน่วยกิต 3 หน่วยกิต (2-2-5)

3. หลักสูตรและประเภทของรายวิชา ศิลปศาสตรบัณฑิต สาขาวิชาการจัดการบริการสังคม เป็นวิชาเลือก

4. คำอธิบายรายวิชา

หลักการและฝึกปฏิบัติการใช้คอมพิวเตอร์ในการประมวลผลสถิติทางสังคมศาสตร์ โดยนำตัวแปรและข้อมูลเกี่ยวกับประชากร ชุมชนและสังคมมาจัดเตรียมข้อมูล บันทึกข้อมูลและวิเคราะห์ตลอดจนการแปลความหมายและนำเสนอรายงาน

5. วัตถุประสงค์ของรายวิชา

1. ผู้เรียนเข้าใจเกี่ยวกับการหาความรู้ด้านสังคมศาสตร์ รวมถึงสาขาวิชาสังคมวิทยา
2. ผู้เรียนเข้าใจวิธีการวิจัยเชิงปริมาณที่เกิดจากความสัมพันธ์ระหว่างทฤษฎี สมมติฐาน และตัวแปร
3. ผู้เรียนเข้าใจและเลือกใช้สถิติแต่ละประเภทในการวิเคราะห์และทดสอบข้อมูลได้ถูกต้องและเหมาะสม
4. ผู้เรียนสามารถแปลงข้อมูลจากแบบสอบถามไปเป็นแฟ้มข้อมูลได้
5. ผู้เรียนรู้จักโปรแกรมคอมพิวเตอร์ด้านสถิติสำหรับวิเคราะห์ข้อมูล สามารถติดตั้งและเริ่มใช้งานเบื้องต้นได้

6. ผู้เรียนสามารถนำข้อมูลจากแบบสอบถามกรอกเข้าโปรแกรม และนำข้อมูลที่เก็บอยู่ในรูปของแฟ้มข้อมูลเข้าโปรแกรม

7. ผู้เรียนสามารถเลือกใช้ข้อมูลและเปลี่ยนรูปแบบข้อมูลให้ตรงกับวัตถุประสงค์ของการวิเคราะห์ข้อมูล

8. ผู้เรียนสามารถวิเคราะห์คุณภาพของเครื่องมือการวิจัยได้

9. ผู้เรียนสามารถตรวจสอบความถูกต้องและความสมบูรณ์ของข้อมูล รวมถึงตรวจสอบลักษณะของข้อมูลตามข้อตกลงเบื้องต้นของสถิติ

10. ผู้เรียนสามารถใช้โปรแกรมประมวลผลการแจกแจงข้อมูลของตัวแปร 1 ตัวแปร ด้วยสถิติเชิงพรรณนา และเขียนบรรยายผลการวิเคราะห์ได้

11. ผู้เรียนสามารถใช้โปรแกรมประมวลผลการเปรียบเทียบข้อมูลของตัวแปร 2 ตัวแปร ด้วยสถิติเชิงพรรณนา และเขียนบรรยายผลการวิเคราะห์ได้

12. ผู้เรียนสามารถใช้โปรแกรมประมวลผลทดสอบข้อมูลของตัวแปรด้วยสถิติเชิงอ้างอิงแบบอิงค่าพารามิเตอร์ และเขียนบรรยายผลการทดสอบได้

13. ผู้เรียนสามารถใช้โปรแกรมประมวลผลทดสอบข้อมูลของตัวแปร 1 ตัวแปร ด้วยสถิติเชิงอ้างอิงแบบไม่อิงค่าพารามิเตอร์ และเขียนบรรยายผลการทดสอบได้

14. ผู้เรียนสามารถใช้โปรแกรมประมวลผลทดสอบข้อมูลของตัวแปร 2 ตัวแปร ด้วยสถิติเชิงอ้างอิงแบบไม่อิงค่าพารามิเตอร์และเขียนบรรยายผลการทดสอบได้

6. อาจารย์ผู้รับผิดชอบรายวิชา ผู้ช่วยศาสตราจารย์ ดร. เรวัต แสงสุริยงค์

7. แผนการสอน

สัปดาห์ที่	หัวข้อ/รายละเอียด	จำนวน (ชั่วโมง)	วิธีการสอน/สื่อ	ผลการเรียนรู้*
1	**แนะนำการเรียนการสอน** **การวิจัยทางสังคมศาสตร์** 1. การหาความรู้ด้านสังคมศาสตร์ - ศาสตร์ คือ อะไร - ความรู้เบื้องต้นเกี่ยวกับสังคมศาสตร์ - วิธีการหาความรู้ 2. ประเภทการวิจัยด้านสังคมศาสตร์ - การจัดประเภทงานวิจัย - หลักเกณฑ์การจัดประเภทงานวิจัย 3. กระบวนการวิจัยด้านสังคมศาสตร์ - การออกแบบการวิจัย - การวิจัยเชิงปริมาณและเชิงคุณภาพ - การวิจัยเชิงผสมผสาน	4	1. วิธีการสอน:- บรรยาย อธิบาย แสดงตัวอย่าง และทำแบบฝึกหัด 2. สื่อ:- เอกสารคำสอน ภาพเลื่อน คอมพิวเตอร์โปรแกรมสอนออนไลน์ (สอนแบบออนไลน์) และเครื่องฉาย (สอนในชั้นเรียน)	2, 3, 6

สัปดาห์ที่	หัวข้อ/รายละเอียด	จำนวน (ชั่วโมง)	วิธีการสอน/สื่อ	ผลการเรียนรู้*
	การวิจัยเชิงปริมาณ 1. ความรู้เบื้องต้นเกี่ยวกับการวิจัยเชิงปริมาณ 2. ทฤษฎี - ทฤษฎี คือ อะไร - ประเภทของทฤษฎี - การเลือกใช้ทฤษีในการวิจัย 3. สมมุติฐาน - สมมติฐาน คือ อะไร - ประเภทของสมมติฐาน 4. ตัวแปร - ตัวแปร คือ อะไร - ประเภทตัวแปร - ระดับการวัดของตัวแปร 5. ความสัมพันธ์ระหว่างสมมุติฐานกับตัวแปร			

สัปดาห์ที่	หัวข้อ/รายละเอียด	จำนวน (ชั่วโมง)	วิธีการสอน/สื่อ	ผลการเรียนรู้*
	- ความสัมพันธ์เชิงสาเหตุ - ความสัมพันธ์เชิงสาเหตุในทิศทางตรงกันข้าม - ความสัมพันธ์เชิงหลายสาเหตุ			
2	**สถิติสำหรับการวิจัย** 1. สถิติ - ความหมายของสถิติ - ประเภทของสถิติ 2. ความสัมพันธ์ระหว่างข้อมูลกับสถิติ - ข้อมูลกับสถิติเชิงพรรณนา - ข้อมูลกับสถิติเชิงอ้างอิง 3. การทดสอบสมมติฐาน - ประเภทการทดสอบสมมติฐาน - กฎเชิงประจักษ์ - การยอมรับและปฏิเสธสมมติฐาน	8	1. วิธีการสอน:- บรรยาย อธิบาย แสดงตัวอย่าง และทำแบบฝึกหัด 2. สื่อ:- เอกสารคำสอน ภาพเลื่อน คอมพิวเตอร์ โปรแกรมสอนออนไลน์ (สอนแบบออนไลน์) และเครื่องฉาย (สอนในชั้นเรียน)	2, 3, 5, 6

สัปดาห์ที่	หัวข้อ/รายละเอียด	จำนวน (ชั่วโมง)	วิธีการสอน/สื่อ	ผลการเรียนรู้*
	4. ความคลาดเคลื่อนของผลการวิจัย - ประเภทของความคลาดเคลื่อนของผลการวิจัย - ปัจจัยที่มีผลต่อความคลาดเคลื่อนของผลการวิจัย **แบบสอบถามและข้อมูล** 1. ความสัมพันธ์ระหว่างแบบสอบถามกับแฟ้มข้อมูล - คำถามแบบปลายปิด - คำถามแบบปลายเปิด - คำถามแบบเลือกตอบ - คำถามแบบประเมินค่า - คำถามแบบหลายคำตอบ - คำถามแบบจัดอันดับ - คำถามแบบเปรียบเทียบ - คำถามแบบถูกผิด 2. การสร้างตัวแปร			

สัปดาห์ที่	หัวข้อ/รายละเอียด	จำนวน (ชั่วโมง)	วิธีการสอน/สื่อ	ผลการเรียนรู้*
	- ชื่อตัวแปร - ประเภทตัวแปร - จำนวนค่าตัวแปร - คำอธิบายตัวแปร - คำอธิบายค่าตัวแปร - ค่าขาดหลาย - ความกว้างของช่องค่าตัวแปรหรือสดมภ์ - การจัดตำแหน่งค่าตัวแปร - ระดับการวัด - บทบาทของตัวแปร			
3-4	**ความรู้เบื้องต้นเกี่ยวกับโปรแกรมทางสถิติ** 1. โปรแกรมสำเร็จรูปทางสถิติ - โปรแกรมเชิงพาณิชย์ - โปรแกรมแบบสาธารณะทั่วไป	8	1. วิธีการสอน:- บรรยาย อธิบาย แสดงตัวอย่าง แนะนำการติดตั้งโปรแกรม และทำแบบฝึกหัดท้ายบท	3, 5

สัปดาห์ที่	หัวข้อ/รายละเอียด	จำนวน (ชั่วโมง)	วิธีการสอน/สื่อ	ผลการเรียนรู้*
	2. การติดตั้งโปรแกรม PSPP - การค้นหาและดาวน์โหลด - วิธีการติดตั้ง 3. การใช้โปรแกรม PSPP เบื้องต้น - การเปิดโปรแกรม - การเปิดแฟ้มข้อมูล - การปิดโปรแกรม 4. ส่วนประกอบของโปรแกรม PSPP - ระบบจัดการข้อมูล - ระบบแสดงผลลัพธ์ - ระบบคำสั่งแบบข้อความ 5. เมนูการใช้โปรแกรม PSPP - เมนูแบบรายการคำสั่ง - เมนูแบบปุ่มภาพ		2. สื่อ:- เอกสารคำสอน ภาพเลื่อน คอมพิวเตอร์ โปรแกรม PSPP โปรแกรมสอนออนไลน์ (สอนแบบออนไลน์) และเครื่องฉาย (สอนในชั้นเรียน)	

สัปดาห์ที่	หัวข้อ/รายละเอียด	จำนวน (ชั่วโมง)	วิธีการสอน/สื่อ	ผลการเรียนรู้*
5	**การนำข้อมูลเข้าโปรแกรม** 1. การนำข้อมูลเข้าจากแบบสอบถาม - การสร้างแฟ้มข้อมูล - การนำข้อมูลเข้าโปรแกรมแบบกรอกข้อมูล 2. การนำข้อมูลเข้าจากแฟ้มข้อมูล - แฟ้มข้อมูลแบบตัวอักษร - แฟ้มข้อมูลจากโปรแกรมตารางคำนวณ **การจัดการข้อมูล** 1. การปรับข้อมูล - การจัดเรียงข้อมูล - การไขว้ตัวแปร - การรวมข้อมูล - การแบ่งข้อมูล - การเลือกข้อมูล	4	1. วิธีการสอน:- บรรยาย อธิบาย แสดงตัวอย่าง และทำแบบฝึกหัดท้ายบท 2. สื่อ:- เอกสารคำสอน ภาพเลื่อน คอมพิวเตอร์ โปรแกรม PSPP โปรแกรมสอนออนไลน์ (สอนแบบออนไลน์) และเครื่องฉาย (สอนในชั้นเรียน)	2, 3, 5, 6

สัปดาห์ที่	หัวข้อ/รายละเอียด	จำนวน (ชั่วโมง)	วิธีการสอน/สื่อ	ผลการเรียนรู้*
	- การกำหนดน้ำหนักข้อมูล 2. การเปลี่ยนรูปแบบข้อมูล - การเปลี่ยนค่าของตัวแปร - การสร้างตัวแปรใหม่			
6	**การวิเคราะห์เครื่องมือการวิจัย** 1. การวิเคราะห์ความถูกต้อง - การวิเคราะห์ความเที่ยงตรงด้วยดัชนีความสอดคล้อง ระหว่างข้อคำถาม-วัตถุประสงค์ของการวิจัยค่า - การวิเคราะห์ความเที่ยงตรงด้วยค่าความเห็นสอดคล้องกัน 2. การวิเคราะห์ความน่าเชื่อถือ - การวัดความสอดคล้องภายใน - การวัดความสอดคล้องภายนอก **การวิเคราะห์ข้อมูลเชิงสำรวจ** 1. การตรวจสอบความถูกต้องและความสมบูรณ์ของข้อมูล	4	1. วิธีการสอน:- บรรยาย อธิบาย แสดงตัวอย่าง และทำแบบฝึกหัดท้ายบท 2. สื่อ:- เอกสารคำสอน ภาพเลื่อน คอมพิวเตอร์ โปรแกรม PSPP โปรแกรมสอนออนไลน์ (สอนแบบออนไลน์) และเครื่องฉาย (สอนในชั้นเรียน)	2, 3, 5, 6

สัปดาห์ที่	หัวข้อ/รายละเอียด	จำนวน (ชั่วโมง)	วิธีการสอน/สื่อ	ผลการเรียนรู้*
	- การตรวจสอบการออกแบบข้อมูล - การตรวจสอบความถูกต้องของข้อมูล - การตรวจสอบค่าสูญหาย - การตรวจสอบค่าผิดปกติและค่าต่างสุด 2. การตรวจสอบข้อมูลตามข้อตกลงเบื้องต้น - การแจกแจงปกติ - ความเท่ากันของความแปรปรวน - ความเป็นเส้นตรง			
7-8	**การวิเคราะห์ข้อมูลด้วยสถิติเชิงพรรณนา** **การวิเคราะห์ข้อมูลเชิงพรรณนา 1 ตัวแปร** 1. การวิเคราะห์ข้อมูลแบบค่าจัดประเภท - การวิเคราะห์ข้อมูลด้วยค่าร้อยละ - การวิเคราะห์ข้อมูลด้วยค่าฐานนิยม 2. การวิเคราะห์ตัวแปรแบบค่าต่อเนื่อง	8	1. วิธีการสอน: บรรยาย อธิบาย แสดงตัวอย่าง และทำแบบฝึกหัดท้ายบท 2. สื่อ:- เอกสารคำสอน ภาพเลื่อน คอมพิวเตอร์ โปรแกรม PSPP โปรแกรมสอนออนไลน์ (สอนแบบออนไลน์) และเครื่องฉาย (สอนในชั้นเรียน)	2, 3, 5, 6

สัปดาห์ที่	หัวข้อ/รายละเอียด	จำนวน (ชั่วโมง)	วิธีการสอน/สื่อ	ผลการเรียนรู้*
	- การวิเคราะห์ข้อมูลด้วยค่าเฉลี่ย (รายตัวแปร) - การวิเคราะห์ข้อมูลค่าเฉลี่ย (ตัวแปรดัชนี) **การวิเคราะห์ข้อมูลเชิงพรรณนา 2 ตัวแปร** 1. การวิเคราะห์ข้อมูลระหว่างตัวแปรแบบค่าแบ่งประเภท - การวิเคราะห์การเปรียบเทียบข้อมูลระหว่างตัวแปรแบบค่าจัดประเภท 2 ตัวแปร - การวิเคราะห์เปรียบเทียบข้อมูลระหว่างตัวแปรแบบค่าแบ่งประเภท 2 ตัวแปร โดยมีตัวแปรคุม 2. การวิเคราะห์ข้อมูลระหว่างตัวแปรแบบค่าแบ่งประเภทกับตัวแปรแบบค่าต่อเนื่อง - การวิเคราะห์เปรียบเทียบค่าเฉลี่ย 2 กลุ่ม - การวิเคราะห์เปรียบเทียบค่าเฉลี่ยมากกว่า 2 กลุ่ม - การวิเคราะห์เปรียบเทียบค่าเฉลี่ยโดยมีตัวแปรคุม			
9	สอบกลางภาค	4	สอบ	1, 2, 3, 5, 6

สัปดาห์ที่	หัวข้อ/รายละเอียด	จำนวน (ชั่วโมง)	วิธีการสอน/สื่อ	ผลการเรียนรู้*
10-12	**การวิเคราะห์ข้อมูลด้วยสถิติเชิงอ้างอิง** **การทดสอบข้อมูลแบบอิงค่าพารามิเตอร์** 1. การทดสอบค่าเฉลี่ย 1 กลุ่มตัวอย่าง 2. การเปรียบค่าเฉลี่ยของกลุ่มตัวอย่างที่ไม่เป็นอิสระจากกัน 2 กลุ่ม 3. การเปรียบเทียบค่าเฉลี่ยของกลุ่มตัวอย่างที่เป็นอิสระจากกัน 2 กลุ่ม 4. การเปรียบเทียบค่าเฉลี่ยของกลุ่มตัวอย่างที่เป็นอิสระจากกันมากกว่า 2 กลุ่ม 5. การทดสอบความสัมพันธ์รายคู่แบบเพียร์สัน	12	1. วิธีการสอน:- บรรยาย อธิบาย แสดงตัวอย่าง และทำแบบฝึกหัดท้ายบท 2. สื่อ:- เอกสารคำสอน ภาพเลื่อน คอมพิวเตอร์ โปรแกรม PSPP โปรแกรมสอนออนไลน์ (สอนแบบออนไลน์) และเครื่องฉาย (สอนในชั้นเรียน)	2, 3, 5, 6
13	**การทดสอบข้อมูลแบบไม่อิงค่าพารามิเตอร์ (1 ตัวแปร)** 1. การทดสอบสัดส่วนของกลุ่มตัวอย่าง 2. การทดสอบการแจกแจงของกลุ่มตัวอย่าง	4	1. วิธีการสอน:- บรรยาย อธิบาย แสดงตัวอย่าง และทำแบบฝึกหัดท้ายบท	2, 3, 5, 6

สัปดาห์ที่	หัวข้อ/รายละเอียด	จำนวน (ชั่วโมง)	วิธีการสอน/สื่อ	ผลการเรียนรู้*
			2. สื่อ:- เอกสารคำสอน ภาพเลื่อน คอมพิวเตอร์ โปรแกรม PSPP โปรแกรมสอนออนไลน์ (สอนแบบออนไลน์) และเครื่องฉาย (สอนในชั้นเรียน)	
14-16	**การทดสอบข้อมูลแบบไม่อิงค่าพารามิเตอร์ (2 ตัวแปร)** 1. การทดสอบความเป็นอิสระของกลุ่มตัวอย่าง 2. การทดสอบค่ากลางของกลุ่มตัวอย่างที่มีความสัมพันธ์กัน 2 กลุ่ม 3. การทดสอบค่ากลางของกลุ่มตัวอย่างที่มีความสัมพันธ์กันมากกว่า 2 กลุ่ม 4. การทดสอบค่ากลางของกลุ่มตัวอย่างที่เป็นอิสระจากกัน 2 กลุ่ม 3. การทดสอบกลุ่มตัวอย่างที่เป็นอิสระจากกันมากกว่า 2 กลุ่ม 6. การทดสอบความสัมพันธ์รายคู่แบบสเปียร์แมน	12	1. วิธีการสอน:- บรรยาย อธิบาย แสดงตัวอย่าง และทำแบบฝึกหัดท้ายบท 2. สื่อ:- เอกสารคำสอน ภาพเลื่อน คอมพิวเตอร์ โปรแกรม PSPP โปรแกรมสอนออนไลน์ (สอนแบบออนไลน์) และเครื่องฉาย (สอนในชั้นเรียน)	2, 3, 5, 6

สัปดาห์ที่	หัวข้อ/รายละเอียด	จำนวน (ชั่วโมง)	วิธีการสอน/สื่อ	ผลการเรียนรู้*
17	**สอบปลายภาค**	3	สอบ	1, 2, 3, 5, 6

ผลการเรียนรู้ของรายวิชา*

1. ด้านคุณธรรมจริยธรรม:- มีจรรยาบรรณทางวิชาการหรือวิชาชีพ (1.5)

2. ด้านความรู้:- สามารถประยุกต์ความรู้จากทฤษฎีสู่การปฏิบัติ (2.2)

3. ด้านทักษะทางปัญญา:- สามารถคิดวิเคราะห์อย่างเป็นระบบ (3.3)

4. ด้านทักษะความสัมพันธ์ระหว่างบุคคลและความรับผิดชอบ:- –

5. ด้านทักษะการวิเคราะห์เชิงตัวเลข การสื่อสาร และการใช้เทคโนโลยีสารสนเทศ:- สามารถเลือกหรือประยุกต์เทคนิคทางคณิตศาสตร์ หรือสถิติที่เกี่ยวข้องในการศึกษาค้นคว้า เสนอแนะแนวทางในการแก้ปัญหาอย่างเหมาะสม (5.3)

6. ด้านทักษะพิสัย:- มีทักษะในการศึกษาและวิเคราะห์สถานการณ์ชุมชน/ องค์กรด้วยเครื่องมือทางการจัดการบริการสังคม (6.2)

8. การวัดผลและการประเมินผล

8.1 การประเมินผลการเรียนรู้

ผลการเรียนรู้*	วิธีการประเมิน	สัปดาห์ที่ประเมิน	สัดส่วนของการประเมิน
1	เวลาเรียน	1-16	10
1, 2, 3, 5, 6	การทำแบบฝึกหัดท้ายบท	1-16	20
1, 3, 5	สอบระหว่างภาค	9	30
1, 3, 5, 6	สอบปลายภาค	17	40

***หมายเหตุ** 1. คุณธรรมจริยธรรม 2. ความรู้ 3. ทักษะทางปัญญา 4. ทักษะความสัมพันธ์ระหว่างบุคคลและความรับผิดชอบ 5. ทักษะการวิเคราะห์เชิงตัวเลข การสื่อสาร และการใช้เทคโนโลยีสารสนเทศ 6. ทักษะพิสัย

8.2 ระดับคะแนน พิจารณาตามความเหมาะสม ดังนี้

8.2.1 อิงเกณฑ์ ตามระดับคะแนนดังนี้

ระดับคะแนน	ผลการเรียน
80 – 100	A
75 – 79	B+
70 – 74	B
65 – 69	C+
60 – 64	C
55 – 59	D+
50 – 54	D
0 – 49	F

8.2.2 หรืออิงกลุ่ม แบบ T-score หรือ MEAN-SD

9. ทรัพยากรประกอบการเรียนการสอน

9.1 ตำราและเอกสารหลัก

เรวัต แสงสุริยงค์. *เอกสารคำสอน รายวิชา 26234160 การใช้โปรแกรมคอมพิวเตอร์ทางสถิติในด้านสังคมศาสตร์*. ภาควิชาสังคมวิทยา คณะมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยบูรพา.

9.2 ตำราและเอกสารแนะนำ

วรภูริ มูลสิน (2555). *หลักการวิจัยและวิเคราะห์ข้อมูลทางสถิติด้วย PSPP*. เข้าถึงได้จาก http://www.edu.ru.ac.th/coved/pdf/PSPP/คู่มือ_pspp-book-2555.pdf

GNU. (2019). *PSPP Users' Guide*. Retrieved from https://www.gnu.org/software/pspp/manual/pspp.pdf

GNU. (2020). *PSPP*. Retrieved from https://www.gnu.org/software/pspp/manual/pspp.html#Introduction

Halter, C. P. (2017). The PSPP Guide: An Introduction to Statistical Analysis. 2nd ed. CreativeMinds Press Group.

สารบัญ

คำนำ ..... **ก**

แผนการสอน ..... **ค**

สารบัญ ..... **ธ**

สารบัญตาราง ..... **ม**

สารบัญภาพ ..... **ย**

ส่วนที่ 1 ความรู้เบื้องต้นเกี่ยวกับการวิจัยเชิงปริมาณด้านสังคมศาสตร์

บทที่ 1 การวิจัยด้านสังคมศาสตร์ **3**

เนื้อหา 4

1. การหาความรู้ด้านสังคมศาสตร์ 4

1.1 ศาสตร์ คือ อะไร 4

1.2 ความรู้เบื้องต้นเกี่ยวกับสังคมศาสตร์ 5

1.3 วิธีการหาความรู้ 7

2. ประเภทการวิจัยด้านสังคมศาสตร์ 9

2.1 การจัดประเภทงานวิจัย 9

2.2 หลักเกณฑ์การจัดประเภทงานวิจัย 13

3. กระบวนการวิจัยด้านสังคมศาสตร์ 14

3.1 การออกแบบการวิจัย 14

3.2 การวิจัยเชิงปริมาณและการวิจัยเชิงคุณภาพ 16

3.3 การวิจัยเชิงผสมผสาน 18

สรุป 19

บทที่ 2 การวิจัยเชิงปริมาณ **21**

เนื้อหา.....22

1. ความรู้เบื้องต้นเกี่ยวกับการวิจัยเชิงปริมาณ.....22

1.1 ความเป็นมาของการวิจัยเชิงปริมาณด้านสังคมศาสตร์.....22

1.2 ประเภทของการวิจัยเชิงปริมาณ.....23

2. ทฤษฎี.....24

2.1 ทฤษฎี คือ อะไร.....24

2.2 ประเภทของทฤษฎี.....26

2.3 การเลือกใช้ทฤษฎีในการวิจัย.....26

3. สมมติฐาน.....28

3.1 สมมติฐาน คือ อะไร.....28

3.2 ประเภทของสมมติฐาน.....29

4. ตัวแปร.....30

4.1 ตัวแปร คือ อะไร.....30

4.2 ประเภทตัวแปร.....31

4.3 ระดับการวัดของตัวแปร.....36

5. ความสัมพันธ์ระหว่างสมมติฐานและตัวแปร.....40

สรุป.....42

แบบฝึกหัดท้ายบท.....42

บทที่ 3 สถิติสำหรับการวิจัย.....43

เนื้อหา.....44

1. สถิติ.....44

1.1 ความหมายของสถิติ.....44

1.2 ประเภทของสถิติ.....45

2. ความสัมพันธ์ระหว่างข้อมูลกับสถิติ.....48

2.1 ข้อมูลกับสถิติเชิงพรรณนา 48
2.2 ข้อมูลกับสถิติเชิงอ้างอิง 49
3. การทดสอบสมมติฐาน.....51
3.1 ประเภทการทดสอบสมมติฐาน.....53
3.2 กฎเชิงประจักษ์.....54
3.3 การยอมรับและปฏิเสธสมมติฐาน.....55
4. ความคลาดเคลื่อนของผลการวิจัย.....58
4.1 ประเภทของความคลาดเคลื่อนของผลการวิจัย 58
4.2 ปัจจัยที่มีผลต่อความคลาดเคลื่อนของผลการวิจัย 58
สรุป.....60
แบบฝึกหัดท้ายบท 60

ส่วนที่ 2 ความรู้เบื้องต้นเกี่ยวกับการใช้โปรแกรมทางสถิติ

บทที่ 4 แบบสอบถามและข้อมูล.....63
เนื้อหา.....64
1. ความสัมพันธ์ระหว่างแบบสอบถามกับแฟ้มข้อมูล.....64
1.1 คำถามแบบเลือกตอบ 67
1.2 คำถามแบบปลายเปิด 67
1.3 คำถามแบบประเมินค่า.....68
1.4 คำถามแบบหลายคำตอบ 69
1.5 คำถามแบบจัดอันดับ 69
1.6 คำถามแบบเปรียบเทียบ 70
1.7 คำถามแบบถูกผิด 71
2. การสร้างตัวแปร.....71
สรุป.....79
บทที่ 5 ความรู้เบื้องต้นเกี่ยวกับโปรแกรมทางสถิติ 81

1. โปรแกรมสำเร็จรูปทางสถิติ....82

2. การติดตั้งโปรแกรมทางสถิติ....83

3. การใช้โปรแกรมทางสถิติเบื้องต้น....88

4. ส่วนประกอบของโปรแกรม PSPP....91

4.1 ระบบจัดการข้อมูล....91

4.2 ระบบแสดงผลลัพธ์....92

4.3 ระบบคำสั่งแบบข้อความ....93

4.4 การใช้ผลลัพธ์....93

5. เมนูการใช้โปรแกรม PSPP....95

5.1 หน้าต่างระบบ DATA EDITOR....95

5.2 หน้าต่างระบบ SYNTAX EDITOR....100

5.3 หน้าต่างระบบ OUTPUT VIEWER....102

สรุป....103

แบบฝึกหัดท้ายบท....103

บทที่ 6 การนำข้อมูลเข้าโปรแกรม....105

เนื้อหา....106

1. การนำข้อมูลเข้าจากแบบสอบถาม....106

1.1 การสร้างแฟ้มข้อมูล....106

1.2 การนำข้อมูลเข้าโปรแกรมแบบกรอกข้อมูล....112

2. การนำเข้าข้อมูลจากแฟ้มข้อมูล....113

2.1 การนำข้อมูลเข้าจากโปรแกรมตารางคำนวณ....114

2.2 การนำข้อมูลเข้าจากแบบสอบถามออนไลน์....119

สรุป....123

แบบฝึกหัดท้ายบท....123

บทที่ 7 การจัดการข้อมูล.....125

เนื้อหา.....126

1. การปรับข้อมูล 127

1.1 การจัดเรียงตัวอย่าง 127

1.2 การไขว้ตัวแปรและตัวอย่าง 128

1.3 การรวมกลุ่มตัวอย่าง..... 129

1.4 การแบ่งกลุ่มตัวอย่าง..... 131

1.5 การเลือกตัวอย่าง..... 132

1.6 การกำหนดน้ำหนักตัวอย่าง..... 133

2. การแปลงข้อมูล 135

2.1 การสร้างตัวแปรแบบค่าต่อเนื่อง 135

2.2 การนับจำนวนข้อมูล 136

2.3 การจัดลำดับตัวอย่าง 137

2.4 การสร้างตัวแปรแบบอัตโนมัติ..... 139

2.5 การเปลี่ยนค่าตัวแปรในตัวแปรเดิม 140

2.6 การสร้างตัวแปรแบบค่าจัดกลุ่ม 141

2.7 การสั่งประมวลผลต่อ 143

สรุป.....143

แบบฝึกหัดท้ายบท 144

ส่วนที่ 3 การวิเคราะห์เครื่องมือการวิจัย และตรวจสอบข้อมูล

บทที่ 8 การวิเคราะห์เครื่องมือการวิจัย.....147

เนื้อหา.....148

1. การวิเคราะห์ความเที่ยงตรง.....148

1.1 การวิเคราะห์ความสอดคล้องระหว่างข้อคำถาม-วัตถุประสงค์ของการวิจัย..... 149

1.2 การวิเคราะห์ความเห็นสอดคล้องกัน 154
สรุป 162
แบบฝึกหัดท้ายบท 162
บทที่ 9 การวิเคราะห์ข้อมูลเชิงสำรวจ 163
เนื้อหา 164
1. การตรวจสอบความถูกต้องและความสมบูรณ์ของข้อมูล 164
1.1 การตรวจสอบการออกแบบข้อมูล 164
1.2 การตรวจสอบความถูกต้องของข้อมูล 167
1.3 การตรวจสอบค่าสูญหาย 173
1.4 การตรวจสอบค่าผิดปกติและค่าต่างสุด 178
2. การตรวจสอบข้อมูลตามข้อตกลงเบื้องต้น 181
2.1 การแจกแจงแบบปกติ 181
2.2 ความเท่ากันของความแปรปรวน 186
2.3 ความเป็นเส้นตรง 188
สรุป 192
แบบฝึกหัดท้ายบท 192

ส่วนที่ 4 การวิเคราะห์ข้อมูลด้วยสถิติเชิงพรรณนา

บทที่ 10 การวิเคราะห์ข้อมูลเชิงพรรณนา 1 ตัวแปร 195
เนื้อหา 196
1. การวิเคราะห์ข้อมูลแบบค่าจัดประเภท 196
1.1 การวิเคราะห์ข้อมูลด้วยค่าร้อยละ 196
1.2 การวิเคราะห์ข้อมูลด้วยค่าฐานนิยม 199
2.1 การวิเคราะห์ข้อมูลค่าเฉลี่ย (รายตัวแปร) 201
2.2 การวิเคราะห์ข้อมูลด้วยค่าเฉลี่ย (ตัวแปรดัชนี) 205
สรุป 208

แบบฝึกหัดท้ายบท 208

บทที่ 11 การวิเคราะห์ข้อมูลเชิงพรรณนา 2 ตัวแปร 211

เนื้อหา 212

1. การวิเคราะห์ข้อมูลระหว่างตัวแปรแบบค่าจัดประเภท 212

1.1 การวิเคราะห์แบบตารางไขว้ 2 ตัวแปร 212

1.2 การวิเคราะห์แบบตารางไขว้ 2 ตัวแปร โดยมีตัวแปรคุม 217

2. การวิเคราะห์ข้อมูลระหว่างตัวแปรแบบค่าแบ่งประเภทกับตัวแปรแบบค่าต่อเนื่อง 223

2.1 การเปรียบเทียบค่าเฉลี่ย 2 กลุ่ม..... 223

2.2 การเปรียบเทียบค่าเฉลี่ยมากกว่า 2 กลุ่ม..... 227

2.3 การเปรียบเทียบค่าเฉลี่ยโดยมีตัวแปรคุม..... 231

สรุป.....234

แบบฝึกหัดท้ายบท 234

ส่วนที่ 5 การวิเคราะห์ข้อมูลด้วยสถิติเชิงอ้างอิง

บทที่ 12 การทดสอบข้อมูลแบบอิงค่าพารามิเตอร์ 237

เนื้อหา.....238

1. การทดสอบค่าเฉลี่ยของกลุ่มตัวอย่าง 1 กลุ่ม 238

2. การทดสอบค่าเฉลี่ยของกลุ่มตัวอย่างที่ไม่เป็นอิสระจากกัน 2 กลุ่ม 242

3. การทดสอบค่าเฉลี่ยของกลุ่มตัวอย่างที่เป็นอิสระจากกัน 2 กลุ่ม 247

4. การทดสอบค่าเฉลี่ยของกลุ่มตัวอย่างที่เป็นอิสระจากกันมากกว่า 2 กลุ่ม 252

5. การทดสอบความสัมพันธ์รายคู่แบบเพียร์สัน.....264

5.1 การทดสอบความสัมพันธ์แบบง่าย..... 266

5.2 การทดสอบความสัมพันธ์แบบหพุ 269

สรุป.....273

แบบฝึกหัดท้ายบท 274

บทที่ 13 การทดสอบข้อมูลแบบไม่อิงค่าพารามิเตอร์ (1 ตัวแปร) 275

เนื้อหา 276

1. การทดสอบการแจกแจงของข้อมูล 277

1.1 การทดสอบการแจกแจงของข้อมูลแบบรันส์ 277

1.2 การทดสอบการแจกแจงของข้อมูลแบบคอลมอโกรอฟ-สเมอร์นอฟ 280

2. การทดสอบสัดส่วนของกลุ่มตัวอย่าง 284

2.1 การทดสอบความแนบสนิทแบบไคสแควร์ 284

2.2 การทดสอบสัดส่วนของกลุ่มตัวอย่างแบบทวินาม 288

สรุป 291

แบบฝึกหัดท้ายบท 291

บทที่ 14 การทดสอบข้อมูลแบบไม่อิงค่าพารามิเตอร์ (2 ตัวแปร) 293

เนื้อหา 294

1. การทดสอบความเป็นอิสระของกลุ่มตัวอย่าง 294

2. การทดสอบค่ากลางของกลุ่มตัวอย่างที่เป็นอิสระจากกัน 2 กลุ่ม 303

3. การทดสอบค่ากลางของกลุ่มตัวอย่างที่เป็นอิสระจากกันมากกว่า 2 กลุ่ม 307

4. การทดสอบค่ากลางของกลุ่มตัวอย่างที่มีความสัมพันธ์กัน 2 กลุ่ม 311

5. การทดสอบค่ากลางของกลุ่มตัวอย่างที่มีความสัมพันธ์กันมากกว่า 2 กลุ่ม 315

6. การทดสอบความสัมพันธ์รายคู่แบบสเปียร์แมน 320

6.1 การทดสอบความสัมพันธ์อันดับที่แบบง่าย 322

6.2 การทดสอบความสัมพันธ์อันดับที่แบบพหุ 326

สรุป 331

แบบฝึกหัดท้ายบท 332

บรรณานุกรม **333**

ภาษาไทย 333

ภาษาอังกฤษ 333

ดัชนี **338**

สารบัญตาราง

ตารางที่ 1.1 เปรียบเทียบขั้นตอนการวิจัยเชิงปริมาณและเชิงคุณภาพ 16
ตารางที่ 2.1 ประเภทตัวแปรแบบตัวอักษรและตัวแปรแบบตัวเลข 39
ตารางที่ 3.1 ความสัมพันธ์ระหว่างประเภทข้อมูลกับสถิติเชิงพรรณนา 48
ตารางที่ 3.2 เปรียบเทียบข้อตกลงเบื้องต้นของการทดสอบแบบอิงค่าพารามิเตอร์และไม่อิงค่าพารามิเตอร์ 50
ตารางที่ 3.3 เปรียบเทียบการทดสอบแบบอิงค่าพารามิเตอร์และไม่อิงค่าพารามิเตอร์ 51
ตารางที่ 3.4 ขั้นตอนการทดสอบสมมติฐาน 53
ตารางที่ 3.5 การสรุปสมมติฐาน 59
ตารางที่ 4.1 การสร้างตัวแปรจากแบบสอบถาม 66
ตารางที่ 4.2 คำอธิบายประเภทข้อมูล 73
ตารางที่ 4.3 ค่าเริ่มต้นค่าตัวแปรของตัวแปรแต่ละประเภท 73
ตารางที่ 4.4 การสร้างคำอธิบายตัวแปร 74
ตารางที่ 4.5 การสร้างคำอธิบายค่าตัวแปร 75
ตารางที่ 4.6 ประเภทของค่าตัวแปร 75
ตารางที่ 4.7 ประเภทตัวแปรและระดับการวัด 78
ตารางที่ 4.8 ความหมายของบทบาทของตัวแปร 78
ตารางที่ 12.1 การแปลค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน 265
ตารางที่ 14.1 การแปลค่าสัมประสิทธิ์สหสัมพันธ์แบบสเปียร์แมน 321

สารบัญภาพ

รูปที่ 1.1 กระบวนการวิจัย....15
รูปที่ 2.1 ความเกี่ยวข้องระหว่างตัวแปรเหตุ ตัวแปรผล และตัวแปรบริบท....36
รูปที่ 3.1 ข้อมูลภายใต้ระฆังคว่ำ....54
รูปที่ 3.2 เขตวิกฤต....55
รูปที่ 3.3 ค่าวิกฤติการทดสอบแบบหางเดียวทางขวา....56
รูปที่ 3.4 ค่าวิกฤติการทดสอบแบบหางเดียวทางซ้าย....56
รูปที่ 3.5 ค่าวิกฤติการทดสอบแบบ 2 หาง....57
รูปที่ 4.1 กระบวนการประมวลผลของคอมพิวเตอร์....64
รูปที่ 4.2 ส่วนประกอบของชุดข้อมูล....65
รูปที่ 4.3 การเปลี่ยนรูปแบบสอบถามเป็นแฟ้มข้อมูล....65
รูปที่ 4.4 ข้อมูลในแฟ้มข้อมูลนามสกุล CSV....66
รูปที่ 9.1 ค่าผิดปกติและค่าต่างสุด....178
รูปที่ 9. 2 ความเบ้และความโด่ง....182
รูปที่ 9.3 ความแปรปรวน....186
รูปที่ 9.4 ความสัมพันธ์เชิงเส้นตรงและเส้นโค้ง....188
รูปที่ 13.1 การแจกแจงความถี่....280

ส่วนที่ 1

ความรู้เบื้องต้นเกี่ยวกับการวิจัยเชิงปริมาณด้านสังคมศาสตร์

บทที่ 1

การวิจัยด้านสังคมศาสตร์

ความมุ่งหมายของบทเรียน

เพื่อให้ผู้เรียนเข้าใจเกี่ยวกับการหาความรู้ด้านสังคมศาสตร์ รวมถึงสาขาวิชาสังคมวิทยา

เนื้อหา

1. การหาความรู้ด้านสังคมศาสตร์
2. ประเภทการวิจัยด้านสังคมศาสตร์
3. ขั้นตอนการวิจัยด้านสังคมศาสตร์

วิธีการสอนและกิจกรรมการเรียนการสอน

บรรยาย อธิบาย แสดงตัวอย่าง และทำแบบฝึกหัดท้ายบท

สื่อและอุปกรณ์การสอน

เอกสารคำสอน ภาพเลื่อน คอมพิวเตอร์ โปรแกรมสอนออนไลน์ (สอนแบบออนไลน์) และเครื่องฉาย (สอนในชั้นเรียน)

งานที่มอบหมาย

ทำแบบฝึกหัดท้ายบท

การวัดผลและประเมินผล

การเข้าชั้นเรียน สังเกตความสนใจจากการเรียน สอบถามความเข้าใจจากการบรรยาย และการทำแบบฝึกหัดท้ายบท

เนื้อหา

การตอบคำถามที่เกิดจากข้อสงสัยของมนุษย์มีการพัฒนาการควบคู่ไปกับวิวัฒนาการของสังคมมนุษย์อย่างต่อเนื่อง กล่าวคือ เริ่มจากการตอบข้อสงสัยด้วยความเชื่อ/ความศรัทธา (fictitious stage) ต่อมามีการพัฒนามาเป็นการตอบข้อสงสัยด้วยหลักเหตุผล (metaphysical stage) และที่ได้รับการยอมรับอย่างแพร่หลายในปัจจุบันก็คือ การตอบปัญหาหรือข้อสงสัยต่าง ๆ ด้วยหลักวิทยาศาสตร์ (positive stage) มีการตั้งสมมุติฐาน (hypothesis) เก็บรวบรวมข้อมูลมาทำวิเคราะห์ว่าจะยอมรับ (accept) หรือปฏิเสธ (reject) ข้อสมมุติฐาน

1. การหาความรู้ด้านสังคมศาสตร์

การค้นหาความรู้ที่ไม่ใช้วิธีการที่เป็นกระบวนการทางวิทยาศาสตร์ (nonscientific method) เช่น การพยากรณ์โดยเทพ (oracles) การเข้าฌาน (mysticism) การบอกเล่าโดยผู้วิเศษ (magic) การทำนายด้วยโหราศาสตร์ (astrology) คำผีบอก (spirit) มิใช่ไม่มีอยู่ แต่เป็นวิธีการหาความรู้ที่อยู่นอกขอบเขตการหาความรู้ที่เรียกว่า “ศาสตร์” และได้รับการยอมรับจากคนในสังคมสมัยใหม่น้อยมาก (Neuman, 2014, p. 8)

1.1 ศาสตร์ คือ อะไร

คำว่า “science” หรือ “ศาสตร์” หมายถึง แบบแผนทางสังคม (social institution) และกระบวนการหาความรู้ (produce knowledge) ประเภทหนึ่ง อันเป็นประดิษฐ์กรรม (invention) ที่เกิดมาจากการเปลี่ยนแปลงทางความคิดของมนุษย์ ในยุคเหตุผล (age of reason) หรือยุคแสงสว่างทางปัญญา (enlightenment period) ของประวัติศาสตร์ยุโรปตะวันตกในช่วงปลาย ค.ศ. 1600 ถึงก่อนปลาย ค.ศ. 1800 ที่ให้ความสำคัญกับการคิดหาเหตุผลเชิงตรรกศาสตร์ (logical reasoning) ประสบการณ์ (experience) ในโลกของวัตถุ (material world) ความเชื่อ (belief) ในความก้าวหน้าของมนุษย์ (human progress) และการท้าทายด้วยการตั้งคำถามเกี่ยวกับอำนาจของศาสนาที่มีมาแต่โบราณ (Neuman, 2014, p. 7-8) การหาความรู้ที่เป็นศาสตร์ มีจุดเน้นที่สำคัญ 2 ส่วน ดังนี้ (Punch, 1998, p. 8)

1. ด้านข้อมูล ให้ความสำคัญกับความคิดที่สามารถนำไปทดสอบได้กับข้อมูลจากโลกความจริง (real-world data) หรือข้อมูลเชิงประจักษ์ (empirical data)

2. ด้านทฤษฎี ที่ถือได้ว่าเป็นหัวใจของศาสตร์ จึงต้องเน้นการอธิบาย (explain) ด้วยข้อมูล ไม่ใช่รวบรวมข้อมูลแล้วไม่ใช้ข้อมูลในการอธิบาย

ดังนั้น “ศาสตร์” จึงเป็นการรวบรวมข้อมูลที่เกี่ยวกับโลก สร้างทฤษฎีโดยใช้ข้อมูลในการอธิบายและทดสอบทฤษฎีด้วยข้อมูล

คนส่วนใหญ่เมื่อได้ยินคำว่า “science” มักจะนึกถึงภาพและเข้าใจว่าเป็นเรื่องของการทดลองในห้องปฏิบัติการคอมพิวเตอร์ การส่งกระสวยอวกาศไปในจักรวาล และคนที่ใส่เสื้อคลุมสีขาวทำงานอยู่ในห้องทดลอง ซึ่งในความเป็นจริงแล้วก็ไม่ใช่เรื่องที่ผิด เพียงแต่ว่า ภาพและความเข้าใจดังกล่าวนั้นเป็นเพียงส่วนหนึ่งของการค้นหาความรู้ที่มีพื้นฐานมาจากวิทยาศาสตร์ธรรมชาติเท่านั้นเอง (Neuman, 2014, p. 7)

ปัจจุบันวิธีการหาความรู้หรือที่เรียกว่า ศาสตร์ (science) แบ่งออกได้เป็น 2 ศาสตร์ใหญ่ๆ คือ วิทยาศาสตร์ธรรมชาติ (natural sciences) และสังคมศาสตร์ (science: social) ความจริงแล้วทั้ง 2 ศาสตร์เป็นทั้งเพื่อนรักและคู่แข่งกันมาอย่างต่อเนื่อง มีทั้งค้นหาความรู้ที่คาบเกี่ยวกัน หยิบยืมวิธีการหาความรู้ (methodology) มาใช้งานร่วมกัน และโต้แย้ง/คัดค้าน/ปฏิเสธความรู้ระหว่างกันอยู่เสมอ

1.2 ความรู้เบื้องต้นเกี่ยวกับสังคมศาสตร์

สังคมศาสตร์ สามารถมองได้หลายมิติ กล่าวคือ เป็นสาขาวิชาที่มีขอบข่ายกว้าง เนื่องจากเป็นการศึกษาเกี่ยวกับสังคมของมนุษย์ (human society) ครอบคลุมตั้งแต่การศึกษาปัจเจกบุคคล (individual) เช่น จิตวิทยาสังคม (social psychology) จนถึงการศึกษาธรรมชาติ (nature) เช่น ชีววิทยาเชิงสังคม (social biology) และภูมิศาสตร์ทางสังคม (social geography) มีระเบียบวิธีการที่ใช้ในการศึกษา (methodologically) หาความรู้ทั้งแบบการวิเคราะห์จากมุมมองที่ควรจะเป็น (normative approaches) เช่น กฎหมาย (law) ปรัชญาสังคมศาสตร์ (social philosophy) ทฤษฎีการเมือง (political theory) และการศึกษาด้วยหลักฐานทางประวัติศาสตร์ (historical approaches) เช่น ประวัติศาสตร์ทางสังคม (social history) และประวัติศาสตร์ทางเศรษฐกิจ (economic history) และก็มีบ่อยครั้งที่ สังคมศาสตร์ หมายถึง สังคมวิทยา (sociology) หรือ การวิเคราะห์ทฤษฎีสังคม (synthetic

social theory) เท่านั้น (Kuper and Kuper, 2003) ซึ่งก็สอดคล้องกับพจนานุกรมฉบับราชบัณฑิตยสถาน พ.ศ. 2525 (ราชบัณฑิตยสถาน, 2539, น. 817) ที่ได้ให้ความหมายของ “สังคมศาสตร์” ว่า เป็นศาสตร์ว่าด้วยความรู้เกี่ยวกับสังคม มีหมวดใหญ่ๆ เช่น ประวัติศาสตร์ (เกี่ยวกับสังคม) มานุษยวิทยา สังคมวิทยา เศรษฐศาสตร์ รัฐศาสตร์ จิตวิทยาสังคม

การศึกษาด้านสังคมศาสตร์มีมาตั้งแต่โบราณ แต่แนวคิดการศึกษาทางด้านสังคมศาสตร์สมัยใหม่ (modern social sciences) ที่ใช้วิธีการที่เป็นวิทยาศาสตร์ (scientific method) ได้เริ่มขึ้นในยุครุ่งเรืองทางความคิด (enlightenment) ในศตวรรษที่ 18 โดยเริ่มต้นในสาขาวิชาเศรษฐศาสตร์ ต่อมาในศตวรรษที่ 19 ได้แยกแนวคิดออกไปพัฒนาในสาขาวิชามานุษยวิทยา รัฐศาสตร์ จิตวิทยา และสังคมวิทยา ในลักษณะของทฤษฎีระดับกว้าง (wide-ranging theories) เช่นงานของ ออกุสต์ กองต์ (Auguste Comte) คาร์ล มาร์กซ์ (Karl Marx) และเฮอร์เบิร์ต สเปนเซอร์ (Herbert Spencer) และในศตวรรษที่ 20 ได้มีการพัฒนาไปสู่การนำเอาวิธีเชิงปริมาณ (quantitative method) และเทคนิคเชิงสถิติ (statistical techniques) และรวมถึงวิธีเชิงประจักษ์ (empirical method)[1] มาใช้มากขึ้น

โดยทั่วไปแล้ว คำว่า “สังคมศาสตร์ (social science)” หมายถึง การศึกษาพฤติกรรมของมนุษย์ด้วยกระบวนการทางวิทยาศาสตร์ ที่เกิดมาจากคำว่า สังคม (social) คือ ประชาชนและพฤติกรรมของประชาชนที่เกิดจากบริบททางสังคม (social context) กับคำว่า ศาสตร์ (science) คือ วิธีการในการศึกษาประชาชนและพฤติกรรมของประชาชน เมื่อเป้าหมายของศาสตร์ คือ การสร้างทฤษฎีที่อธิบายเกี่ยวกับประชาชนและพฤติกรรมของประชาชน ดังนั้น ทฤษฎีด้านสังคมศาสตร์จึงวางอยู่บนฐานของพฤติกรรมของมนุษย์ และการทดสอบด้วยข้อมูลจากโลกความจริง (Punch, 1998, p. 9)

สังคมวิทยา (sociology) คือ วิชาที่ว่าด้วยการศึกษาสังคมมนุษย์และการกระทำระหว่างกันทางสังคม (social interaction) อย่างเป็นระบบ (systematic) และเป็นการศึกษาที่มีลักษณะเป็นวัตถุวิสัย (objective) แบบวิทยาศาสตร์ (David Popenoe 1993 , p. 2) สาขาวิชาสังคมวิทยานอกจาเป็นวิชาที่มี

---

[1] ปัจจุบันนักสังคมศาสตร์ส่วนใหญ่มักจะใช้วิธีเชิงประจักษ์ 2 ประเภทด้วยกันในการศึกษาหาความรู้ นั่นก็คือ

1. การบรรยาย (descriptive studies) ขั้นแรกของการค้นหาความรู้ ก็คือ การค้นหาว่าเหตุการณ์ที่เกิดขึ้นนั้นเป็นอย่างไร เกิดกับใคร ที่ไหน และเมื่อใด

2. การอธิบาย (explanatory studies) เมื่อได้ทำการบรรยายเหตุการณ์ที่เกิดขึ้นแล้ว นักสังคมวิทยาก็จะอธิบายว่าเหตุการณ์นั้นเกิดขึ้นได้อย่างไร และเกิดขึ้นเพราะอะไร

ความใกล้เคียงและสัมพันธ์กับวิทยาศาสตร์ธรรมชาติตลอดจนวิธีการที่ใช้ในการศึกษาแล้ว สาขาวิชาสังคมวิทยายังมีความสัมพันธ์กับสังคมศาสตร์สาขาอื่น ๆ อีกด้วย โดยในการศึกษาทางสังคมวิทยาจะมองว่าพฤติกรรมของมนุษย์ คือรูปแบบของพฤติกรรมที่ถูกกำหนดโดยสังคม และสภาพแวดล้อมทางสังคม (social circumstance) นักสังคมวิทยาส่วนใหญ่จะอธิบายว่า สิ่งที่มนุษย์กระทำขึ้นในสังคมนั้นจะมีผลกระทบต่อบุคคลต่าง ๆ ด้วย ดังนั้นปรากฏการณ์ต่าง ๆ ที่เกิดขึ้นในสังคมย่อมจะมีผลกระทบซึ่งกันและกันและมีผลกระทบต่อกันอย่างสลับซับซ้อนด้วยกระบวนการกระทำระหว่างกันทางสังคม (เรวัต แสงสุริยงค์, 2541)

มีการกล่าวว่า กลุ่มวิชาทางสังคมศาสตร์ใช้ระเบียบวิธีทางวิทยาศาสตร์เพื่อศึกษาโลกในเรื่องที่เกี่ยวข้องกับมนุษย์ วิชาในสังคมศาสตร์แตกต่างจากวิชาในกลุ่มมนุษยศาสตร์ เนื่องจากหลายๆ สาขาวิชาในสังคมศาสตร์เน้นการหาความรู้ด้วยกระบวนการทางวิทยาศาสตร์ อย่างไรก็ตามหลายๆ สาขาก็อาจจะเน้นระเบียบวิธีเชิงคุณภาพมากกว่าก็ได้ นั่นอาจเพราะมีสาเหตุมาจากที่ทั้งสังคมศาสตร์และวิทยาศาสตร์ธรรมชาติต่างก็มีฐานะเป็นส่วนหนึ่งของปรัชญาวิทยาศาสตร์ (philosophy of science) ที่แยกตัวออกมาจากปรัชญาความรู้ (epistemology)[2] แต่ด้วยเหตุที่ความก้าวหน้าของวิทยาศาสตร์ธรรมชาติที่มีบทบาทและครอบงำแนวคิดและวิธีการศึกษาทางวิทยาศาสตร์อย่างมากและรวมถึงวิธีการศึกษาทางสังคมศาสตร์ด้วย จึงทำให้นักคิดทางสังคมศาสตร์มีการต่อต้านวิธีการของวิทยาศาสตร์ธรรมชาติและพัฒนาวิธีการศึกษาทางสังคมศาสตร์มาใช้ศึกษาปรากฏการณ์ทางสังคม คือ การวิเคราะห์ในเชิงตีความ (interpretive analysis) อันเป็นแนวการศึกษาปรากฏการณ์ทางสังคมที่มาจากปรัชญาการตีความ (hermeneutics or interpretative philosophy) (อนุสรณ์ ลิ่มมณี, 2542, น. 4-57)

1.3 วิธีการหาความรู้

แม้ว่าการหาความรู้จะมีการสร้างและพัฒนาวิธีการการศึกษาที่แตกต่างกันไปตามธรรมชาติแต่ละสาขาวิชา แต่วิธีการหาความรู้ทั้งหลายตกอยู่ภายใต้ 2 แนวทางด้วยกัน ดังนี้

[2] เป็นปรัชญาสาขาหนึ่งที่เน้นศึกษาเกี่ยวกับลักษณะของความรู้และข้ออ้างด้านหลักการตลอดจนแนวคิดเกี่ยวกับความรู้ในแง่มุมต่าง ๆ โดยทั่วไป อาทิ ที่มา ฐานคติ สภาพ ขอบเขต และความถูกต้อง (อนุสรณ์ ลิ่มมณี, 2542, p. 5)

1. การใช้วิธีเชิงประจักษ์ (Empirical method) ซึ่งเป็นวิธีที่นักวิทยาศาสตร์ รวมทั้งนักสังคมศาสตร์นิยมใช้กันเป็นส่วนมาก การหาความรู้ด้วยวิธีเชิงประจักษ์ เป็นการใช้การรับรู้ของมนุษย์ เช่น การได้เห็น และการได้ยิน หรือที่เราเรียกกันว่า การสังเกต (observation)

2. การใช้วิธีไม่ใช่เชิงประจักษ์ (Non-empirical method) มี 2 วิธี ดังนี้

2.1 การใช้หลักเหตุผล (Logic) เป็นการหาความรู้ด้วยวิธีการใช้เหตุผลเพื่อให้ได้มาซึ่งคำตอบนั้น มี 4 วิธี คือ

2.1.1 หลักนิรนัย/อนุมาน (Deductive) คือ การหาความรู้โดยใช้ข้อสรุปทั่วไปหรือทฤษฎีอธิบายเรื่องเฉพาะเรื่องใดเรื่องหนึ่ง (ผล) อธิบายปรากฏการณ์ที่เกิดขึ้น (เหตุ) โดยผลที่เกิดขึ้นนั้นมาจากเหตุเดียวเท่านั้น

2.1.2 หลักอุปนัย (Inductive) คือ การหาความรู้โดยใช้ปรากฏการณ์ที่เกิดขึ้น (เหตุ) อธิบายหรือสร้างข้อสรุปทั่วไปหรือทฤษฎี (ผล) แต่ผลที่เกิดขึ้นนั้นไม่จำเป็นต้องเกิดมาจากเหตุเดียว อาจเกิดมาจากหลายเหตุด้วยกันก็ได้

2.1.3 หลักจารนัย (Abductive) คือ การหาความรู้โดยใช้หลัก การอ้างถึงปรากฏการณ์ที่เกิดขึ้น (เหตุ) ที่เป็นสิ่งที่ใช้อธิบายข้อสรุปทั่วไปหรือทฤษฎี (ผล) เป็นการใช้เหตุผลในทิศทางตรงข้ามกับหลักนิรนัย

2.1.4 หลักกลนัย (Retroductive) คือ การหาความรู้โดยใช้หลักเหตุผลเหมือนกับการอุปนัย แต่การอธิบาย/ทำนายข้อสรุปทั่วไปหรือทฤษฎี (ผล) มีความเกี่ยวข้องกับหลายกฎ (rules) หรือหลายสิ่งที่ได้จากการสังเกต (observations) หรืออาจกล่าวได้ว่า ตัวอธิบาย/ทำนายมีหลายกฎ

2.2 การใช้ญาณหรือการหยั่งรู้ (Intuition) คือ การหาความรู้ที่ได้จากประสาทสัมผัส เช่น การเห็น การได้ยิน ความรู้สึกจากสิ่งต่าง ๆ

อาจกล่าวได้ว่าการหาความรู้ของมนุษย์ที่กล่าวมาวางอยู่บนฐานของ 2 ขั้วระหว่างการสังเกต (observation) กับการใช้เหตุผล (reason) หรือที่บางครั้งเราอาจได้ยินและได้เห็นในรูปของคำพูดที่ว่า ความรู้เชิงประจักษ์ (empirical knowledge) คือ ความรู้ที่ได้มาจากการสังเกตและได้ฟังเกี่ยวกับสิ่งต่างๆ และความรู้เชิงเหตุผล (logical/rational knowledge) คือ ความรู้ที่ได้มาจากการให้เหตุผลโดยไม่ได้เห็นหรือไม่ได้ยินมา

การหาความรู้อาจขาดความน่าเชื่อถือและไม่ได้รับการยอมรับ หากเป็นการหาความรู้ที่ไม่ได้มาจากการวิจัย (research) ด้วยการค้นคว้าเพื่อหาข้อมูลอย่างถี่ถ้วนตามหลักวิชา หรือเป็นกระบวนการค้นหาความรู้อย่างเป็นระบบ (systematic process) จากเหตุการณ์ (events) พฤติกรรม (behaviors) ทฤษฎี (theories) และกฎ (laws) ต่าง ๆ เพื่อการค้นพบ (discovering) การตีความ (interpreting) และแก้ไขข้อเท็จจริง (revising facts) ดังนั้น นักสังคมศาสตร์ จึงต้องใช้วิธีการหาความรู้ด้วยการวิจัย เพื่อตอบคำถามหรือข้อสงสัยเกี่ยวกับสังคมมนุษย์

2. ประเภทการวิจัยด้านสังคมศาสตร์

สังคมศาสตร์มีขอบเขตของการหาความรู้ที่กว้างและมีบริบทที่เกี่ยวข้องที่ซับซ้อน คือ มีการศึกษาทั้งเรื่องในอดีต ปัจจุบัน อนาคต ธรรมชาติของมนุษย์ สิ่งที่สร้างโดยมนุษย์ และสภาพแวดล้อมของมนุษย์ ดังนั้นนักสังคมศาสตร์ (social scientist) จึงต้องใช้หลากหลายวิธีการในการอธิบาย (describe) ค้นหา (explore) และทำความเข้าใจ (understand) ความเป็นอยู่ของคนในสังคม (social life) และเหตุการณ์ต่างๆ ที่เกิดขึ้นในสังคมตามความเหมาะสม จึงทำให้งานวิจัยทางสังคมศาสตร์ถูกสร้างและพัฒนาขึ้นมาเป็นจำนวนมาก แต่หากพิจารณาให้ดีจะพบว่างานวิจัยเหล่านั้นมีคุณลักษณะบางประการร่วมกันอยู่ เช่น วิธีการใช้เหตุผลในการหาความรู้ การใช้องค์ความรู้ในการวิจัย ลักษณะของข้อมูลที่เก็บรวบรวม วิธีการออกแบบการวิจัย เป็นต้น

2.1 การจัดประเภทงานวิจัย

นักวิชาการนำเอาคุณลักษณะของงานวิจัยมาแบ่งประเภทของงานวิจัยให้เป็นหมวดหมู่ แต่ก็ไม่ได้หมายความว่านักวิชาการแต่ละคนจะมีมุมมองที่เหมือนกัน จึงเป็นสาเหตุที่ทำให้เราเห็นการจำแนกหมวดหมู่ของงานวิจัยที่แตกต่างกันหลายรูปแบบดังต่อไปนี้ (Babbie, 2007, p. 25-26; Neuman, 2014, p. 24; Keith, 1998, p. 39-41)

1. การจำแนกตามประโยชน์ของการวิจัย (Use of research) มี 2 ประเภท ดังนี้

1.1 การวิจัยพื้นฐาน (Basic research) เป็นการวิจัยเพื่อพัฒนากรอบแนวคิด ทฤษฎี และองค์ความรู้ทางสังคมศาสตร์เอง เน้นการทดสอบทฤษฎีเพื่อสนับสนุนหรือปฏิเสธทฤษฎี การวิจัยประเภทนี้อาจเรียกได้ว่าเป็นการวิจัยเชิงยืนยัน (confirmatory research) หรือนักวิชาการบางท่านอาจเรียกว่า

การวิจัยบริสุทธ์ (pure research) คือ การวิจัยที่ใช้องค์ความรู้ที่มีอยู่ทำให้เกิดองค์ความรู้ใหม่ตามมา (knowledge for knowledge's sake)

1.2 การวิจัยประยุกต์ (Applied research) เป็นการวิจัยเพื่อตอบปัญหาและแก้ไขปัญหาทางสังคม เน้นการค้นหาความสัมพันธ์ระหว่างเหตุและผลของปรากฏการณ์ทางสังคมหรือเหตุการณ์ต่างๆ ที่เกิดขึ้นในสังคม การวิจัยประเภทนี้อาจเรียกได้ว่าเป็นการวิจัยเชิงสำรวจ (exploratory research) เช่น การวิจัยเพื่อประเมินผลโครงการและการใช้นโยบายของรัฐและเอกชน การทดลองใช้โครงการต่าง ๆ ของสังคมก่อนที่จะนำไปใช้จริง การคาดคะเนสถานการณ์ทางสังคม การสำรวจความต้องการของผู้บริโภค เพื่อประชาสัมพันธ์และโฆษณาขายสินค้าต่าง ๆ งานวิจัยประเภทนี้มีหลายลักษณะ แต่ในที่นี้จะจำแนกเพื่อเป็นตัวอย่างดังนี้

1.2.1 การวิจัยประเมินผล (Evaluation research) เป็นการวิจัยเกี่ยวกับประสิทธิภาพและประสิทธิผลในการดำเนินโครงการ นโยบาย และการรณรงค์ต่างๆ

1.2.2 การวิจัยปฏิบัติการ (Action research) เป็นการวิจัยที่ให้ความสำคัญกับการมีส่วนร่วมของหน่วยวิเคราะห์ เพราะเชื่อว่าการพัฒนาความรู้ต้องมาจากประสบการณ์ของสามัญชนและประชาชนทั่วไป และสามารถนำไปใช้ปรับปรุงและแก้ไขปัญหาสถานการณ์ได้เป็นอย่างดี

1.2.3 การวิจัยประเมินผลกระทบทางสังคม (Social impact assessment research) เป็นการวิจัยที่ต้องการคาดการณ์ผลที่อาจเกิดจากการเปลี่ยนแปลงในการดำเนินโครงการ และการเลือกดำเนินนโยบายใดนโยบายหนึ่ง

2. การจำแนกตามลักษณะการรวบรวมข้อมูลและเทคนิคการวิเคราะห์ (Data collection and analysis techniques) สามารถแบ่งย่อยได้เป็น 2 ประเภทกว้างๆ ดังนี้

2.1 การวิจัยเชิงปริมาณ (Quantitative method) เป็นวิธีการวิจัยที่พยายามจำแนกประเภทปรากฏการณ์ทางสังคม แล้วทำการรวบรวมและวิเคราะห์ (analyze) ด้วยข้อมูลที่เป็นตัวเลข (numerical data) โดยมุ่งเน้นการหาความเชื่อมโยงที่เกิดขึ้นจำนวนเล็ก ๆ น้อยๆ (a smaller number) ในบรรดาคุณลักษณะ (attribute) ของสิ่งที่ทำการศึกษาจากหลายๆ กรณี (many cases) ดังนั้น สิ่งต่างๆ ที่ทำการศึกษาจึงต้องสามารถวัดหรือนับได้เป็นตัวเลข มีการออกแบบการวิจัยในรูปแบบต่างๆ ดังนี้

2.1.1 การวิจัยแบบมีการปฏิสัมพันธ์กับหน่วยวิเคราะห์ (Reactive research) เป็นวิธีการวิจัยที่มีการติดต่อกันระหว่างนักวิจัยและหน่วยวิเคราะห์ในรูปแบบใดรูปแบบหนึ่ง มี 2 ประเภท ดังนี้

2.1.1.1 การวิจัยเชิงทดลอง (Experimental research) เป็นการออกแบบการวิจัยด้วยการจำลองหรือสร้างสอดแทรก (Intervention) หรือเลือกหน่วยวิเคราะห์ในห้องปฏิบัติการหรือชีวิตจริงที่มีสภาพแวดล้อมที่แตกต่างกันมาทำการศึกษาเปรียบเทียบความแตกต่างที่เกิดขึ้นระหว่างกลุ่มต่างๆ

2.1.1.2 การวิจัยเชิงสำรวจ (Survey research) เป็นการออกแบบการวิจัยโดยใช้แบบสอบถาม (questionnaires) หรือการสัมภาษณ์ (interviews) ในการค้นหาคำตอบจากหน่วยวิเคราะห์จำนวนมากๆ ด้วยวิธีการสุ่มตัวอย่าง (sampling)

2.1.2 การวิจัยแบบไม่มีการปฏิสัมพันธ์กับหน่วยวิเคราะห์ (Nonreactive research) เป็นวิธีการวิจัยที่มีการนำเอาข้อมูลของผู้ที่มีส่วนร่วมหรือหน่วยวิเคราะห์มาใช้โดยที่เขาไม่รู้ว่าข้อมูลถูกนำไปในการวิจัย มี 2 ประเภท ดังนี้

2.1.2.1 การวิเคราะห์เนื้อหา (Content analysis) เป็นการใช้วิธีการวิเคราะห์ข้อมูลจากเนื้อหาหรือสารสนเทศที่มีการจัดเก็บไว้แล้ว เช่น หนังสือ บทเพลง วรรณกรรม ภาพยนตร์ ภาพถ่าย โฆษณา

2.1.2.2 การใช้สถิติที่มีอยู่แล้ว (Existing statistics) เป็นการใช้ข้อมูลที่มีการเก็บรวบรวมไว้แล้ว ทั้งจากงานวิจัยที่ทำไว้แล้วหรือข้อมูลจากหน่วยงานภาครัฐและเอกชน เช่น รายได้ อายุ การศึกษา

2.2 การวิจัยเชิงคุณภาพ (Qualitative method) เป็นวิธีวิจัยที่ให้ความสำคัญกับประสบการณ์ของบุคคล (personal experiences) และการตีความ (interpretation) เพื่อทำความเข้าใจ (understanding) ความหมาย (meaning) ของปรากฏการณ์ทางสังคม มากกว่าการจำแนกประเภท และมุ่งเน้นไปที่การหาความเชื่อมโยงที่เกิดขึ้นเป็นส่วนใหญ่ (a larger number) ในบรรดาคุณลักษณะ (attribute) ของสิ่งที่ทำการศึกษาจาก 2-3 กรณี (few case) ที่มีความเกี่ยวข้องกัน โดยใช้การพรรณนา (descriptive) หรือการอธิบาย (explanatory) ด้วยคำพูดหรือภาษาบรรยายเป็นหลัก มีการออกแบบการวิจัยในรูปแบบต่างๆ ดังนี้

2.2.1 การวิจัยภาคสนาม (Field research) เป็นการวิจัยที่นักวิจัยทำการเก็บรวบรวมข้อมูลโดยการเข้าไปสังเกตด้วยตนเองและจดบันทึกเกี่ยวกับหน่วยวิเคราะห์และบริบทด้านต่างๆ ที่เกิดขึ้นตามความเป็นจริง โดยใช้ระยะเวลาที่ยาวนานจนกว่าจะได้ข้อมูลครบถ้วนและสมบูรณ์

2.2.2 การวิจัยเชิงประวัติศาสตร์เปรียบเทียบ (Historical-comparative research) เป็นการวิจัยที่นักวิจัยทำการเก็บรวบรวมข้อมูลเกี่ยวกับเหตุการณ์ต่างๆ ในเชิงประวัติศาสตร์ที่เกิดขึ้นในอดีต หรือเปรียบเทียบความแตกต่างในเชิงวัฒนธรรมระหว่างสังคม

3. การจำแนกตามเป้าหมายของการวิจัย (Purpose of research) มี 2 ประเภท ดังนี้

3.1 การวิจัยเชิงสำรวจเบื้องต้น (Exploratory research) เป็นการวิจัยที่มีเป้าหมายหลักอยู่ที่การตรวจสอบหรือค้นหาคำตอบสำหรับทำความเข้าใจประเด็นปัญหาหรือปรากฏการณ์เล็กๆ จากคำถามว่า อะไร (what) เพื่อพัฒนาความคิดและคำถามการวิจัยในระดับต่อไป

3.2 การวิจัยเชิงพรรณนา (Descriptive research) เป็นการวิจัยที่มีเป้าหมายหลักอยู่ที่การบรรยายให้เห็นภาพสิ่งที่ศึกษาด้วยตัวเลขและตัวอักษร ตอบคำถามตามเค้าโครงที่กำหนดไว้ว่าใคร (who) เมื่อไหร่ (when) ที่ไหน (where) อย่างไร (how) และนำเสนอการวิจัยในรูปของรายละเอียดและการจำแนกประเภท

3.3 การวิจัยเชิงอธิบาย (Explanatory research) เป็นการวิจัยที่มีเป้าหมายหลักอยู่ที่การอธิบายเหตุการณ์ต่างๆ ว่าเกิดขึ้นได้อย่างไร (why) รวมถึงการสร้าง ต่อเติม พัฒนา ขยาย และทดสอบทฤษฎี

4. การจำแนกตามประเภทกรณีศึกษา (Within or across cases)

4.1 การวิจัยแบบศึกษาเป็นรายกรณี (Case-study research) เป็นการศึกษาเชิงลึก (in-depth) อย่างละเอียดจากหน่วย (unites) หรือสิ่งที่ศึกษา (cases) ประมาณ 2-3 กรณี ในช่วงเวลาหนึ่งหรือเปรียบเทียบหลายๆช่วงเวลา

4.2 การวิจัยแบบศึกษาเปรียบเทียบกรณี (Across-case research) ส่วนใหญ่เป็นการศึกษาเชิงปริมาณ จากสิ่งที่ศึกษาที่มีจำนวนมาก (30-3,000 กรณี) เพื่อเปรียบเทียบข้อมูลระหว่างกรณีศึกษาประมาณ 2-3 คุณลักษณะหรือตัวแปร มากกว่าที่จะสังเกตในรายละเอียดแต่ละกรณี

5. การจำแนกตามมิติของเวลาในการวิจัย (Time dimension in research) มี 3 ประเภท ดังนี้

5.1 การวิจัยแบบตัดขวาง (Cross-sectional research) เป็นการวิจัยที่ต้องการค้นหาคำตอบโดยการรวบรวมข้อมูลมาจากหน่วยวิเคราะห์ในช่วงเวลาใดเวลาหนึ่ง (at one time) แต่การวิจัยแบบนี้มีข้อดีที่ประหยัดเวลาและงบประมาณ แต่ไม่สามารถแสดงให้เห็นถึงกระบวนการหรือการเปลี่ยนแปลงทางสังคมได้

5.2 การวิจัยแบบระยะยาว (Longitudinal research) เป็นการวิจัยที่ต้องการค้นหาคำตอบเชิงความสัมพันธ์โดยการรวบรวมข้อมูลจากหน่วยวิเคราะห์ที่อยู่ในช่วงเวลาที่แตกต่างกัน มี 3 รูปแบบ คือ

5.2.1 แบบช่วงเวลาต่อเนื่องกัน (Time-series) เป็นการรวบรวมข้อมูลจากหน่วยวิเคราะห์ที่มีคุณลักษณะแตกต่างกันกลุ่มหนึ่งหรือหลายกลุ่มในช่วงเวลาหนึ่งแบบติดต่อเป็นลำดับ

5.2.2 แบบช่วงเวลาที่ต่างกัน (Panel study) เป็นการเก็บรวบรวมข้อมูลจากหน่วยวิเคราะห์ที่มีคุณลักษณะเหมือนกันตั้งแต่ 2 ช่วงเวลาขึ้นไป

5.2.3 แบบติดตามไปข้างหน้า (Cohort) ทำการเก็บรวบรวมข้อมูลจากหน่วยวิเคราะห์ที่มีคุณลักษณะหรือประสบการณ์เหมือนกัน ตั้งแต่ 2 ช่วงเวลาขึ้นไป ทั้งแบบใช้ข้อมูลย้อนหลังและไม่ย้อนหลัง

5.3 แบบรายกรณี (Case study) เป็นการวิจัยที่ต้องการค้นหาข้อมูลเชิงลึก (in-depth) จากหน่วยวิเคราะห์จำนวนไม่มาก หรือหนึ่งกรณีในช่วงเวลาหนึ่งหรือเปรียบเทียบในหลายช่วงเวลา

6. การจำแนกตามประเภทของคำถามการวิจัย (Research question) มี 2 ประเภท ดังนี้

6.1 แบบไม่มีสมมติฐาน (without hypotheses) เป็นการวิจัยที่ออกแบบและวางแผนการวิจัยโดยเน้นการรวบรวมข้อมูลและวิเคราะห์ข้อมูลเพื่อนำไปใช้ตอบคำถามการวิจัยเท่านั้น

6.2 แบบมีสมมติฐาน (with hypotheses) เป็นการวิจัยที่ออกแบบและวางแผนที่เน้นการรวบรวมข้อมูลและวิเคราะห์ข้อมูลเพื่อนำไปตอบถามการวิจัยด้วยวิธีการทดสอบสมมติฐานที่ได้มาจากทฤษฎี

2.2 หลักเกณฑ์การจัดประเภทงานวิจัย

ภายใต้ความแตกต่างของการจำแนกประเภทงานวิจัยมีรูปแบบของการนิยามประเภทการวิจัยที่แตกต่างกันพอสรุปได้ ดังนี้

1. นิยามการแบ่งประเภทและแบ่งประเภทงานวิจัยเหมือนกัน เช่น จำแนกตามวิธีการรวบรวมข้อมูล แบ่งเป็น การวิจัยเชิงปริมาณ และการวิจัยเชิงคุณภาพ

2. นิยามการแบ่งประเภทเหมือนกันแต่มีประเภทงานวิจัยไม่เหมือนกัน เช่น จำแนกตามจุดมุ่งหมาย แบ่งเป็นการวิจัยพื้นฐาน การวิจัยเชิงทฤษฎี การวิจัยประยุกต์ การวิจัยแบบมีสมมติฐาน และการวิจัยแบบไม่มีสมมติฐาน

3. นิยามการแบ่งประเภทไม่เหมือนกันแต่แบ่งประเภทงานวิจัยเหมือนกันหรือคาบเกี่ยวกัน เช่น จำแนกตามวิธีการวิจัย แบ่งเป็นการวิจัยเชิงสำรวจ การวิจัยเชิงพรรณนา การวิจัยเชิงอธิบาย การวิจัยเชิงทำนาย

4. นิยามไม่เหมือนกันและมีประเภทงานวิจัยที่ไม่เหมือนกัน เช่น จำแนกตามสาขาการวิจัย แบ่งเป็น การวิจัยทางสังคมวิทยา การวิจัยทางรัฐศาสตร์ การวิจัยทางการศึกษา

กล่าวโดยสรุป มีวิธีการแบ่งประเภทงานวิจัยที่หลากหลาย แต่ที่พบอย่างแพร่หลายและนิยมมากที่สุด คือ แบ่งเป็นการวิจัยเชิงปริมาณ และการวิจัยเชิงคุณภาพ โดยมีนิยามการแบ่งประเภทที่แตกต่างกัน เช่น แบ่งตามลักษณะข้อมูล แบ่งตามวิธีการวิเคราะห์ข้อมูล แบ่งตามระเบียบวิธี แบ่งตามการตอบคำถามการวิจัย เป็นต้น

3. กระบวนการวิจัยด้านสังคมศาสตร์

การวิจัยทุกสาขาวิชาและทุกประเภทมีขั้นตอนการทำวิจัยที่ไม่แตกต่างกันมากนัก แต่โดยทั่วไปอาจแบ่งออกเป็นขั้นตอนหลักๆ ได้ดังนี้ (Scanlon cited in Wilkinson, 2000, p. 9)

1. ขั้นคัดเลือกความสนใจ (Choosing a focus) คือ การเลือกและกำหนดปัญหาที่จะทำการวิจัย

2. ขั้นออกแบบการวิจัย (Research design) คือ การเตรียมและตัดสินใจเกี่ยวกับการใช้วิธีการรวบรวมข้อมูลและการวิเคราะห์ข้อมูล

3. ขั้นรวบรวมข้อมูล (Data collection) คือ การเก็บข้อมูลด้วยการสัมภาษณ์ด้วยตนเองหรือส่งแบบสอบถามไปให้กลุ่มเป้าหมายตอบเอง

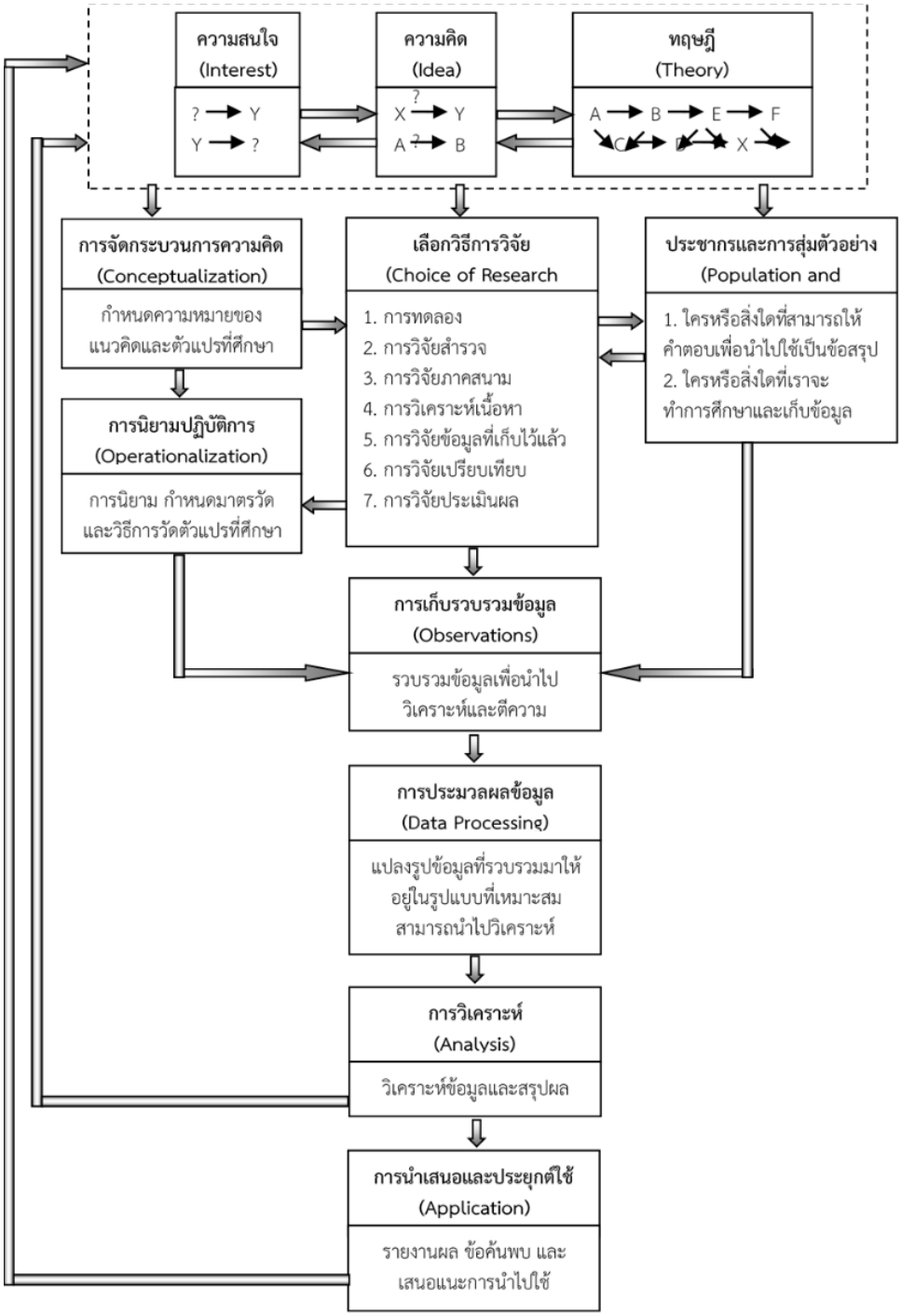
4. ขั้นวิเคราะห์ข้อมูล (Data analysis) คือ การวิเคราะห์และตีความข้อมูลที่ได้มาจากการเก็บรวบรวมข้อมูล

5. ขั้นเขียนผลการวิจัย (Writing up the results) คือ การนำเสนอผลการวิจัยด้วยพรรณนาและอธิบายจากผลการวิเคราะห์ข้อมูล

3.1 การออกแบบการวิจัย

การค้นหาความรู้ด้วยการวิจัยมีกระบวนการที่เป็นระบบ (systematic) การเริ่มทำการวิจัยจึงต้องมีการออกแบบการวิจัย (research design) อย่างเป็นขั้นตอน (step-by-step) ตามระเบียบวิธีของ

การทำวิจัย การออกแบบการวิจัยตามกระบวนการวิจัยทางสังคมศาสตร์สามารถแสดงให้เห็นเป็นแผนผังดังรูปด้านล่าง



รูปที่ 1.1 กระบวนการวิจัย (ปรับปรุงมาจาก Babbie, 2007, p. 108)

การเริ่มดำเนินโครงการวิจัยอาจมีจุดเริ่มต้นมาจากความสนใจ หรือความคิด หรือทฤษฎี เป็นไปได้ทั้งสิ้น หรืออาจเริ่มมาจากความสนใจที่ทำให้เกิดความคิดซึ่งเหมาะกับทฤษฎีหนึ่ง หรือทฤษฎีอาจทำให้เกิดความคิดใหม่ๆ และสร้างความสนใจใหม่ๆ ตามมาก็เป็นไปได้เช่นกัน (Babbie, 2007, p. 109)

3.2 การวิจัยเชิงปริมาณและการวิจัยเชิงคุณภาพ

การทำวิจัยของนักสังคมศาสตร์ต้องปฏิบัติการวิจัยตามลำดับขั้นของกระบวนการวิจัยให้ถูกต้อง แต่เนื่องจากการวิจัยทางสังคมศาสตร์ 2 ประเภทใหญ่ คือ การวิจัยเชิงปริมาณและการวิจัยเชิงคุณภาพ ต่างมีปรัชญาและแนวคิดในการค้นหาความรู้ที่แตกต่างกัน ดังนั้นการวิจัยทั้งสองประเภทจึงมีลำดับและขั้นตอนของการทำวิจัยที่แตกต่างกันอยู่บ้าง ดังตารางด้านล่าง (Neuman, 2014, p. 14-15)

ตารางที่ 1.1 เปรียบเทียบขั้นตอนการวิจัยเชิงปริมาณและเชิงคุณภาพ

ขั้นที่	การวิจัยเชิงปริมาณ	การวิจัยเชิงคุณภาพ
1	เลือกหัวข้อ (Select topic)	ทำความเข้าใจสังคม (Acknowledge social self)
2	ตั้งคำถามให้ชัดเจน (Focus question)	เกิดมุมมองและความเข้าใจ (Adopt perspective)
3	ออกแบบการศึกษา (Design study)	ออกแบบการศึกษา (Design study)
4	รวบรวมข้อมูล (Collect data)	รวบรวมข้อมูล (Collect data)
5	วิเคราะห์ข้อมูล (Analyze data)	วิเคราะห์ข้อมูล (Analyze data)
6	ตีความข้อมูล (Interpret data)	ตีความข้อมูล (Interpret data)
7	นำเสนอการวิจัย (Inform others)	นำเสนอการวิจัย (Inform others)

นักวิจัยที่เน้นวิธีการวิจัยเชิงคุณภาพ เริ่มต้นการวิจัยด้วยการใช้ตัวนักวิจัยเองทำการประเมินและไตร่ตรอง (self-assessment and reflections) สถานการณ์จากบริบทของสังคมเชิงประวัติศาสตร์ (sociohistorical context) เพื่อรับรู้ในเรื่องตัวตน (self-aware) ของสังคมหรือการวางตัวของนักวิจัยในสังคม ไม่มุ่งเน้นที่การตั้งคำถามเฉพาะเรื่องเหมือนกับการวิจัยเชิงปริมาณ แต่ให้ความสำคัญกับกระบวนทัศน์ทางทฤษฎีและปรัชญา (theoretical-philosophical paradigm) ของการอยากรู้อยากเห็น (inquisitive) และมุ่งเน้นการสร้างคำถามที่ต้องการคำตอบที่เปิดกว้าง (open-ended) อันจะนำไปสู่การเกิดมุมมองและความเข้าใจ สำหรับขั้นการออกแบบการศึกษา การรวบรวมข้อมูล การวิเคราะห์ข้อมูล และการตีความข้อมูล

การวิจัยเชิงปริมาณและการวิจัยเชิงคุณภาพมีขั้นตอนที่เหมือนกัน แต่มีความแตกต่างกันในรายละเอียดของแต่ละขั้นตอน กล่าวคือ นักวิจัยเชิงคุณภาพจะเน้นการกระบวนการวิจัยระหว่างขั้นตอนการรวบรวมข้อมูล การวิเคราะห์ข้อมูล และการตีความข้อมูลไปที่การสร้างทฤษฎีใหม่และสรุปผลตามทฤษฎีที่มีอยู่แล้ว ขณะที่นักวิจัยเชิงปริมาณส่วนใหญ่จะเน้นขั้นตอนการตีความไปที่การทดสอบสมมติฐาน (test hypothesis) ที่สร้างไว้อยู่ก่อนแล้ว ส่วนนักวิจัยเชิงคุณภาพจะเน้นขั้นตอนการตีความไปที่การตีความเพื่อการสร้างแนวคิดใหม่และการสร้างทฤษฎี (constructing theoretical) สำหรับขั้นตอนสุดท้าย คือ การนำเสนอการวิจัย การวิจัยทั้งสองประเภทมีการนำเสนอการวิจัยเหมือนกัน แต่อาจแตกต่างกันในรูปแบบของรายงานตามแนวทางของการวิจัยแต่ละประเภท (Neuman, 2006, p. 14-15)

การวิจัยแต่ละวิธีมีทั้งจุดอ่อนและจุดแข็ง บางสถานการณ์อาจต้องเลือกใช้การวิจัยเชิงผสมผสาน (mixed methods research) ทั้งการวิจัยเชิงปริมาณและการวิจัยเชิงคุณภาพ ดังนี้

1. ต้องการความเข้าใจที่สมบูรณ์ (More complete understanding) เพราะการใช้เครื่องมือวัดแนวคิดด้วยวิธีการเชิงปริมาณ ค่าคะแนนที่ได้จากการวัดอาจไม่สามารถบอกเรื่องราวต่างๆได้อย่างสมบูรณ์ ดังนั้นจึงต้องสอบถามหรือสนทนากับประชาชนที่เกี่ยวข้องจำนวนหนึ่งอาจทำให้เกิดความเข้าใจได้ดีขึ้น

2. ต้องการตรวจสอบบริบทตามความแตกต่างของเครื่องมือ (Context specific instruments) การสร้างเครื่องมือเพื่อศึกษาแนวคิด นำไปใช้กับกลุ่มตัวอย่าง/ประชากรที่แตกต่างกัน การทำความเข้าใจกลุ่มตัวอย่าง/ประชากรก่อนการพัฒนาเครื่องมือ ทำให้การพัฒนาเครื่องมือเป็นไปด้วยดีและเหมาะสมมากกว่า

3. ต้องการอธิบายผล (Explain results) การใช้เครื่องมือรวบรวมข้อมูลเกี่ยวกับตัวแปรปัจจัยเพื่อการทำนายแนวคิดนั้น เป็นการให้ความสำคัญกับตัวทำนาย สามารถอธิบายว่าอะไร (what) และทำไม (why) เป็นผลทำให้เกิดปรากฏการณ์ แต่การวิจัยแบบผสมผสม สามารถบอกได้ว่า ปรากฏการณ์ที่ทำการศึกษามีกลไกในการเกิดอย่างไร (how)

4. ต้องการตรวจสอบการกำหนดปัจจัยแทรก (Determine if an intervention will work) การค้นหาปัจจัยแทรกที่นักวิจัยผู้อื่นพัฒนาไว้ แต่ไม่สามารถค้นหาคำตอบได้จากตัวอย่างที่ทำการศึกษา ดังนั้นจึงต้องใช้วิธีการแบบผสมผสานในการตรวจสอบก่อนที่จะกำหนดตัวแปรแทรก

5. ต้องการเชื่อมโยงขั้นตอนต่างๆ ในกระบวนการประเมินผล (Tie together several steps in an evaluation process) เช่น การประเมินประสิทธิภาพขององค์กร เพื่อทำความเข้าใจผลลัพธ์ขององค์กร การใช้เครื่องมือวัดประสิทธิภาพให้คำตอบเฉพาะผลลัพธ์ แต่การวิจัยแบบผสมผสาน จะช่วยอธิบายว่า ทำไม (why) ผลลัพธ์จึงเกิดขึ้น

3.3 การวิจัยเชิงผสมผสาน

การวิจัยเชิงผสมผสาน (mixed method research) เป็นทั้งวิธีการ (method) และระเบียบวิธี (methodology) เพื่อดำเนินการวิจัย (conducting research) เกี่ยวกับการรวบรวม (collecting) การวิเคราะห์ (analyzing) และการบูรณาการ (integrating) ด้วยวิธีการวิจัยเชิงปริมาณ (quantitative research) และการวิจัยเชิงคุณภาพ (qualitative research) ในการค้นหาคำตอบ (inquiry) อาจเป็นการศึกษาเรื่องเดียว (single study) หรือ โครงการต่อเนื่องระยะยาว (longitudinal) โดยมีจุดมุ่งหมายคือ ทำให้เกิดความเข้าใจ (understanding) ปัญหาหรือประเด็นการวิจัยที่ดีกว่าการใช้วิธีการใดวิธีการหนึ่งขึ้นเพียงวิธีการเดียว แบ่งได้เป็น 3 ประเภท ดังนี้

1. วิธีการเปรียบเทียบ (Triangulation mixed methods design) นำข้อมูลและผลการวิจัยที่ได้จากวิธีการวิจัยเชิงปริมาณและเชิงคุณภาพมาทำการเปรียบเทียบและตีความ (interpretation)

Quantitative Research (data and results)	-> Interpretation <-	Qualitative Research (data and results)

2. วิธีการอธิบาย (Explanatory mixed methods design) เริ่มทำการวิจัยด้วยวิธีการเชิงปริมาณ และทำการติดตามหรือตรวจสอบ (follow up) เพื่ออธิบายด้วยวิธีการเชิงคุณภาพ

Quantitative Research (data and results)	-> Follow Up	Qualitative Research (data and results)

3. วิธีการสำรวจเบื้องต้น (Exploratory mixed methods design) ใช้วิธีการเชิงคุณภาพสำรวจข้อมูลเบื้องต้นก่อนก่อร่างหรือสร้าง (building) การวิจัยด้วยวิธีการเชิงปริมาณ

Qualitative Research (data and results)	-> Building	Quantitative Research (data and results)

สรุป

การวิจัยทางสังคมศาสตร์ เป็นวิธีการหนึ่งในการหาความรู้ของมนุษย์ที่ได้รับอิทธิพลมาจากวิทยาศาสตร์ธรรมชาติและตามแนวทางที่พัฒนาขึ้นมาเองจากนักคิดด้านสังคมศาสตร์ โดยมีการออกแบบอย่างเป็นขั้นตอนและเป็นระบบตามหลักวิชาการ (methodology) ใช้วิธีเชิงประจักษ์ทั้งทางตรงและทางอ้อมในการเก็บข้อมูลและสารสนเทศ (information) ทั้งเชิงปริมาณ และ/หรือเชิงคุณภาพ นำไปวิเคราะห์ อธิบาย ทำนาย สร้างข้อสรุป และสนับสนุนทฤษฎี รวมถึงตอบคำถามที่เกิดจากข้อสงสัยเกี่ยวกับสังคมมนุษย์ในทุกด้าน (social phenomena) และนำไปประยุกต์ใช้ในการแก้ไขปัญหาที่เกี่ยวของในสังคม

แบบฝึกหัดท้ายบท

ให้ผู้เรียนค้นหางานวิจัยด้านสังคมศาสตร์จากฐานข้อมูลของห้องสมุดที่สนใจแล้วตอบคำถามดังนี้

1. งานวิจัยที่ค้นได้ เป็นงานวิจัยเชิงปริมาณ เชิงคุณภาพ หรือเชิงผสมผสาน
2. งานวิจัยดังกล่าว มีขั้นตอนการวิจัยกี่ขั้นและอะไรบ้าง
3. งานวิจัยดังกล่าว มีและไม่มีขั้นตอนการวิจัยตามที่บรรยายไว้หรือไม่

บทที่ 2
การวิจัยเชิงปริมาณ

ความมุ่งหมายของบทเรียน

เพื่อให้ผู้เรียนเข้าใจวิธีการวิจัยเชิงปริมาณที่เกิดจากความสัมพันธ์ระหว่างทฤษฎี สมมติฐาน และตัวแปร

เนื้อหา

1. ความรู้เบื้องต้นเกี่ยวกับการวิจัยเชิงปริมาณ
2. ทฤษฎี
3. สมมุติฐาน
4. ตัวแปร
5. ความสัมพันธ์ระหว่างสมมติฐานและตัวแปร

วิธีการสอนและกิจกรรมการเรียนการสอน

บรรยาย แสดงตัวอย่าง สืบค้นข้อมูลจากฐานข้อมูลห้องสมุด และทำแบบฝึกหัดท้ายบท

สื่อและอุปกรณ์การสอน

เอกสารคำสอน ภาพเลื่อน คอมพิวเตอร์ โปรแกรมสอนออนไลน์ (สอนแบบออนไลน์) และเครื่องฉาย (สอนในชั้นเรียน)

งานที่มอบหมาย

ทำแบบฝึกหัดท้ายบท

การวัดผลและประเมินผล

การเข้าชั้นเรียน สังเกตความสนใจจากการเรียน สอบถามความเข้าใจจากการบรรยาย และการทำแบบฝึกหัดท้ายบท

เนื้อหา

การทำวิจัยแต่ละวิธีการ (approach) มีกระบวนการที่แตกต่างกัน การทำวิจัยด้วยวิธีการนิรนัย (deductive) มีการตั้งสมมติฐานจากทฤษฎีและเก็บรวบรวมข้อมูล[3]มาทดสอบสมมติฐาน ส่วนการทำวิจัยด้วยวิธีการอุปนัย (inductive) ใช้การรวบรวมข้อมูลเพื่อค้นหากฎเกณฑ์หรือรูปแบบ (pattern) แล้วจึงสร้างเป็นข้อเสนอเชิงทฤษฎีชั่วคราว (tentative conclusion or proposition) สำหรับการค้นหาข้อมูลมาสนับสนุน

1. ความรู้เบื้องต้นเกี่ยวกับการวิจัยเชิงปริมาณ

การวิจัยด้านสังคมศาสตร์รวมถึงสาขาสังคมวิทยาจำแนกแบบกว้างๆ หรือตามวิธีการ (techniques) การรวบรวมข้อมูล (data collection) และการวิเคราะห์ข้อมูล (data analysis) ได้เป็น 2 ประเภทใหญ่ คือ การวิจัยเชิงปริมาณ (quantitative research) และการวิจัยเชิงคุณภาพ (qualitative research) การวิจัยทั้งสองประเภทต่างมีความเชื่อเกี่ยวกับความจริงที่เรียกว่า ภาววิทยา (ontology) และวิธีการค้นหาความรู้หรือการเข้าถึงความจริงที่เรียกว่า ญาณวิทยา (epistemology) จึงทำให้เกิดมุมมองที่แตกต่างกันในวิธีการรวบรวบข้อมูลและการวิเคราะห์ข้อมูล (Blaikie, 2010, p. 204) แต่การวิจัยทั้ง 2 วิธี มีเป้าหมายที่เหมือนกัน คือ การได้ความรู้หรือข้อสรุปผลการวิจัยที่มีความถูกต้องและมีความน่าเชื่อถือ

1.1 ความเป็นมาของการวิจัยเชิงปริมาณด้านสังคมศาสตร์

การวิจัยเชิงปริมาณเป็นกลยุทธ์หนึ่ง (a strategy) ในการหาความรู้ (inquiry) เกี่ยวกับโลกและธรรมชาติด้วยวิธีการที่เป็นวิทยาศาสตร์ (scientific method) ของกลุ่มนักวิจัยที่เชื่อ (belief) ในความเป็นจริง (true) หรือที่เรียกกันว่า นักปฏิฐานนิยม (positivist)[4] และนักปฏิฐานนิยมยุคหลัง (postpositivist)[5] การท้าทายแนวคิดนักปฏิฐานนิยมของนักปฏิฐานนิยมยุคหลังที่เป็นนักสังคมวิทยายุค

---

[3] ส่วนใหญ่ใช้ข้อมูลเชิงปริมาณ แต่ก็สามารถใช้ข้อมูลเชิงคุณภาพได้ในกรณีที่การทดสอบสมมุติฐานนั้นต้องการให้ข้อคิดเห็นหรือให้เหตุผลที่ได้มาจากหลักฐาน

[4] แนวคิดเก่าที่เน้นการหาความรู้ที่เป็นความจริงแท้ (absolute truth) คือ เป็นจริงตลอดเวลาและทุกสถานที่ และเป็นจริงเสมอไม่ว่าสถานการณ์จะเป็นอย่างไร

[5] แนวคิดใหม่ที่ไม่เชื่อว่า ความจริงแท้มีอยู่จริง

บุกเบิก เช่น ออกุสต์ กองต์ (Auguste Comte) อีมีล เดอร์ไคม์ (Emile Durkheim) ในศตวรรษที่ 19 ที่เห็นว่า แนวคิดเกี่ยวกับความจริงแท้ไม่สามารถนำมาใช้ในการศึกษาพฤติกรรม (behavior) และการกระทำ (actions) ของมนุษย์ได้ ทำให้มีการนำเอาวิธีการหาความรู้เชิงเหตุและผลบนฐานความคิดของนักปฏิฐานนิยมไปพัฒนาใหม่โดยใช้วิธีการกำหนดความสัมพันธ์เชิงสาเหตุ (deterministic) เพื่อพิสูจน์ (identify) สาเหตุ (causes) ที่มีอิทธิพล (influence) ทำให้เกิดผลลัพธ์ (outcomes) เช่น วิธีการทดลอง หรือใช้วิธีการลดทอน (reductionistic) โดยการนำเอาสิ่งที่เป็นความคิด (ideas) เช่น ทฤษฎี สมมติฐาน และคำถามการวิจัย มาลดทอนให้มีขนาดเล็กเป็นชุดความคิดหรือตัวแปรที่สามารถนำไปทดสอบได้ แต่ยังคงรักษาวิธีการเข้าถึงความรู้ที่เป็นจริงเชิงวัตถุวิสัย (objective reality) ด้วยวิธีการสังเกตุ (observation) และการวัด (measurement) จากความจริงที่มีอยู่ภายนอก (out there) ที่ดำรงอยู่ หลังจากนั้นเป็นต้นมาทำให้เกิดการพัฒนาการวัดเชิงตัวเลข (numeric measurement) ด้วยวิธีการสังเกตและศึกษาพฤติกรรมของปัจเจกบุคคลอย่างแพร่หลายในบรรดานักปฏิฐานนิยมยุคหลัง รวมถึงนำมาใช้ทดสอบ (tested) และตรวจสอบ (verified) กฎหรือทฤษฎีให้เกิดความเข้าใจเกี่ยวกับโลกให้ดีขึ้น ทำให้วิธีการวิจัยที่เริ่มจากทฤษฎีใดทฤษฎีหนึ่ง แล้วทำรวบรวมข้อมูลมาเพื่อสนับสนุน (supports) หรือปฏิเสธ (refutes) ตามวิธีการของนักปฏิฐานนิยมยุคหลังได้รับการยอมรับว่าเป็นวิธีการที่เป็นวิทยาศาสตร์ (Creswell, 2009, p. 6-7)

ตั้งแต่ช่วงปลายศตวรรษที่ 19 เป็นต้นมา การวิจัยเชิงทดลองด้านจิตวิทยาเริ่มนำเอาวิธีการวิจัยเชิงปริมาณที่วางอยู่บนฐานความคิดแบบปฏิฐานนิยมมาใช้เป็นเครื่องมือค้นหาความรู้กันอย่างแพร่หลาย (Creswell, 2014, p. 12) ข้อมูลในการวิจัยเชิงปริมาณด้านสังคมส่วนใหญ่เมื่อเก็บรวบรวมมาแล้วนิยมทำให้เป็นรหัส (coding) ในรูปแบบของตัวเลข (numerical) สำหรับนำไปใช้กับโปรแกรมคอมพิวเตอร์เพื่อวิเคราะห์ด้วยวิธีการทางสถิติ (statistical analysis) ด้วยการพรรณนา (describing) และการอธิบาย (explaining) ก่อนนำไปนำเสนอตามจุดมุ่งหมาย (Babbie, 2013, p. 413-414)

1.2 ประเภทของการวิจัยเชิงปริมาณ

การวิจัยเชิงปริมาณอาจแบ่งได้เป็นหลายประเภท แต่หากแบ่งตามวัตถุประสงค์ของการวิจัยสามารถแบ่งได้เป็น 2 วิธีการ (approaches) ดังนี้

1. การวิจัยพื้นฐาน (Basic research) คือ การวิจัยที่เน้นการทดสอบทฤษฎีเพื่อสนับสนุนหรือปฏิเสธทฤษฎี การวิจัยประเภทนี้อาจเรียกได้ว่าเป็นการวิจัยเชิงยืนยัน (confirmatory research)

2. การวิจัยเชิงประยุกต์ (Applied research) คือ การวิจัยที่เน้นการค้นหาความสัมพันธ์ระหว่างเหตุและผลของปรากฏการณ์ทางสังคมหรือเหตุการณ์ต่างๆที่เกิดขึ้นในสังคม การวิจัยประเภทนี้อาจเรียกได้ว่าเป็นการวิจัยเชิงสำรวจ (exploratory research)

การวิจัยเชิงปริมาณแม้ว่าจะมีวัตถุประสงค์ที่แตกต่างกัน แต่เป้าหมายที่สำคัญ คือ การทดสอบทฤษฎี (theories) ที่เป็นข้อสรุปที่เป็นนามธรรมจากความเป็นจริงมาทำให้เป็นรูปธรรม (operationalization) ด้วยการสร้างเป็นสมมติฐาน (hypothesis) ทำการนิยาม (identification) หรือพรรณนา (description) ตัวแปร (variable) สำหรับสร้างเป็นเครื่องมือ (instrument) วัดหรือรวบรวมข้อมูลที่เป็นตัวเลข (numbered data) จากกลุ่มตัวอย่าง (sample) หรือประชากร (population) ที่มีจำนวนมากหรือขนาดใหญ่ และนำไปวิเคราะห์ด้วยวิธีการทางสถิติ (statistical procedures) เพื่อการทดสอบหานัยสำคัญเชิงสถิติ (statistical significance) ของข้อมูล (Creswell, 2009, p. 4; Wagner & Okeke cited by Garner, M., Wagner, C., & Kawulich, B., 2009, p. 62; Williams, 2016, p. 88)

2. ทฤษฎี

คำว่า ทฤษฎี (theory) มาจากภาษากรีก คือ “theoria” ซึ่งหมายถึง “looking at” เป็นคำกว้างๆ และเป็นประโยคเชิงข้อความ คล้ายกับสมมติฐาน แต่ทฤษฎีมีลักษณะที่แตกต่างกันมาก บางทฤษฎีอาจอยู่ในรูปของทฤษฎีชาวบ้าน (folk theory) ซึ่งไม่ต่างไปจากความเชื่อในชีวิตประจำวัน (everyday beliefs) บางทฤษฎีอาจเป็นทฤษฎีระดับกว้าง (grand theory) อธิบายเค้าโครงหรือภาพรวม (explanatory schema) ของปรากฏการณ์ทางสังคม (social phenomena) และบางทฤษฎีอาจเป็นทฤษฎีระดับกลาง (middle range theory) ที่เป็นประโยคหรือข้อความ (statement) ที่สามารถทดสอบได้ (testable) และมีขอบเขตการอธิบายเฉพาะปรากฏการณ์ทางสังคมภายใต้บริบท (context) เฉพาะของสังคม-ประวัติ/วัฒนธรรม (socio-historical/cultural) (Williams, 2016, p. 218-224)

2.1 ทฤษฎี คือ อะไร

ทฤษฎี (theory) เป็นความคิด (notion) หรือคำกล่าว (statement) ที่สรุป (abstract) ให้เห็นถึงทิศทางของความสัมพันธ์ระหว่างแนวคิด (concept) เช่น เหตุการณ์หนึ่งเกิดจากอะไร มีกระบวนการ

อย่างไร อะไรเกิดขึ้นก่อน อะไรเกิดขึ้นหลัง อะไรเป็นสาเหตุที่ทำให้เกิดและไม่ทำให้เกิด บางทฤษฎีได้รับการยอมรับอย่างเป็นทางการ (formal) แต่บางทฤษฎีอาจยังไม่ได้รับการยอมรับอย่างเป็นทางการ (informal) บางทฤษฎีมีความละเอียด (elaborate) และความซับซ้อน (complex) แต่ลักษณะทฤษฎีที่ดีต้องมีความเรียบง่าย (simplicity) และความชัดเจน (clarity) (Bouma and Ling, 2004, p. 19)

ทฤษฎีอาจเป็นแนวคิด (a concept) ชื่อของปรากฏการณ์ (a naming of phenomena) ประเภทของความคิด (an ideal type) การพรรณนารายละเอียดของการกระทำ (a thick description of an action) การบรรยายในลักษณะของตัวแบบ (a narrative as well as a model) ทฤษฎีมีลักษณะการอธิบายข้อเท็จจริงหรือจารนัย (abductive) กระบวนการ (process) มากกว่าการชี้แนะ (suggested) และมีความสำคัญมากในการทำวิจัย เพราะเป็นเครื่องมือหรือเลนส์ (lens) ที่ช่วยให้นักวิจัยมองปรากฏการณ์ (phenomena) แม้ว่าจะไม่ให้ภาพที่สมบูรณ์แต่ช่วยให้นักวิจัยตัดสินใจว่าจะมองที่อะไร รวมถึงสร้างคำถามการวิจัย ออกแบบการวิจัย สร้างกรอบแนวคิด การวิจัยแบบนิรนัย (deduction approach) ทฤษฎีจะถูกใช้เป็นแนวทางในการสร้างกรอบแนวคิด (conceptual framework) ของการวิจัย เพื่อนำไปสู่การสร้างสมมติฐาน (hypothesis) และการทดสอบ (tested) นอกจากนั้นทฤษฎียังช่วยเป็นกรอบของการอ้างอิงในการวิเคราะห์ ช่วยให้นักวิจัยเข้าใจว่าเกิดอะไรขึ้น (what happened) และเกิดผลที่ตามมาจากสิ่งที่เกิดขึ้น (consequences for what happened) อธิบายว่าอะไรคือสาเหตุของสิ่งที่เกิดขึ้น กรอบแนวคิดเชิงทฤษฎี (theoretical frameworks) เป็นสะพานเชื่อมระหว่างความเชื่อเกี่ยวกับญาณวิทยา (epistemology) และภาวะวิทยา (ontology) และการวิเคราะห์ (analysis) และการตีความของข้อมูล (interpretation of data) (Hammond and Wellington, 2013, p.141-144)

ทฤษฎีเป็นการรวมความคิดหลายๆความคิดเข้าด้วยกัน (a collection of ideas) หรือคำบรรยายสรุปความเป็นจริง (reality) ที่ได้มาจากการสังเกตจากโลกเชิงประจักษ์ (empirical world) หรือเป็นการอธิบายอย่างเป็นระบบ (systematic explanation) เกี่ยวกับกฎเกณฑ์ (regularities) สำหรับนำไปใช้ในการสังเกต (observation) (Bryman, 2012, p. 21; Aneshensel, 2013, p. 20, 32 Babbie, 2013, p. 8)

ทฤษฎีเป็นข้อเสนอหรือประพจน์ (propositions) หรือสมมติฐาน (hypotheses) ที่เป็นชุดความเกี่ยวพันระหว่างกัน (interrelated) ของมโนทัศน์หรือสิ่งที่เป็นนามธรรม (constructs) หรือตัวแปร (variables) แสดงถึงความสัมพันธ์ระหว่างตัวแปรในรูปของขนาด (magnitude) หรือทิศทาง (direction)

หรืออาจปรากฏอยู่งานวิจัยในรูปของข้อโต้แย้ง (argument) การอภิปราย (discussion) เหตุผล (rationale) ที่ใช้ในการอธิบาย (explain) หรือทำนาย (prediction) ปรากฏการณ์ที่เกิดขึ้นในโลก (Creswell, 2009, p. 51)

2.2 ประเภทของทฤษฎี

นักปรัชญาและนักวิชาการมีการสร้างทฤษฎีสำหรับใช้ในการอธิบายและศึกษาสังคมไว้จำนวนมาก การนำเอาทฤษฎีมาใช้ในการค้นหาความรู้หรือทำวิจัยวิจัยขึ้นอยู่กับความสนใจของนักวิจัยและปรากฏการณ์ทางสังคมที่ศึกษา บางเรื่องอาจที่มีขอบเขตกว้างและมีขนาดใหญ่ระดับสังคม เช่น การต่อสู้ระหว่างชนชั้นของสังคม ความสัมพันธ์ระหว่างประเทศ ความสัมพันธ์ระหว่างสถาบันหลักทางสังคม (รัฐบาล ศาสนา ครอบครัว) บางเรื่องเป็นวิถีชีวิตของคนหรือกลุ่มคนในสังคม เช่น พฤติกรรมการหาคู่ การพิจารณาคดีของผู้พิพากษา ความสัมพันธ์เพื่อนร่วมงาน แต่บางเรื่องอาจเป็นทั้งระดับมหภาคและจุลภาค เช่น ความขัดแย้ง (conflict) โดยทั่วไปมีการจำแนกทฤษฎีที่ใช้ในการศึกษาสังคมออกเป็น 3 ระดับ ดังนี้ (Babbie, 2013, p. 33)

1. ทฤษฎีระดับมหภาค (Macro theory) เป็นทฤษฎีที่ใช้อธิบายลักษณะโดยรวมของสังคม (aggregate entities of society) หรือทั้งสังคม (whole societies) เช่น ทฤษฎีโครงสร้าง-หน้าที่ (structural-functional) ทฤษฎีเครือข่ายทางสังคม (social network)

2. ทฤษฎีระดับกลาง (Meso theory) มีนักวิจัยบางกลุ่มนิยมเอาทฤษฎีระดับมหภาคมาประยุกต์ใช้ระหว่างระดับมหภาคกับระดับจุลภาค เพื่อใช้ศึกษาองค์การ (organization) ชุมชน (community) และประเภทของสังคม (social category)

3. ทฤษฎีระดับจุลภาค (Micro theory) เป็นทฤษฎีที่ใช้อธิบายวิถีชีวิตของคนในสังคม (social life) ที่มีความใกล้ชิดและมีปฏิสัมพันธ์กันในระดับบุคคล (individual) หรือกลุ่มขนาดเล็ก (small group) เช่น ทฤษฎีการกระทำระหว่างกันเชิงสัญลักษณ์ (symbolic interactionism)

2.3 การเลือกใช้ทฤษฎีในการวิจัย

การเลือกใช้ทฤษฎีในการอธิบายปรากฏการณ์ทางสังคมในแต่ละสาขาวิชามีความเชื่อแตกต่างกัน นักสังคมวิทยาส่วนใหญ่เชื่อว่า การศึกษาพฤติกรรม (behaviors) และการตัดสินใจ (decisions) ของบุคคล (individuals) ไม่สามารถนำไปใช้อธิบายปรากฏการณ์ทางสังคมที่มีความซับซ้อนได้อย่างสมบูรณ์

เพราะปรากฏการณ์ทางสังคม (social phenomena) เกิดมาจากปฏิสัมพันธ์ของประชาชน (social interaction) ในองค์การ (organizations) สถาบัน (institutions) และสังคม (societies) ซึ่งตรงข้ามกับนักจิตวิทยาและนักเศรษฐศาสตร์ส่วนใหญ่ที่มีแนวโน้มจะยอมรับว่าปรากฏการณ์ทางสังคมนั้นสามารถลดทอน (reducible) ไปสู่การอธิบายการกระทำ (actions) ของบุคคล เพราะบุคคลมีทางเลือกที่เป็นอิสระ (independent choices) ดังนั้นจึงสามารถเอามารวมกันเพื่ออธิบายกระบวนการทางสังคม (social processes) ที่ซับซ้อนได้ (Sawyer, 2005)

การนำเอาทฤษฎีไปใช้เป็นกรอบในการมองหรืออธิบายปรากฏการณ์ทางสังคมต้องเลือกใช้ทฤษฎีให้เหมาะกับประเด็นที่ต้องการศึกษา บางเรื่องอาจใช้ทฤษฎีระดับใดระดับหนึ่งเท่านั้น แต่บางเรื่องอาจนำเอาทฤษฎีในระดับที่แตกต่างกันมาอธิบายร่วมกัน (link) เพื่อให้เห็นความเกี่ยวข้องระหว่างกัน

แม้ว่าแนวคิดการศึกษาและอธิบายปรากฏการณ์ทางสังคมด้วยการเชื่อมโยงระหว่างแนวคิดที่มีระดับแตกต่างกันของอีมีล เดอร์ไคม์ (Emile Durkheim) มีนักสังคมศาสตร์จำนวนไม่น้อยที่ไม่เห็นด้วย แต่ก็มีนักวิชาการจำนวนไม่น้อยที่นำเอาข้อเสนอของเดอร์ไคม์ไปใช้ในการอธิบายข้อเท็จจริงทางสังคม (social facts) อีกทั้งยังเป็นตัวอย่างหนึ่งที่สามารถนำมาอธิบายเพื่อให้เข้าใจวิธีใช้ทฤษฎีที่มีระดับแตกต่างกันเชื่อมโยงความสัมพันธ์ระหว่างกลไกระดับจุลภาค (microlevel mechanism) กับกลไกระดับมหภาค (macrolevel mechanism) โดยเดอร์ไคม์ได้เสนอไว้ว่า ความแตกต่างของอัตราการฆ่าตัวตาย (suicide) ของคนในสังคมเกิดมาจากระดับความเป็นปึกแผ่นทางสังคม (social solidarity) ในระดับที่สูงกว่า พร้อมกับยกตัวอย่างให้เห็นว่า ประเทศที่ประชาชนส่วนใหญ่เป็นคาทอลิกมีอัตราการฆ่าตัวตายน้อยกว่าประเทศที่ประชาชนส่วนใหญ่เป็นโปรเตสแตนต์ เพราะสังคมคาทอลิกมีความเป็นปึกแผ่นทางสังคมสูงจึงช่วยป้องกันไม่ให้ประชาชนฆ่าตัวตายและทำให้มีอัตราการฆ่าตัวน้อยกว่า (Babones, 2014, p. 7-8) ปรากฏการณ์ทางสังคมในโลกเสมือนจริง (virtual world) ที่อุบัติ/ผุดบังเกิด (emergence) อย่างหลากหลายและรวดเร็ว อาจต้องเชื่อมโยงทั้งทฤษฎี/แนวคิดในระดับจุลภาคกับมหภาค (micro-macro link) เช่น การปฏิสัมพันธ์ทางสังคม กลุ่มสังคม เครือข่ายสังคม สื่อสังคมออนไลน์ ในการทำความเข้าใจและอธิบายสิ่งที่เกิดขึ้น

ทฤษฎีและสมมติฐานมีความสัมพันธ์กันอย่างมาก เราสามารถนำเอาทฤษฎีที่เป็นนามธรรมและเป็นคำกล่าวกว้างๆ ไปทดสอบได้ (testable) โดยการลดทอน (deductive) ให้มีลักษณะเฉพาะหรือแคบลงให้เป็นประโยคในเชิงสมมติฐาน (Williams, 2016, p. 88)

3. สมมติฐาน

มีการกล่าวกันว่า สมมติฐาน ไม่มีความจำเป็นหรือไม่เหมาะกับการทำวิจัยด้วยวิธีการอุปนัย (inductive) ที่เน้นการสังเกตและทำความเข้าใจปรากฏการณ์ทางสังคมแล้วสร้างเป็นทฤษฎี การวิจัยแบบจารนัย (abductive) อาจใช้สมมติฐานเป็นแนวทางในการสร้างทฤษฎี (generating theory) แต่แตกต่างจากวิธีการนิรนัย (deductive) คือ เป็นการสร้างสมมติฐานที่เกิดขึ้นในระหว่างทำการวิจัยและถูกนำมาใช้เป็นขั้นตอนย่อย (subsequent stage) ของการวิจัย ส่วนการวิจัยแบบกลนัย (retroductive) ไม่มีการสร้างสมมุติฐาน หน้าที่ของนักวิจัยคือ การออกแบบ (postulate) โครงสร้าง (structure) หรือกลไก (mechanism) ที่มีอยู่ และดำเนินการในรูปของการเสนอแนะ (Blaikie, 2010, p. 10,147)

สมมติฐานโดยทั่วไปสร้างมาจากทฤษฎีเพื่อนำไปใช้ในการรวบรวมข้อมูลและตรวจสอบ (verify) หรืออาจสร้างมาจากผลการวิจัยในอดีต การทบทวนวรรณกรรมที่เกี่ยวข้อง และรวมถึงอาจได้มาจากการคิดขึ้นเอง (thin air) (Flynn, Sakakibara, Schroeder, Bates & Fynn, 1990, p. p. 253)

สมมติฐานที่ได้มาจากทฤษฎีหรือแนวคิดต้องนำมาประยุกต์ใช้ให้เหมาะสมกับสถานการณ์ อาจใช้เป็นคำตอบ (answer) ที่เป็นไปได้ในการตอบคำถาม (question) ทำไม (why) และอย่างไร (how)[6] และอาจไม่จำเป็นต้องมีความสัมพันธ์เชิงเหตุและผล (causal) แต่ต้องมีความชัดเจน มีความเข้าใจง่าย ไม่มีความคลุมเครือ และสามารถทำไปทดสอบได้อย่างเป็นขั้นเป็นตอน (Black, 1999, p. 45, 53; Babbie, 2008, p. 48)

3.1 สมมติฐาน คือ อะไร

สมมติฐาน (Hypothesis) คือ ประโยคหรือข้อความที่เป็นคำตอบล่วงหน้าของปัญหาวิจัยที่ระบุถึงความสัมพันธ์ระหว่างตัวแปรตั้งแต่สองตัวแปรขึ้นไปที่ได้มาจากทฤษฎี การคาดคะเนจากความรู้ และประสบการณ์ของนักวิจัย เมื่อกำหนดเป็นสมมติฐานเชิงวิจัยแล้ว ในการทดสอบสมมติฐานด้วยสถิติจะมีการแปลงสมมติฐานเชิงวิจัยไปเป็นสมมติฐานเชิงสถิติก่อน

สมมติฐานไม่เพียงแต่จะเป็นตัวกำหนดตัวแปร แต่มีบทบาทสำคัญในการนำไปสู่ความเกี่ยวข้องกันของสิ่งที่ทำการศึกษา แม้ว่าตัวแปรนั้นไม่ได้อยู่ในความสนใจที่ต้องกำหนดและนิยามความหมาย

[6] คำถามการวิจัยประเภท อะไร (what) ไม่ต้องตั้งสมมติฐาน

เพราะบางประโยคแสดงให้เห็นหรือหมายถึงความสัมพันธ์เชิงเหตุและผลอย่างมีเหตุผลพร้อมกับลำดับเวลาของเหตุการณ์ และสะท้อนถึงความเป็นไปได้ของปรากฏการณ์ที่เกิดขึ้นอันเป็นผลมาจากหลายสาเหตุ หรือมีความต่อเนื่องของสาเหตุของการเกิดเหตุการณ์ (Black, 1999, p. 34-35)

สมมติฐานจะเขียนในลักษณะประโยคบอกเล่าทั่วไป (basic statement) เพื่ออธิบายข้อเท็จจริงที่ยังคลุมเครือและกำหนดทิศทางในการทำวิจัยให้เกิดความชัดเจน (Fisher, 1949, p. p. 13 cited by Cohen, 1977, p. 1) โดยแสดงการคาดคะเนของผลลัพธ์ (outcome) ที่เกิดจากความสัมพันธ์ระหว่างตัวแปร 2 ตัวแปรหรือมากกว่า (Babbie, 2008, p. 48) หรือเป็นความสัมพันธ์ระหว่างสองแนวคิด (concept) ประโยคหนึ่งเป็นเหตุ (cause) และอีกประโยคหนึ่งเป็นผล (outcome) ทั้งเหตุและผล คือ ตัวแปร ที่ใช้เป็นแนวทางในการรวบรวมข้อมูลของนักวิจัยและนำไปใช้ในการทดสอบสมมติฐาน (Blaikie, 2010, p. 10; Field, 2013, p. 7)

3.2 ประเภทของสมมติฐาน

สมมติฐานมีลักษณะเหมือนกับทฤษฎีมาก แต่ถูกนำมาใช้ในการวิจัยด้านสังคมใน 2 แนว คือ เป็นประโยคเชิงข้อความ (propositional statement) อาจเป็นประโยคบอกเล่าหรือประโยคปฏิเสธที่เป็นจริงหรือเป็นเท็จอย่างใดอย่างหนึ่ง และเป็นประโยคเชิงตรรกะ (logical form) เพื่อใช้ในการทดสอบนัยยะสำคัญเชิงสถิติ (statistical significance) ของข้อมูล (Williams, 2016, p. 88) ดังนั้นจึงมีการเขียนสมมติฐานในงานวิจัยแต่ละเรื่องหรือสาขาวิชาให้เห็นใน 2 รูปแบบ ดังนี้

1. สมมติฐานเชิงวิจัย (Research hypothesis) คือ คำตอบที่ผู้วิจัยเขียนไว้ในตอนแรกเริ่มการวิจัย เป็นการคาดเดาผลไว้ล่วงหน้าว่าจะเป็นไปตามนั้น เป็นลักษณะของการใช้คำบรรยายหรือพรรณนาลักษณะความสัมพันธ์ของตัวแปร มักอยู่ในลักษณะการคาดคะเนผลว่าไม่แตกต่างกัน แตกต่างกันมากกว่า หรือน้อยกว่า เช่น ชาวต่างชาติที่ทำงานอยู่ในนิคมอุตสาหกรรมแหลมฉบังและนิคมอุตสาหกรรมมาบตาพุดมีการดำเนินชีวิตประจำวันไม่แตกต่างกัน ประชาชนในเขตชนบทมีระดับการศึกษาแตกต่างกับประชาชนในเขตเมือง ประชาชนในเขตชนบทมีรายได้น้อยกว่าประชาชนในเขตเมือง ครัวเรือนในเขตชนบทมีจำนวนสมาชิกมากกว่าครัวเรือนในเขตเมือง

2. สมมติฐานเชิงสถิติ (Statistical hypothesis) คือ สมมติฐานทางวิจัยที่ถูกแทนที่ด้วยการใช้สัญลักษณ์ในทางคณิตศาสตร์ หรือสถิติที่บ่งบอกความสัมพันธ์ของตัวแปร ในรูปของเครื่องหมายแทน

คุณลักษณะของตัวแปรนั้นๆ เป็นสมมติฐานที่ใช้ในการวิเคราะห์ข้อมูลเพื่อการทดสอบสมมติฐานการวิจัย เช่น $\mu = \mu_0$ หรือ $\mu \neq \mu_0$ หรือ $\mu > \mu_0$ หรือ $\mu < \mu_0$

ในการวิจัยเชิงปริมาณที่มีการทดสอบสมมติฐาน สมมติฐานทางวิจัยจะถูกแปลงให้เป็นสมมติฐานทางสถิติใน 2 ลักษณะ ดังนี้

1. สมมติฐานศูนย์หรือสมมติฐานหลักหรือสมมติฐานที่เป็นกลางหรือสมมติฐานไร้นัยสำคัญ (Null hypothesis: H_o หรือ H_0) หมายถึง สมมติฐานที่มิได้บ่งชี้ความแตกต่างหรือความสัมพันธ์ของตัวแปรหรือเป็นประโยคที่กล่าวถึง ความไม่แตกต่างของประชากร (no difference in the population) เช่น ชาวต่างชาติที่ทำงานอยู่ในนิคมอุตสาหกรรมแหลมฉบังและนิคมอุตสาหกรรมมาบตาพุดมีการดำเนินชีวิตประจำวันไม่แตกต่างกัน

2. สมมติฐานทางเลือกหรือสมมติฐานรอง หรือสมมติฐานที่เหลือ (Alternative hypothesis: H_a หรือ H_1) หมายถึง สมมติฐานที่ระบุความสัมพันธ์ของตัวแปร ทั้งในลักษณะกำหนดทิศทางบวก หรือลบ และไม่กำหนดทิศทาง หรือเป็นประโยคที่กล่าวถึง ความแตกต่างของประชากร (difference) เช่น ชาวต่างชาติที่ทำงานอยู่ในนิคมอุตสาหกรรมแหลมฉบังและนิคมอุตสาหกรรมมาบตาพุดมีการดำเนินชีวิตประจำวันแตกต่างกัน

4. ตัวแปร

ตัวแปร (variables) เป็นแนวคิดที่เป็นนามธรรมสามารถสังเกต (observable) และวัด (measurable) ได้ด้วยการทำให้แนวคิดสามารถวัดได้ด้วยตัวแปร (Bouma and Ling, 2004, p. 41)

ตัวแปรได้มาจากหลายแหล่งทั้งแนวคิด ทฤษฎี ปัญหาการวิจัย สมมติฐาน การทบทวนวรรณกรรม งานวิจัยที่เกี่ยวข้อง หน่วยวิเคราะห์ (ใคร/อะไร) และการสร้างของนักวิจัย (นิยามตัวแปร/การวิเคราะห์) มีตำราทั้งด้านการวิจัยและการวิเคราะห์ข้อมูลให้คำนิยามหรือความหมายเกี่ยวกับตัวแปรไว้หลากหลาย จึงนำเอามานำเสนอเป็นตัวอย่างพอสังเขปดังต่อไปนี้

4.1 ตัวแปร คือ อะไร

ตัวแปร คือ คุณลักษณะ (a characteristic) ของสิ่งต่างๆ (objects) ในการวัด (measurement) คุณลักษณะของสิ่งต่างๆนั้นสามารถที่จะสมมติ (assume) ให้เป็นสิ่งใดก็ได้ (any one) ตามที่กำหนดให้

เป็นผลลัพธ์ (outcomes) ไว้ตั้งแต่ต้น เช่น เพศ ส่วนสูง น้ำหนัก เชื้อชาติ สัญชาติ และรายได้ของบุคคล ผลลัพธ์ที่ได้จากการวัดคุณลักษณะจากสิ่งประเภทเดียวกันหรือคนกลุ่มเดียวกันจะมีค่าที่แตกต่างกันในแต่ละสิ่งหรือแต่ละคน (Traub, 1994, p. 5)

ตัวแปร คือ แนวคิดประเภทหนึ่ง (a type of concept) ซึ่งมีความแตกต่างกัน (vary) ในเรื่องของจำนวน (amount) หรือคุณภาพ (quality) ซึ่งอาจมากหรือน้อย หรืออาจมีความแตกต่างกันเป็นประเภท (categories) ตัวแปรที่สนใจไม่ใช่เรื่องของความแตกต่างของจำนวนหรือประเภท แต่เป็นเรื่องของความสามารถในการวัดได้ (measurable) (Bouma and Ling, 2004, p. 19)

ตัวแปร คือ ทุกสิ่งที่สามารถทำการวัดได้ (measured) และมีลักษณะหรือคุณสมบัติที่แตกต่างกัน (across entities) หรืออาจเป็นสิ่งเดียวกันแต่อยู่ในช่วงเวลาที่แตกต่างกัน (across time) (Field, 2013)

อาจกล่าวได้ว่า ตัวแปร คือ สัญลักษณ์ (symbol) ที่นักวิจัยสร้างมาจากแนวคิดหรือทฤษฎีเพื่อแทนสรรพสิ่งที่ต้องการศึกษา (objects) และสามารถนำไปนับ คำนวณ และสรุปเป็นปริมาณด้วยตัวเลข กราฟ แผนภูมิ แผนภาพ และคำอธิบาย

การวัดอาจทำได้โดยตรงจากประสาทสัมผัสของนักวิจัย เช่น ตา หู จมูก หรือใช้สิ่งประดิษฐ์เป็นเครื่องมือช่วยวัดจากสิ่งที่อยู่นอกความสามารถการเข้าถึงด้วยประสาทสัมผัสของมนุษย์ เช่น ดาวเทียม กล้องจุลทรรศ์ ยานสำรวจ และอาจต้องสร้างตัวแปรเทียมโดยอาศัยวัดหลายๆดัชนีที่สามารถสังเกตได้ เช่น โครงสร้างของสังคม สถาบันทางสังคม ความศรัทธา ความเชื่อ

4.2 ประเภทตัวแปร

ตัวแปรไม่เพียงแต่จำแนกสรรพสิ่งที่วัดให้มีแตกต่างกันเท่านั้น ในการทดสอบหรือหาความสัมพันธ์ระหว่างเหตุ (cause) และผล (outcome) ผู้วิจัยยังต้องกำหนดให้ตัวแปรมีหน้าที่ (functions) แตกต่างกันไปตามตรรกะและความสัมพันธ์ตามทฤษฎีและสมมติฐาน บางตำราแบ่งเป็น 3 ประเภท เช่น ตัวแปรอิสระ ตัวแปรตาม และตัวแปรคุม แต่บางตำราแบ่งเป็น 6 ประเภท เช่น ตัวแปรอิสระ ตัวแปรตาม ตัวแปรคุม ตัวแปรแทรกแซง ตัวแปรกำกับ และตัวแปรเกินหรือตัวแปรแทรกซ้อน

สำหรับในที่นี้ได้รวบรวมชื่อตัวแปรที่ปรากฏในตำราต่างๆ มาแบ่งตามลักษณะของการทำหน้าที่ความสัมพันธ์กับตัวแปรอื่นๆ และการสร้างตัวแปร โดยมีรายละเอียดดังต่อไปนี้

1. กลุ่มตัวแปรสาเหตุ (Cause Variables) คือ ตัวแปรที่ผู้วิจัยกำหนดให้เป็นต้นเหตุของปรากฏการณ์ ซึ่งมีหลายชื่อ แต่ชื่อที่นิยมใช้มากที่สุด คือ ตัวแปรอิสระ โดยมีรายละเอียดดังต่อไปนี้

1.1 ตัวแปรอิสระ (Independent variable) คือ ตัวแปรที่เป็นสาเหตุ (cause) ที่มีผลกระทบ (impact) หรือทำให้เกิดการเปลี่ยนแปลงต่อตัวแปรอื่น ตัวแปรอิสระนอกจากจะมีชื่อเรียกว่า ตัวแปรต้น เพราะเป็นต้นเหตุทำให้ตัวแปรอื่นมีเกิดการเปลี่ยนแปลง

1.2 ตัวแปรอธิบาย (Explanatory/Explaining variable) หรืออาจรู้จักในชื่อว่า ตัวแปรตัวทำนาย (predictor variable) คือ ตัวแปรอิสระหรือตัวแปรเหตุประเภทหนึ่งที่ใช้อธิบายการเปลี่ยนแปลงในตัวแปรตอบสนอง (respond variables) มักใช้ในการศึกษาเชิงความสัมพันธ์ (correlational studies) การวิเคราะห์เชิงทดถอย (regression) และการศึกษาแบบไม่ทดลอง (non-experimental study) เช่น ระดับการศึกษาเป็นตัวแปรอธิบายรายได้ที่แตกต่างกัน

1.3 ตัวแปรทดลอง (Treatment variable) เพราะผู้วิจัยสามารถกำหนดค่าให้กลุ่มตัวอย่างมีลักษณะที่แตกต่างกันเพื่อทดลองเปรียบเทียบผลลัพธ์ที่เกิดขึ้น เช่น การทดสอบความเร็วในการให้บริการด้วยระบบอัตโนมัติกับระบบเจ้าหน้าที่

1.4 ตัวแปรถูกจัดการ (Manipulated variable) หากผู้วิจัยจัดการหรือกำหนดให้มีค่าเพิ่มขึ้นหรือลดลง หรือควบคุมให้มีค่าคงที่ (constant) เพื่อทดลองการเกิดผลลัพธ์ในกระบวนการวิเคราะห์หรือทดลอง

2. กลุ่มตัวแปรผล (Result Variables) คือ ตัวแปรที่ผู้วิจัยกำหนดให้แสดงผลของปรากฏการณ์ มีหลายชื่อเช่นกัน แต่ชื่อที่นิยมใช้มากที่สุด คือ ตัวแปรตาม โดยมีรายละเอียดดังต่อไปนี้

2.1 ตัวแปรตาม (Dependent variable) คือ ตัวแปรที่มีค่าหรือผลลัพธ์เปลี่ยนแปลงไปตามตัวแปรอิสระหรือตัวแปรอื่นๆ อาจมีชื่อเรียกอื่น เช่น ตัวแปรที่ถูกอธิบาย (explained variables) ตัวแปรถูกทำนาย (predicted variables) รวมถึงมีชื่อเรียกที่แตกต่างกันไปอีกหลายชื่อดังต่อไปนี้

2.2 ตัวแปรผลลัพธ์ (Outcome variable คือ ตัวแปรที่แสดงค่าที่เป็นผล (results) หรือผลลัพธ์ (outcome) ที่ได้มาจากการศึกษา ส่วนใหญ่ใช้ในการศึกษาแบบไม่ทดลอง (non-experimental study)

2.3 ตัวแปรเกณฑ์ (Criterion variable) ส่วนใหญ่นิยมเรียกในการวิเคราะห์ตัวแบบเชิงสถิติ (statistical modeling) เช่น การวิเคราะห์ถดถอย (regression analysis) การวิเคราะห์ความแปรปรวน (analysis of variance) การวิเคราะห์จำแนกกลุ่ม (discriminant analysis) และการวิเคราะห์สหสัมพันธ์คาโนนิคอล (canonical correlation) ที่กำหนดให้ตัวแปรตามเป็นตัวแปรที่ถูกพยากรณ์ (being predicted) ผลลัพธ์ที่ได้ (outcome) จะขึ้นอยู่กับ (depend on) หลายสาเหตุ

2.4 ตัวแปรตอบสนอง (Response variable) เป็นชื่อที่นิยมเรียกกันในการวิจัยสาขาวิชาจิตวิทยาที่เป็นการทดลองเกี่ยวกับการแสดงออกของพฤติกรรมเมื่อได้รับการกระตุ้นโดยสิ่งเร้า

2.5 ตัวแปรภายใน (Endogenous variable) คือ ตัวแปรเชิงสาเหตุที่กำหนดขึ้นในระบบหรือตัวแบบเชิงสาเหตุ (a causal model) ที่ทำการศึกษา เช่น การวิเคราะห์การถดถอยเชิงเส้นตรง (linear regression) และมีค่าที่ถูกกำหนดโดยตัวแปรอื่นๆ จึงมีความผันแปรไปตามสาเหตุจากตัวแปรที่เป็นต้นเหตุของการเกิดตัวแปร (คล้ายกับตัวแปรตาม แต่ไม่เหมือนกันทั้งหมด)

3. กลุ่มตัวแปรบริบท (Context Variables) คือ ตัวแปรที่ผู้วิจัยกำหนดให้มีหรือไม่มีอิทธิพล (influence) ต่อความสัมพันธ์ระหว่างตัวแปรเหตุกับตัวแปรผล มีหลายชื่อเรียก โดยมีรายละเอียดดังต่อไปนี้

3.1 ตัวแปรคุม (Control variable) คือ ตัวแปรที่ผู้วิจัยไม่ต้องการให้อยู่หรือต้องกำจัดออกไปในการทดสอบความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตาม เพื่อแสดงให้เห็นว่าความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตามไม่ได้เกิดจากตัวแปรอื่นหรือตัวแปรที่นำมาใช้เป็นตัวแปรคุม ดังนั้นตัวแปรคุมจึงมีชื่อเรียกในชื่ออื่น เช่น ตัวแปรเกินหรือตัวแปรแทรกซ้อนหรือตัวแปรภายนอก (extraneous variable) หรือตัวแปรร่วม (covariate variable)

3.2 ตัวแปรแทรกแซง (Intervening variable) คือ ตัวแปรที่ไม่สามารถสังเกตได้โดยตรง (not direct observable) แต่มีความเกี่ยวโยง (link) กับตัวแปรอิสระและตัวแปรตาม อาจมีการใช้ชื่อเรียกอื่น เช่น ตัวแปรส่งผ่าน (mediating variable) หรือตัวแปรคั่นกลาง (intermediary variable) เช่น มีความสัมพันธ์ทางสถิติเชิงลบระหว่างความเป็นเมืองกับขนาดครอบครัว ความเป็นเมืองไม่ใช่ปัจจัยที่มีอิทธิพลต่อขนาดครอบครัวเพียงอย่างเดียว แต่ความเป็นเมืองมีแนวโน้มทำให้คนมีการศึกษาและความรู้ในการวางแผนครอบครัวที่ดี ทำให้มีบุตรจำนวนที่เหมาะสมกับสภาพเศรษฐกิจ ตัวแปรด้านการศึกษาและเศรษฐกิจ คือ ตัวแปรแทรกที่คั่นกลางความสัมพันธ์ระหว่างความเป็นเมืองกับขนาดครอบครัว หรือสถานภาพทางสังคมของพ่อแม่มีผลต่อทางตรงต่อสถานภาพทางสังคมของลูก แต่สถานภาพทางสังคมตัว

แปรที่มีผลทางอ้อมต่อสถานภาพทางสังคมของลูก คือ การศึกษาของลูก (Parent's status -> child's education -> child's status)

3.3 ตัวแปรคั่นกลาง (Mediator variable) คือ ตัวแปรที่มีอิทธิพล (influences) หรือคั่นกลางความสัมพันธ์ (relation) ระหว่าง 2 ตัวแปร และทำให้เกิดกระทบระหว่างกัน (interaction effect)

3.4 ตัวแปรกำกับ (Moderator variable) หรืออาจมีชื่อเรียกเป็นอย่างอื่น เช่น ตัวแปรปรับตัวแปรขยาย ตัวแปรกลาง คือ ตัวแปรที่มีปฏิสัมพันธ์ (interaction) กับตัวแปรอิสระ และมีผลทำให้เกิดการเปลี่ยนทิศทางของความสัมพันธ์ระหว่างตัวแปรอิสระกับตัวแปรตาม เช่น ทฤษฎีบูรณาการการยอมรับและใช้เทคโนโลยี (The Unified Theory of Acceptance and Use of Technology: UTAUT) เพศ อายุ และประสบการณ์ เป็นตัวแปรคั่นกลางและมีอิทธิพลระหว่างตัวแปรอิทธิพลทางสังคม (social influence) และความคาดหวัง (effort expectation) กับพฤติกรรมความตั้งใจ (behavioral intention) ตัวแปรกำกับอาจเป็นตัวแปรแบบค่าแบ่งกลุ่มหรือตัวแปรแบบค่าต่อเนื่องก็ได้ เช่น สถานะภาพทางเศรษฐกิจและสังคมมีความสัมพันธ์กับความพร้อมด้านเทคโนโลยีดิจิทัล ตัวแปรกำกับที่เป็นตัวแปรแบบค่าแบ่งกลุ่ม คือ ระดับการศึกษา หรือตัวแปรกำกับที่เป็นตัวแปรแบบค่าต่อเนื่อง คือ รายได้

3.5 ตัวแปรเกินหรือตัวแปรแทรกซ้อน (Extraneous variable) คือ ตัวแปรที่อยู่ในสภาพแวดล้อม (environment) ของการวิจัยแต่ไม่ได้ทำการควบคุม หรืออาจกล่าวได้ว่า ตัวแปรทั้งหมดที่ไม่ใช่ตัวแปรอิสระ และเป็นตัวแปรไม่ได้ตั้งใจนำมาศึกษาหรือวิเคราะห์ แต่อาจส่งผลหรือไม่มีผลต่อตัวแปรตามหรือผลลัพธ์ ตัวแปรอิสระกับตัวแปรตามอาจไม่ได้มีความสัมพันธ์เชิงสาเหตุระหว่างกัน แต่ความสัมพันธ์เกิดจากตัวแปรเกิน เช่น ศึกษาความสัมพันธ์ระหว่างสถานะภาพทางเศรษฐกิจและสังคมกับความพร้อมด้านเทคโนโลยีดิจิทัล แต่ไม่ได้นำเอาประเภทของสังคมมาศึกษาด้วย

3.6 ตัวแปรพัวพัน/รบกวน (Confounding variable) คือ ตัวแปรภายนอกในแบบจำลองทางสถิติ ที่ไม่ใช่ตัวแปรหลักที่ศึกษา (ตัวแปรอิสระและตัวแปรตาม) และมีอิทธิพลต่อตัวแปรอิสระและตัวแปรตาม ทำให้ความสัมพันธ์ระหว่างตัวแปรหลักไม่ชัดเจน (อาจสูงกว่าความเป็นจริง) หรือเป็นความสัมพันธ์ปลอม เช่น คนในชนบทมีและใช้เทคโนโลยีดิจิทัลน้อยกว่าคนในเมือง ปัจจัยภายนอกที่มีซ่อนอยู่และมีอิทธิพลต่อการมีและการใช้เทคโนโลยีดิจิทัลของคนในเมืองและชนบท คือ ความพร้อมด้านโครงสร้างพื้นฐานด้านเทคโนโลยีดิจิทัล

3.7 ตัวแปรภายนอก (Exogenous variable) คือ ตัวแปรปัจจัย (factor variables) ในระบบหรือตัวแบบเชิงสาเหตุ (a causal model) ที่ทำการศึกษา ค่าของตัวแปรเป็นอิสระจากตัวแปรที่อยู่

ภายในระบบหรือตัวแบบ แต่ถูกกำหนด (determined) จากปัจจัยหรือตัวแปรภายนอกระบบหรือตัวแบบที่ทำการศึกษา และเป็นปัจจัยที่ไม่นำเอามาศึกษา ในเชิงสถิติคือตัวแปรที่กำหนดไม่ให้มีความสัมพันธ์กัน (non-correlation) ระหว่างตัวแปรภายนอกกับตัวแปรอิสระที่อยู่ภายในระบบหรือตัวแบบ เช่น สถานภาพทางเศรษฐกิจและสังคมมีหลายสาเหตุเป็นตัวกำหนดสถานภาพทางเศรษฐกิจและสังคมแต่ไม่นำเอามาเป็นปัจจัยในการอธิบายด้วย เช่น ระดับความเป็นเมือง ชาติตระกูล มรดก

4. กลุ่มตัวแปรจำลอง (Pseudo Variables) คือ ตัวแปรที่ผู้วิจัยสร้างหรือกำหนดให้มีค่าจากดัชนีต่างๆ เพื่อทำหน้าที่เป็นตัวแปรเหตุ ตัวแปรผล หรือตัวแปรบริบท โดยมีรายละเอียดดังต่อไปนี้

4.1 ตัวแปรเสริม (Auxiliary variable) คือ ตัวแปรที่ช่วยให้ตัวแปรที่มีข้อมูลที่ไม่สมบูรณ์สามารถวิเคราะห์ได้ เช่น การวิเคราะห์ตัวแปรที่มีข้อมูลสูญหายต้องใช้วิธีการรวมตัวแปรเสริมเข้าช่วย หรืออาจสร้างตัวแปรพิเศษ (extra variables) เพื่อใส่ค่าเพิ่มทำให้การวิเคราะห์มีความแม่นยำและถูกต้องมากขึ้น

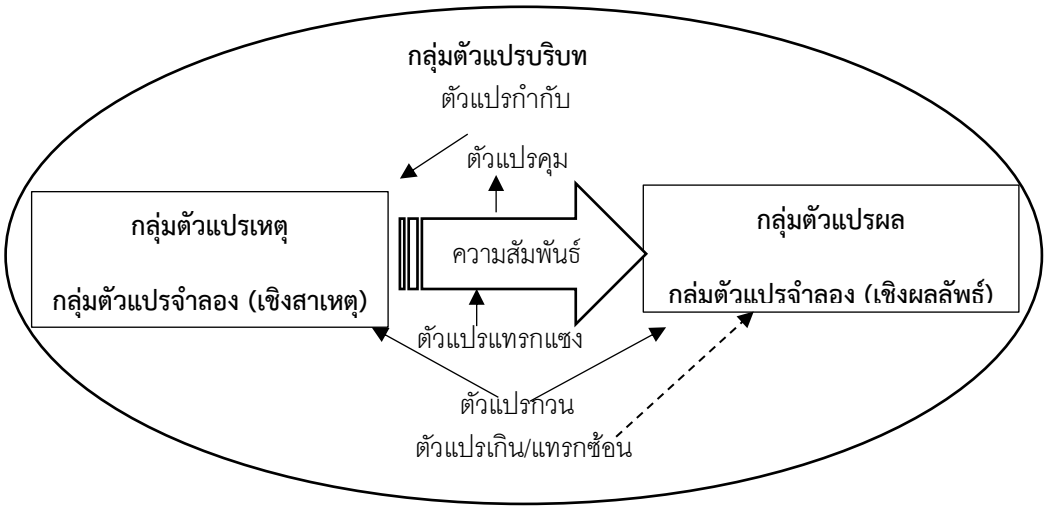
4.2 ตัวแปรเชิงโครงสร้าง (Construct variable) คือ ตัวแปรที่สร้างจากแนวคิดหรือทฤษฎี มีลักษณะเป็นตัวแปรนามธรรม (abstract variables) และไม่สามารถสังเกตได้โดยตรง (latent variables)

4.3 ตัวแปรหุ่น (Dummy variable) คือ ตัวแปรที่สร้างจากการแปลงค่า (recoding) ตัวแปรแบบค่าแบ่งกลุ่ม (categorial variables) ที่มีค่าแบ่งกลุ่มมากกว่า 2 กลุ่ม ให้มี 2 ค่า (binary data) เพื่อใช้ในการวิเคราะห์การถดถอย (regression analysis) เช่น ตัวแปรที่อยู่ (residence) มีค่าตัวแปร 1 = เมือง (urban) 2 = ชานเมือง (suburb) และ 3 = ชนบท (rural) สร้างเป็นตัวแปรหุ่นชื่อใหม่ คือ เขตบ้านพักอาศัย (home_area) มีค่าตัวแปร 2 ค่า คือ 1 = เมือง (urban) 0 = ไม่ใช่เมือง (non-urban) หรือ 1 = ชนบท (rural) 0 = ไม่ใช่ชนบท (non-rural) เพื่อใช้ในการวิเคราะหการถดถอย (regression analysis)

4.4 ตัวแปรสังเกตได้ (Manifest variable) คือ ตัวแปร (variable) หรือปัจจัย (factor) ที่สามารถวัด (measured) หรือสังเกตได้ (observed) โดยตรง และเป็นตัวแปรดัชนี (indicator variable) สำหรับสร้างตัวแปรแฝง (latent variable) เพื่อให้สามารถทดสอบหรือวัดได้ว่ามีอยู่หรือไม่ เช่น สถาบันครอบครัวไม่สามารถวัดหรือสังเกตุได้โดยตรงต้องวัดหรือสังเกตจากการแต่งงาน อำนาจการตัดสินใจของสมาชิกในครอบครัว ขนาดครอบครัว ประเภทครอบครัว สายเลือด

4.5 ตัวแปรแฝง (Latent variable) คือ ตัวแปร (variable) หรือปัจจัย (factor) ที่ไม่สามารถวัดหรือสังเกตได้โดยตรง ต้องสร้างหรืออ้างอิงจากตัวแปรหรือปัจจัยที่สามารถสังเกตได้ เช่น ความศรัทธาในศาสนาต้องสร้างหรือวัดจากตัวแปรที่สามารถสังเกตได้ เช่น การถือศีล การฟังธรรม การทำบุญ

ความหลากหลายของประเภทตัวแปรที่จำแนกตามการทำหน้าที่ของตัวแปรที่กล่าวมาเป็นผลมาจากการแบบการวิจัยและการวิเคราะห์ข้อมูล หากนำตัวแปรดังกล่าวมาจัดกลุ่มใหม่สามารถแสดงให้เห็นได้ดังรูปด้านล่าง



รูปที่ 2.1 ความเกี่ยวข้องระหว่างตัวแปรเหตุ ตัวแปรผล และตัวแปรบริบท

การวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรเหตุและตัวแปรผล เป็นความพยายามค้นหาต้นเหตุที่แท้จริง หากทราบพฤติกรรมหรือผลกระทบของตัวแปรบริบทแต่ละตัวแปร จะทำให้กำจัดปัจจัยที่ไม่เกี่ยวข้องออกได้ถูกต้อง และได้คำตอบที่เป็นจริง

4.3 ระดับการวัดของตัวแปร

ตัวแปรมีความแตกต่างกันอย่างซับซ้อนทั้งด้านรูปแบบ (forms) และระดับ (levels) ความสัมพันธ์ระหว่างการวัดและจำนวนที่แสดงได้จากการวัดถูกเรียกว่า ระดับการวัด (level of measurement) หรือมาตรการวัด (scale of measurement)

ระดับการวัดตัวแปรมีความเกี่ยวข้องกับเลือกใช้ประเภทของสถิติเพื่อการพรรณนา (describe) หรือการอ้างอิง (inference) ไปสู่ประชากร โดยทั่วไปนักวิจัยจะใช้ตัวอักษร (string) หรือเลข (numeric)

เป็นค่าในการจัดกลุ่มหรือจัดประเภท (classify or categorize) จัดตำแหน่งหรืออันดับ (rank or order) หรือกำหนดให้เป็นคะแนนหรืออัตรา (score or rating)

ใน ปี ค.ศ. 1932 ผลของการประชุมของคณะกรรมการของสมาคมความก้าวหน้าทางวิทยาศาสตร์แห่ง สหราชอาณาจักร (The British Association for the Advancement of Science) ที่ประกอบไปด้วยนักคณิตศาสตร์ นักฟิสิกส์ และนักจิตวิทยา ได้เกิดประเด็นคำถามว่า “มันเป็นไปได้หรือไม่ที่จะวัดความรู้สึกของมนุษย์?” ทำให้ต้องมีการประชุมเพื่อถกเถียงกันอย่างต่อเนื่องยาวนานถึง 7 ปี จึงได้ข้อสรุปในปี ค.ศ. 1940 ว่า การวัดสามารถแบ่งได้เป็น 2 ประเภท คือ การวัดด้วยวิธีการเชิงประจักษ์ (empirical operations) และการวัดที่เป็นมาตราตามหลักคณิตศาสตร์ (mathematical) ต่อมาในปี ค.ศ. 1946 สแตนลี่ย์ สมิทธิ์ สตีเว่นส์ (Stanley Smith Stevens) จึงเสนอแนวคิดในการแบ่งกลุ่มมาตราการวัดเป็น 4 ประเภท โดยมาตราส่วนการวัดแต่ละประเภทจะมีคุณลักษณะและประเภทที่แตกต่างกันไปตามการนิยามจากข้อมูลเชิงประจักษ์ ดังนี้ (Stevens, 1946; Osborn, 2008, p. 50-51)

1. มาตราส่วนแบบจัดกลุ่ม (Nominal scale) คือ ตัวแปรที่กำหนดให้สิ่งที่ศึกษา (objects) มีลักษณะเป็นฉลากหรือชื่อ (label) ที่เป็นคำที่สามารถนับจำนวนได้ (numeral) เช่น เพศ อาชีพ ศาสนา ภูมิลำเนา เขตที่อยู่อาศัย

2. มาตราส่วนแบบจัดอันดับ (Ordinal scale) คือ ตัวแปรที่กำหนดให้สิ่งที่ศึกษา (objects) มีลักษณะเป็นอันดับตำแหน่ง (rank order) ที่เป็นคำที่สามารถนับจำนวนได้ (numeral) เช่น ระดับการศึกษา (ประถม มัธยมศึกษา และอุดมศึกษา) ระดับรายได้ (น้อย ปานกลาง และมาก) ชนชั้น (ต่ำ กลาง และสูง)

3. มาตราส่วนแบบช่วง (Interval scale) คือ ตัวแปรที่ได้จากการวัดวัตถุให้มีขนาดจำนวนที่เท่ากัน (same amount) และกำหนดเป็นตัวเลข (numerical) ที่แตกต่างกัน เพื่อแสดงให้เห็นสิ่งที่ศึกษามีปริมาณที่แตกต่างกัน เช่น ขั้นเงินได้ในการเสียภาษี 0-150,000 (0%), เกิน 150,000-300,000 (5%), เกิน 300,000-500,000 (10%), เกิน 500,000-750,000 (15%), เกิน 7500,000-1,000,000 (20%), เกิน 1,000,000-2,000,000 (25%), เกิน 2,000,000-5,000,000 (30%) และเกิน 5,000,000 บาท ขึ้นไป (35%)

4. มาตราส่วนแบบอัตราส่วน (Ratio scale) คือ ตัวแปรที่มีการวัดเหมือนกับช่วงมาตรา แต่มีมาตราวัดที่เป็นศูนย์แท้ (absolute zero) เช่น อายุ เงินเดือน น้ำหนัก ส่วนสูง คะแนนสอบ

ปัจจุบันมาตราส่วนการวัดของ สตีเว่นส์ ถูกนำมาใช้อย่างแพร่หลายทั้งการวิจัยด้านวิทยาศาสตร์และสังคมศาสตร์ แต่ในการวิจัยของแต่ละสาขามีธรรมชาติของการวัดและการใช้มาตราส่วนการวัดที่แตกต่างกัน เช่น การสังเกตของนักวิทยาศาสตร์กายภาพส่วนใหญ่ข้อมูลเป็นมาตราส่วนแบบอัตราส่วน (น้ำหนัก ส่วนสูง ความเร็ว) ส่วนการสังเกตของนักสังคมศาสตร์ส่วนใหญ่ข้อมูลเป็นมาตราส่วนแบบจัดอันดับ (ทัศนคติ ความฉลาด สมรรถนะ) นอกจากที่กล่าวมาในเอกสาร ตำรา และงานวิจัยต่างๆ มีการแบ่งมาตราส่วนการวัดของตัวแปรในมิติต่างๆ ดังนี้

1. ตัวแปรแบบค่าจัดประเภท (Categorical variable) เป็นตัวแปรที่มีค่าที่แบ่งออกจากกันอย่างเด็ดขาดหรือแบ่งเป็นกลุ่มประเภทอย่างชัดเจน ได้แก่ ตัวแปรแบบจัดกลุ่ม (nominal variable) และตัวแปรแบบจัดอันดับ (ordinal variable)

2 ตัวแปรแบบค่าต่อเนื่อง (Continuous variable) เป็นตัวแปรที่มีค่าที่มีความต่อเนื่องกัน (continuous) ได้แก่ ตัวแปรแบบช่วง (interval variable) และตัวแปรแบบอัตราส่วน (ratio variable)

3. ตัวแปรแบบค่าไม่ต่อเนื่อง (Discontinuous variable) เป็นตัวแปรที่มีการแบ่งค่าเป็นกลุ่ม ได้แก่ ตัวแปรแบบจัดกลุ่ม ตัวแปรแบบจัดอันดับ ตัวแปรแบบทวิลักษณะ (binary variables) และตัวแปรแบบค่าวิมุต (discrete variable)

4. ตัวแปรแบบค่าสองประเภท (Dichotomous variable) เป็นตัวแปรที่มีค่าแบ่งกลุ่ม 2 กลุ่ม ได้แก่ ตัวแปรแบบสองประเภท/ทวิลักษณะ (binary variables) และตัวแปรแบบค่าวิมุต/ไม่ต่อเนื่อง (discrete variable) รวมถึงตัวแปรแบบจัดกลุ่มที่มีค่าแบ่งกลุ่ม 2 กลุ่ม

5. ตัวแปรแบบค่าทวิลักษณะ (Binary variable) มีลักษณะเป็นทั้งตัวแปรแบบจัดกลุ่ม ตัวแปรแบบจัดอันดับ และตัวแปรแบบค่าไม่ต่อเนื่อง ค่าของตัวแปรมีลักษณะเป็นคำหรือข้อความ (string) ที่มีลักษณะตรงข้ามหรือเป็นสิ่งที่อยู่คู่กัน เช่น ชาย/หญิง เมือง/ชนบท ไม่ใช่/ใช่ ผิด/ถูก ไม่สำเร็จ/สำเร็จ หรือเป็นเลขที่มีค่าไม่ต่อเนื่องกันแต่สามารถนับได้ (counting) เช่น 0 กับ 1

6. ตัวแปรแบบค่าวิมุต (Discrete variable) มีลักษณะเป็นทั้งตัวแปรแบบค่าต่อเนื่องและตัวแปรแบบทวิลักษณะ ค่าของตัวแปรมีลักษณะเป็นตัวเลข (numeric) แบบจำนวนเต็ม (integer) ที่สามารถคำนวณได้ (calculated) เช่น 0 กับ/ถึง 1 หรือเป็นตัวแปรสองประเภทที่แปลงค่ามาเป็นคะแนน 0 กับ 1 เช่น 0 ไม่ใช่ชาย กับ 1 = ชาย หรือ 0 = ไม่ใช่เมือง กับ 1 = เมือง

7. ตัวแปรแบบค่าหลายประเภท (Polychotomous variable) เป็นตัวแปรแบบจัดประเภท (categorical variable) ที่มีค่าของตัวแปรมากกว่า 2 ค่า จึงเป็นประเภทตัวแปรที่มีคุณลักษณะครอบคลุมตัวแปรแบบแบ่งกลุ่มเกือบทุกประเภท ยกเว้นตัวแปรแบบทวิลักษณะ (binary variables)

การแบ่งประเภทตัวแปรตามระดับการวัดหรือมาตราการวัดตามที่กล่าวมา ขึ้นอยู่กับวัตถุประสงค์หรือวิธีการของการจำแนก จึงทำให้เห็นลักษณะของตัวแปรในหลายมิติ แต่ในกระบวนการในการวิเคราะห์ข้อมูลด้วยโปรแกรมคอมพิวเตอร์ต้องบอกให้โปรแกรมหรือคอมพิวเตอร์รู้ว่า แต่ละตัวแปรเป็นตัวแปรแบบตัวอักษร (string variables) หรือตัวแปรแบบตัวเลข (numeric variables) เพราะตัวแปรแต่ละประเภทต้องใช้วิธีการประมวลผลในการคำนวณที่เหมาะสมตามหลักสถิติ คือ ตัวแปรแบบตัวอักษรต้องใช้การนับ (counting) ไม่สามารถนำไปคำนวณหาค่าเฉลี่ย (mean) ส่วนตัวแปรแบบตัวเลขอาจใช้การนับหรือการหาค่าเฉลี่ยก็ได้ ดังนั้นจึงควรทำความเข้าใจและจัดตัวแปรในทุกระดับการวัดให้เป็น 2 กลุ่มตัวแปร ดังตารางด้านล่าง

ตารางที่ 2.1 ประเภทตัวแปรแบบตัวอักษรและตัวแปรแบบตัวเลข

ตัวแปรแบบตัวอักษร (String Variables)	**ตัวแปรแบบตัวเลข (Numeric Variables)**
ตัวแปรเชิงคุณภาพ/คุณลักษณะ (qualitative variables)	ตัวแปรเชิงปริมาณ (quantitative variables)
ตัวแปรแบบค่าจัดประเภท (categorical variables)	ตัวแปรค่าต่อเนื่อง (continuous variables)
ตัวแปรมาตราแบบจัดกลุ่มและจัดอันดับ (nominal scale and ordinal scale variables)	ตัวแปรมาตราแบบช่วงและอัตราส่วน (interval scale and ratio scale variables)
ตัวแปรแบบค่าวิมุต (discrete variables)	ตัวแปรแบบค่าวิมุต (discrete variables)
ตัวแปรแบบค่าสองประเภทและตัวแปรแบบค่าทวิลักษณะ (dichotomous and binary variables)	
ตัวแปรแบบค่าหลายประเภท (polychotomous variables)	

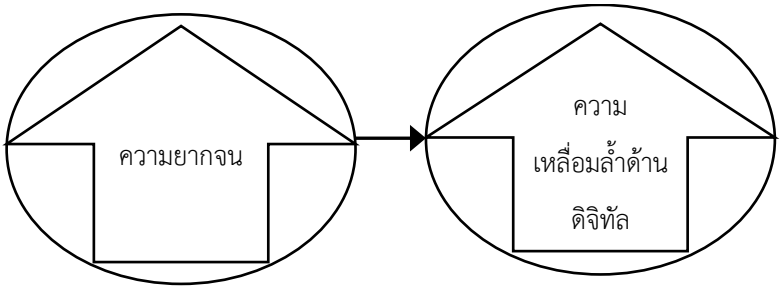
การแบ่งประเภทการวัดของตัวแปรตามที่กล่าวมามีความเกี่ยวข้องกับการตัดสินใจเบื้องต้นในการเลือกใช้สถิติแต่ละประเภทในการวิเคราะห์ข้อมูล การเข้าใจและสามารถแบ่งประเภทการวัดของตัว

แปรได้ถูกต้องจะทำให้การเลือกใช้สถิติถูกต้องก่อนทำตรวจสอบลักษณะของข้อมูลตามข้อตกลงเบื้องต้นของสถิติแต่ละประเภทในขั้นต่อไป

5. ความสัมพันธ์ระหว่างสมมติฐานและตัวแปร

ทฤษฎีมีความเป็นนามธรรมมาและกล่าวถึงความสัมพันธ์ระหว่างปรากฏการณ์แบบกว้างๆ การลดทอนทฤษฎีมาสร้างเป็นสมมติฐานในรูปแบบของความสัมพันธ์ระหว่างตัวแปรสามารถแบ่งรูปแบบของความสัมพันธ์ได้ดังต่อไปนี้

1. ความสัมพันธ์เชิงสาเหตุ (Causal relationship) เป็นความสัมพันธ์ที่ตัวแปรเหตุและตัวแปรผลมีการเปลี่ยนแปลงไปในทิศทางเดียวกัน เมื่อตัวแปรเหตุมีค่าเพิ่มขึ้น ตัวแปรผลจะเปลี่ยนค่าเพิ่มขึ้นตามไปด้วย

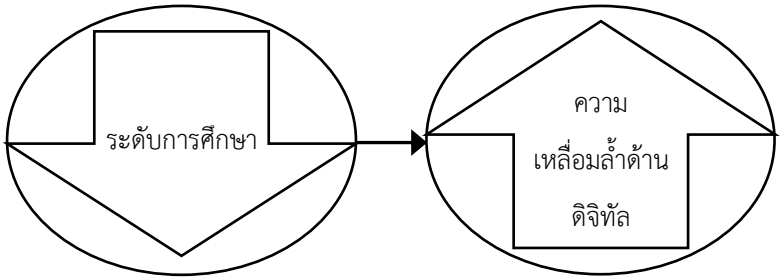


ความยากจนเป็นสาเหตุของความเหลื่อมล้ำด้านดิจิทัล

H_0 = ความยากจนไม่ใช่สาเหตุของความเหลื่อมล้ำด้านดิจิทัล

H_1 = ความยากจนเป็นสาเหตุของความเหลื่อมล้ำด้านดิจิทัล

2. ความสัมพันธ์เชิงผกผัน (Inverse causal relationship) เป็นความสัมพันธ์ในรูปของตัวแปรเหตุมีค่าเพิ่มขึ้นหรือลดลง ตัวแปรตามจะค่าเปลี่ยนไปในทางตรงข้าม คือ ตัวแปรเหตุมีค่าเพิ่มขึ้น ตัวแปรผลจะเปลี่ยนค่าลดลง แต่ถ้าตัวแปรเหตุมีค่าลดลง ตัวแปรผลจะเปลี่ยนค่าเพิ่มขึ้น

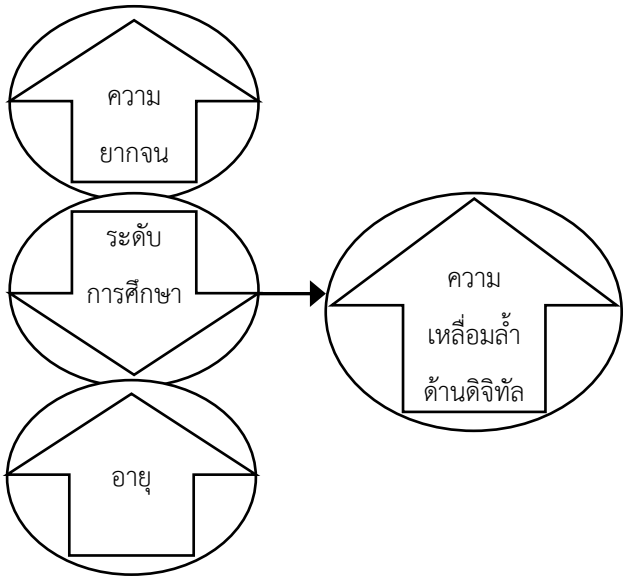


ระดับการศึกษาเป็นสาเหตุของความเหลื่อมล้ำด้านดิจิทัล

H_0 = ระดับการศึกษาไม่ใช่สาเหตุของความเหลื่อมล้ำด้านดิจิทัล

H_1 = ระดับการศึกษาเป็นสาเหตุของความเหลื่อมล้ำด้านดิจิทัล

3. ความสัมพันธ์เชิงหลายสาเหตุ (Multiple-cause relationship) เป็นความสัมพันธ์ของตัวแปรผลที่เกิดมาจากหลายสาเหตุและอาจมีทิศทางที่ไม่เหมือนกัน



มีหลายสาเหตุร่วมกันที่ทำให้เกิดความเหลื่อมล้ำด้านดิจิทัล

H_0 = ความยากจน การศึกษา และอายุ ไม่ใช่สาเหตุร่วมกันของการเกิดความเหลื่อมล้ำด้านดิจิทัล

H_1 = ความยากจน การศึกษา และอายุ เป็นสาเหตุร่วมกันของการเกิดความเหลื่อมล้ำด้านดิจิทัล

สรุป

การวิจัยเชิงปริมาณเป็นวิธีการหาความรู้ที่เน้นการวัดและการวิเคราะห์ข้อมูลที่เก็บรวบรวมมาแล้วนิยมทำให้เป็นรหัสในรูปแบบของตัวเลขสำหรับนำไปวิเคราะห์ด้วยวิธีการทางสถิติ โดยการนำเอาทฤษฎีมาลดทอนเป็นสมมติฐานหรือชุดตัวแปรขนาดเล็กที่สามารถทำการทดสอบได้ด้วยข้อมูลเชิงประจักษ์ ใช้ตัวแปรเป็นเครื่องมือวัดและรวบรวมข้อมูลที่เป็นตัวเลขจากประชากรหรือกลุ่มตัวอย่าง และทำการสรุปข้อมูลด้วยสถิติ เพื่อพรรณนา อธิบาย หาความสัมพันธ์ และทำนายปรากฏการณ์ระหว่างตัวแปร

แบบฝึกหัดท้ายบท

ให้ผู้เรียนค้นหางานวิจัยเชิงปริมาณด้านสังคมศาสตร์จากฐานข้อมูลของห้องสมุดที่สนใจ แล้วตอบคำถาม ดังนี้

1. งานวิจัยที่ค้นได้ มีทฤษฎีอะไรบ้าง
2. งานวิจัยที่ค้นได้ สมมติฐานประเภทใด
3. งานวิจัยที่ค้นได้ เขียนสมมติฐานแต่ละประเภทไว้ว่าอย่างไร
4. ทดลองเขียนสมมติฐานเชิงสถิติทั้งสมมติฐานหลักและสมมติฐานรองจากงานวิจัยที่ค้นได้
5. ทดลองเขียนสมมติฐานเพื่อทดสอบแบบหางเดียวและสองหางจากงานวิจัยที่ค้นได้
6. งานวิจัยที่ค้นได้ มีตัวแปรอะไรบ้าง และแต่ละตัวแปรมีระดับการวัดแบบใด

บทที่ 3
สถิติสำหรับการวิจัย

ความมุ่งหมายของบทเรียน

ผู้เรียนเข้าใจและเลือกใช้สถิติแต่ละประเภทในการวิเคราะห์และทดสอบข้อมูลได้ถูกต้องและเหมาะสม

เนื้อหา

1. สถิติ
2. ความสัมพันธ์ระหว่างข้อมูลกับสถิติ
3. การทดสอบสมมติฐาน
4. ความคลาดเคลื่อนของผลการวิจัย

วิธีการสอนและกิจกรรมการเรียนการสอน

บรรยาย แสดงตัวอย่าง สืบค้นข้อมูลจากฐานข้อมูลห้องสมุด และทำแบบฝึกหัดท้ายบท

สื่อและอุปกรณ์การสอน

เอกสารคำสอน ภาพเลื่อน คอมพิวเตอร์ โปรแกรมสอนออนไลน์ (สอนแบบออนไลน์) และเครื่องฉาย (สอนในชั้นเรียน)

งานที่มอบหมาย

ทำแบบฝึกหัดท้ายบท

การวัดผลและประเมินผล

การเข้าชั้นเรียน สังเกตความสนใจจากการเรียน สอบถามความเข้าใจจากการบรรยาย และการทำแบบฝึกหัดท้ายบท

เนื้อหา

ช่วงปลายคริสต์ศตวรรษที่ 17 เริ่มมีการใช้คำว่า ความน่าจะเป็น (probability) จนถึงต้นคริสต์ศตวรรษที่ 19 คำว่า โอกาส (chance) น่าจะเริ่มครอบงำโลก นักวิทยาศาสตร์สมัยใหม่ยุคแรกๆไม่มีการใช้สถิติและก่อนทศษวรรษที่ 1840 ก็ไม่มีสิ่งที่บงบอกให้เห็นถึงเรื่องการวัด (measurement) จนถึงประมาณกลางศตวรรษที่ 19 เป็นต้นมา สถิติ (statistics) ถูกนำมาใช้โดยรัฐ (state) และเริ่มเข้ามามีบทบาทในวิทยาศาสตร์ธรรมชาติ (natural science) รวมถึงสังคมศาสตร์สมัยใหม่ (Williams, 2016, p. 210-211)

1. สถิติ

สถิติเป็นเรื่องเกี่ยวกับการรวบรวม (collection) การวิเคราะห์ (analysis) และแปลความหมาย (interpretation) ข้อมูลที่เป็นตัวเลข รวมถึงเป็นเครื่องมือเชิงประจักษ์ (empirical tool) และเครื่องมือเชิงญาณวิทยา (epistemology tool) เพราะทำให้เราสามารถใช้อธิบาย (explain) และทำนาย (predict) สิ่งต่างๆในโลกนี้ได้ หากเราไม่มีสถิติเราก็ไม่สามารถที่จะอธิบายหรือทำนายได้

1.1 ความหมายของสถิติ

สถิติ (statistics) คือ การรวมกลุ่มของข้อมูลเชิงตัวเลข (numerical data) ที่ได้มาจากการสำรวจทางสังคม (social survey) หรือการสำมะโน (census) (Jary and Jary, 1995, p. 654)

สถิติ คือ วิธีการหนึ่งทางคณิตศาสตร์ (a set of mathematical techniques) ที่ใช้ในการจัดการ (organize) และเปลี่ยนแปลง (manipulate) ข้อมูลเพื่อนำไปใช้ตอบคำถามและทดสอบทฤษฎี (Healey, 1999, p. 1)

สถิติเป็นคำที่มีสองความหมาย ความหมายแรก คือ การอธิบาย (describing) ข้อมูลของปรากฏการณ์ใดปรากฏการณ์หนึ่ง และความหมายที่สอง คือ การรวบรวม (collecting) และการวิเคราะห์ (analyzing) ข้อมูล (Greenfield and Greener, 2016, p. 289)

นิยามในเชิงการวิจัย ส่วนใหญ่อธิบายว่า สถิติ คือ ค่าสรุปรวมที่คำนวณมาจากลักษณะหรือตัวแปรของตัวอย่าง เพื่อนำไปใช้บรรยาย (describe) คุณลักษณะ (characteristic) ของตัวอย่าง (sample) และอ้างอิง (inference) ไปยังประชากร (population)

1.2 ประเภทของสถิติ

โดยทั่วไปและตำราหลายเล่มแบ่งประเภทสถิติตามเทคนิคทางสถิติเป็น 3 ประเภท คือ สถิติเชิงพรรณนา สถิติเชิงอ้างอิง และสถิติเชิงความสัมพันธ์และทำนาย (Healey, 1999, p. 259-260; Field, 2009, p. 540; Yang, 2010, p. 7; Evans, 2014, p. 2-5) ดังนี้

1. สถิติเชิงพรรณนา (Descriptive statistic) คือ สถิติสาขาหนึ่งที่มีวัตถุประสงค์ในการใช้บรรยาย (describes) ลักษณะของข้อมูลพื้นฐานของสิ่งที่ใช้ในการศึกษา โดยใช้วิธีสรุป (summary indictors) หรือลดจำนวนข้อมูลขนาดใหญ่จากกลุ่มตัวอย่างหรือการวัดให้อยู่ในรูปของดัชนีหรือค่าสรุปขนาดเล็กที่มีความกระชับหรือสามารถเข้าใจได้ ไม่สามารถนำผลไปอ้างอิงหรือพยากรณ์ค่าของกลุ่มอื่น ๆ เช่น จากการศึกษารายจ่ายต่อเดือนของประชาชนในชุมชนแห่งหนึ่ง พบว่า จากประชาชนที่สุ่มมาทั้งหมด 100 คน มีรายจ่ายรายเดือนเฉลี่ยคนละ 3,500 บาท โดยทั่วไปมีการแบ่งเป็น 5 ประเภท ดังนี้

1.1 การวัดการแจกแจงความถี่ (Frequency distribution) ได้แก่ ค่าความถี่ (frequency) ค่าจำนวนนับ (count) และค่าร้อยละ (percentage)

1.2 การวัดแนวโน้มสู่ส่วนกลาง (Central tendency) ได้แก่ ค่าฐานนิยม (mode) ค่ามัธยฐาน (median) และค่าเฉลี่ย (mean)

1.3 การวัดการกระจายหรือความผันแปร (Dispersion/Variation) ได้แก่ ค่าพิสัย (range) ค่าส่วนเบี่ยงเบนมาตรฐาน (standard deviation) ค่าความแปรปรวน (variance) ค่าสัมประสิทธิ์ความแปรปรวน (coefficient of variation) ค่าส่วนเบี่ยงเบนควอไทล์ (quartile deviation) ค่าส่วนเบี่ยงเบนเฉลี่ย (mean or average deviation) และค่าพิสัยระหว่างควอไทล์ (interquartile range)

1.4 การวัดตำแหน่ง (Position/Location) ได้แก่ ค่าควอไทล์ (quartiles) ค่าเดไซล์ (decile) ค่าเปอร์เซ็นไทล์ (percentiles) ค่า N-ไทล์ (N-tiles) และคะแนนมาตรฐาน (standardized scores: z-scores)

1.5 การวัดรูปร่าง (Shape) ได้แก่ ค่าความเบ้ (skewness) และค่าความโด่ง (kurtosis)

2. สถิติเชิงอ้างอิง (Inferential statistic) คือ สถิติกลุ่มตัวแบบเชิงเส้นทั่วไป (general linear models) ที่คำนวณข้อมูลจากกลุ่มตัวอย่างหนึ่งกลุ่มตัวอย่าง หรือมากกว่าหนึ่งกลุ่มตัวอย่าง แล้วให้ค่าสถิติ (เช่น ค่าเฉลี่ยของประชากรทั้งหมด) ที่สามารถใช้อ้างอิงหรือทำนายลักษณะของประชากร เพื่อตัดสินใจเชิงความน่าจะเป็น (probability) เช่น จากประชาชนที่สุ่มมาทั้งหมด 100 คน มีรายจ่ายรายเดือนเฉลี่ยคนละ 3,500 บาท และโดยอาศัยทฤษฎีความน่าจะเป็น ทำการประมาณรายจ่ายโดยเฉลี่ยรายเดือนของประชาชนทั้งหมดในชุมชนได้เป็นระหว่าง 3,300 และ 3,700 บาท โดยทั่วไปมีการแบ่งเป็น 2 ประเภท ดังนี้

2.1 การทดสอบแบบอิงค่าพารามิเตอร์ (Parametric test) คือ สถิติที่มีค่าตัวเลขทางประชากรที่ใช้ในการอ้างอิงไปสู่คุณลักษณะ (characteristics) ของประชากร ข้อตกลงเบื้องต้นข้อมูลที่ใช้ในการวิเคราะห์มีระดับการวัดเป็นแบบค่าต่อเนื่อง ได้แก่ ช่วงมาตรา (interval) และสัดส่วนมาตรา (ratio) และข้อมูลมีการแจกแจงปกติ (normal distribution) เช่น การทดสอบค่าเฉลี่ยของกลุ่มตัวอย่าง 1 กลุ่ม (One-sample t-test) การทดสอบค่าเฉลี่ยของกลุ่มตัวอย่างที่อิสระจากกัน 2 กลุ่ม (Two-independent samples t-test) การทดสอบค่าเฉลี่ยของกลุ่มตัวอย่างที่เป็นอิสระจากกันมากกว่า 2 กลุ่ม (การทดสอบความแปรปรวนแบบทางเดียว: One-way ANOWA) การทดสอบค่าเฉลี่ยของกลุ่มตัวอย่างที่ไม่อิสระจากกัน 2 กลุ่ม (Two-dependent sample t-test) การทดสอบค่าเฉลี่ยของกลุ่มตัวอย่างที่ไม่อิสระจากกันมากกว่า 2 กลุ่ม หรือที่เรียกว่า การทดสอบความแปรปรวนแบบวัดซ้ำ (Repeated Measures ANOVA) การทดสอบความแปรปรวนแบบสองทาง (Two-way ANOVA) การทดสอบความสัมพันธ์รายคู่แบบเพียร์สันหรือการทดสอบสหสัมพันธ์แบบเพียร์สัน (Pearson Correlation) การทดสอบการถดถอยเชิงเส้นตรง (Linear Regression) การวิเคราะห์องค์ประกอบ (Factor Analysis) และการวิเคราะห์ตัวแปรพหุ (Multivariate Analysis)[7]

2.2 การทดสอบแบบไม่อิงค่าพารามิเตอร์ (Non-parametric test) คือ สถิติที่ไม่มีค่าตัวเลขทางประชากรที่ใช้ในการอ้างอิงไปสู่คุณลักษณะ (characteristics) ของประชากร จึงทำให้บางตำราจัดรวมอยู่ในกลุ่มของสถิติเชิงพรรณนา ข้อมูลที่ใช้ในการวิเคราะห์มีระดับการวัดเป็นแบบข้อมูลแบ่งกลุ่ม ได้แก่ มาตราแบบจัดกลุ่ม (nominal scale) และจัดอันดับ (ordinal/rank-order scale) บางตำราอาจ

[7] การวิเคราะห์ที่มีเงื่อนไขต้องใช้ตัวอย่างขนาดใหญ่และข้อมูลมีการแจกแจงแบบปกติ เช่น ตัวแบบสมการโครงสร้างแบบใช้ค่าความแปรปรวนร่วม (Covariance Based Structural Equation Modeling: CB-SEM) ที่เน้นการวิเคราะห์เชิงยืนยัน (Awang, Afthanorhan, & Asri 2015)

เรียกสถิติประเภทนี้ว่า การทดสอบที่มีการแจกแจงแบบอิสระ (distribution-free test)[8] หรือ การทดสอบที่เป็นอิสระจากข้อตกลง (assumption-free test) เพราะสถิติที่ใช้ทดสอบข้อมูลอาจมีข้อตกลงเพียงเล็กน้อยหรือไม่มีข้อตกลงเลยเกี่ยวกับลักษณะการแจกแจงของข้อมูล (data distributions) เช่น การทดสอบความสอดคล้องของข้อมูลด้วยไคสแควร์ (X^2 goodness of fit) การทดสอบความเป็นอิสระของข้อมูลด้วยไคสแควร์ (X^2 Test for Independence) การทดสอบค่ากลางของข้อมูล 1 กลุ่มตัวอย่าง (One-sample Test) การทดสอบค่ากลางของกลุ่มตัวอย่างที่มีความสัมพันธ์กัน (2 Related Samples Test) การทดสอบค่ากลางของกลุ่มตัวอย่างที่เป็นอิสระจากกัน (Independence Sample Test) การทดสอบสหสัมพันธ์แบบสเปียร์แมน (Spearman Correlation Test) การวิเคราะห์การถดถอยแบบไม่อิงค่าพารามิเตอร์ (Non-parametric Regression Analysis) และการวิเคราะห์พหุแบบไม่อิงค่าพารามิเตอร์ (Non-parametric Multivariate Analysis)[9]

3. สถิติเชิงความสัมพันธ์และการทำนาย (Correlational and predictive statistic) คือ สถิติที่ใช้ในการบอก/พรรณนา (describing) ความสัมพันธ์ระหว่างเหตุการณ์ต่างๆ หรือทำนายจากเหตุการณ์หนึ่งไปสู่อีกเหตุการณ์หนึ่ง โดยมีรายละเอียด ดังนี้

3.1 สถิติเชิงความสัมพันธ์ เป็นเครื่องมือที่ใช้ในการวัดความเข้มข้น (strength) และทิศทาง (direction) ของความสัมพันธ์ (relationships) โดยรวมระหว่าง 2 ชุดหรือกลุ่มข้อมูลที่ได้จากการสังเกต เช่น ความสัมพันธ์ระหว่างความยากจนกับอาชญากรรม และนำไปใช้ในเชิงการอ้างอิงจากกลุ่มตัวอย่างไปสู่กลุ่มประชากร ได้แก่ การวิเคราะห์สหสัมพันธ์แบบเพียร์สัน (Pearson Correlation) และการวิเคราะห์สหสัมพันธ์แบบสเปียร์แมน (Spearman Correlation)

3.2 สถิติเชิงการทำนาย เป็นเครื่องมือที่ใช้ในการทำนาย (prediction) ผลลัพธ์ของเหตุการณ์หนึ่งจากอีกเหตุการณ์หนึ่ง เช่น การทำนายความพร้อมด้านเทคโนโลยีสารสนเทศและการสื่อสารจากสถานะภาพทางเศรษฐกิจและสังคม การทำนายจะมีความถูกต้องมากน้อยเพียงใดขึ้นกับว่าข้อมูลที่ได้มาจากการสังเกตมีความสัมพันธ์กันมากน้อยเพียงใด ได้แก่ การวิเคราะห์การถดถอยเชิงเส้นตรง (linear regression) และการวิเคราะห์การถดถอยเชิงโลจิสติก (logistic regression)

---

[8] เป็นการทดสอบที่ข้อมูลมีลักษณะการแจกแจงแบบใดก็ได้ หรืออาจเรียกว่า การทดสอบที่ไม่มีข้อตกลงหรือการทดสอบที่เป็นอิสระจากข้อตกลง (free assumption)

[9] การวิเคราะห์ที่สามารถใช้ตัวอย่างขนาดเล็กและข้อมูลมีการแจกแจงแบบอิสระ เช่น การวิเคราะห์ตัวแบบสมการโครงสร้างแบบใช้ค่าความแปรปรวน (Variance Based Structural Equation Modeling: VB-SEM) ที่เน้นการวิเคราะห์เชิงสำรวจ (exploratory analysis) (Awang, Afthanorhan and Asri 2015)

จากที่กล่าวมาแสดงให้เห็นว่า สถิติที่ใช้ในการวิเคราะห์ข้อมูลแต่ละประเภทมีความคาบเกี่ยวกัน การจำแนกประเภทต้องบอกหลักเกณฑ์ให้ชัดเจน เพราะสถิติเชิงความสัมพันธ์และเชิงการทำนายยังสามารถจำแนกได้เป็น 2 ประเภท คือ สถิติแบบอิงค่าพารามิเตอร์และสถิติแบบไม่อิงค่าพารามิเตอร์ หรือสถิติเชิงพรรณาสามารถจัดเป็นสถิติแบบไม่อิงค่าพารามิเตอร์ได้เช่นกัน เนื่องจากเป็นสถิติที่ใช้พรรณนาเฉพาะกลุ่มตัวอย่างหรือประชากรเท่านั้น โดยไม่อ้างอิงจากกลุ่มตัวอย่างไปสู่ประชากร

2. ความสัมพันธ์ระหว่างข้อมูลกับสถิติ

ข้อมูลที่ใช้ในการวิจัยมีหลายระดับการวัด หากเป็นข้อมูลแบบเมตริก (metric data) เป็นค่าต่อเนื่อง (continuous data) เช่น อายุ เงินเดือน น้ำหนัก ส่วนสูง สามารถนำไปคำนวณหาค่าเฉลี่ยได้ แต่การวิจัยด้านสังคมศาสตร์ส่วนใหญ่ข้อมูลเป็นแบบไม่ใช่เมตริก (nonmetric data) โดยเฉพาะข้อมูลแบบค่าจัดประเภท (categorical data) เช่น เพศ วัย เขตที่อยู่อาศัย ไม่สามารถนำไปคำนวณหาค่าเฉลี่ยได้เหมือนกับข้อมูลแบบเมตริก ต้องใช้ค่าความถี่หรือค่าร้อยละ

สถิติแต่ละประเภทมีข้อตกลงเบื้องต้น (basic assumptions) ในการวิเคราะห์ข้อมูลแต่ละประเภท หากฝืนข้อตกลง (violate assumptions) มีความเสี่ยงทำให้เกิดความคลาดเคลื่อนของผลการวิเคราะห์และนำไปสู่การตีความผิดตามมา

2.1 ข้อมูลกับสถิติเชิงพรรณนา

การบรรยายลักษณะทางประชากรของกลุ่มตัวอย่างมีข้อตกลงเกี่ยวกับข้อมูลเช่นกัน การเลือกใช้สถิติวัดแนวโน้มเข้าสู่ส่วนกลางและการกระจายมีแนวทางในการเลือกใช้สถิติให้ตรงกับข้อตกลงดังตารางด้านล่าง (Verma, & Abdel-Salam, 2019, p. 3-4)

ตารางที่ 3.1 ความสัมพันธ์ระหว่างประเภทข้อมูลกับสถิติเชิงพรรณนา

ประเภทของข้อมูล	ลักษณะของข้อมูล	แนวโน้มสู่ส่วนกลาง	การกระจาย
ไม่ใช่เมตริก	จัดกลุ่ม	ค่าฐานนิยม	ค่าความถี่/ค่าร้อยละ
(จัดประเภท)	จัดอันดับ	ค่ามัธยฐาน	ค่าพิสัย
เมตริก	ช่วง/อัตราส่วน	ค่าเฉลี่ย	

ประเภทของข้อมูล	ลักษณะของข้อมูล	แนวโน้มสู่ส่วนกลาง	การกระจาย
(ต่อเนื่อง)	(ถ้ามีการแจกแจงแบบปกติหรือใกล้เคียง)	ค่ามัธยฐาน	ค่าเบี่ยงเบนมาตรฐาน
	ช่วง/อัตราส่วน (ถ้ามีการแจกแจงแบบเบ้)		

2.2 ข้อมูลกับสถิติเชิงอ้างอิง

ข้อมูลที่ได้มาจากประชากรหรือตัวอย่างการวิจัยมีหลากหลายลักษณะหรือระดับการวัด หากข้อมูลเป็นแบบเมตริก (metric data) หรือเป็นข้อมูลแบบค่าต่อเนื่อง (continuous data) เช่น อายุ เงินเดือน น้ำหนัก ส่วนสูง หรือเป็นมาตราแบบช่วง (interval scale) หรือมาตราแบบอัตราส่วน (ratio scale) สามารถเลือกใช้สถิติแบบอิงค่าพารามิเตอร์ (parametric statistical) ในการวิเคราะห์ข้อมูลที่รวบรวมมาได้ แต่ในการวิจัยด้านสังคมศาสตร์ส่วนใหญ่ข้อมูลเป็นแบบไม่ใช่เมตริก (nonmetric data) หรือเป็นข้อมูลแบบค่าจัดประเภท (categorical data) เช่น ทัศนคติ ความฉลาด สมรรถนะ หรือเป็นมาตราแบบจัดกลุ่ม (nominal scale) หรือมาตราส่วนแบบจัดอันดับ (ordinal scale) ไม่เหมาะกับการวิเคราะห์ข้อมูลด้วยสถิติแบบอิงค่าพารามิเตอร์ หากนักวิจัยฝืนใช้สถิติแบบอิงค่าพารามิเตอร์ เช่น ใช้ ANOVA วิเคราะห์หาอิทธิพลรวมด้วยข้อมูลมาตราแบบจัดอันดับจะได้ผลการวิเคราะห์ที่เป็นปฏิสัมพันธ์ลวง (spurious interactions) และได้ค่าการประมาณการของขนาดผลกระทบต่ำกว่าที่ควรจะเป็น (underestimation of effect sizes) (Osborn, 2008, p. 50-51)

การเลือกใช้สถิติอ้างอิงหรือสถิติเชิงอนุมานทุกประเภท มีข้อตกลงเบื้องต้นร่วมกัน คือ ตัวอย่างได้มาจากประชากรด้วยวิธีการสุ่มแบบสุ่ม (random sample) และแต่ละประเภทมีข้อตกลงเกี่ยวกับระดับการวัดและลักษณะของข้อมูลดังนี้ (Allison, 2002, p. V; Hair & Others, 2006, p. 49-50; Field, 2009; Garson, 2012; Halter, 2017, p. 10-11; Verma, & Abdel-Salam, 2019, p. 65-188)

1. การทดสอบแบบอิงค่าพารามิเตอร์ (Parametric test) ข้อมูลต้องเป็นข้อมูลตัวเลขหรือเป็นค่าต่อเนื่องหรือมีการวัดตั้งแต่ระดับช่วงมาตรา (interval scale) ขึ้นไป ข้อมูลมีการแจกแจงปกติ (normal distribution) มีความเป็นแบบเดียวกันของความแปรปรวน (homogeneity of variance)

ความเป็นอิสระของข้อมูล (independence) และ/หรือข้อมูลมีความเป็นเส้นตรงในการวิเคราะห์การถดถอยเชิงเส้นตรง (linear regression analysis)

2. การทดสอบแบบไม่อิงค่าพารามิเตอร์ (Non-parametric test) เป็นสถิติที่มีข้อตกลงเพียงเล็กน้อยหรือเป็นการทดสอบที่เป็นอิสระจากข้อตกลง (assumption-free test) ข้อมูลมีการวัดระดับใดก็ได้ ข้อมูลไม่จำเป็นต้องมีการแจกแจงปกติ ข้อมูลแต่ละกลุ่มไม่จำเป็นต้องมีการกระจายเท่ากัน และ/หรือไม่มีความเป็นเส้นตรงในการวิเคราะห์การถดถอยไม่ใช่เชิงเส้นตรง (nonlinear regression analysis)

ตารางที่ 3.2 เปรียบเทียบข้อตกลงเบื้องต้นของการทดสอบแบบอิงค่าพารามิเตอร์และไม่อิงค่าพารามิเตอร์

ข้อมูล	อิงค่าพารามิเตอร์	ไม่อิงค่าพารามิเตอร์
ความเป็นสุ่ม (randomness)	/	/
ความเป็นอิสระ (independence)	/	/
ระดับการวัด (level of measurement)		
- ค่าต่อเนื่อง (อัตราส่วน/ช่วง)	/	X
- ค่าจัดประเภท (จัดอันดับ/จัดกลุ่ม)	/	/
การแจกแจงแบบปกติ (normality)	/	X
ความเท่ากันของความแปรปรวน (homogeneity)	/	X
ความสัมพันธ์เชิงเส้นตรง (linearity)	/	X

การทดสอบสมมติฐานด้วยสถิติเชิงอ้างอิงแบบเดิม (traditional test) คือ การทดสอบแบบอิงค่าพารามิเตอร์ (parameter test) โดยอิงอยู่กับการแจกแจงแบบความน่าจะเป็น (probability distribution) เช่น การแจกแจกปกติ (normal distribution) และนักวิจัยส่วนใหญ่ยอมรับกันอย่างแพร่หลายว่าเป็นสถิติที่มีอำนาจจำแนก (statistical power) ทำผลการทดสอบหรือข้อสรุปมีความถูกต้องมาก อย่างไรก็ตามหากข้อมูลที่ใช้ในการทดสอบไม่ตรงกับข้อตกลงเบื้องต้นของสถิติเชิงอ้างอิงแบบอิงค่าพารามิเตอร์ สามารถใช้การทดสอบแบบไม่อิงค่าพารามิเตอร์ (non-parametric test) ทดแทนได้ แม้ว่าจะมีประสิทธิภาพไม่เท่ากับการทดสอบแบบอิงค่าพารามิเตอร์ เพราะจะช่วยลดความคลาดเคลื่อนของผลการวิเคราะห์ได้

ก่อนการวิเคราะห์ข้อมูลด้วยสถิติแบบอิงค่าพารามิเตอร์จึงต้องทำการตรวจสอบคุณลักษณะของข้อมูล (exploratory data analysis) ให้ตรงกับข้อตกลง หากตรวจสอบแล้วพบว่าไม่ตรงกับข้อตกลงให้เลือกใช้สถิติแบบไม่อิงค่าพารามิเตอร์แทนดังนี้

ตารางที่ 3.3 เปรียบเทียบการทดสอบแบบอิงค่าพารามิเตอร์และไม่อิงค่าพารามิเตอร์

ประเภทการทดสอบ	อิงค่าพารามิเตอร์	ไม่อิงค่าพารามิเตอร์
กลุ่มตัวอย่าง 1 กลุ่ม	1-sample t test	Sign test, Wilcoxon signed rank test
กลุ่มตัวอย่าง 2 กลุ่ม	2-sample t test	Mann-Whitney test
กลุ่มตัวอย่างมากกว่า 2 กลุ่ม	One-Way ANOVA	Kruskal-Wallis test
ทดสอบซ้ำ 2 เหตุการณ์	Paired sample t-test	Wilcoxon Test
ทดสอบซ้ำมากกว่า 2 เหตุการณ์	One-Way repeated-measures ANOWA	Friedman test
ความสัมพันธ์รายคู่	Pearson’ s correlation	Spearman’ s correlation

3. การทดสอบสมมติฐาน

มีตัวอย่างของทฤษฎีจำนวนมากที่ได้รับการยอมรับมาอย่างยาวนาน แต่ต่อมาขาดความน่าเชื่อถือ เพราะความจริง (truths) ที่อยู่บนฐานของความเป็นจริงถูกปฏิเสธ ถ้าเราไม่มั่นใจว่า ทฤษฎี (theory) นั้น ถูกต้อง (true) วิธีที่ดี คือ เราต้องตัดสินด้วยความน่าจะเป็น (probability) ว่า ทฤษฎี นั้นถูกต้องหรือไม่

การทดสอบสมมติฐาน คือ วิธีการหนึ่งในกระบวนการสรุป (generates) การกระทำ (actions) ของประชากรที่ได้รับอิทธิพลมาจากตัวแปรบางตัว อีกทั้งเป็นกระบวนการที่เป็นวิทยาศาสตร์ในการทดสอบว่า สมมติฐานนั้นเป็นไปได้หรือน่าเชื่อถือหรือไม่

การทดสอบสมมติฐานที่ใช้วิธีการอ้างอิงเชิงสถิติ (statistical inference) มี 2 สำนักความคิดและมีฐานความเชื่อที่แตกต่างกัน ดังนี้ (Orloff & Bloom, 2014)

1. สำนักความคิดแบบเบย์ (Bayesian) มีความเชื่อว่า ความน่าจะเป็น (probability) ที่สมมติฐานจะมีความเป็นจริงก่อน (priori) มีข้อมูลมาสนับสนุน ความน่าจะเป็นที่สมมติฐานจะมีความเป็น

จริงหลัง (posterior) มีข้อมูลมาสนับสนุนแล้ว และการพิสูจน์สมมติฐานต้องอยู่ภายใต้เงื่อนความควรจะเป็น (likelihood) ที่ได้จากข้อมูล และความน่าจะเป็นรวม (total probability) ของข้อมูลทำให้สมมติฐานมีความเป็นไปได้ทั้งหมด (all possible)

2. สำนักความคิดแบบฟรีเควนทิสต์ (Frequentist) มีความเชื่อว่า ความน่าจะเป็นของสมมติฐานไม่มีความเป็นจริงทั้งก่อนและหลัง แต่ความน่าจะเป็นของสมมติฐานขึ้นอยู่กับความควรจะเป็น (likelihood) ของข้อมูลที่ได้มาจากทั้งการสังเกตและไม่ได้ทำการสังเกต (observed and unobserved data) ภายใต้เงื่อนไขเฉพาะแต่ละสมมติฐาน สมมติฐานที่เป็นจริงต้องได้ข้อมูลมาจากการสุ่มตัวอย่างที่มีลักษณะของการกระจายหรือแจกแจก (distributions)

การทดสอบสมมติฐานแบบเบย์มีอิทธิพลต่อการวิเคราะห์ข้อมูลในช่วงก่อนศตวรรษที่ 20 แม้ว่าจะมีประโยชน์ในการค้นหาความรู้ใหม่ในหลายเรื่องที่ไม่รู้ความน่าจะเป็นของสมติฐาน แต่การมุ่งมั่นอยู่กับการค้นหาความรู้ก่อนมีข้อมูลที่ได้จากการสังเกตแบบกระจาย (priori distributions) มาสนับสนุน เป็นความรู้ที่ให้ความสำคัญกับความคิดเห็นส่วนตัวหรือความรู้ก่อนประสบการณ์เชิงจิตพิสัย (subjective priori) ที่เรียกสั้นๆ ว่า ความเชื่อ (belief) การทดสอบสมมติฐานด้วยความไม่แน่นอนของความน่าจะเป็นจากกระจายของข้อมูล และการอ้างอิงจากสิ่งหนึ่งด้วยระดับความเชื่อมั่นจากสิ่งหนึ่งที่มีอยู่ก่อน ทำให้ในช่วงศตวรรษที่ 20 การทดสอบแบบฟรีเควนทิสต์เข้ามามีอิทธิพลแทน เนื่องจากไม่ใช้ความรู้ก่อนประสบการณ์เชิงจิตพิสัย (subjective priori) ที่มีความไม่แน่นอนขึ้นอยู่กับแต่ละบุคคล แต่ใช้หลักความรู้จากประสบการณ์เชิงวัตถุ (objective posterior) ทดสอบความน่าจะเป็นของสมมติฐานจากข้อมูลที่ได้มาจากการสำรวจตรวจสอบเชิงวิทยาศาสตร์ (scientific investigations) กำหนดระดับนัยสำคัญทางสถิติ (significance level) และวิเคราะห์หาค่าความน่าจะเป็น (probability value: p-value) เพื่อใช้ตัดสินใจที่จะปฏิเสธ (reject) สมมติฐานหลัก (null hypothesis) ผลการทดสอบอาจมีความแตกต่างกันจากการกำหนดระดับนัยสำคัญทางสถิติที่แตกต่างกัน และรวมถึงการเกิดความคลาดเคลื่อนแบบที่ 1 หรือ 2 (Orloff & Bloom, 2014)

การวิจัยเชิงปริมาณที่มีการตั้งสมมติฐาน นักวิจัยจะใช้ผลทดสอบสมมติฐานในกระบวนการการตัดสินใจ (determine) ประเมินข้อมูลเชิงประจักษ์ (empirical value) ในการทดสอบ/ยืนยันทฤษฎี (theory)

การทดสอบสมมติฐาน พัฒนามาจากวิธีการตรวจสอบ (proof) ความขัดแย้งทางสถิติ (statistical contradiction) จากหลักฐาน (evidence) ที่รวบรวมมาจากข้อสนับสนุนทางทฤษฎี (theory) โดยการอธิบายว่า (demonstrating) ข้อมูล (data) มีความแตกต่าง (unlikely) จากการสังเกต (observed) ถ้าตัวแบบทางทฤษฎีที่สร้างมาจากสมมติฐานเป็นเท็จ (false) โดยมีขั้นตอนดังต่อไปนี้

ตารางที่ 3.4 ขั้นตอนการทดสอบสมมติฐาน

ขั้นที่ 1	สร้างสมมติฐาน H_0 และ H_1	สมมติฐานแบบหางเดียวหรือสองหาง
ขั้นที่ 2	กำหนดระดับนัยสำคัญทางสถิติ	0.05 หรือ 0.025 หรือ 0.00
ขั้นที่ 3	รวบรวมข้อมูล	ข้อมูลแบบค่าแบ่งกลุ่มหรือค่าต่อเนื่อง
ขั้นที่ 4	วิเคราะห์ข้อมูล	อิงค่าพารามิเตอร์หรือไม่อิงค่าพารามิเตอร์
ขั้นที่ 5	ทดสอบสมมติฐาน	ปฏิเสธหรือยอมรับสมมติฐาน H_0 หรือ H_1
ขั้นที่ 6	สรุปและตีความ	ตัดสินใจถูกต้องหรือเกิดความคาดเคลื่อนแบบที่ 1 หรือ 2

3.1 ประเภทการทดสอบสมมติฐาน

การตัดสินใจเกี่ยวกับสมมติฐานโดยอาศัยเกณฑ์บางอย่างเข้าช่วย โดยทั่วไปจะใช้วิธีการทางสถิติกับข้อมูลที่ได้มาจากตัวอย่าง แล้วพิจารณาความน่าจะเป็นของเหตุการณ์ที่เกิดขึ้น 2 แบบ ดังนี้

1. การทดสอบแบบ 2 หาง (Two tailed Test) คือ การทดสอบว่าค่าพารามิเตอร์มีค่าเท่ากับค่าที่กำหนดหรือไม่ (สังเกตจากใน H_1 หรือ H_a จะมีเครื่องหมาย =) เขียนเป็นสัญลักษณ์ทางคณิตศาสตร์ดังนี้

H_0 $\mu = \mu_0$ (ประชาชนในเขตชนบทมีระดับการศึกษาไม่แตกต่าง (เท่ากับ) กับประชาชนในเขตเมือง)

H_1 $\mu \neq \mu_0$ (ประชาชนในเขตชนบทมีระดับการศึกษาแตกต่าง (ไม่เท่ากับ) กับประชาชนในเขตเมือง)

2. การทดสอบแบบหางเดียว (One tailed Test) คือ การทดสอบว่าค่าพารามิเตอร์มีค่ามากกว่าหรือน้อยกว่าค่าที่กำหนด (สังเกตได้จากใน H_1 หรือ H_a จะมีเครื่องหมาย > หรือ <) มี 2 แบบ คือ

2.1 ทดสอบแบบหางเดียวด้านขวา (right tailed) เขียนเป็นสัญลักษณ์ทางคณิตศาสตร์ดังนี้

H_0: $\mu = \mu_0$ (ครัวเรือนในเขตชนบทมีจำนวนสมาชิกเท่ากับครัวเรือนในเขตเมือง)

H_1: $\mu > \mu_0$ (ครัวเรือนในเขตชนบทมีจำนวนสมาชิกมากกว่าครัวเรือนในเขตเมือง)

2.2 ทดสอบแบบหางเดียวด้านซ้าย (left tailed) เขียนเป็นสัญลักษณ์ทางคณิตศาสตร์ดังนี้

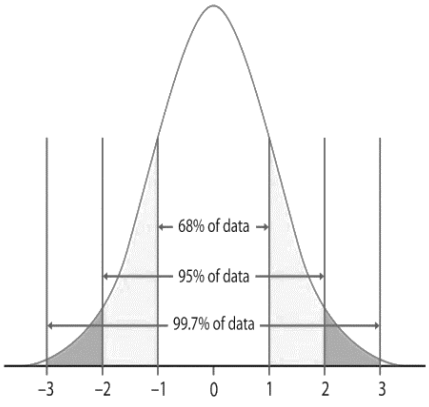
H_0 $\mu = \mu_0$ (ประชาชนในเขตชนบทมีรายได้เท่ากับประชาชนในเขตเมือง)

H_1 $\mu < \mu_0$ (ประชาชนในเขตชนบทมีรายได้น้อยกว่าประชาชนในเขตเมือง)

3.2 กฎเชิงประจักษ์

การวิจัยเชิงปริมาณด้านสังคมศาสตร์และสาขาสังคมวิทยาวางอยู่บนฐานของปรัชญาวิทยาศาสตร์ธรรมชาติสำนักประจักษ์นิยม (empiricism) ที่ยอมรับเฉพาะความรู้เชิงประสบการณ์ที่มีหลักฐานเชิงประจักษ์สนับสนุน จึงสร้างกฎเชิงประจักษ์ (empirical rule) สำหรับใช้ทดสอบสมมติฐานหลัก (null hypothesis) ด้วยข้อมูลที่สร้างมาจากตัวแปรเชิงสุ่ม (random variable) บนฐานความเชื่อที่ว่า ปรากฏการณ์ที่เกิดขึ้นตามธรรมชาติมีลักษณะสุ่ม

ข้อมูลจากตัวแปรเชิงสุ่มตามกฎเชิงประจักษ์มีลักษณะการกระจายหรือแจกแจงแบบปกติ (normal distribution) โดยกำหนดให้ข้อมูลทั้งหมดที่อยู่ภายใต้โค้งปกติ (normal curve) เท่ากับ 100% หรือ 1 แบ่งข้อมูลด้วยกฎสามซิกม่า (three sigma rule) เป็น 3 ส่วนจากจุดกลางที่เป็นค่าเฉลี่ยด้วยค่าเบี่ยงเบนมาตรฐาน (three standard deviations of the mean) หรือเป็นกฎ 68-95-99.7 (68-95-99.7 rule) ข้อมูลภายใต้ระฆังคว่ำ (bell curve) ถูกแบ่งข้อมูลออกเป็น 3 ส่วน ดังต่อไปนี้



รูปที่ 3.1 ข้อมูลภายใต้ระฆังคว่ำ

68% ตกอยู่ภายในพื้นที่ที่ 1 ของความเบี่ยงเบนมาตรฐาน (+/-1σ) ของค่าเฉลี่ย

95% ตกอยู่ภายในพื้นที่ที่ 2 ของความเบี่ยงเบนมาตรฐาน (+/-2σ) ของค่าเฉลี่ย

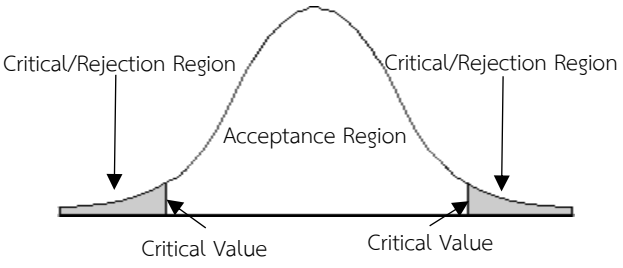
99.7% ตกอยู่ภายในพื้นที่ที่ 3 ของความเบี่ยงเบนมาตรฐาน (+/-3σ) ของค่าเฉลี่ย

การนำเอากฎเชิงประจักษ์มาประยุกต์ใช้ทดสอบความน่าจะเป็นของสมมติฐานหลักโดยใช้ข้อมูลที่ได้มาจากประสบการณ์หรือการสำรวจ หากข้อมูลตกอยู่ในพื้นที่ที่ 1 ถึง 3 หรือช่วงร้อยละ 68 ถึง 99.7 แสดงว่า ข้อมูลมีลักษณะของการกระจายหรือแจกแจงใกล้เคียงกับกับปรากฏการณ์หรือข้อมูลที่เกิดขึ้นตามธรรมชาติ จึงยอมรับสมมติฐานหลัก

การอ้างอิงข้อมูลจากตัวอย่างในการทดสอบความน่าจะเป็นของสมมติฐานในแต่ละศาสตร์มีข้อตกลงหรือการยอมรับแตกต่างกัน ด้านวิทยาศาสตร์ข้อมูลส่วนใหญ่ได้มาจากห้องปฏิบัติการที่มีการควบคุมและจัดสภาพแวดล้อมของตัวแปรให้มีความใกล้เคียงกับธรรมชาติได้มากที่สุด จึงใช้ค่าร้อยละ 99.7 หรือกำหนดค่านัยสำคัญทางสถิติ (significance level) = 0.01 ในการทดสอบสมมติฐานหลัก ส่วนด้านสังคมศาสตร์ข้อมูลส่วนใหญ่ได้มาจากสังคมที่มีความไม่แน่นอนอีกทั้งไม่สามารถควบคุมบริบทให้ปราศจากการรบกวนจากปัจจัยต่างๆได้ จึงใช้ค่าร้อยละ 95 หรือกำหนดค่านัยสำคัญทางสถิติ (significance level) = 0.05 ในการทดสอบสมมติฐานหลัก

3.3 การยอมรับและปฏิเสธสมมติฐาน

การตัดสินใจที่จะปฏิเสธหรือยอมรับสมมติฐานหลัก (H_0) ขึ้นอยู่กับค่าสถิติที่เราคำนวณได้จากข้อมูลที่สังเกตุได้จากตัวอย่าง (เช่น ค่าของ X^2, Z, t) โดยนำค่าสถิติที่คำนวณได้ไปเทียบกับ ค่าวิกฤต (critical value)[10] เพื่อดูว่าค่าสถิตินั้นตกอยู่ในเขตวิกฤต (critical region) หรือไม่ (ค่าวิกฤตมักหาโดยการเปิดตารางสถิติ แต่ปัจจุบันผลที่ได้จากการวิเคราะห์ข้อมูลด้วยโปรแกรมคอมพิวเตอร์จะให้ค่าสถิติสำหรับใช้ตัดสินใจปฏิเสธหรือยอมรับได้เลย)

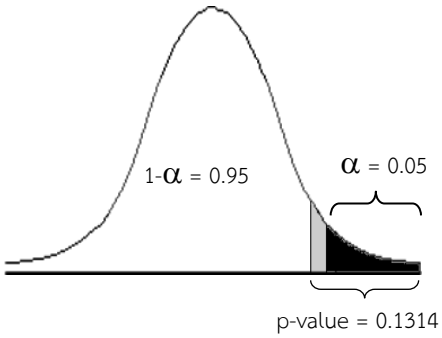


รูปที่ 3.2 เขตวิกฤต

[10] ค่าวิกฤต คือ ค่าที่แบ่งพื้นที่ใต้โค้งความน่าจะเป็นออกเป็น 2 ส่วน คือ ส่วนแรกเรียกว่า เขตวิกฤต (critical/rejection region) และส่วนที่สองเรียกว่า เขตการยอมรับ (acceptance region) ซึ่งเป็นบริเวณที่ทำให้ยอมรับ H_0 ถ้าค่าสถิติที่คำนวณจากข้อมูลของตัวอย่างมีค่าตกอยู่ในเขตวิกฤต (critical region) เราจะปฏิเสธ H_0 และยอมรับ H_1 แต่ถ้ามีค่าตกอยู่ในเขตการยอมรับ (acceptance region) เราจะยอมรับ H_0

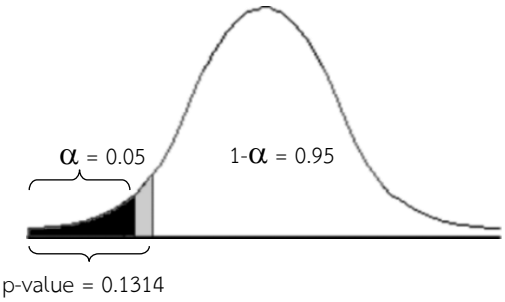
ค่าสถิติที่คำนวณและนำไปเปรียบเทียบกับค่าวิกฤตว่าตกอยู่ในเขตวิกฤติหรือไม่ คือ ค่าความน่าจะเป็น (probability value: p-value) ของข้อมูลที่อยู่นอกพื้นที่ของข้อมูลที่ได้จากการสังเกต (observed) หรือ Acceptance Region ตามที่คาดคะเนไว้ว่าจะยอมรับสมมติฐานหลัก (H_0) โดยค่าความน่าจะเป็นมีค่าอยู่ระหว่าง 0 ถึง 1 ก่อนการทดสอบสมมติฐานผู้วิจัยต้องกำหนดค่าวิกฤติหรือค่านัยสำคัญทางสถิติ (significance: Sig.)[11] หรือค่า α ที่จะปฏิเสธสมมติฐานหลักก่อน เช่น α = 0.01 (1%) หรือ α = 0.05 (5%) ถ้าคำนวณได้ค่าความน่าจะเป็นอยู่นอกเขตวิกฤติหรือมากกว่าค่า α เราจะยอมรับสมมติฐานหลัก (H_0) แต่ถ้าคำนวณได้ค่าความน่าจะเป็นตกอยู่ในเขตวิกฤติหรือได้ค่าความน่าจะเป็นเท่ากับหรือน้อยกว่าค่า α เราจะปฏิเสธสมมติฐานหลัก (H_0) และยอมรับสมมติฐานทางเลือก (H_1) ดังนี้

1. การทดสอบแบบหางเดียวทางขวา -> ยอมรับ (H_0 เพราะ P-value > α)



รูปที่ 3.3 ค่าวิกฤติการทดสอบแบบหางเดียวทางขวา

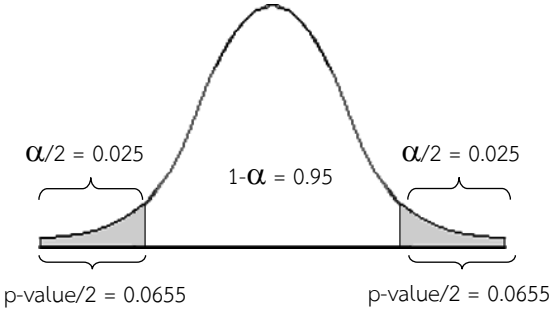
2. การทดสอบแบบหางเดียวทางซ้าย -> ยอมรับH_0 เพราะ P-value > α



รูปที่ 3.4 ค่าวิกฤติการทดสอบแบบหางเดียวทางซ้าย

[11] นักสังคมศาสตร์ส่วนใหญ่ยอมรับหรือนิยมใช้ค่านัยสำคัญทางสถิติที่ระดับ 0.05 (Halter, 2017, p. 14)

3. การทดสอบแบบ 2 หาง -> ยอมรับ H_0 เพราะ P-value > α



รูปที่ 3.5 ค่าวิกฤติการทดสอบแบบ 2 หาง

การใช้โปรแกรมคอมพิวเตอร์ทดสอบสมมติฐานด้วยสถิติ ผลการวิเคราะห์สำหรับใช้ตัดสินใจยอมรับหรือปฏิเสธสมมติฐานหลักมีการแสดงผลลัพธ์หลายรูปแบบ เช่น แสดงด้วยคำว่า p-value หรือบางโปรแกรมแสดงด้วยคำว่า Sig. เพื่อให้นักวิจัยนำไปเทียบกับค่านัยสำคัญทางสถิติ (α) ที่กำหนด (0.05, 0.25, 0.00) เพื่อเป็นจุดแบ่งระหว่างพื้นที่ยอมรับ (acceptance region) กับพื้นที่วิกฤติ/ปฏิเสธ (critical/rejection region) ที่เขียนเป็นเกณฑ์ไว้ก่อนการวิเคราะห์ข้อมูล 2 รูปแบบดังต่อไปนี้

เกณฑ์ในการทดสอบสมมุติฐาน คือ หากการทดสอบสัดส่วนของข้อมูลด้วยวิธี.....ได้ค่าความน่าจะเป็น (p-value) มากกว่าค่านัยสำคัญทางสถิติ (α = 0.05) จะยอมรับสมมติฐานหลัก (H_0) แต่หากได้ค่าความน่าจะเป็น (p-value) น้อยกว่าหรือเท่ากับค่านัยสำคัญทางสถิติ (α = 0.05) จะปฏิเสธสมมติฐานหลัก (H_0) และยอมรับสมมติฐานทางเลือก (H_1)

เกณฑ์ในการทดสอบสมมุติฐาน คือ หากการทดสอบสัดส่วนของข้อมูลด้วยวิธี.....ได้ค่าความน่าจะเป็น (p-value > 0.05) ตกอยู่ในพื้นที่การยอมรับ จะยอมรับสมมติฐานหลัก (H_0) แต่หากได้ค่าความน่าจะเป็น (p-value =< 0.05) ตกอยู่ในพื้นที่การปฏิเสธ จะปฏิเสธสมมติฐานหลัก (H_0) และยอมรับสมมติฐานทางเลือก (H_1)

การแปลความหมายค่าความน่าจะเป็น (p-value) ต้องระมัดระวังการตีความผิดพลาด ระดับนัยสำคัญทางสถิติ (significance level) = 0.05 หรือ 0.01 หมายถึง ความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนแบบที่ 1 (type I error) ร้อยละ 5 หรือร้อยละ 1 หรือตีความได้ว่า ถ้ายอมรับสมมติฐาน

หลักหรือสมมติฐานหลักเป็นจริง มีโอกาสแบบสุ่ม (randomness) ที่จะเกิดการปฏิเสธสมมติฐานหลัก 5 หรือ 1 ครั้ง (Orloff & Bloom, 2014)

4. ความคลาดเคลื่อนของผลการวิจัย

ผู้วิจัยที่เลือกใช้วิธีการวิจัยเชิงปริมาณนอกจากมีความรู้ในระเบียบวิธีการวิจัยและการสร้างเครื่องมือในการรวบรวมข้อมูลเป็นอย่างดีแล้ว ยังต้องมีความรู้และความเข้าใจในการใช้สถิติเป็นอย่างดี รวมถึงต้องนำเอาข้อมูลมาใช้ให้ถูกต้องตามข้อตกลงเบื้องต้น (basic assumptions) ของแต่ละสถิติด้วย เพื่อป้องกันไม่ให้เกิดความคลาดเคลื่อนในการอ้างอิง (inferential errors) ไปสู่แนวคิดหรือทฤษฎีที่ต้องการพิสูจน์หรือตรวจสอบ

4.1 ประเภทของความคลาดเคลื่อนของผลการวิจัย

ความคลาดเคลื่อนในการวิจัยเชิงปริมาณเป็นห่วงโซ่ต่อเนื่องกันและเกิดมาจากหลายสาเหตุแต่แบ่งได้เป็น 2 กลุ่ม ดังนี้ (Lash, Fox, & Fink, 2009: 13)

1. ความคลาดเคลื่อนเชิงสถิติ (Statistical error) หรือที่เรียกว่า ความคลาดเคลื่อนจากการเลือกตัวอย่าง (sampling error) มีสาเหตุเกิดมาจากตัวอย่างที่ได้มาจากประชากรทั้งด้านขนาดตัวอย่าง (sample size) และการเลือกตัวอย่าง (sampling)

2. ความคลาดเคลื่อนเชิงระบบ (Systematic error) หรือที่เรียกว่า ความคลาดเคลื่อนจากเครื่องมือ (instrument error) มีสาเหตุเกิดมาจากคุณภาพของเครื่องมือวัดทั้งด้านความเที่ยงตรง (validity) และความน่าเชื่อถือ (reliability)

4.2 ปัจจัยที่มีผลต่อความคลาดเคลื่อนของผลการวิจัย

ความคลาดเคลื่อนที่เกิดขึ้นอาจไม่สามารถกำจัดออกไปได้จากการวิจัย แต่สามารถลดความเสี่ยงของการเกิดความคลาดเคลื่อนได้โดยการเพิ่มอำนาจของสถิติ (statistical power) หรือที่เรียกว่า อำนาจในการทดสอบสมมติฐาน (power of a hypothesis test) ที่เป็นความน่าจะเป็น (probability) ของความถูกต้องที่จะปฏิเสธสมมติฐานหลักที่เป็นเท็จ (H_0=false) หรือปฏิเสธสมมติฐานหลักเมื่อสมมติฐานทางเลือกเป็นจริง (H_1=true) หรือเป็นดัชนีความน่าจะเป็นที่จะไม่เกิดความความคลาดเคลื่อนแบบที่ 2

หากอำนาจของสถิติเพิ่มขึ้นหรือเข้าใกล้ 1 โอกาสที่จะเกิดความคลาดเคลื่อนแบบที่ 2 จะลดลง ซึ่งอำนาจของสถิติมีความเกี่ยวข้องกับ 3 ปัจจัย ดังนี้

1. ระดับนัยสำคัญทางสถิติ (Statistical significance) คือ ค่าความน่าจะเป็น (p-value) ที่ใช้กำหนดเกณฑ์ในการปฏิเสธสมมติฐานหลัก (H_0) เช่น 0.05 หรือ 0.01 หรือ 0.001

2. ขนาดของอิทธิพล (Effect size) คือ ขนาด (magnitude) ของค่าทางสถิติที่ได้จากทดสอบเปรียบเทียบหรือทดสอบความสัมพันธ์ (relationship) ระหว่าง 2 ตัวแปร เช่น ค่าเฉลี่ย ค่าสัมประสิทธิ์สหสัมพันธ์ ค่าสัมประสิทธิ์การถดถอย หากขนาดอิทธิพลมีมากแสดงว่าปรากฏการณ์ที่ศึกษามีความสัมพันธ์กันมากกว่าขนาดอิทธิพบที่น้อยกว่า หรืออาจกล่าวได้ว่า มีโอกาสมากที่จะไม่เกิดทับกัน (nonoverlap) ของคะแนนระหว่างกลุ่มที่ทดสอบ

3. ขนาดตัวอย่าง (Sample size) การเลือกตัวอย่างมีวัตถุประสงค์เพื่อใช้ในการอ้างอิง (inference) จากตัวอย่างไปสู่ประชากรที่ทราบจำนวน (finite population) ด้วยวิธีการประมาณ (estimate) จากค่าพารามิเตอร์ของประชากร (population parameters) เช่น จำนวนรวม (total) ค่าเฉลี่ย (mean) และสัดส่วน (proportion) ขนาดตัวอย่างจึงเป็นองค์ประกอบที่สำคัญในการออกแบบการวิจัย จึงต้องมีจำนวนที่เพียงพอและมีผลทำให้เกิดผลกระทบ (effect) ต่ออำนาจของสถิตหรือระดับของนัยสำคัญทางสถิติ

ความคลาดเคลื่อนในการสรุปผลการทดสอบสมมติฐานหรือผลการวิจัยมีความเป็นไปได้สองผลลัพธ์ (two possible outcomes) คือ สมมติฐานหลักอาจได้รับการยอมรับหรือปฏิเสธ ดังนี้

1. ความคลาดเคลื่อนแบบที่ 1 (Type I error) คือ ปฏิเสธสมมติฐานหลักที่เป็นจริง (H_0 = true)

2. ความคลาดเคลื่อนแบบที่ 2 (Type II error) คือ ยอมรับสมมติฐานหลักที่เป็นเท็จ (H_0 = false)

ตารางที่ 3.5 การสรุปสมมติฐาน

	ยอมรับ (Accept) H_0	**ปฏิเสธ (Reject) H_0**
H_0 = จริง (True)	สรุปถูกต้อง	ความคลาดเคลื่อนแบบที่ 1 (Type II Error)
H_0 = เท็จ (False)	ความคลาดเคลื่อนแบบที่ 2 (Type II Error)	สรุปถูกต้อง

สรุป

สถิติที่ใช้ในการวิเคราะห์ข้อมูลการวิจัยประกอบด้วยสถิติเชิงพรรณนาและสถิติเชิงอ้างอิง ข้อตกลงเบื้องต้นที่สำคัญของการเลือกใช้สถิติแต่ละประเภท คือ ประเภทของข้อมูล แม้ว่าสถิติเชิงพรรณนาจะไม่สามารถนำผลไปอ้างอิงหรือพยากรณ์ค่าของกลุ่มอื่น ๆ แต่หากเลือกใช้ไม่ตรงกับประเภทของข้อมูลก็ทำให้ได้ผลการวิเคราะห์ที่ผิดได้ ส่วนการทดสอบแบบอิงค่าพารามิเตอร์แม้ว่าจะมีอำนาจจำแนกเชิงสถิติ (statistical power) ที่ทำให้มั่นใจได้ว่า ผลการทดสอบหรือข้อสรุปมีความถูกต้องมาก แต่ถ้าฝืนใช้การทดสอบแบบอิงค่าพารามิเตอร์โดยละเมิดข้อตกลง (violate assumptions) จะมีผลทำให้การค้นพบหรือข้อสรุปที่ได้มีความไม่ถูกต้อง หากข้อมูลไม่ตรงกับข้อตกลงของการทดสอบแบบอิงค่าพารามิเตอร์ควรเลือกใช้การทดสอบแบบไม่อิงค่าพารามิเตอร์แทน

แบบฝึกหัดท้ายบท

ให้ผู้เรียนค้นหางานวิจัยเชิงปริมาณด้านสังคมศาสตร์จากฐานข้อมูลของห้องสมุดที่สนใจ แล้วตอบคำถาม ดังนี้

1. งานวิจัยที่ค้นได้วิเคราะห์ข้อมูลด้วยสถิติอะไรบ้าง
2. สถิติแต่ละตัวในงานวิจัยที่ค้นได้จัดเป็นสถิติประเภทใด

ส่วนที่ 2

ความรู้เบื้องต้นเกี่ยวกับการใช้โปรแกรมทางสถิติ

บทที่ 4
แบบสอบถามและข้อมูล

ความมุ่งหมายของบทเรียน

ผู้เรียนสามารถแปลงข้อมูลจากแบบสอบถามไปเป็นแฟ้มข้อมูลได้

เนื้อหา

1. ความสัมพันธ์ระหว่างแบบสอบถามกับแฟ้มข้อมูล
2. การสร้างตัวแปร

วิธีสอนและกิจกรรมการเรียนการสอน

บรรยาย แสดงตัวอย่าง แนะนำการแปลงแบบสอบถามให้เป็นข้อมูลที่คอมพิวเตอร์สามารถอ่านได้ และทำแบบฝึกหัดท้ายบท

สื่อและอุปกรณ์การสอน

เอกสารคำสอน ภาพเลื่อน คอมพิวเตอร์ โปรแกรมสอนออนไลน์ (สอนแบบออนไลน์) และเครื่องฉาย (สอนในชั้นเรียน)

งานที่มอบหมาย

ทำแบบฝึกหัดท้ายบท

การวัดผลและประเมินผล

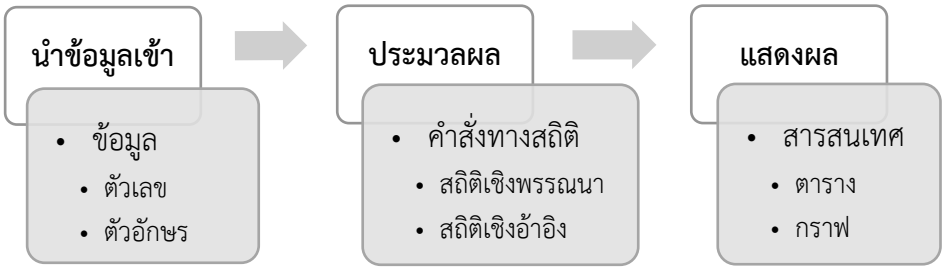
การเข้าชั้นเรียน สังเกตความสนใจจากการเรียน สอบถามความเข้าใจจากการบรรยาย และการทำแบบฝึกหัดท้ายบท

เนื้อหา

การนำข้อมูลที่เก็บรวบรวมด้วยแบบสอบถามหรือแบบฟอร์มมาประมวลผลด้วยคอมพิวเตอร์ ต้องแปลงรูปข้อมูลที่เก็บรวบรวมมาให้อยู่ในรูปข้อมูลที่คอมพิวเตอร์สามารถอ่านได้ (machine-readable data) ข้อมูลที่สามารถใช้ในกระบวนการประมวลผลด้วยโปรแกรมคอมพิวเตอร์ ต้องเป็นข้อมูลดิจิทัลหรือข้อมูลเชิงเลข (digital) ดังนั้นการนำข้อมูลที่ได้จากแบบสอบถามไปประมวลผลหรือวิเคราะห์ข้อมูล ต้องทำการเปลี่ยนรูปข้อมูลในแบบสอบถามให้เป็นข้อมูลในรูปแบบที่โปรแกรมสามารถนำไปประมวลผลได้

1. ความสัมพันธ์ระหว่างแบบสอบถามกับแฟ้มข้อมูล

คอมพิวเตอร์มีกระบวนการประมวลผล คือ นำข้อมูลเข้า (input) จากแฟ้มข้อมูล -> ประมวลผล (processing) ตามเงื่อนไข -> แสดงผล (display) ตามคำสั่ง ดังนี้



รูปที่ 4.1 กระบวนการประมวลผลของคอมพิวเตอร์

การวิเคราะห์ข้อมูลจากแบบสอบถาม (questionnaire) ต้องแปลงข้อมูลจากแบบสอบถามไปเก็บไว้ในรูปของแฟ้มข้อมูล (file) ที่เครื่องจักร (คอมพิวเตอร์) สามารถอ่าน (machine readable) และนำไปประมวลผลได้ โดยแฟ้มข้อมูลต้องมีส่วนประกอบที่สำคัญ คือ กรณีศึกษา/ตัวอย่าง (case) ตัวแปร (variable) และค่าตัวแปร (value)

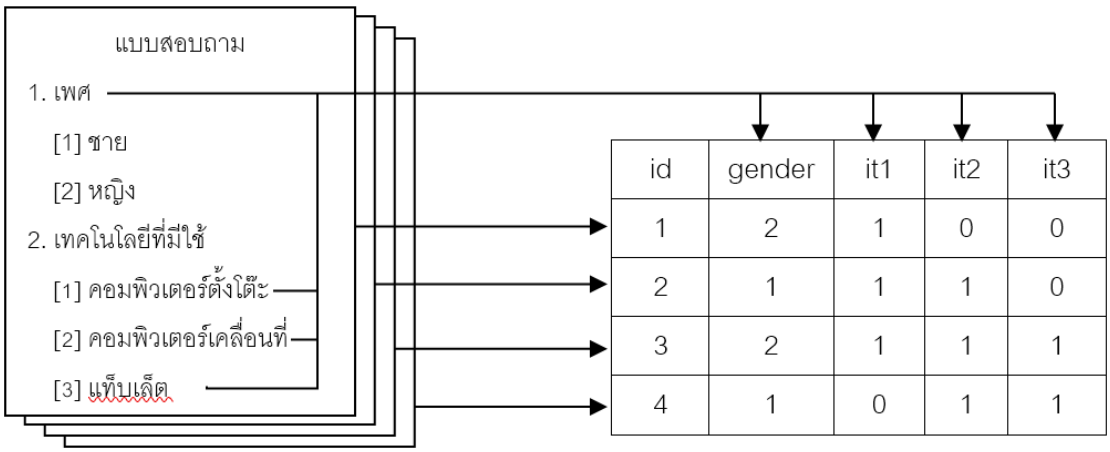
id	gender	age	ICT1	ICT2	ICT3
1	2	20	1	0	0
2	1	25	1	1	0
3	2	22	1	1	1
4	1	18	0	1	1

รูปที่ 4.2 ส่วนประกอบของชุดข้อมูล (data set)

ก่อนออกแบบแฟ้มข้อมูลต้องทำความเข้าใจความสัมพันธ์ระหว่างส่วนประกอบของแบบสอบถามกับส่วนประกอบของแฟ้มข้อมูล ดังนี้

1. ข้อมูลของแบบสอบถามแต่ละชุด คือ กรณีศึกษา/ตัวอย่าง (case)
2. ข้อคำถามแต่ละคำถาม คือ ตัวแปร (variable)
3. คำตอบของแต่ละข้อคำถาม คือ ค่าตัวแปร (value)

เพื่อให้เห็นภาพการเปลี่ยนรูปจากแบบสอบถามไปเป็นแฟ้มข้อมูล ในรูปของตาราง (table) โดยมีแถว (row) เป็นชุดข้อมูลของแบบสอบถามแต่ละฉบับ มีหัวของแต่ละสดมภ์ (column) เป็นชื่อตัวแปร และมีช่อง (cell) เป็นข้อมูลของแต่ละตัวแปร จึงเขียนเป็นภาพได้ดังนี้



id	gender	it1	it2	it3
1	2	1	0	0
2	1	1	1	0
3	2	1	1	1
4	1	0	1	1

รูปที่ 4.3 การเปลี่ยนรูปแบบสอบถามเป็นแฟ้มข้อมูล

แฟ้มข้อมูลที่คอมพิวเตอร์สามารถอ่านและนำเข้าโปรแกรมอาจไม่อยู่ในรูปของตารางก็ได้ เพราะงานวิจัยบางเรื่องได้ข้อมูลจากฐานข้อมูล (database) หรือรวบรวมด้วยเครื่องจักร (machine) แต่ต้อง

จัดเก็บข้อมูลเป็นแถวๆ และมีตัวคั่น (delimiter) เช่น , (comma) ; (semicolon) " " (quotation marks) (tab) เพื่อแบ่งข้อมูลออกเป็นส่วนๆ เช่น แฟ้มข้อมูล *.csv (comma-separated value) หรือ *.txt (text) ดังนั้นจึงต้องเปิดดูรูปแบบของข้อมูลก่อนด้วยโปรแกรมแก้ไขข้อความ (Text Editor) เช่น Notepad ดังตัวอย่างในภาพ

```
id,location,gender,region,national,race,v3,v4,v5,
001,1,1,2,British,Chinese,83,2,2,1,1, , ,3,2,70,2,1
002,1,1,1,Japanese,Japan,20,1,2,1,1, , ,3,2,10,3
```

รูปที่ 4.4 ข้อมูลในแฟ้มข้อมูลนามสกุล CSV แบบมีตัวคั่นด้วยเครื่องหมายจุลภาค

การรวบรวมข้อมูลในการวิจัยเชิงสำรวจส่วนใหญ่นิยมออกแบบข้อคำถามเป็นแบบปลายปิด (close end) มากกว่าคำถามแบบปลายเปิด (open end) เพราะคำถามแบบปลายปิดแปลงข้อมูลไปวิเคราะห์ง่ายกว่าคำถามแบบปลายเปิด เพราะมีคำตอบหรือค่าตัวแปร (value) ที่กำหนดไว้แล้ว แต่เนื่องจากคำถามมีหลายรูปแบบ ดังนั้นการสร้างตัวแปรจากแบบสอบถามจึงมีวิธีการที่ไม่เหมือนกัน ดังนี้

ตารางที่ 4.1 การสร้างตัวแปรจากแบบสอบถาม

เครื่องมือรวบรวมข้อมูล	**โปรแกรมวิเคราะห์ข้อมูล/แฟ้มข้อมูล**
	แบบสอบถามแบบปลายปิด (คำตอบเดียว)
แบบสอบถาม	-> กรณีศึกษา/ตัวอย่าง (case)
ข้อคำถาม	-> ชื่อตัวแปร (variable name) และคำอธิบายตัวแปร (variable label)
คำตอบ	-> ค่าตัวแปร (value)
	แบบสอบถามแบบปลายปิด (หลายคำตอบ)
แบบสอบถาม	-> กรณีศึกษา/ตัวอย่าง (case)
ข้อคำถาม	-> -
คำตอบ	-> ชื่อตัวแปร (variable name) หรือ ค่าตัวแปร (value)
	แบบสอบถามแบบปลายเปิด
แบบสอบถาม	-> กรณีศึกษา/ตัวอย่าง (case)
ข้อคำถาม	-> -
คำตอบ	-> ชื่อตัวแปร (variable name) หรือ ค่าตัวแปร (value)

การบันทึกข้อมูลลงในแฟ้มข้อมูล ถ้าเป็นข้อคำถามที่มีคำตอบเดียว ก็นำเอาคำตอบมาสร้างเป็นตัวแปรได้เลย แต่หากเป็นข้อคำถามที่มีหลายคำตอบ ต้องใช้คำตอบมาสร้างเป็นตัวแปร ดังนี้

1. คำถามแบบคำตอบเดียว

ตัวอย่าง เพศ สร้างเป็นตัวแปรได้ 1 ตัวแปร คือ gender

ตัวอย่าง อายุ....ปี สร้างเป็นตัวแปรได้ 1 ตัว คือ age

2. คำถามแบบหลายคำตอบ

ตัวอย่าง เทคโนโลยีที่มีใช้ [1] คอมพิวเตอร์ [2] โทรศัพท์ [3] อินเทอร์เน็ต ต้องสร้างเป็นตัวแปรหลายตัวแปร คือ ICT1 ICT2 และ ICT3

ตัวอย่าง เชื่อชาติ .....สัญชาติ..... สร้างเป็นตัวแปรได้ 2 ตัว คือ Nation และ Race

แม้ว่าแบบสอบถามจะมีข้อคำถามแบบปลายปิดและปลายเปิด มีคำตอบเดียวและหลายคำตอบตามที่กล่าวมาแล้ว แต่แบบสอบถามยังมีข้อคำถามที่หลากหลายรูปแบบ จึงรวบรวมข้อคำถามที่นิยมใช้กันมาออกแบบเป็นแฟ้มข้อมูลตัวอย่าง ดังต่อไปนี้

1.1 คำถามแบบเลือกตอบ

คำถาม	ชื่อตัวแปร	คำอธิบายตัวแปร	ค่าตัวแปรและคำอธิบาย
1. เพศ [1] ชาย [2] หญิง	Gender	Gender	1 = Male 2 = Female
2. ที่อยู่ [1] เมือง [2] ชานเมือง [3] ชนบท	Residence	Residence	1 = Urban 2 = Suburb 3 = Rural

ตัวอย่างแฟ้มข้อมูลคำถามแบบเลือกตอบ

ID	Gender	Residence
1	1	2
2	2	1

1.2 คำถามแบบปลายเปิด

คำถาม	ชื่อตัวแปร	คำอธิบายตัวแปร	ค่าตัวแปรและคำอธิบาย
1. สัญชาติ .....และเชื้อชาติ.....ของท่าน (โปรดระบุ)	Nation	Nation of Responder	-
	Race	Race of Responder	-
2. ระยะเวลาที่อาศัยอยู่ในประเทศไทย.....ปี.....เดือน	Time	Stay in Thai (month)	1-999

ตัวอย่างแฟ้มข้อมูลคำถามแบบปลายเปิด

Nation	Race	Time
British	Chinese	6
Japan	Japan	10
Indian	Hindo	36

1.3 คำถามแบบประเมินค่า

คำถาม	ชื่อตัวแปร	คำอธิบายตัวแปร	ค่าตัวแปรและคำอธิบาย
ความพึงพอใจที่มีต่อการใช้บริการผ่านเว็บไซต์	eService1	Expect of Services	0 = No
			1 = Least
	eService2	Convenience of Services	2 = Little
			3 = Nature
	eService3	Trust of Services	4 = More
			5 = Most

ข้อ	การใช้บริการ	ระดับคะแนน					
		5	4	3	2	1	0
1	มีบริการผ่านเว็บไซต์ตรงกับที่คาดหวังไว้						
2	มีความสะดวกเข้าไปใช้งานในส่วนต่าง ๆ ของเว็บไซต์						
3	มีความไว้วางใจเมื่อทำธุรกรรมผ่านเว็บไซต์						

ตัวอย่างแฟ้มข้อมูลคำถามแบบประเมินค่า

eService1	eService2	eService3
5	2	4
2	1	3

1.4 คำถามแบบหลายคำตอบ

คำถาม	ชื่อตัวแปร	คำอธิบายตัวแปร	ค่าตัวแปรและคำอธิบาย
เทคโนโลยีสารสนเทศและการสื่อสารที่ท่านมีใช้ในบ้าน (ตอบได้มากกว่า 1 ข้อ)			
ข้อ / คำถาม / มี / ไม่มี			
1 คอมพิวเตอร์แบบเคลื่อนที่	ICT1	Laptop PC	
2 คอมพิวเตอร์แบบแท็บเล็ต	ICT2	Tablet PC	0 = No
3 คอมพิวเตอร์แบบสวมใส่	ICT3	Wearable PC	1 = Yes
4 โทรศัพท์แบบอัจฉริยะ	ICT4	Smartphone	

ตัวอย่างแฟ้มข้อมูลคำถามแบบหลายคำตอบ

ICT1	ICT2	ICT3	ICT4	ICT5
0	1	1	0	1
1	1	1	1	1

1.5 คำถามแบบจัดอันดับ

คำถาม	ชื่อตัวแปร	คำอธิบายตัวแปร	ค่าตัวแปรและคำอธิบาย
ท่านใช้อินเทอร์เน็ตที่ใดมากที่สุด 3 อันดับ	Place1	Home	1 = Least
	Place2	Relation	2 = More
	Place3	Office	3 = Most
	Place4	School	
	หรือ		
	Use1	Last Use	1 = Home
	Use2	More Use	2 = Relation
	Use3	Most Use	3 = Office
			4 = School

ข้อ	คำถาม	อันดับ		
		มากที่สุด	น้อย	น้อยที่สุด
1	บ้านตนเอง			
2	บ้านญาติ/เพื่อน			
3	ที่ทำงาน			
4	โรงเรียน/มหาวิทยาลัย			

ตัวอย่างแฟ้มข้อมูลคำถามแบบจัดอันดับ

Place1	Place 2	Place 3	Place 4		USE1	USE2	USE3
3		2	1		2	3	1
2		3		หรือ	2	1	3
0	1		3		3	1	4

1.6 คำถามแบบเปรียบเทียบ

ข้อคำถาม												ชื่อตัวแปร	คำอธิบายตัวแปร	ค่าตัวแปรและคำอธิบาย
สำนักงาน						คำถาม	ออนไลน์							
ระดับคะแนน							ระดับคะแนน							
5	4	3	2	1	0		5	4	3	2	1	0			
						1. ความง่ายในการดำเนินการ							ServOff1 ServOn1	Ease of Office Ease of Web	0 = No 1 = Least 2 = Little 3 = Nature 4 = More 5 = Most
						2. ความประหยัดในการใช้บริการ							ServOff2 ServOn2	Saving to Office Saving to Web	

ตัวอย่างแฟ้มข้อมูลคำถามแบบเปรียบเทียบ

ServOff1	ServOn1	ServOff2	ServOn2
3	5	2	4
2	3	3	5

1.7 คำถามแบบถูกผิด

ข้อคำถาม	ชื่อตัวแปร	คำอธิบายตัวแปร	ค่าตัวแปรและคำอธิบาย
1. ความรู้เกี่ยวกับกฎและระเบียบที่เกี่ยวกับการใช้เทคโนโลยีสารสนเทศ			
ข้อ: 1 — คำถาม: ข้อมูลอิเล็กทรอนิกส์ไม่สามารถใช้เป็นพยานหลักฐานในการพิจารณาคดีทางกฎหมายได้ — ถูก / ไม่แน่ใจ / ผิด	Know1	Evident in Law	0 = 0 คะแนน 1 = 1 คะแนน
ข้อ: 2 — คำถาม: การยอมรับในความเป็นเจ้าของข้อมูลอิเล็กทรอนิกส์ต้องมีลายมือชื่ออิเล็กทรอนิกส์ด้วย — ถูก / ไม่แน่ใจ / ผิด	Know2	eSignature of Owner	
ข้อ: 3 — คำถาม: หน่วยงานของรัฐสามารถปฏิเสธความรับผิดชอบทางกฎหมายจากข้อความในข้อมูลอิเล็กทรอนิกส์ได้ — ถูก / ไม่แน่ใจ / ผิด	Know3	Disclaimer of Government	

ตัวอย่างแฟ้มข้อมูลคำถามแบบถูกผิด

Know1	Know2	Know3
1	1	0
0	0	0
1	1	1

เฉลยคำตอบ

ข้อ 1 ถูก = 1 ไม่แน่ใจ = 0 และผิด = 0

ข้อ 2 ถูก = 1 ไม่แน่ใจ = 0 และผิด = 0

ข้อ 3 ถูก = 0 ไม่แน่ใจ = 0 และผิด = 1

แบบสอบถาม 1 ชุด อาจมีข้อคำถามเป็นแบบคำตอบเดียว หรือเป็นแบบสอบถามที่มีทั้งข้อคำถามแบบคำตอบเดียวและหลายคำตอบรวมกัน การออกแบบแบบสอบถามต้องคำนึงถึงการนำเอาข้อคำถามและคำตอบไปใช้ในการสร้างตัวแปรและค่าตัวแปรในแต่ละแฟ้มข้อมูล รวมถึงระดับการวัดตัวของตัวแปร (variable scale) สำหรับใช้สถิติในการวิเคราะห์ด้วย

จากตัวอย่างแบบสอบถาม ข้อคำถามแบบประเมินค่า (rating scale) ส่วนใหญ่กำหนดการประเมินค่าเป็น มากที่สุด มาก ปานกลาง น้อย น้อยที่สุด แม้ว่าจะมีการกำหนดตัวเลขควบคู่ไปกับระดับการประเมิน เช่น 5 4 3 2 1 ด้วย และในการวิเคราะห์ข้อมูลก็มีการอนุโลมให้คำตอบที่ได้จากการประเมินค่าแบบนี้เป็นระดับการวัดแบบมาตราวัด (scale measurement) หรือเป็นค่าต่อเนื่อง (continuous data) เพื่อไม่ให้ละเมิดข้อตกลง (violate assumptions) และสามารถวิเคราะห์ด้วยสถิติสหสัมพันธ์ (correlation) ได้ แต่หากการออกแบบคำตอบให้เป็นคะแนน (score) ตั้งแต่ 0-5 น่าจะทำให้ระดับการวัดของตัวแปรเป็นไปตามข้อตกลง (meet assumptions) จะดีกว่าการอนุโลมเพื่อให้เป็นไปตามเงื่อนไขทางสถิติ

2. การสร้างตัวแปร

การสร้างตัวแปร (variable definition) คือ กำหนดคุณลักษณะของตัวแปร (variable attribute) ให้คอมพิวเตอร์รู้รายละเอียดเกี่ยวกับตัวแปรและสามารถนำข้อมูลไปใช้วิเคราะห์ได้ด้วยโปรแกรมคอมพิวเตอร์ โดยมีรายละเอียดและเงื่อนไขดังต่อไปนี้

1. ชื่อตัวแปร (Variable Name) คือ สัญลักษณ์ (symbol) ที่กำหนดขึ้นเพื่ออธิบายข้อมูล (data) ให้โปรแกรมคอมพิวเตอร์รู้จักและนำไปใช้ในฟังก์ชั่น (function) ในกระบวนการประมวลผล โดยทั่วไปมีเงื่อนไขหรือกฎที่คล้ายกัน ดังนี้

1.1 ชื่อตัวแปรต้องเริ่มต้นด้วยตัวอักษรเท่านั้น (โดยทั่วไปใช้ภาษาอังกฤษ) หลังจากนั้นอาจเป็นตัวอักษร (letter) ตัวเลข (number) อักขระพิเศษ (special character) เช่น @, #, _ หรือ $ ก็ได้

1.2 ชื่อตัวแปรห้ามปิดท้ายด้วย จุด หรือ เส้นใต้อักขระ (underscore)

1.3 ชื่อตัวแปรห้ามเว้นวรรค (blank) และห้ามใช้อักษรพิเศษ เช่น !, ?, ‘ และ *

1.4 ชื่อตัวแปรในแฟ้มข้อมูลเดียวกันต้องไม่เหมือนกันหรือซ้ำกัน

1.5 ชื่อตัวแปรโดยทั่วไปควรมีความยาวไม่เกิน 64 ตัวอักษร หรือ ไบต์ (bye) ภาษาที่ใช้ไบต์เดียว เช่น อังกฤษ เยอรมัน ฝรั่งเศส และสเปน เป็นต้น บางโปรแกรมสามารถกำหนดชื่อตัวแปรได้ยาวมากกว่า 64 ตัวอักษร

1.6 ชื่อตัวแปรห้ามใช้คำสงวน (reserved keyword) เช่น ALL, AND, BY, EQ, GE, GT, LE, LT, NE, NOT, OR, TO และ WITH)

1.7 ชื่อตัวแปรอาจเป็นอักษรใหญ่เล็กหรือตัวเล็ก (upper or lower case) เช่น GENDER หรือ Gender หรือ gender ก็ได้ แต่การใช้อักษรตัวเลขจะสะดวกในการนำไปใช้กับโปรแกรมอื่นๆ

1.8 ชื่อตัวแปรที่มีความยาวมาก จะมีผลกระทบต่อการแสดงผล เช่น แสดงผลไม่ครบ หรือแสดงผลหลายบรรทัด ดังนั้นไม่ควรกำหนดชื่อตัวแปรยาวมากเกิน ควรใช้คำสั้นๆ ที่มีความหมายในตัว หรือใช้คำย่อที่นิยมใช้กันอย่างแพร่หลายและเข้าใจได้ง่ายแทนข้อคำถามจากแบบสอบถามที่เป็นประโยคยาวๆ เช่น เทคโนโลยีสารสนเทศและการสื่อสารที่ท่านมีใช้ในบ้าน ใช้ชื่อตัวแปรว่า ICT หากเป็นเรื่องเดียวกันและมีความเกี่ยวข้องกัน อาจใช้ตัวเลขต่อท้ายตามลำดับ เช่น ICT1 ICT2 ก็ได้

2. ประเภทตัวแปร (Variable Type) คือ ประเภทของข้อมูลหรือค่าตัวแปร (value) ที่ใช้ในการประมวลผล เนื่องจากข้อมูลที่ได้จากกระบวนการรวบรวมข้อมูล (data collection) มีหลายรูปแบบ แต่ในกระบวนการประมวลผล ข้อมูลที่เป็นตัวอักษร (character/text/string/alphanumeric) ต้องทำการนับ (count) เป็นตัวเลข (numeric) แล้วจึงนำไปประมวลผลด้วยตัวดำเนินการทางคณิตศาสตร์ (arithmetic operator) เช่น + - x ÷ ส่วนข้อมูลที่เป็นตัวเลข (numeric) สามารถคำนวณ (calculate) ด้วยตัวดำเนินการทางคณิตศาสตร์ได้เลย อีกทั้งบางโปรแกรมมีระบบแบ่งประเภทตัวแปรตามประเภทของข้อมูลที่กำหนดไว้ตั้งแต่ขั้นตอนนำข้อมูลเข้า (data input) ให้เลือกใช้ให้ตรงกับข้อตกลงเกี่ยวกับประเภทของข้อมูลของแต่ละคำสั่งทางสถิติอีกด้วย ดังนั้นในที่นี้จึงต้องทำความเข้าใจเกี่ยวกับข้อมูลของตัวแปร 2 ประเภทดังต่อไปนี้

ตารางที่ 4.2 คำอธิบายประเภทข้อมูล

ประเภทข้อมูล	คำอธิบาย
Numeric	ข้อมูลที่เป็นตัวเลข มีค่าเป็นตัวเลข มีค่าเป็น + หรือ – ก็ได้
String	ข้อมูลที่ไม่ใช่ตัวเลข เป็นได้ทั้ง ตัวอักษร ข้อความ และตัวอักษรเลข เป็นตัวแปรที่มีค่าเป็นตัวอักษรหรือตัวเลขหรือเครื่องหมายต่างๆ แม้ว่าค่าตัวแปรจะเป็นตัวเลข ก็ไม่สามารถนำไปคำนวณได้ เพราะเป็นตัวอักษรเลข (alphanumeric)

3. จำนวนค่าตัวแปรและทศนิยม (Width and Decimals) ตัวแปรแต่ละตัวจากแบบสอบถามมีค่าตัวแปรมาจากคำตอบที่แตกต่างกัน แต่โปรแกรมกำหนดค่าเริ่มต้น (default) ไว้ให้แล้ว ดังต่อไปนี้

ตารางที่ 4.3 ค่าเริ่มต้นค่าตัวแปรของตัวแปรแต่ละประเภท

ประเภทตัวแปร	จำนวนค่าตัวแปร	ทศนิยม
Numeric	8	2

ประเภทตัวแปร	จำนวนค่าตัวแปร	ทศนิยม
Comma	8	2
Dot	8	2
Scientific notation	8	2
Date	11	0
Dollar	8	2
Custom currency	8	2
String	8	0

4. คำอธิบายตัวแปร (Variable Label) คือ คำอธิบายหรือข้อความขยายชื่อตัวแปร เพื่อให้เข้าใจความหมายของชื่อตัวแปร เพราะโดยทั่วไปมักสร้างชื่อตัวแปรด้วยคำย่อหรือคำที่มีความหมายสั้นๆ แต่ชื่อตัวแปรบางตัวอาจไม่ต้องให้ความหมายก็ได้ เพราะชื่อตัวแปรสามารถสื่อความหมายหรือเข้าใจได้ในตัวเองว่า ตัวแปรนั้นหมายถึงอะไร เช่น age แต่ชื่อตัวแปรที่สร้างขึ้นเองบางครั้งอาจไม่สื่อความหมายจึงต้องให้คำอธิบายตัวแปรไว้ เพื่อให้ผู้ใช้แฟ้มข้อมูลที่ไม่ได้เป็นผู้สร้างแฟ้มข้อมูลหรือการอ่านผลลัพธ์ (output) สามารถเข้าใจได้ว่า ตัวแปรนั้นหมายถึงอะไร ดังตัวอย่างต่อไปนี้

ตารางที่ 4.4 การสร้างคำอธิบายตัวแปร

ชื่อตัวแปร (Variable Name)	คำอธิบายตัวแปร (Variable Label)
Id	Questionnaire Identification
Gender	*กำหนดหรือไม่กำหนดก็ได้*
ICT1	Have the Desktop PC in Home
ICT2	Have the Laptop PC in Home
Place	Place for Surf Internet
Cost	Amount spent on communication per month

5. คำอธิบายค่าตัวแปร (Value Label) การบันทึกข้อมูลลงในแฟ้มข้อมูลที่ง่ายและสะดวกที่สุด คือ การใช้ตัวเลขหรือตัวอักษรเป็นรหัสแทนคำตอบจากแบบสอบถาม แต่การใช้รหัสที่เป็นตัวเลขเป็นสิ่งที่นิยมมากกว่าการใช้ตัวอักษร เพราะขณะทำการบันทึกข้อมูลแป้นพิมพ์ตัวเลขที่อยู่กันเป็นกลุ่มบริเวณเดียวกันสามารถทำการบันทึกข้อมูลได้ง่ายกว่าการใช้แป้นพิมพ์ตัวอักษร เช่น Gender หรือ เพศ

กำหนดให้เลข 1 เป็นรหัสแทน Male หรือ เพศชาย และ 2 เป็นรหัสแทน Female หรือ เพศหญิง ดังนั้น รหัสที่เป็นตัวเลขหรือตัวอักษรจึงต้องให้คำอธิบาย เพื่อความเข้าใจการในการแปลผลลัพธ์

การให้คำนิยมค่าตัวแปร ไม่จำเป็นต้องให้คำอธิบายทุกค่าตัวแปร ค่าตัวแปรที่ควรให้นิยาม คือ ค่าของตัวแปรแบบค่าแบ่งประเภท (ระดับการวัดแบบจัดกลุ่มและจัดอันดับ) ส่วนค่าตัวแปรแบบค่าต่อเนื่อง (ระดับการวัดแบบช่วงและสัดส่วน) เป็นค่าตัวแปรที่มีความหมายอยู่ในตัวอยู่แล้ว ไม่จำเป็นต้องให้คำอธิบาย เช่น ค่าใช้จ่าย มีค่าตัวแปรที่บันทึกไว้เป็น 10,000 เป็นต้น การให้คำอธิบายค่าตัวแปรสามารถแสดงให้เห็นได้เป็นตัวอย่าง ดังต่อไปนี้

ตารางที่ 4.5 การสร้างคำอธิบายค่าตัวแปร

ชื่อตัวแปร (Variable Name)	คำอธิบายค่าตัวแปร (Value Label)
Id	*ไม่ต้องให้ความหมายเพราะลำดับที่แบบสอบถามมีความหมายอยู่แล้ว*
Gender	1 = Male, 2 = Female
ICT1	0 = No, 1 = Yes
ICT2	0 = No, 1 = Yes
Place	1 = Home, 2 = Relation, 3 = Office, 4 = School 5 = Telecenter, 6 = Internet Café
Cost	*ไม่ต้องให้ความหมายเพราะจำนวนค่าใช้จ่ายมีความหมายอยู่ในตัวแล้ว*

ค่าตัวแปรที่ได้จากคำตอบเรื่องเดียวกัน อาจมีค่าตัวแปรที่แตกต่างกัน ขึ้นอยู่กับวัตถุประสงค์ในการเก็บข้อมูล เช่น อายุ ค่าใช้จ่าย สามารถเก็บข้อมูลได้ 2 แบบ ดังนี้

ตารางที่ 4.6 ประเภทของค่าตัวแปร

ชื่อตัวแปร	ค่าตัวแปร	
	ค่าแบ่งกลุ่ม	ค่าต่อเนื่อง
อายุ	(1) ต่ำกว่า 15 ปี (2) 15 – 60 ปี (3) มากกว่า 60 ปี	.....ปี

ชื่อตัวแปร	ค่าตัวแปร	
	ค่าแบ่งกลุ่ม	ค่าต่อเนื่อง
ค่าใช้จ่ายด้านการสื่อสารรายเดือน	(1) น้อยกว่า 1,000 บาท (2) 1,001 – 5,000 บาท (3) มากกว่า 5,000 บาท	.....บาท

ค่าตัวแปรอายุและค่าใช้จ่าย เป็นค่าแบ่งกลุ่มต้องให้คำอธิบาย ส่วนค่าตัวแปรที่เป็นค่าต่อเนื่องไม่จำเป็นต้องให้คำอธิบาย เพราะมีความหมายอยู่ในตัวเองแล้ว

6. ค่าขาดหาย (Missing Value) หรือ ค่าว่าง หรือ ค่าสูญหาย อาจอยู่ในรูปของไม่ค่าตัวแปรในช่องใส่ค่าตัวแปร (cell) หรือ มีค่าตัวแปรแต่มีการกำหนดให้ไม่นำมาใช้ในการคำนวณ

ค่าขาดหาย เกิดขึ้นได้ในทุกๆ งานวิจัย แม้ว่ามีการวางแผนการออกแบบแบบสอบถามให้รองรับคำตอบที่จะทำให้เกิดค่าขาดหายไว้แล้วก็ตาม เช่น ไม่ทราบ ไม่ตอบ ไม่มีความคิดเห็น เป็นต้น คำตอบเหล่านี้ส่วนใหญ่จะไม่นำเข้าไปรวมอยู่ในการประมวลผลข้อมูล เพราะมีผลทำให้ผลลัพธ์ที่ได้จากการคำนวณผิดไป การสร้างค่าขาดหายในแฟ้มข้อมูลสามารถทำได้ 2 แบบ คือ

6.1 ไม่ใส่ค่าใดๆ ลงไปในช่องใส่ค่าตัวแปร (blank) แต่โปรแกรมจะให้ค่าเป็นจุด (.) วิธีนี้โปรแกรมจะกำหนดให้เป็นค่าขาดหายของระบบ (system missing value) และไม่นำเอาเข้ามารวมในการคำนวณโดยอัตโนมัติ

6.2 ใส่รหัสที่กำหนดให้เป็นค่าขาดหาย (user missing value) รหัสตัวเลขที่นิยมนำมาใช้เป็นค่าขาดหาย คือ เลข 9 วิธีการนี้ผู้ใช้โปรแกรมจะต้องกำหนดให้โปรแกรมรู้ว่าเลข 9 คือ ค่าขาดหาย ไม่ต้องนำมารวมในการคำนวณ

การกำหนดรหัสค่าขาดหายต้องสัมพันธ์กับประเภทของข้อมูล หากประเภทข้อมูลเป็นตัวเลข รหัสค่าขาดหายก็ต้องเป็นตัวเลขเช่นกัน

7. ความกว้างของช่องค่าตัวแปรหรือสดมภ์ (Columns) ค่าเริ่มต้น (default) ความกว้างของสดมภ์ที่โปรแกรมกำหนดไว้ คือ 8 ตำแหน่ง หากชื่อตัวแปรมีความยาวไม่เกินที่โปรแกรมกำหนดไว้ การแสดงผลตัวแปรในหน้าต่างแสดงข้อมูล (data view) จะแสดงผลอยู่ในบรรทัดเดียว แต่หากชื่อตัวแปรมีความยาวมาก โปรแกรมจะแสดงผลในบรรทัดใหม่ในลักษณะซ้อนต่อกัน ดังนั้นผู้ใช้งานสามารถกำหนดให้การแสดงผลตัวแปรอยู่ในบรรทัดเดียวกันได้โดยกำหนดความกว้างของสดมภ์ให้มากขึ้น แต่การสร้างชื่อ

ตัวแปรยาวๆ เป็นสิ่งที่ควรหลีกเลี่ยง เพราะไม่เพียงแต่จะทำให้การแสดงผลไม่สวยงามแล้ว การเพิ่มความกว้างของสดมภ์ จะมีผลทำให้การแสดงผลตัวแปรในหน้าต่างทำได้น้อย การดูตัวแปรและค่าตัวแปลในแนวนอนทำได้ไม่สะดวก เพราะต้องใช้แถบเลื่อนหน้าจอ (scroll bar) เลื่อนดูข้อมูลไปมา

8. การจัดตำแหน่งค่าตัวแปร (Align) โปรแกรมมีการออกแบบการแสดงค่าตัวแปรให้เห็นถึงความแตกต่างของประเภทตัวแปรที่สร้างไว้เป็นค่าปริยาย ดังต่อไปนี้

8.1 ตัวแปรที่มีค่าเป็นตัวเลข การแสดงผลค่าตัวแปรในหน้าต่างแสดงข้อมูล ค่าตัวแปรจะแสดงผลชิดด้านซ้ายในช่องใส่ค่าตัวแปร

8.2 ตัวแปรที่มีค่าเป็นตัวอักษร ค่าตัวแปรจะแสดงผลชิดด้านขวาในช่องใส่ค่าตัวแปร

การเปลี่ยนแปลงการแสดงผลตำแหน่งค่าตัวแปรใหม่ สามารถทำได้ 3 แบบ จากรายการเมนู คือ แสดงผลชิดด้านซ้าย (left) แสดงผลชิดด้านขวา (right) และแสดงผลกึ่งกลาง (center) ตามที่ต้องการ แต่การคงค่าปริยายไว้น่าจะช่วยให้การจำแนกประเภทตัวแปรทำได้ง่ายกว่าการจัดตำแหน่งใหม่

9. ระดับการวัด (Measure) โดยทั่วไปตัวแปรในการวิจัยแบ่งระดับการวัดเป็น 4 ระดับ คือ มาตราแบบจัดกลุ่ม มาตราแบบจัดอันดับ มาตราแบบช่วง และมาตราแบบสัดส่วน นำเอาทั้ง 4 ระดับมาจำแนกเป็นตัวแปรเชิงคุณภาพและเชิงปริมาณ หรือตัวแปรแบ่งกลุ่มและตัวแปรต่อเนื่อง เป็นต้น แต่การใช้โปรแกรม PSPP สร้างตัวแปร แบ่งระดับการวัดตัวแปรเป็น 3 แบบ ดังต่อไปนี้

9.1 ระดับการวัดแบบสเกล (Scale)

9.2 ระดับการวัดแบบจัดกลุ่ม (Nominal)

9.3 ระดับการวัดแบบจัดอันดับ (Ordinal)

การสร้างตัวแปรที่มีค่าเป็นตัวเลขและตัวอักษรตามที่กล่าวมาแล้วเป็นการกำหนดประเภทตัวแปรตามคุณลักษณะของตัวแปร แต่ไม่สามารถจำแนกได้ว่า ตัวแปรใดมีระดับการวัดที่เหมาะสมกับการนำเอาไปวิเคราะห์ด้วยสถิติใด และโปรแกรมได้กำหนดค่าปริยายระดับการวัดของตัวแปรแต่ละประเภทไว้ดังต่อไปนี้

ตารางที่ 4.7 ประเภทตัวแปรและระดับการวัด

ประเภทตัวแปร	ระดับการวัด
Numeric	Scale
Comma	Scale
Dot	Scale
Scientific notation	Scale
Date	Scale
Dollar	Scale
Custom currency	Scale
String	Nominal

การกำหนดระดับการวัดตัวแปรไม่เพียงแต่จะช่วยให้โปรแกรมทำการคัดกรองตัวแปรแต่ละตัวเข้าไปในกระบวนการเลือกคำสั่งประมวลผลที่เหมาะสมเท่านั้น แต่ยังช่วยแสดงสัญลักษณ์ระดับการวัดให้ผู้วิเคราะห์ตัดสินใจเลือกใช้ตัวแปรในการเลือกใช้คำสั่งในการคำนวณได้อย่างเหมาะสมอีกด้วย

10. บทบาทของตัวแปร (Role) คือ การกำหนดคุณสมบัติของตัวแปรที่เตรียมไว้ใช้กับบางคำสั่งที่ต้องมีการเตรียมข้อมูล (data preparation) การแทนค่าแบบพหุ (multiple imputation) และการวิเคราะห์แบบจำลอง (model techniques) เพื่อนำไปใช้ในระบบสนทนา (dialog box) ในการติดต่อกับระบบผู้ใช้งานแบบกราฟิก (graphical user interfaces) ในการวิเคราะห์ข้อมูล ดังนี้

ตารางที่ 4.8 ความหมายของบทบาทของตัวแปร

บทบาทของตัวแปร	ความหมาย
Input	ตัวแปรข้อมูลเข้าหรือตัวแปรอิสระ
Output	ตัวแปรผลลัพธ์หรือตัวแปรตาม
None	ไม่กำหนดบทบาท (เป็นค่าเริ่มต้น)
Partition	ตัวแปรที่มีการแบ่งข้อมูลเป็นกลุ่มๆ สำหรับการสอน (training) หรือ การทดสอบ (testing) หรือการตรวจสอบ (validation)
Split	ตัวแปรที่เตรียมไว้สำหรับโปรแกรมอื่นๆ หรือ third party software

สรุป

ข้อมูลที่เก็บรวบรวมด้วยแบบสอบถามและต้องการนำมาประมวลผลด้วยโปรแกรมคอมพิวเตอร์ทางด้านสถิติ ต้องแปลงรูปให้เป็นข้อมูลที่คอมพิวเตอร์สามารถอ่านได้ในรูปของแฟ้มข้อมูลที่มีส่วนประกอบที่สำคัญ คือ กรณีศึกษาหรือตัวอย่าง ตัวแปร และค่าตัวแปร

โดยทั่วไปแฟ้มข้อมูลที่นำข้อมูลเข้าโปรแกรมอาจจัดเก็บข้อมูลเป็นตารางที่ประกอบด้วยแถวที่เป็นชุดข้อมูลของแบบสอบถามแต่ละฉบับ มีหัวแบ่งเป็นสดมภ์เพื่อใช้เป็นชื่อตัวแปร และมีช่องให้ใส่ข้อมูลหรือค่าของแต่ละตัวแปร นอกจากแฟ้มข้อมูลแบบตารางแล้วยังมีแฟ้มข้อมูลอีกประเภทหนึ่งที่คอมพิวเตอร์สามารถนำไปประมวลผลได้ คือ แฟ้มข้อมูลแบบตัวอักษรที่เก็บข้อมูลเป็นเป็นแถวๆ และมีตัวคั่นเป็นเครื่องหมายลูกน้ำ (,) เครื่องหมายอัญประกาศ (" ") เว้นระยะ (tab) เช่น แฟ้มข้อมูล *.csv หรือ *.txt

แบบฝึกหัดท้ายบท

ให้ผู้เรียนฝึกแปลงแบบสอบถามเป็นตัวแปรและข้อมูลดังต่อไปนี้

1. ให้ผู้เรียนสร้างแบบสอบถามมีทั้งข้อคำถามเลือกตอบแบบจัดกลุ่มและจัดอันดับ และข้อคำถามแบบกรอกข้อมูลเป็นตัวเลขและตัวอักษร

2. ให้ผู้เรียนแปลงแบบสอบถามให้เป็นตัวแปรและค่าตัวแปร

บทที่ 5
ความรู้เบื้องต้นเกี่ยวกับโปรแกรมทางสถิติ

ความมุ่งหมายของบทเรียน

ผู้เรียนรู้จักโปรแกรมคอมพิวเตอร์ด้านสถิติสำหรับวิเคราะห์ข้อมูล สามารถติดตั้งและเริ่มใช้งานเบื้องต้นได้

เนื้อหา

1. โปรแกรมสำเร็จรูปทางสถิติ
2. การติดตั้งโปรแกรม PSPP
3. การใช้โปรแกรม PSPP เบื้องต้น
4. ส่วนประกอบของโปรแกรม PSPP
5. เมนูการใช้โปรแกรม PSPP

วิธีสอนและกิจกรรมการเรียนการสอน

บรรยาย แนะนำการติดตั้งและใช้โปรแกรม และทำแบบฝึกหัดท้ายบท

สื่อและอุปกรณ์การสอน

เอกสารคำสอน ภาพเลื่อน คอมพิวเตอร์ โปรแกรม PSPP โปรแกรมสอนออนไลน์ (สอนแบบออนไลน์) และเครื่องฉายภาพโปรเจคเตอร์ (สอนในชั้นเรียน)

งานที่มอบหมาย

ติดตั้งโปรแกรมวิเคราะห์ข้อมูลให้เสร็จสมบูรณ์ และทำแบบฝึกหัดท้ายบท

การวัดผลและประเมินผล

การเข้าชั้นเรียน สังเกตความสนใจจากการเรียน สอบถามความเข้าใจจากการบรรยาย ประเมินจากการติดตั้งโปรแกรม การใช้โปรแกรม และการทำแบบฝึกหัดท้ายบท

เนื้อหา

การวิจัยเชิงปริมาณ ส่วนใหญ่มีการเก็บรวบรวมข้อมูลจำนวนมากระดับหนึ่งตามขนาดของประชากร ข้อมูลที่เก็บรวบรวมมาแล้ว หากไม่สามารถคำนวณและวิเคราะห์อย่างรวดเร็ว ข้อมูลที่เก็บมาอาจล้าสมัยและไม่มีประโยชน์ต่อการวิจัย เพราะคำตอบและผลการทดสอบสมมติฐานการวิจัยไม่ใกล้เคียงสถานการณ์จริงที่เปลี่ยนไปตามการเปลี่ยนแปลงของสังคม ดังนั้นนักวิจัยจึงต้องมีเครื่องมือที่สามารถประมวลผลได้รวดเร็วและวิเคราะห์ข้อมูลได้อย่างซับซ้อน นั่นก็คือ โปรแกรมคอมพิวเตอร์สำหรับการคำนวณและวิเคราะห์ข้อมูล

1. โปรแกรมสำเร็จรูปทางสถิติ

ปัจจุบันโปรแกรมสำเร็จรูปสำหรับใช้วิเคราะห์ข้อมูลมีให้เลือกใช้จำนวนมาก มีทั้งโปรแกรมแบบให้ใช้ฟรีและมีค่าใช้จ่าย และมีการจัดกลุ่มเป็นโปรแกรมรหัสเปิดหรือโอเพนซอร์ซ (open source) โปรแกรมสาธารณะ (public domain) โปรแกรมแจกฟรี (freeware) และโปรแกรมมีลิขสิทธิ์ (proprietary) (Wikipedia, 2020) สำหรับในที่นี้แบ่งเป็น 2 กลุ่ม และมีตัวอย่างโปรแกรมที่นิยมใช้งานกันอย่างแพร่หลาย ดังต่อไปนี้

1. โปรแกรมเชิงพาณิชย์ (Commercial Application) เช่น IBM SPSS (Statistical Package for the Social Sciences) หรือ PASW (Predictive Analytics SoftWare), SAS (Statistical Analysis System)

2. โปรแกรมแบบสาธารณะทั่วไป (Non-commercial Application หรือ GNU (General Public License: GNU-GPL)[12] เช่น PSPP, R, JASP, JAMOVI

โปรแกรมสำเร็จรูปทางสถิติมีการพัฒนาและออกแบบผ่านมาหลายยุค ในอดีตโปรแกรมส่วนใหญ่ที่ติดตั้งอยู่บนเครื่องคอมพิวเตอร์ขนาดใหญ่ (main frame) และคอมพิวเตอร์เครื่องแม่ข่าย (server computer) ผู้ใช้งานต้องส่งข้อมูลและคำสั่งไปทำการประมวลผลผ่านเครื่องคอมพิวเตอร์ลูกข่าย (client computer) แบบข้อความ (command line) จนถึงในระยะแรกของยุคคอมพิวเตอร์ส่วนบุคคล (PCs

[12] ไม่มีลิขสิทธิ์ (copyleft) เปิดเผยและแพร่รายละเอียดของโปรแกรม (open-source) ที่เรียกว่า รหัสต้นฉบับ (source code) เพื่ออนุญาตให้บุคคลทั่วไปสามารถไปทำซ้ำ เผยแพร่ และดัดแปลงซอฟต์แวร์นั้นได้อย่างถูกต้องตามกฎหมาย และโดยเสรี

Era) ต่อมาได้มีการพัฒนาระบบติดต่อกับผู้ใช้งานแบบกราฟิก (graphic user interface: GUI) มีไอคอน (icon) และเมนู (menu) ทำให้สามารถใช้งานได้ง่าย (user friendly) และในยุคหลังคอมพิวเตอร์ส่วนบุคคล (Post-PCs Era) มีหลายโปรแกรมที่สามารถใช้งานจากระบบประมวลผลแบบกลุ่มเมฆ (cloud computing) ด้วยโปรแกรมเรียกดูข้อมูลบนอินเทอร์เน็ต (browser) แบบออนไลน์ (online)

สำหรับเอกสารคำสอนนี้จะแนะนำการใช้เฉพาะโปรแกรม PSPP ที่ติดตั้งบนคอมพิวเตอร์ส่วนบุคคลในระบบปฏิบัติการ Windows เท่านั้น เพื่อให้เรียนรู้และฝึกใช้งานเบื้องต้น โดยมีรายละเอียดดังต่อไปนี้

2. การติดตั้งโปรแกรมทางสถิติ

โปรแกรม PSPP เป็น 1 ในโปรแกรมทางสถิติแบบรหัสเปิด (Open Source Statistical Software: OSSS) ภายใต้สัญญาอนุญาตให้ใช้ซอฟต์แวร์แบบสาธารณะทั่วไปที่เรียกกว่า GNU General Public License (GNU-GPL) มีให้เลือกใช้งานบนเครื่องคอมพิวเตอร์ส่วนบุคคล (PC) หลายระบบปฏิบัติการ (OS)[13] โดยพัฒนาขึ้นมาเพื่อเลียนแบบ (clone) ให้เป็นตัวเลือกใช้งานแทนโปรแกรม SPSS มีทั้งระบบการติดต่อกับผู้ใช้งานแบบกราฟิก (GUI) และระบบคำสั่งแบบข้อความ (command line) ที่คล้ายกัน แต่มีคำสั่งและสถิติให้เลือกใช้วิเคราะห์ข้อมูลน้อยกว่า บางคำสั่งไม่มีเมนูคำสั่งแบบกราฟิกต้องใช้ระบบคำสั่งแบบข้อความแทน เช่น การทดสอบกลุ่มตัวอย่างมากกว่า 2 กลุ่มที่เป็นอิสระจากกัน (K Independent Samples Tests)

ผู้ใช้งานสามารถค้นหาโปรแกรม PSPP จากอินเทอร์เน็ต[14] และทำการติดตั้งโปรแกรมก่อนการใช้งานตามขั้นตอนดังต่อไปนี้

ขั้นที่ 1 การค้นหาและดาวน์โหลดโปรแกรม

1. เปิดโปรแกรมค้นหาข้อมูลบนอินเทอร์เน็ต (browser) เช่น Google Chrome, Microsoft Edge

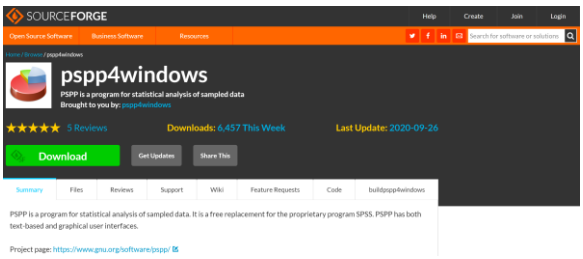
2. พิมพ์คำว่า pspp download

[13] วินโดว์ (Windows) แมคโอเอสเท็น (Max OS X) และยูนิกซ์/ลินุกซ์ (Unix/Linux)

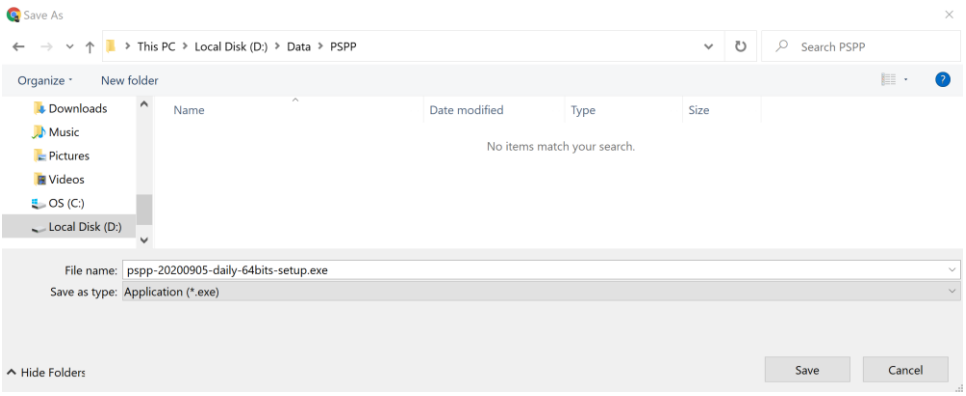
[14] ควรดาวน์โหลด (download) จากเว็บไซต์ GNU Project หรือ SOURCEFORGE เพราะเป็นแหล่งเผยแพร่โปรแกรมที่มีความปลอดภัยมากกว่าแหล่งอื่น และไม่มีการดัดแปลงเพื่อติดตั้งโปรแกรมเสริมให้ติดตั้งโปรแกรมอื่นเพิ่มเติมลงในเครื่องของผู้ใช้งานซึ่งอาจเป็นทำให้คอมพิวเตอร์เสียหายหรือสร้างความรำคาญระหว่างการใช้งาน

3. เพื่อความง่ายและลดขั้นตอนการเข้าถึงไฟล์ติดตั้งให้เลือกลิงค์ pspp4windows download | SourceForge.net (https://sourceforge.net/projects/pspp4windows/)[15]

4. คลิกเมาส์ที่ Download

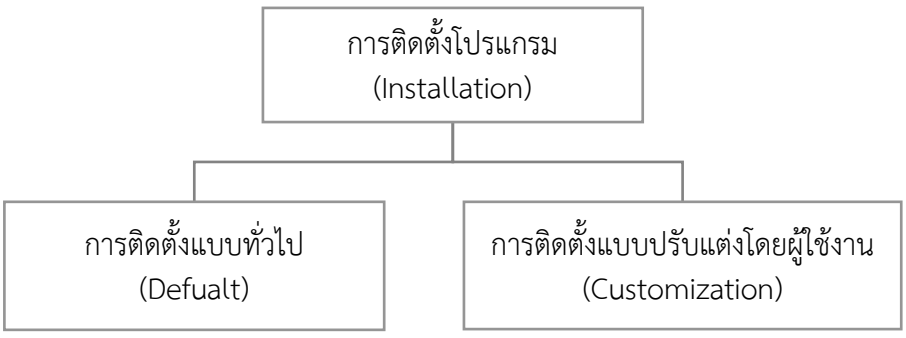


4. Save เก็บไว้ใน Drive หรือ Folder ในเครื่องคอมพิวเตอร์ที่ต้องการติดตั้งโปรแกรม



ขั้นที่ 2 วิธีการติดตั้งโปรแกรม

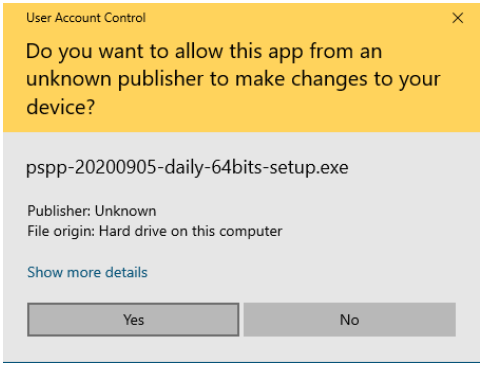
ก่อนการติดตั้งโปรแกรม ควรปิดหรือหยุดการทำงานของโปรแกรมป้องกันไวรัส การติดตั้งโปรแกรมสามารถเลือกทำได้ 2 วิธี ดังนี้



[15] ระบบให้บริการดาวโหลดจะพาไปที่เว็บไซต์ SOURCEFORGE และเลือกไฟล์ติดตั้งเวอร์ล่าสุดและระบบปฏิบัติการ (OS) ที่กำลังใช้งานให้อัตโนมัติ

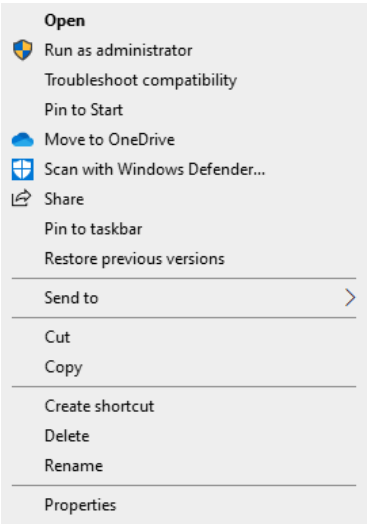
1. การติดตั้งแบบทั่วไป (default) เป็นวิธีการให้โปรแกรมกำหนดคุณสมบัติของโปรแกรมให้เหมาะกับระบบปฏิบัติการเอง มีขั้นตอนดังต่อไปนี้

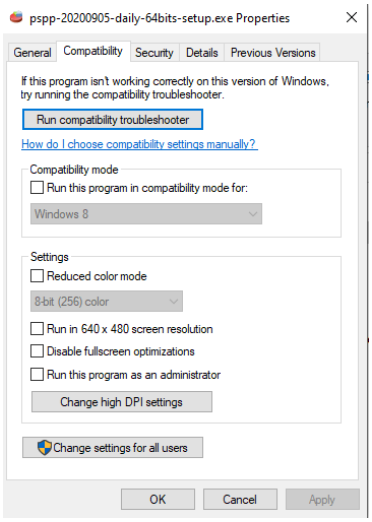
- ดับเบิลคลิกเมาส์ที่ไฟล์ติ pspp-20200905-daily-64bits-setup.exe (ชื่อไฟล์จะแตกต่างไปตามเวอร์ชั่นล่าสุด)
- หากโปรแกรม Windows แสดงหน้าต่าง User Account Control -> กดปุ่ม Yes



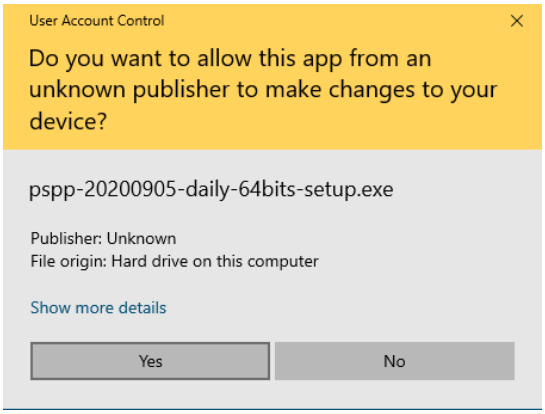
2. การติดตั้งแบบปรับแต่งโดยผู้ใช้งาน (customization) เป็นวิธีที่ผู้ใช้งานสามารถกำหนดคุณสมบัติในการติดตั้งโปรแกรมให้เหมาะสมกับระบบปฏิบัติการที่ใช้งาน โดยมีขั้นตอนดังนี้

- คลิกขวาที่ไฟล์ติดตั้ง (ชื่อไฟล์จะแตกต่างไปตามเวอร์ชั่นล่าสุดที่ปรับปรุง) -> เลือกเมนูคำสั่ง Properties
- เลือกแถบ Compatibility ในหน้าต่าง Properties
- เลือกคลิก / ที่ค่าที่ต้องการกำหนดค่า
- กดปุ่ม Apply และกดปุ่ม OK

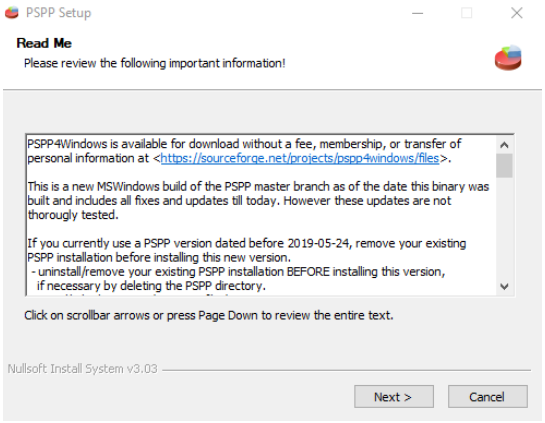




- ดับเบิลคลิกเมาส์ที่ไฟล์ติดตั้ง pspp-20200905-daily-64bits-setup.exe
- หากโปรแกรม Windows แสดงหน้าต่าง User Account Control -> กดปุ่ม Yes

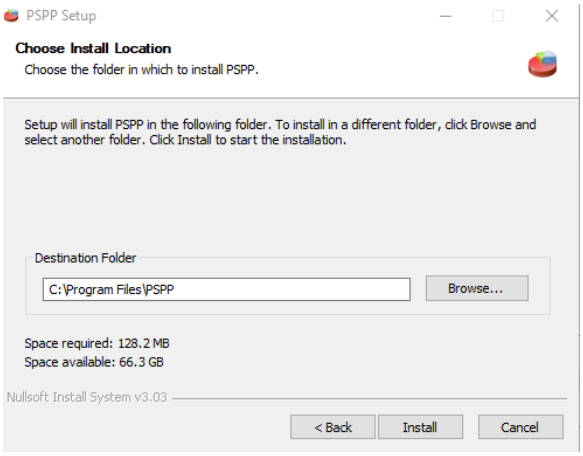


3. โปรแกรมแสดงหน้าต่าง Readme -> กดปุ่ม Next

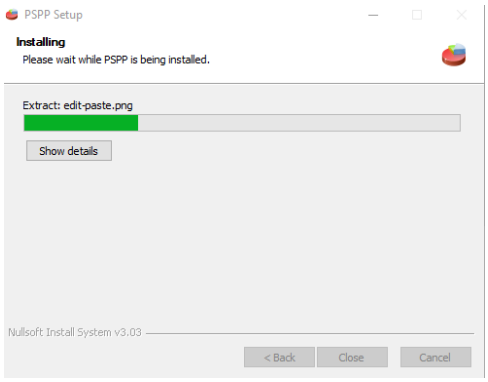


4. โปรแกรมแสดงหน้าต่างให้เลือกพื้นที่ในการติดตั้ง

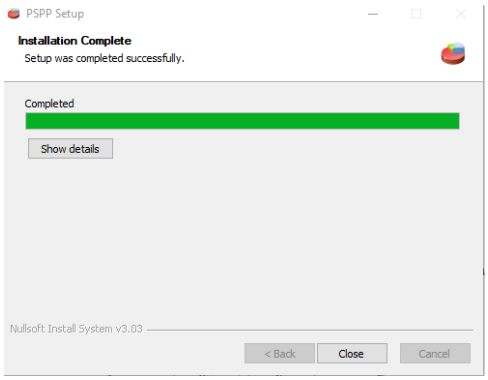
- หากต้องการติดตั้งในพื้นที่ที่โปรแกรมกำหนด -> กดปุ่ม Install
- หากไม่ต้องการติดตั้งในพื้นที่ที่โปรแกรมกำหนด -> กดปุ่ม Browse เลือกพื้นที่ติดตั้งใหม่



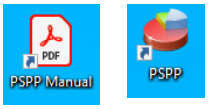
5. ระหว่างการติดตั้ง โปรแกรมจะแสดงข้อความว่า Installing



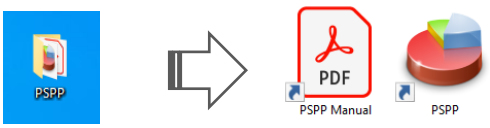
6. เมื่อติดตั้งเสร็จเรียบร้อยแล้ว โปรแกรมจะแสดงข้อความว่า Installation Complete หรือ



7. เมื่อติดตั้งโปรแกรมเสร็จเรียบร้อยแล้ว บนพื้นจอของคอมพิวเตอร์ (desktop) จะแสดงไอคอนหรือสัญรูป (icon) 2 ไอคอน คือ แฟ้มข้อมูลคู่มือการใช้งาน (PSPP Manual) และปุ่มเปิดโปรแกรม (PSPP)



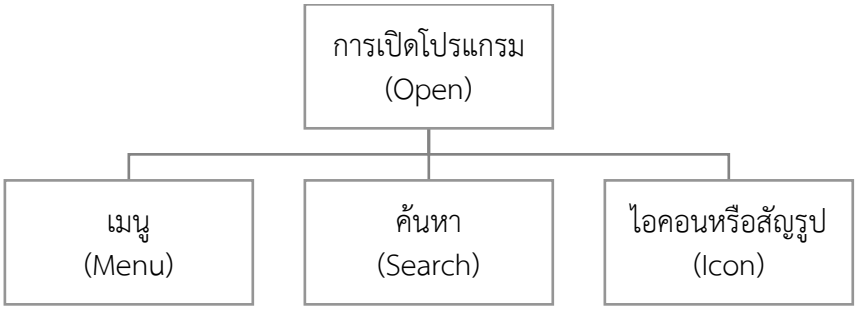
หากต้องการจัดเก็บไอคอนไว้ให้เป็นระเบียบ ไม่ปนกับไอคอนอื่นๆ ทำให้หาง่ายเมื่อต้องการใช้งาน ให้สร้างโฟลเดอร์บนพื้นจอของคอมพิวเตอร์แล้วย้าย 2 ไอคอน ไปเก็บไว้ในโฟลเดอร์ดังกล่าว โดยคลิกขวาบนพื้นจอคอมพิวเตอร์ -> New -> Folder -> พิมพ์ชื่อ PSPP -> ใช้เมาส์ลากไอคอนไปใส่ไว้ในโฟลเดอร์



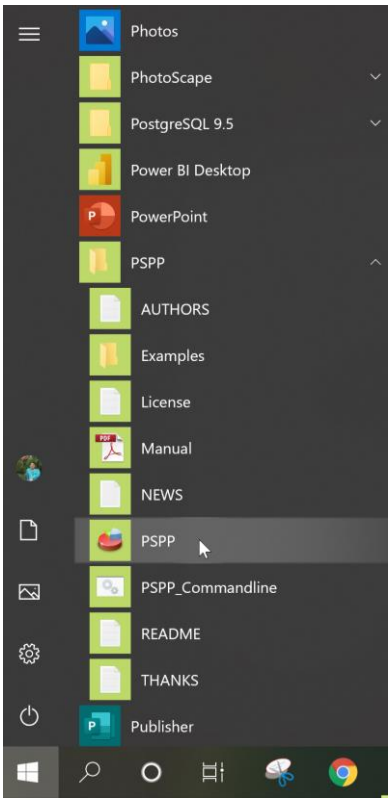
3. การใช้โปรแกรมทางสถิติเบื้องต้น

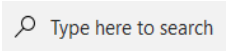
การใช้โปรแกรมเบื้องต้น เป็นการแนะนำวิธีการเปิดและปิดโปรแกรม และให้รู้จักส่วนประกอบที่สำคัญๆของโปรแกรม เช่น เมนู คำสั่ง ระบบนำเข้าข้อมูล และระบบแสดงผล

3.1 การเปิดโปรแกรม มี 3 วิธี ดังนี้

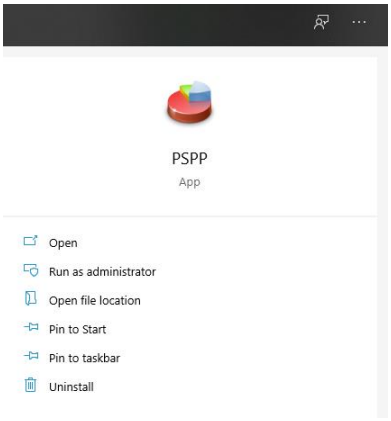


- Start (เมนู) -> PSPP (กลุ่มโปรแกรม) -> PSPP (โปรแกรม)

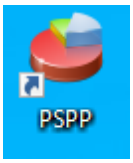


- คลิกที่เครื่องมือค้นหา (search) ด้านล่างซ้ายของจอภาพ  -> พิมพ์คำว่า pspp

- คลิกที่คำว่า Open

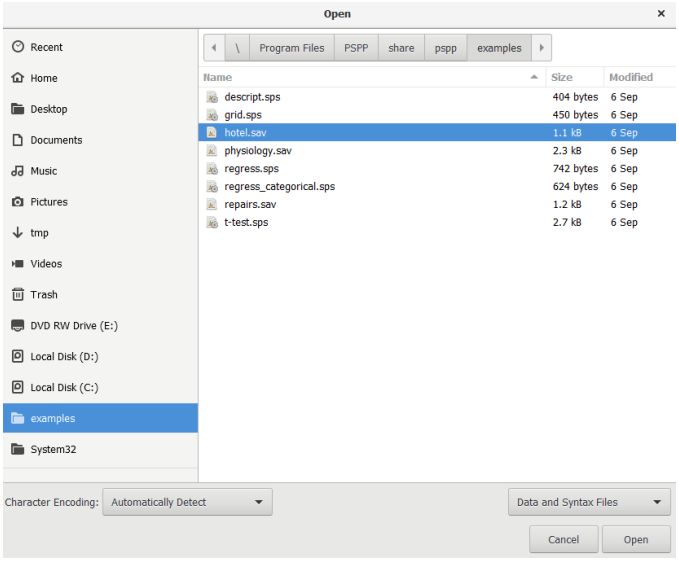


- คลิกที่ไอคอนหรือสัญรูป (icon) PSPP บนพื้นจอของคอมพิวเตอร์ (desktop)

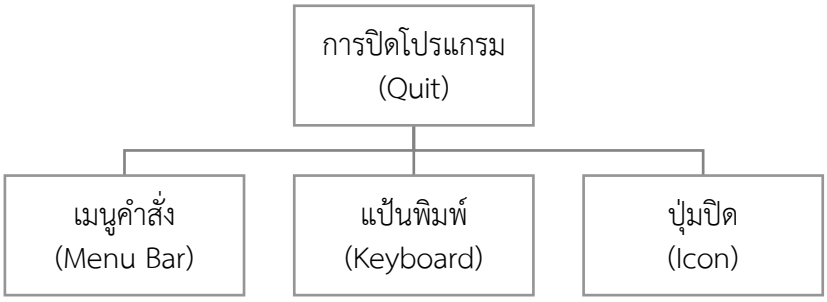


3.2 การเปิดแฟ้มข้อมูล (ทดลองเปิดแฟ้มข้อมูลตัวอย่าง hotel.sav)

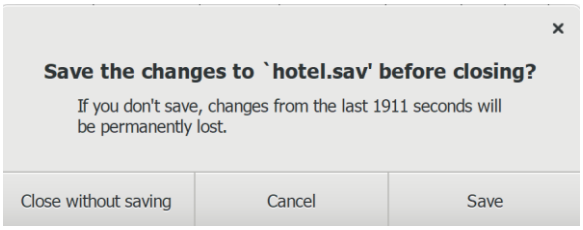
- File -> New -> Data (ไปที่ C:\Program Files\PSPP\share\pspp\examples)
- เลือกไฟล์ hotel.sav
- กดปุ่ม Open



3.3 การปิดโปรแกรม มี 3 วิธี ดังนี้



- เมนู File -> Quit (Ctrl+Q) หรือ
- กดปุ่ม Ctrl+Q บนแป้นพิมพ์
- คลิกเครื่องหมาย X มุมบนด้านซ้ายของโปรแกรม



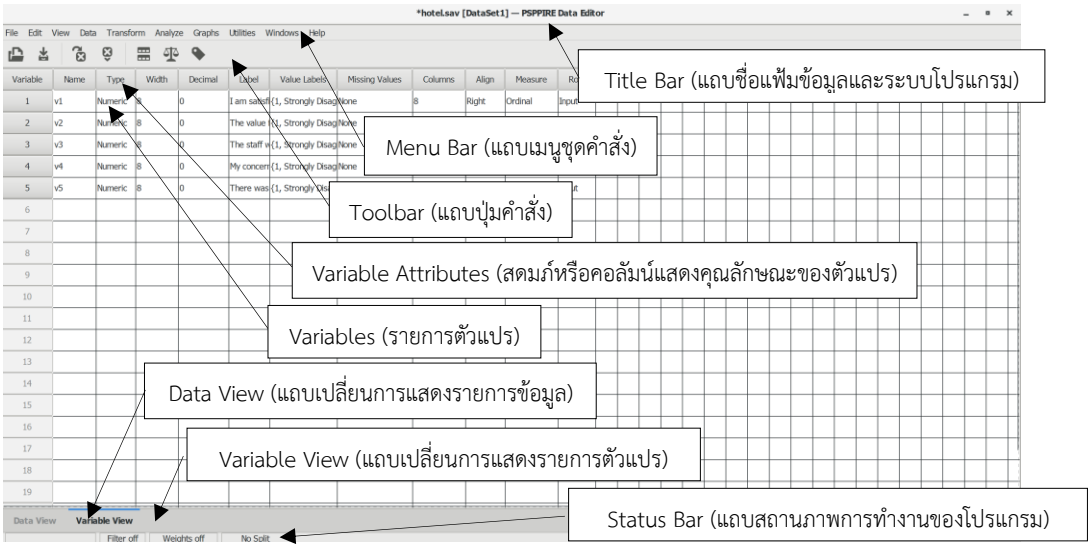
- กดปุ่ม Close without saving (ปิดโดยไม่ทำการบันทึกแฟ้มข้อมูล)

4. ส่วนประกอบของโปรแกรม PSPP

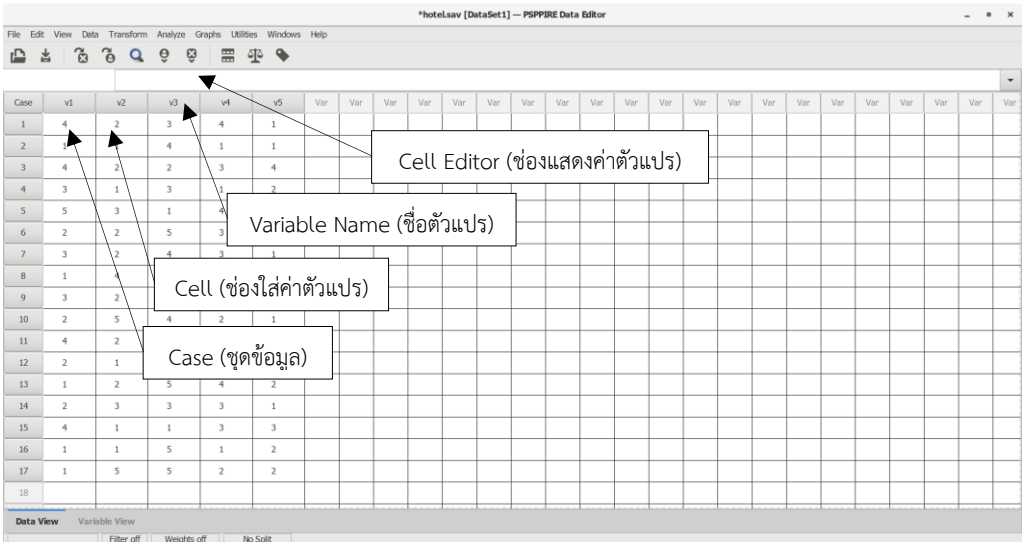
โปรแกรม PSPP มี 3 ระบบที่สำคัญ คือ ระบบจัดการข้อมูล (Data Editor) ระบบแสดงผลลัพธ์ (Output Viewer) และระบบคำสั่งแบบข้อความ (Syntax Editor) โดยมีรายละเอียดดังต่อไปนี้

4.1 ระบบจัดการข้อมูล

ระบบจัดการข้อมูลประกอบด้วยระบบแสดงรายการตัวแปร (Variable View) และระบบแสดงรายการข้อมูล (Data View) โดยมีรายละเอียดตามภาพด้านล่าง



ระบบแสดงรายการตัวแปร (Variable View)

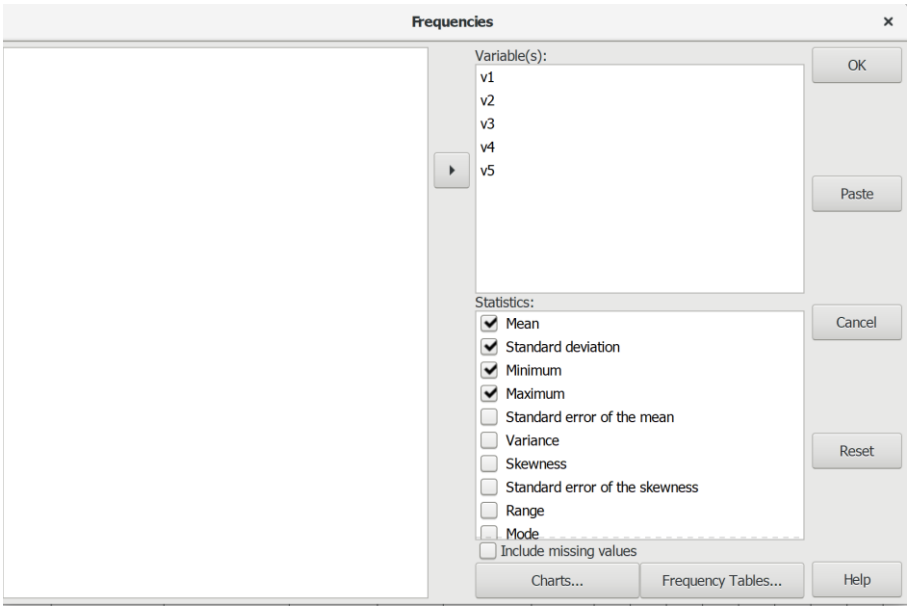


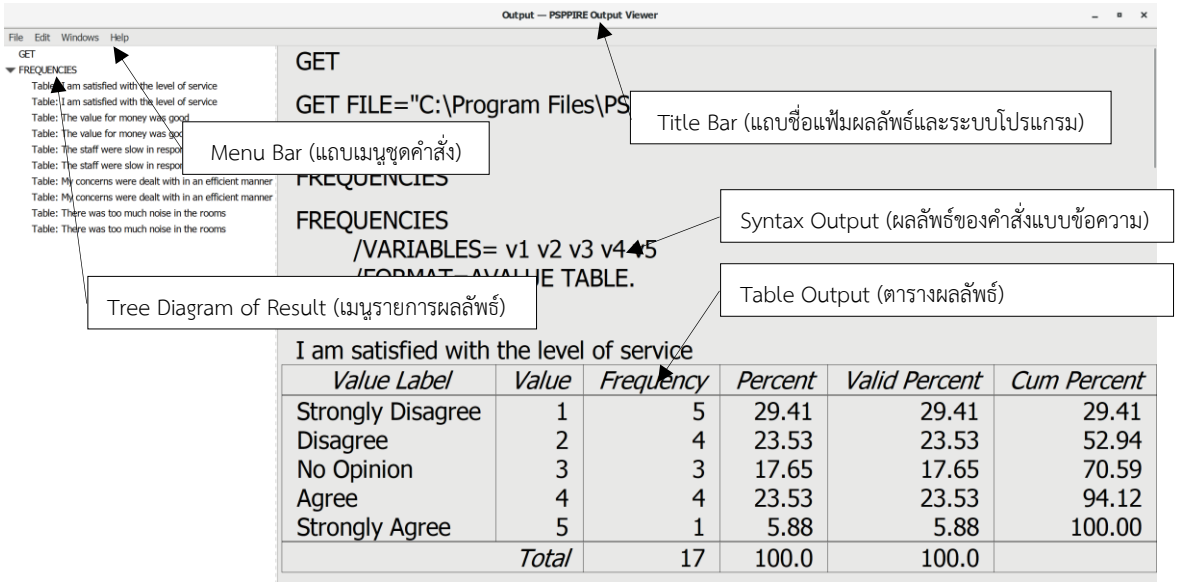
ระบบแสดงรายการข้อมูล (Data View)

4.2 ระบบแสดงผลลัพธ์

ระบบแสดงผลลัพธ์จะปรากฏให้เห็นเมื่อมีการใช้คำสั่งใดคำสั่งหนึ่งในกระบวนการวิเคราะห์ข้อมูล

- ทดลองวิเคราะห์ข้อมูลจากแฟ้มข้อมูล hotel.sav
- เลือกเมนู **Analyze -> Descriptive Statistics -> Frequencies...**
  - **Variables(s):** เลือกตัวแปรทั้งหมด (v1 v2 v3 v4 v5) ใส่ในช่อง Variables (ดับเบิลคลิกเมาส์เมาส์ที่ชื่อตัวแปรหรือใช้เมาส์ชี้ที่ชื่อตัวแปรและ ▶)
  - กดปุ่ม **OK**

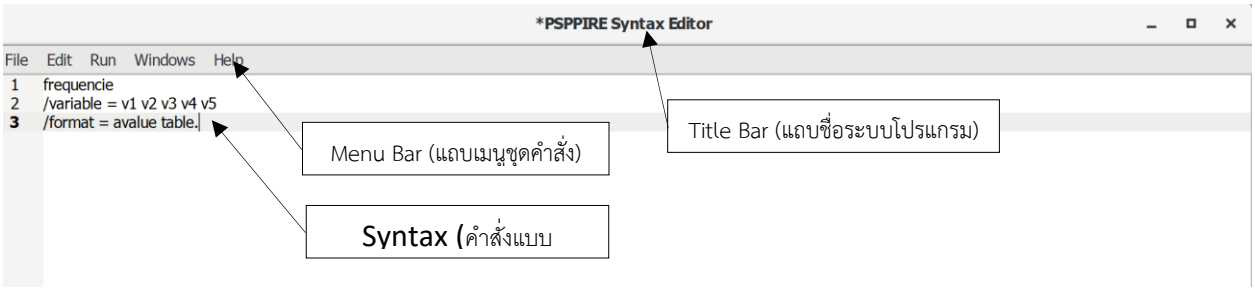




4.3 ระบบคำสั่งแบบข้อความ

- เปิดจากเมนูคำสั่ง File -> New -> Syntax ในระบบจัดการข้อมูล
- ทดลองพิมพ์คำสั่งแบบข้อความใน Syntax Editor ตามตัวอย่างด้านล่าง

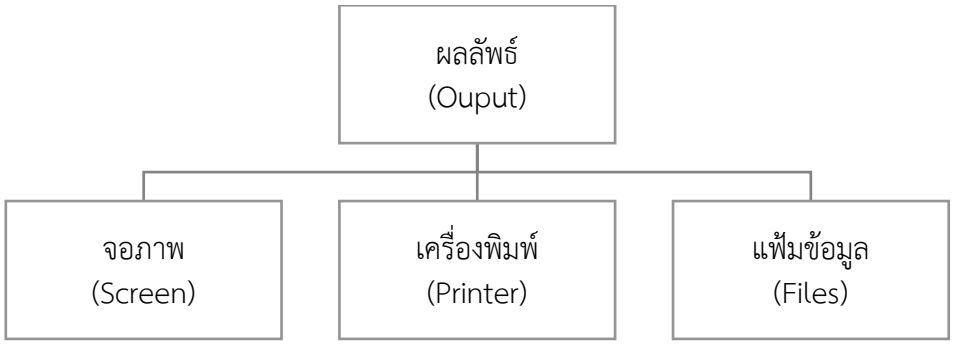
```
frequencie
/variable = v1 v2 v3 v4 v5
/format = avalue table.
```



- เลือกเมนู Run -> All

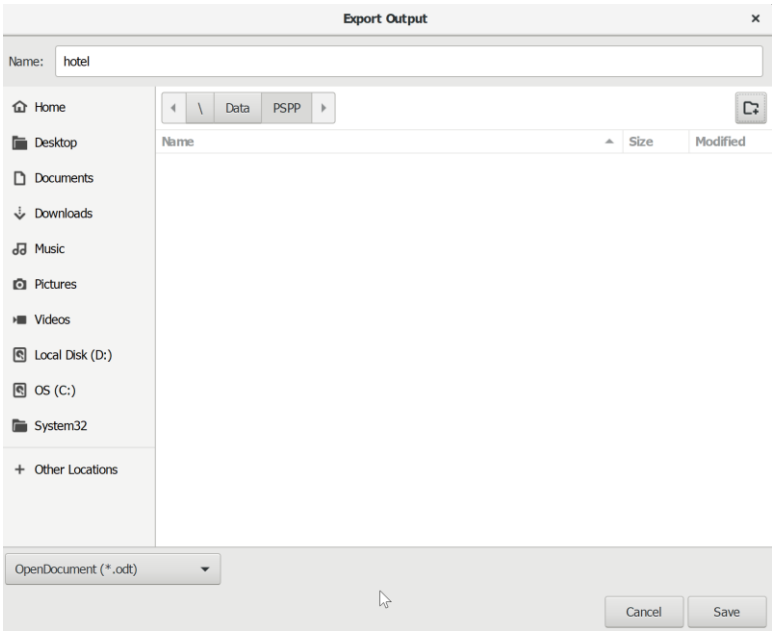
4.4 การใช้ผลลัพธ์

ผลลัพธ์ที่ได้จากการประมวลผลโปรแกรมจะเปิดเป็นหน้าต่างระบบแสดงผลลัพธ์ (Output Viewer) การใช้ผลลัพธ์มี 3 วิธี ดังนี้



- การดูผลลัพธ์จากหน้าจอภาพ
  - นำผลลัพธ์ไปกรอกใส่ตารางที่เตรียมไว้ ผลลัพธ์จากหน้าจอภาพไม่สามารถ Copy ไปใส่ในเอกสารอื่นได้
- การพิมพ์ออกเครื่องพิมพ์
  - เมนู File -> Print
  - เลือกเครื่องพิมพ์

- การส่งออก (export) เป็นแฟ้มข้อมูลนามสกุล PDF (*.pdf) HTML (*.html) OpenDocument (*.odt) Text (*.txt) และ Text [plain] (*.txt) PostScript (*.ps) และ Comma-Separated Value (*.csv)
  - เมนู File -> Export
  - พิมพ์ชื่อไฟล์ ในช่อง Name ด้านบนของหน้าต่าง Export Output
  - เลือกประเภทไฟล์จาก Infer file type from extension
  - เลือก Drive และ Folder ที่ต้องการเก็บข้อมูล
  - กดปุ่ม Save



5. เมนูการใช้โปรแกรม PSPP

โปรแกรม PSPP ออกแบบเมนูสำหรับให้ใช้งานเป็น 2 ส่วน คือ เมนูแบบรายการคำสั่ง (menu bar) และเมนูแบบปุ่มภาพ (toolbar) โดยมีรายละเอียดดังต่อไปนี้

5.1 หน้าต่างระบบ Data Editor

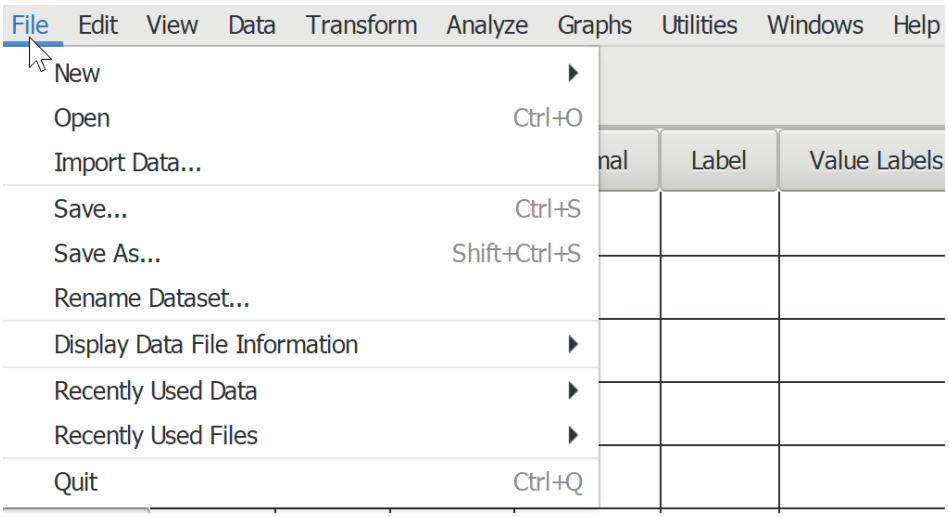
Data Editor เป็นระบบที่ใช้งานมากที่สุดระบบหนึ่งในการใช้โปรแกรมแบบกราฟิก (GUI) มีเมนู 2 แบบดังต่อไปนี้

5.1.1 เมนูแบบรายการคำสั่ง (Menu Bar)

เมนูแบบรายการคำสั่งเป็นเมนูแบบผังหรือต้นไม้ (tree menu) แบ่งเป็นกลุ่มคำสั่ง เมื่อใช้คลิกเมาส์ชี้และคลิกขวาจะมีรายการคำสั่งย่อยให้เลือกใช้ดังต่อไปนี้

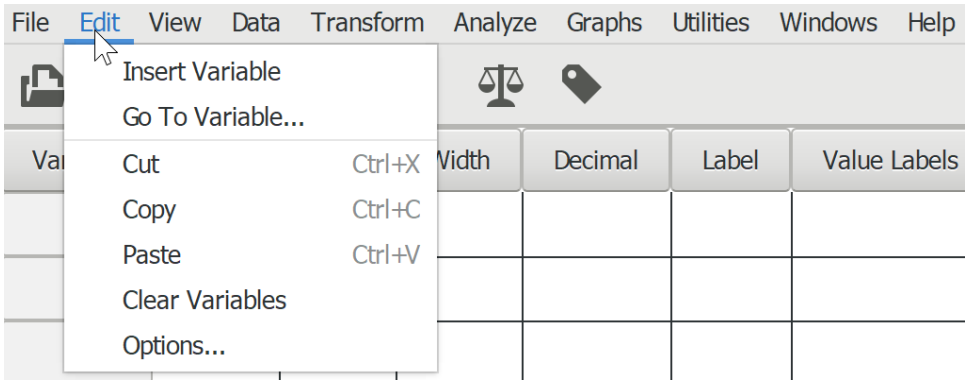
5.1.1.1 เมนู File

เมนู File เป็นเมนูรวมคำสั่งที่เกี่ยวกับการจัดการแฟ้มข้อมูล (file management) ได้แก่ New (สร้างแฟ้มข้อมูลใหม่) Open (เปิดแฟ้มข้อมูลที่มีอยู่แล้ว) Import Data (นำข้อมูลเข้าจากแฟ้มข้อมูลโปรแกรมอื่น) Save (จัดเก็บข้อมูล) Save As (จัดเก็บข้อมูลเป็นแฟ้มข้อมูลชื่อใหม่) Rename Dataset (เปลี่ยนชื่อชุดข้อมูล) Display Data File Information (แสดงรายละเอียดของข้อมูลในแฟ้มข้อมูล) Recently Used Data และ Recently Used Files (เลือกใช้แฟ้มข้อมูลที่เปิดใช้งาน) และ Quit (ปิดหรือออกจากโปรแกรม PSPP)



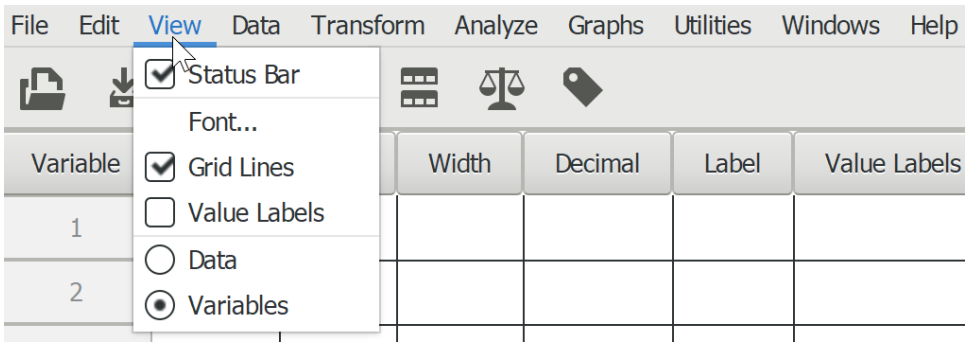
5.1.1.2 เมนู Edit

เมนู Edit เป็นเมนูรวมคำสั่งเกี่ยวกับจัดการตัวแปร (variable management) เพื่อเพิ่มหรือแก้ไขตัวแปร ได้แก่ Insert Variable (เพิ่มตัวแปร) Go To Variable (เลื่อนหรือวิ่งไปยังตัวแปรที่เลือก) Cut (ตัดข้อมูล) Copy (สำเนาข้อมูล) Paste (วางหรือแปะข้อมูลที่ได้จากการตัดหรือสำเนาข้อมูล) Clear Variables (ลบตัวแปร) และ Options (กำหนดรูปแบบการแสดงผลรายการตัวแปร)



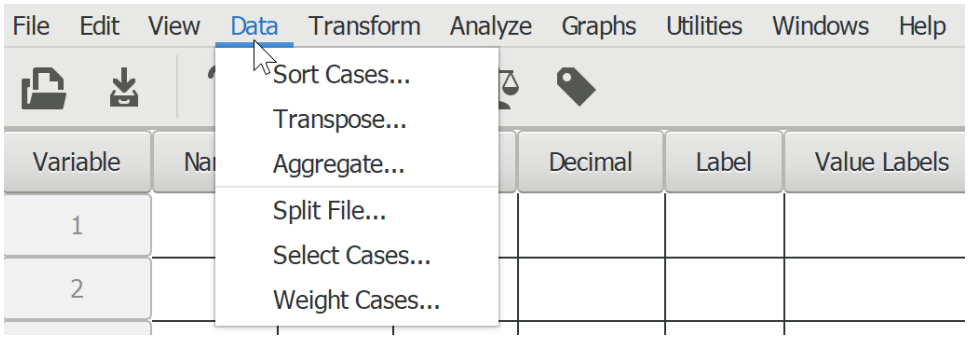
5.1.1.3 เมนู View

เมนู View เป็นเมนูรวมคำสั่งเกี่ยวกับการจัดการการแสดงผล (display management) เพื่อให้แสดงผลหรือหรือไม่แสดงผล และเปลี่ยนแปลงการแสดงผล ได้แก่ Status Bar (ปิด/เปิดแถบสถานภาพการทำงานของโปรแกรม) Font (เปลี่ยนประเภทและขนาดตัวอักษร) Grid Line (แสดง/ไม่แสดงเส้นแบ่งในตารางข้อมูล) Value Labels (แสดง/ไม่แสดงคำอธิบายค่าตัวแปร) Data (แสดงข้อมูล) และ Variables (แสดงรายการตัวแปร)



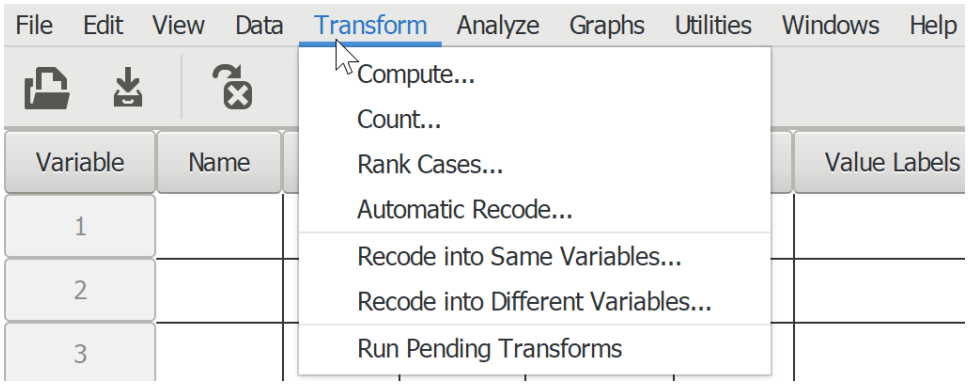
5.1.1.4 เมนู Data

เมนู Data เป็นเมนูรวมคำสั่งเกี่ยวกับปรับข้อมูล (data manipulation) ได้แก่ Sort Cases (การจัดเรียงข้อมูล) Transpose (การไขว้ตัวแปร) Aggregate (การรวมข้อมูล) Split File (การแบ่งข้อมูล) Select Cases (การเลือกข้อมูล) และ Weight Cases (การกำหนดน้ำหนักข้อมูล)



5.1.1.5 เมนู Transform

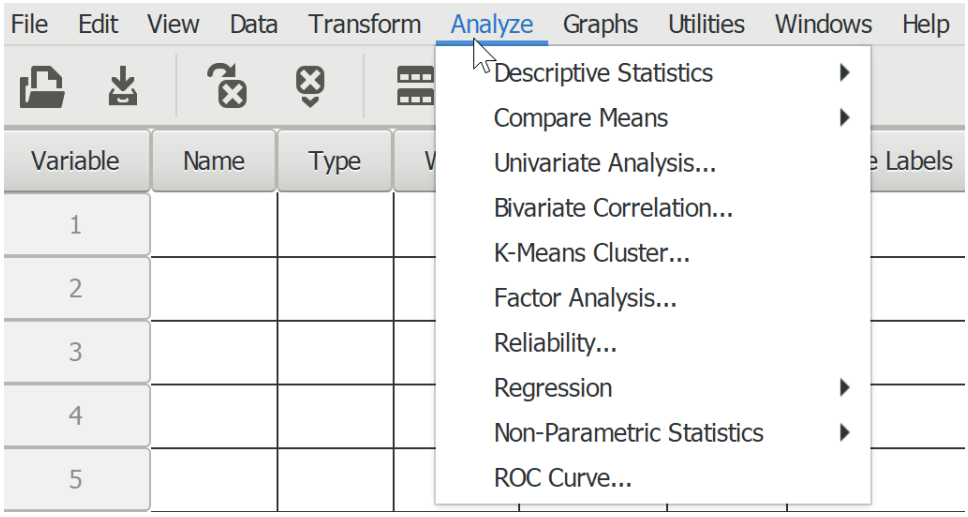
เมนู Transform เป็นเมนูรวมคำสั่งเกี่ยวกับการแปลงรูปแบบข้อมูล (variable transformation) ได้แก่ Compute (การสร้างตัวแปรแบบค่าต่อเนื่อง) Count (การนับจำนวนข้อมูล) (Rank Cases (การจัดลำดับตัวอย่าง) Automatic Recode (การสร้างตัวแปรแบบอัตโนมัติ) Recode into Same Variables (การเปลี่ยนค่าตัวแปรในตัวแปรเดิม) Recode into Different Variables (การสร้างตัวแปรแบบค่าแบ่งกลุ่ม) และ Run Pending Transforms (การสั่งประมวลผลต่อ)



5.1.1.6 เมนู Analyze

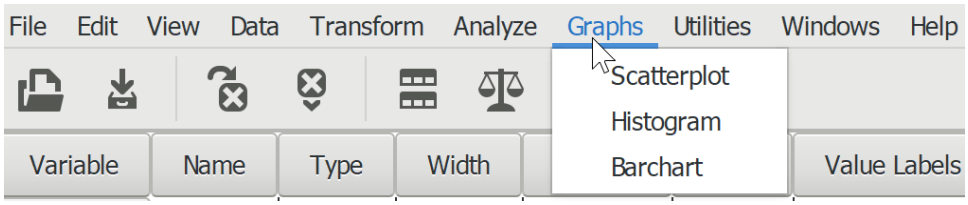
เมนู Analyze เป็นเมนูรวมคำสั่งเกี่ยวกับการวิเคราะห์ข้อมูล (data analysis) ด้วยสถิติประเภทต่างๆ ได้แก่ Descriptive Statistics (การวิเคราะห์ด้วยสถิติเชิงพรรณนา) Compare Means (การ

วิเคราะห์เปรียบเทียบค่าเฉลี่ย) Univariate Analysis (การวิเคราะห์ตัวแปรเดี่ยว) Bivariate Correlation (การวิเคราะห์ความสัมพันธ์รายคู่) Factor Analysis (การวิเคราะห์องค์ประกอบ) Reliability (การวิเคราะห์ความน่าเชื่อถือ) Linear Regression (การวิเคราะห์ถดถอยเชิงเส้นตรง) Non-Parametric Statistics (การวิเคราะห์ด้วยสถิติไม่อิงค่าพารามิเตอร์) และ ROC Curve (การวิเคราะห์พื้นที่ใต้โค้ง)



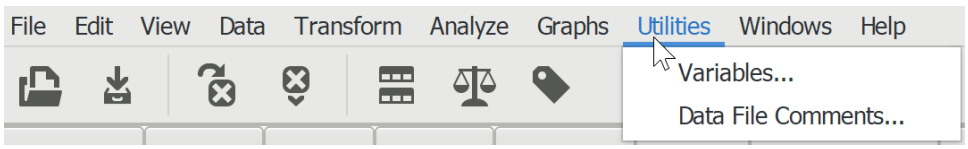
5.1.17 เมนู Graphs

เมนู Graphs เป็นเมนูรวมคำสั่งเกี่ยวกับการวิเคราะห์ข้อมูลด้วยวิทัศน์หรือกราฟ ได้แก่ Scatterplot (กราฟการกระจาย) Histogram (กราฟแจกแจงความถี่) และ Barchart (กราฟแท่ง)



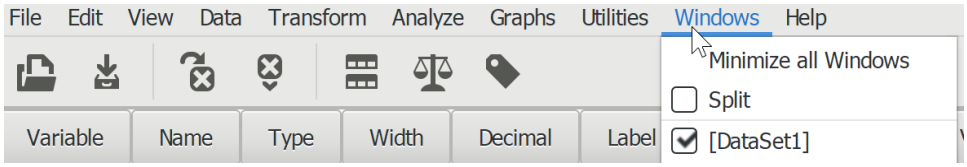
5.1.1.8 เมนู Utilities

เมนู Utilities เป็นเมนูรวมคำสั่งเกี่ยวกับการจัดการตัวแปรและแฟ้มข้อมูล ได้แก่ Variables (แสดงรายละเอียดหรือข้อมูลเกี่ยวกับตัวแปร เลื่อนหรือวิ่งไปยังตัวแปร และแปะข้อมูลเกี่ยวกับตัวแปร) และ Data File Comments (บันทึกหรือเขียนคำแนะนำเกี่ยวกับแฟ้มข้อมูล)



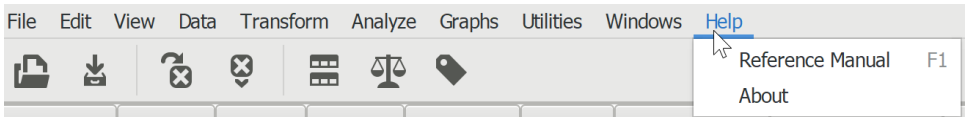
5.1.1.9 เมนู Windows

เมนู Windows เป็นเมนูรวมคำสั่งเกี่ยวกับการจัดการหน้าต่างที่กำลังใช้งาน (windows management) ได้แก่ Minimize all Windows (ยุบ/ปิดหน้าต่างทั้งหมด) Split (แบ่งตารางรายการตัวแปรหรือข้อมูลเป็น 4 ส่วน) และ DataSet (เลือกชุดข้อมูลที่กำลังเปิดใช้งาน)



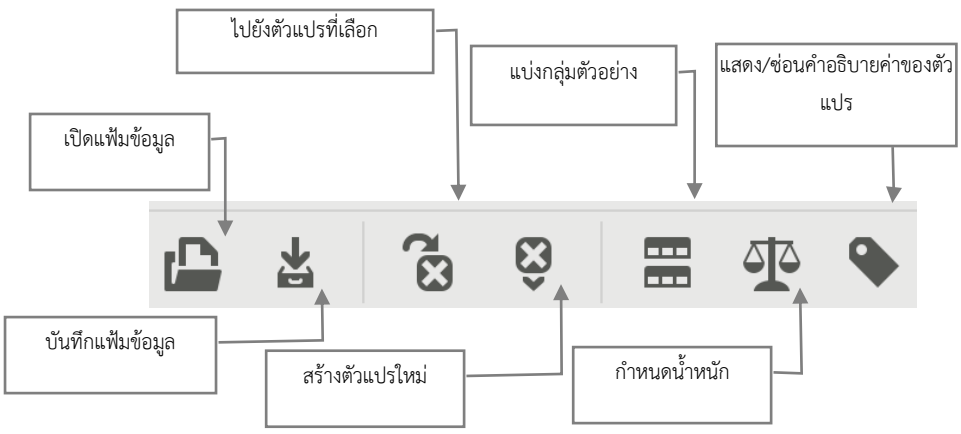
5.1.1.10 เมนู Help

เมนู Help เป็นเมนูรวมคำสั่งเกี่ยวกับความช่วยเหลือ ได้แก่ Reference Manual (คู่มืออ้างอิง/การใช้งาน) และ About (ข้อมูลโดยย่อเกี่ยวกับโปรแกรม)



5.1.2 เมนูแบบปุ่มภาพ (Toolbar)

เมนูแบบปุ่มภาพเป็นเมนูคำสั่งประเภทหนึ่งที่ทีมพัฒนาเลือกคำสั่งที่ใช้งานกันบ่อยมาออกแบบเป็นปุ่มกดและมีภาพสัญลักษณ์ประกอบ ๆ เพื่อให้ใช้งานได้สะดวกและรวดเร็วตามแนวคิด GUI โดยมีคำสั่งดังต่อไปนี้

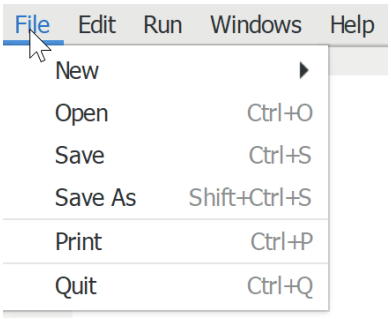


5.2 หน้าต่างระบบ Syntax Editor

การใช้โปรแกรม PSPP แบบ GUI บางครั้งอาจต้องใช้คำสั่งแบบข้อความ (command line) โดยเปิดระบบ Syntax Editor จากเมนู File -> New -> Syntax และเมื่อหน้าต่างระบบ Syntax Editor ปรากฏขึ้นจะมีเมนูให้ใช้งานดังต่อไปนี้

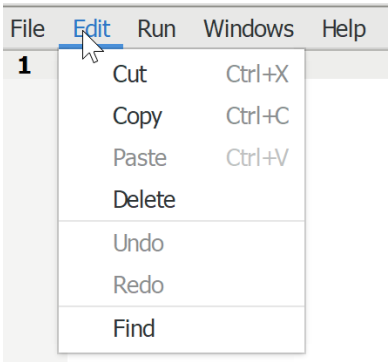
5.2.1 เมนู File

เมนู File เป็นเมนูรวมคำสั่งที่เกี่ยวกับการจัดการแฟ้มข้อมูล (file management) เหมือนกับระบบ Data Editor ได้แก่ New (สร้างแฟ้มข้อมูลใหม่) Open (เปิดแฟ้มข้อมูลที่มีอยู่แล้ว) Save (จัดเก็บข้อมูล) Save As (จัดเก็บข้อมูลเป็นแฟ้มข้อมูลในชื่อใหม่) Print (พิมพ์ Syntax ออกเครื่องพิมพ์) และ Quit (ปิดระบบ Syntax Editor)



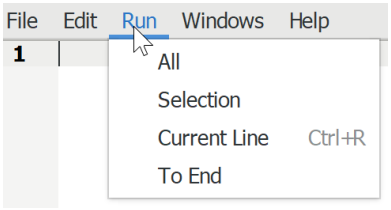
5.2.2 เมนู Edit

เมนู Edit เป็นเมนูรวมคำสั่งที่เกี่ยวกับการจัดการคำสั่งหรือข้อความใน Syntax ได้แก่ Cut (ตัด) Copy (สำเนา) Paste (แปะ) Delete (ลบ) Undo (ทำกลับ/ยกเลิก) Redo (ทำใหม่) และ Find (ค้นหา)



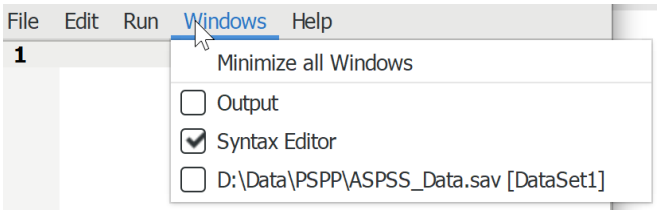
5.2.3 เมนู Run

เมนู Run เป็นเมนูรวมคำสั่งที่เกี่ยวกับการประมวลผล Syntax ได้แก่ All (ประมวลผลทั้งหมด) Selection (ประมวลผลเฉพาะที่เลือก) Current Line (ประมวลผลเฉพาะบรรทัดที่ตัวชี้ตำแหน่งวางอยู่) และ To End (ประมวลผลจากตำแหน่งที่ตัวชี้ตำแหน่งวางอยู่จนสุด Syntax)



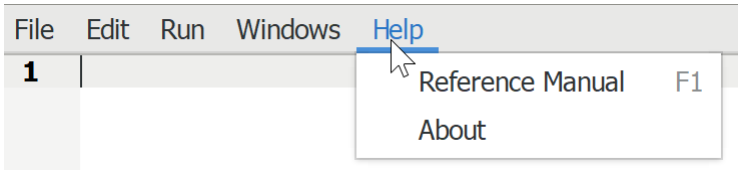
5.2.4 เมนู Windows

เมนู Windows เป็นเมนูรวมคำสั่งที่เกี่ยวกับการจัดการหน้าต่างระบบ Syntax Editor ได้แก่ Minimize all Windows (ยุบ/ปิดหน้าต่างทั้งหมด) Output (เลือกเปิดหน้าต่างผลลัพธ์) และ Syntax Editor (เลือกเปิดหน้าต่าง Syntax Editor)



5.2.5 เมนู Help

เมนู Help เป็นเมนูรวมคำสั่งเกี่ยวกับความช่วยเหลือ ได้แก่ Reference Manual (คู่มืออ้างอิง/การใช้งาน) และ About (ข้อมูลโดยย่อเกี่ยวกับโปรแกรม)

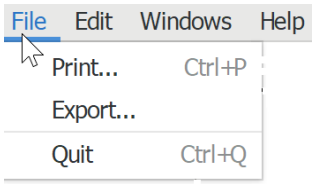


5.3 หน้าต่างระบบ Output Viewer

หน้าต่างระบบ Output Viewer จะแสดงผลทุกครั้งที่มีการใช้คำสั่งของโปรแกรม เช่น เปิดแฟ้มข้อมูล ประมวลผลข้อมูล แก้ไขข้อมูล ฯลฯ และเมื่อหน้าต่างระบบ Output Viewer ปรากฏขึ้นจะมีเมนูให้ใช้งานดังต่อไปนี้

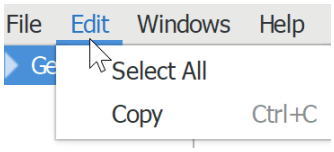
5.3.1 เมนู File

เมนู File เป็นเมนูรวมคำสั่งเกี่ยวกับการพิมพ์และส่งออกผลลัพธ์ ได้แก่ Print (พิมพ์ผลลัพธ์ออกทางเครื่องพิมพ์) Export (ส่งผลลัพธ์ออกเป็นไฟล์ *.spv, *.pdf, *.html, *.odt, *.txt, *.ps และ *.csv) และ Quit (ปิดระบบ Output Viewer)



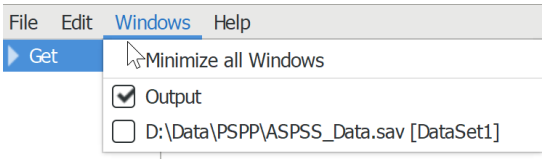
5.3.2 เมนู Edit

เมนู Edit เป็นเมนูรวมคำสั่งเกี่ยวกับการจัดสำเนาผลลัพธ์ ได้แก่ Select All (เลือกผลลัพธ์ทั้งหมด) Copy (สำเนาผลลัพธ์ที่เลือก)



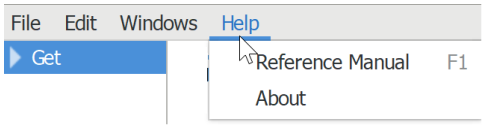
5.3.3 เมนู Windows

เมนู Windows เป็นเมนูรวมคำสั่งที่เกี่ยวกับการจัดการหน้าต่างระบบ Output Viewer ได้แก่ Minimize all Windows (ยุบ/ปิดหน้าต่างทั้งหมด) Output (เลือกเปิดหน้าต่างผลลัพธ์) และ D:\Data\PSPP\ASPSS_Data.sav [DataSet1] (เลือกเปิดหน้าต่างข้อมูล)



5.3.4 เมนู Help

เมนู Help เป็นเมนูรวมคำสั่งเกี่ยวกับความช่วยเหลือ ได้แก่ Reference Manual (คู่มืออ้างอิง/การใช้งาน) และ About (ข้อมูลโดยย่อเกี่ยวกับโปรแกรม)



สรุป

โปรแกรม PSPP เป็นโปรแกรมแบบสาธารณะทั่วไป สามารถดาวน์โหลดจากอินเทอร์เน็ตมาติดตั้งและใช้งานได้ฟรี มีระบบติดต่อกับผู้ใช้งานทั้งแบบใช้คำสั่งที่เป็นข้อความ (command line) และแบบกราฟิก (GUI) เหมาะสำหรับผู้เรียนที่เริ่มเรียนรู้การวิเคราะห์ข้อมูลและเริ่มทำวิจัยเบื้องต้น และยังเป็นโปรแกรมเริ่มต้นสำหรับเรียนรู้การวิเคราะห์ข้อมูลด้วยสถิติก่อนก้าวไปใช้โปรแกรมอื่นต่อไป

แบบฝึกหัดท้ายบท

ให้ผู้เรียนทบทวนการใช้โปรแกรมดังต่อไปนี้

1. ให้ผู้เรียนทดลองเปิดโปรแกรมและปิดโปรแกรม
2. ศึกษาเมนู คำสั่ง และส่วนประกอบของโปรแกรมว่ามีอะไรบ้าง

บทที่ 6
การนำข้อมูลเข้าโปรแกรม

ความมุ่งหมายของบทเรียน

ผู้เรียนสามารถนำข้อมูลจากแบบสอบถามกรอกเข้าโปรแกรม และนำข้อมูลที่เก็บอยู่ในรูปของแฟ้มข้อมูลเข้าโปรแกรม

เนื้อหา

1. การนำข้อมูลเข้าจากแบบสอบถาม
2. การนำข้อมูลเข้าจากแฟ้มข้อมูล

วิธีการสอนและกิจกรรมการเรียนการสอน

บรรยาย แสดงตัวอย่าง แนะนำการกรอกข้อมูลเข้าโปรแกรมและนำข้อมูลเข้าจากแฟ้มข้อมูล และทำแบบฝึกหัดท้ายบท

สื่อและอุปกรณ์การสอน

เอกสารคำสอน ภาพเลื่อน คอมพิวเตอร์ โปรแกรม PSPP โปรแกรมสอนออนไลน์ (สอนแบบออนไลน์) และเครื่องฉาย (สอนในชั้นเรียน)

งานที่มอบหมาย

ทำแบบฝึกหัดท้ายบท

การวัดผลและประเมินผล

การเข้าชั้นเรียน สังเกตความสนใจจากการเรียน สอบถามความเข้าใจจากการบรรยาย และการทำแบบฝึกหัดท้ายบท

เนื้อหา

ข้อมูลที่ใช้ในการวิจัยเชิงปริมาณอาจเป็นข้อมูลปฐมภูมิ (primary data) ที่ได้จากการใช้แบบสอบถามในการรวบรวมข้อมูล มีทั้งการสำรวจด้วยแบบสอบที่สร้างด้วยกระดาษ (non-digital questionnaire) และแบบสอบถามที่สร้างด้วยโปรแกรมคอมพิวเตอร์ (digital questionnaire) และข้อมูลทุติยภูมิ (secondary data) ที่จัดเก็บไว้แล้ว มีทั้งข้อมูลแบบเก่าที่เก็บไว้เป็นเอกสาร (non-digital data) เช่น หนังสือ รายงานประจำปี และข้อมูลแบบใหม่ที่เป็นดิจิทัล (digital data) เช่น แฟ้มข้อมูลคอมพิวเตอร์ และโปรแกรมฐานข้อมูล

การนำข้อมูลเข้าไปวิเคราะห์ด้วยโปรแกรมคอมพิวเตอร์จึงต้องมีการเตรียมข้อมูลตามลักษณะของข้อมูลที่ได้มา ซึ่งอาจต้องเริ่มจากการสร้างแฟ้มข้อมูล นำข้อมูลเข้าแบบกรอกข้อมูล หรือดึงเข้าจากแฟ้มข้อมูล

1. การนำข้อมูลเข้าจากแบบสอบถาม

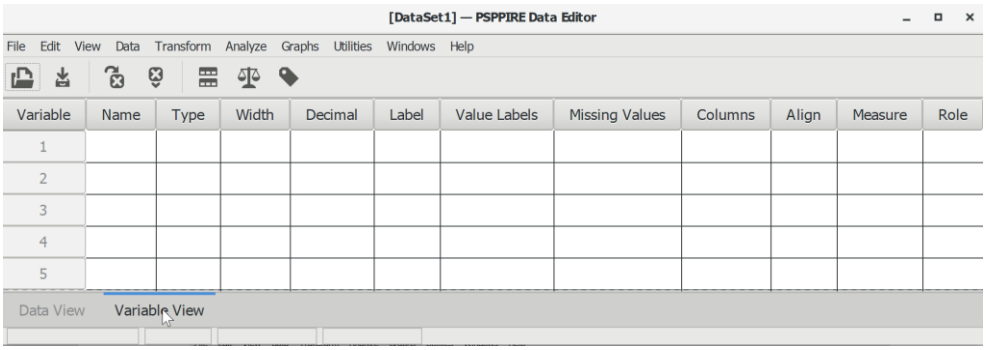
แบบสอบถามหรือแบบสำรวจเป็นเครื่องมือวัดที่นิยมใช้ในการรวบรวมข้อมูลอย่างแพร่หลาย ปัจจุบันนักวิจัยอาจต้องสร้างแบบสอบถามทั้งแบบเก่าโดยใช้การพิมพ์ด้วยกระดาษหรือที่เรียกว่าแบบสอบถามแบบออฟไลน์ และอาจต้องใช้แบบสอบถามออนไลน์ควบคู่กันไป เพราะผู้ตอบแบบสอบถามอาจถนัดหรือมีความพร้อมใช้แบบใดแบบหนึ่ง จึงต้องเตรียมสร้างแบบสอบถามทั้งสองแบบไว้ให้เลือกตอบ ดังนั้นหากมีการเก็บรวบรวมข้อมูลด้วยแบบสอบถามแบบออฟไลน์จึงต้องทำการสร้างแฟ้มข้อมูลและกรอกข้อมูลเข้าโปรแกรมโดยมีวิธีการดังต่อไปนี้

1.1 การสร้างแฟ้มข้อมูล

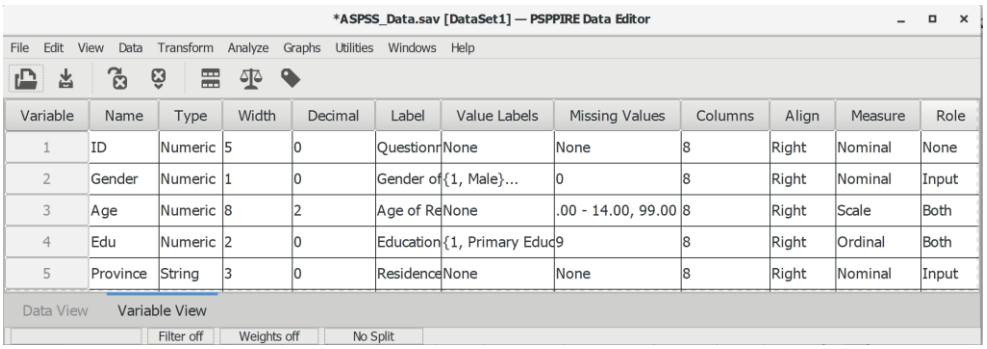
การสร้างแฟ้มข้อมูล คือ การสร้างแบบฟอร์มรับข้อมูจากแบบสอบถาม หรือเป็นการแปลงแบบสอบถามให้เป็นข้อมูลที่โปรแกรมสามารถนำไปประมวลผลได้ เมื่อเข้าใจถึงความสัมพันธ์ระหว่างแบบสอบถามกับแฟ้มข้อมูลแล้ว เราสามารถสร้างแฟ้มข้อมูลจากโปรแกรมได้ดังนี้

1. เปิดโปรแกรม PSPP โปรแกรมจะแสดงหน้าต่างระบบจัดการข้อมูล (Data Editor)

2. คลิกเมาส์ที่แถบ **Variable View** ด้านล่างซ้ายของหน้าต่าง **Data Editor**



3. ตัวอย่างแฟ้มข้อมูล



4. การสร้างตัวแปร

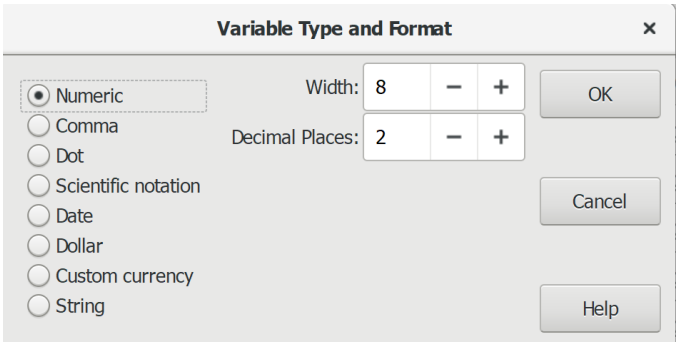
การสร้างตัวแปร คือ การแปลงข้อคำถามให้เป็นตัวแปร บางข้อคำถามอาจต้องแปลงจากคำตอบ รวมถึงการกำหนดคุณลักษณะของตัวแปร เช่น ชื่อตัวแปร ประเภทตัวแปร ลักษณะของข้อมูล คำอธิบายชื่อตัวแปร และคำอธิบายค่าตัวแปร

โดยทั่วไปการเลื่อนตัวชี้ตำแหน่งหรือเคอร์เซอร์ (cursor) ไปยังช่องต่างๆ (cells) ในโปรแกรม PSPP สามารถใช้เมาส์ (mouse) หรือ ทัชแพด (touchpad) หรือ แทร็คแพด (trackpad) เลื่อนได้เหมือนกับโปรแกรมทั่วไป และสามารถใช้คีย์บอร์ด (keyboard) ปุ่มลูกศรซ้ายหรือขวา และปุ่มลูกศรขึ้นหรือลง เลื่อนตัวชี้ตำแหน่งหรือเคอร์เซอร์ไปยังช่องที่อยู่ใกล้เคียงได้เช่นกัน

เนื่องจากมีบางคอลัมน์ที่มีระบบเมนูให้เลือกกำหนดค่าซ่อนอยู่ เช่น Type, Width Decimal, Value Labels, Missing Values, Align, Measure และ Role ทำให้ไม่สามารถใช้คีย์บอร์ดปุ่มลูกศรซ้ายหรือขวาเลื่อนไปทางซ้ายหรือขวาได้ ต้องใช้คีย์บอร์ดปุ่มแท็บ (tab) เลื่อนจากซ้ายไปขวา และใช้ปุ่มชิฟท์ (shift) + แท็บ (tab) เลื่อนจากขวาไปซ้าย

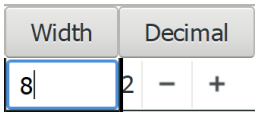
- การกำหนดชื่อตัวแปร (Variable Name)
  - **พิมพ์ชื่อตัวแปร**ในช่องใต้คอลัมน์ **Name**

- การกำหนดประเภทและรูปแบบตัวแปร (Variable Type)
  - คลิกเมาส์ในช่องใต้คอลัมน์ **Type** โปรแกรมจะแสดงหน้าต่างให้กำหนดรูปแบบและประเภทของตัวแปร (ถ้าใช้ตัวเลขเป็นข้อมูลของตัวแปรให้**เลือก Numeric หรือ**ถ้าใช้ตัวอักษรเป็นข้อมูลของตัวแปรให้เลือก **String**)
  - กำหนดจำนวนค่าตัวแปร ในช่อง **Width:** โดย**พิมพ์ตัวเลข**ทับตัวเลขค่าเริ่มต้น **หรือกดปุ่ม - หรือ +**
  - กำหนดจำนวนทศนิยมตัวแปร ในช่อง **Decimal Places:** โดยพิมพ์ตัวเลขทับตัวเลขค่าเริ่มต้น หรือ**กดปุ่ม - หรือ +**
  - กดปุ่ม **OK**



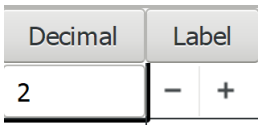
ข้อมูลของตัวแปรด้านสังคมศาสตร์ส่วนใหญ่เป็นตัวเลข (numeric) และตัวอักษร (string) หากข้อมูลเป็นตัวอักษรควรแทนค่าให้เป็นตัวเลข และกำหนดให้เป็นตัวแปรประเภท Numeric เพราะบางคำสั่งไม่แสดงตัวแปรที่เป็นตัวอักษร เช่น Means หากกำหนดค่าตัวแปร ชาย = M และ หญิง = F ตัวแปรจะไม่ปรากฏในรายการตัวแปรให้เลือกเพื่อทำการวิเคราะห์เปรียบเทียบ ต้องแปลงค่าของตัวแปรให้เป็นตัวเลขก่อน)

- การกำหนดจำนวนค่าตัวแปร (Variable Width)
  - **พิมพ์ตัวเลข**จำนวนค่าตัวแปรใต้คอลัมน์ **Width** **หรือกดปุ่ม - หรือ +**



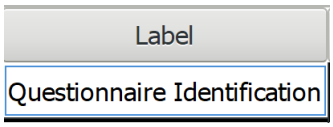
ค่าเริ่มต้น (default) จำนวนค่าตัวแปรของตัวแปรแบบตัวเลขและตัวอักษรกำหนดไว้เป็น 8 ตำแหน่ง หากไม่ต้องการบังคับจำนวนค่าตัวแปรที่ให้กรอกข้อมูล ให้ใช้ค่าเริ่มต้น

- การกำหนดจำนวนทศนิยมของค่าตัวแปร (Variable Decimal)
  - **พิมพ์ตัวเลข**จำนวนทศนิยมใต้คอลัมน์ **Decimal** หรือ**กดปุ่ม - หรือ +**



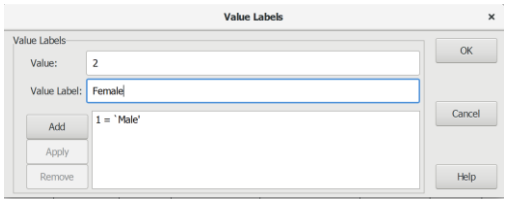
ค่าเริ่มต้น (default) จำนวนทศนิยมของตัวแปรประเภทตัวเลขกำหนดไว้เป็น 2 ตำแหน่ง และตัวแปรประเภทตัวอักษรกำหนดไว้ 0 ตำแหน่ง หากต้องการให้การแสดงข้อมูลในหน้าต่างแสดงข้อมูล (data view) เข้าใจง่าย ดูแล้วสามารถจำแนกประเภทตัวแปรได้ว่าตัวแปรใดเป็นตัวแปรแบบค่าแบ่งกลุ่ม และตัวแปรแบบค่าต่อเนื่อง ให้กำหนดจำนวนทศนิยมของตัวแปรแบบค่าแบ่งกลุ่ม คือ Nominal และ Ordinal เป็น 0 ตำแหน่ง และกำหนดจำนวนทศนิยมของตัวแปรแบบค่าต่อเนื่อง คือ Scale เป็น 2 ตำแหน่ง

- การกำหนดคำอธิบายชื่อตัวแปร (Variable Label)
  - **พิมพ์คำอธิบายชื่อตัวแปร**ใต้คอลัมน์ **Label**



หากชื่อตัวแปรสื่อความหมาย อาจไม่ต้องมีคำอธิบายชื่อตัวแปร เช่น Gender หากใช้ชื่อย่อและไม่สื่อความหมาย เช่น V1 หรือ QoS ควรกำหนดชื่อตัวแปรพอให้เข้าใจได้ ไม่ยาวมาก เพราะหากสร้างชื่อตัวแปรยาวมากจะมีผลในตารางการแสดงผลการวิเคราะห์

- กำหนดคำอธิบายค่าของตัวแปร (Values)
  - คลิกเมาส์ในช่องว่างใต้คอลัมน์ Value Labels (โปรแกรมจะแสดงหน้างต่างให้กำหนดค่าและคำอธิบายค่าของตัวแปร)
  - **พิมพ์ค่าของตัวแปร**ในช่อง **Value:**
  - **พิมพ์ความหมาย**ของค่าของตัวแปรในช่อง **Value Label:**
  - กดปุ่ม **Add**



การแก้ไขค่าและความหมายของค่าของตัวแปร ให้คลิกเมาส์ที่ค่านั้นและทำการแก้ไข เมื่อทำการแก้ไขแล้วให้กดปุ่ม Apply เพื่อเปลี่ยนแปลงค่าและความหมาย และหากต้องการลบความหมายของค่าของตัวแปร ให้คลิกเมาส์ที่ค่านั้นและกดปุ่ม Remove

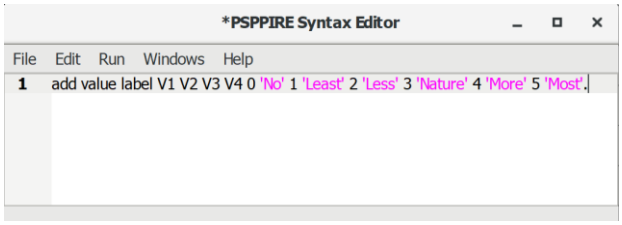
โปรแกรม PSPP ไม่สามารถ Copy ค่าในช่อง Value Labels ได้ หากตัวแปรมี Value Label เหมือนกัน ให้ใช้วิธีการดังนี้

1. สร้างตัวแปรให้ครบ โดยยังไม่ต้องกำหนด Value Labels
2. เลือกเมนู **File -> New -> Syntax**

ตัวอย่าง

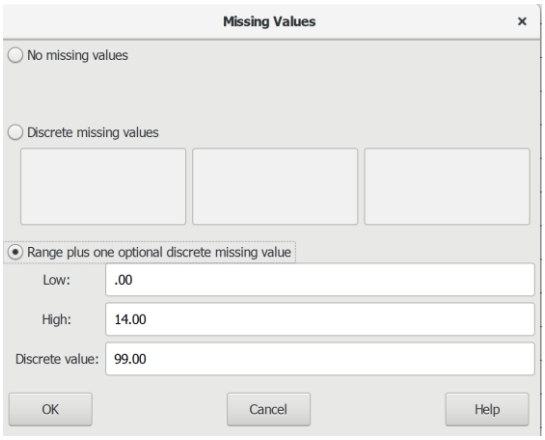
ตัวแปร V1 V2 V3 V4 ให้พิมพ์คำสั่งในหน้าต่าง Syntax Editor ดังนี้

add value label V1 V2 V3 V4 0 'No' 1 'Least' 2 'Less' 3 'Nature' 4 'More' 5 'Most'.



3. เมนู **Run -> Current Line** หรือกด **Cltr+R** ที่ท้ายบรรทัดของคำสั่ง

- การกำหนดค่าขาดหาย (Missing Values)
  - คลิกเมาส์ในช่องว่างใต้คอลัมน์ **Missing Values**
  - เลือก**กำหนดค่าขาดหาย**แบบค่าไม่ต่อเนื่อง (discrete) เช่น 0 (ไม่มีคำตอบ) 9 (ไม่ตอบ) 8 (ไม่ทราบ) หรือ เป็นช่วง (range) เช่น Low: 0 High: 14 Discrete value: 99 (วิเคราะห์เฉพาะค่าตั้งแต่ 15 ขึ้นไปและไม่รวม 99)
  - กดปุ่ม **OK**

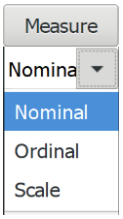


- การกำหนดความกว้างของช่องค่าตัวแปรหรือสดมภ์ (Columns)
  - **พิมพ์ตัวเลข**จำนวนความกว้างของช่องค่าตัวแปรใต้คอลัมน์ **Columns** **หรือกดปุ่ม - หรือ +**

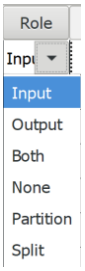
ค่าเริ่มต้น (default) ความกว้างของช่องค่าตัวแปรกำหนดไว้เป็น 8 ตำแหน่ง หากมีจำนวนตัวแปรมากและต้องการลดความกว้างของหน้าจอในการแสดงข้อมูลให้ลดตัวเลขจำนวนความกว้างลง หรือใช้เมาส์ดึงลดขนาดความกว้างในหน้าต่างแสดงข้อมูล (data view)

- การจัดตำแหน่งค่าตัวแปร (Align)
  - คลิกเมาส์ในช่องใต้คอลัมน์ **Align** (โปรแกรมจะแสดงเมนูรายการตำแหน่งค่าตัวแปร)
  - **เลือกตำแหน่ง**ค่าตัวแปร

- การกำหนดระดับการวัด (Measure)
  - คลิกเมาส์ในช่องใต้คอลัมน์ Measure (โปรแกรมจะแสดงเมนูรายการระดับการวัด)
  - เลือกระดับการวัด



- การกำหนดบทบาทของตัวแปร (Variable Role)
  - คลิกเมาส์ในช่องใต้คอลัมน์ **Role** (โปรแกรมจะแสดงเมนูรายการจุดมุ่งหมายของตัวแปร)
  - **เลือกประเภทที่มา**ของข้อมูลหรือตัวแปร

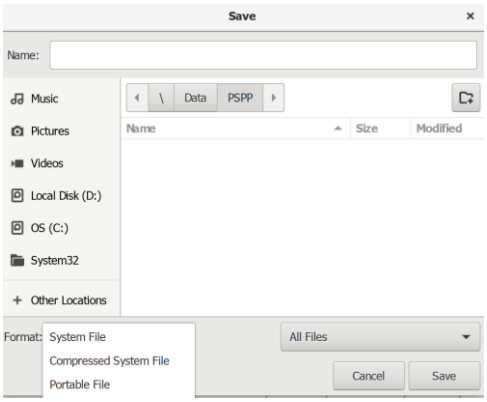


ผู้ใช้โปรแกรมอาจสร้างตัวแปรโดยใช้ค่าเริ่มต้นทั้งหมด แต่ที่สำคัญในการวิเคราะห์ข้อมูลต้องทราบว่า ตัวแปรใดเป็นตัวแปรแบบค่าแบ่งกลุ่มหรือตัวแปรแบบค่าต่อเนื่อง เพราะในการเลือกวิเคราะห์ด้วยสถิติแต่ละประเภทต้องเลือกใช้ตัวแปรให้ถูกต้องและตรงกับข้อตกลงเบื้องต้น (basic assumptions)

5. การบันทึกแฟ้มข้อมูล (Save)

แฟ้มข้อมูล (data files) ที่สร้างด้วยโปรแกรม PSPP ใช้โครงสร้างระบบแฟ้มข้อมูล (system file) แบบเดียวกับแฟ้มข้อมูลของโปรแกรม SPSS วิธีการบันทึกแฟ้มข้อมูลมีขั้นตอนดังต่อไปนี้

- เลือกเมนู **File -> Save... (Ctrl + S)** หรือ **Save As... (Shift + Ctrl + S)** หรือคลิกเมาส์ที่ไอคอน
  - เลือกพื้นที่ (folder) เก็บแฟ้มข้อมูล
  - พิมพ์ชื่อแฟ้มข้อมูลในช่อง Name (ทดลองพิมพ์ **ASPSS_Data** โดยไม่ต้องใส่นามสกุลของแฟ้มข้อมูล)
  - กดปุ่ม **Save**



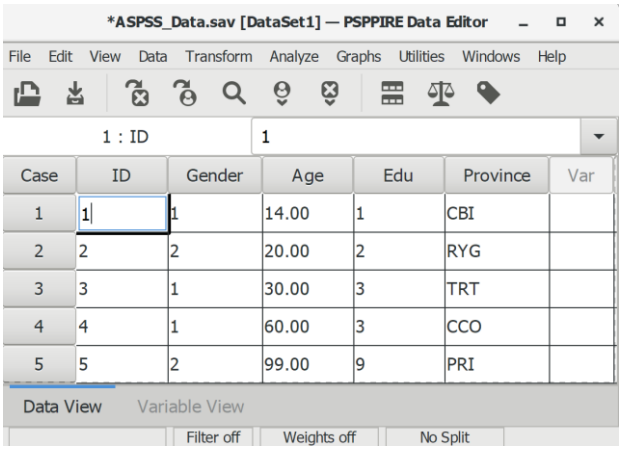
รูปแบบ (format) ของแฟ้มข้อมูล หากไม่เลือกรูปแบบใดโปรแกรมจะบันทึกให้เป็น System File เป็นแฟ้มข้อมูลแบบเก่าที่ไม่มีการบีบอัดข้อมูล ได้แฟ้มข้อมูลที่มีนามสกุลเป็น sav หากต้องการบันทึกเป็นแฟ้มข้อมูลแบบใหม่ที่มีการบีบอัดข้อมูลให้เลือกเป็น Compressed System File จะได้แฟ้มข้อมูลที่มีนามสกุล zsav และถ้าต้องการบันทึกเป็นแฟ้มข้อมูลแบบใช้ได้หลายระบบ สำหรับใช้แลกเปลี่ยนข้อมูล ให้เปิดอ่านได้อย่างเดียว ไม่สามารถแก้ไขได้ ให้เลือก Portable File จะได้แฟ้มข้อมูลที่มีนามสกุล por

1.2 การนำข้อมูลเข้าโปรแกรมแบบกรอกข้อมูล

ข้อมูลที่นำมาใช้ในการวิจัยมีที่มาจากหลายแหล่งข้อมูลและมีหลายรูปแบบ แต่ข้อมูลทั้งหลายต้องผ่านกระบวนการแปลงข้อมูลให้เป็นข้อมูลดิจิทัล (digital data) และมีโครงสร้างข้อมูล (data structure) เป็นตารางข้อมูลที่โปรแกรมสามารถนำไปประมวลผลได้

การวิจัยที่ใช้การสำรวจข้อมูลด้วยแบบสอบถามแบบเก่าที่เป็นกระดาษ หรือที่ปัจจุบันนิยมเรียกว่า แบบสอบถามแบบออฟไลน์ (offline questionnaire) ก่อนการกรอกข้อมูลต้องทำการสร้างแฟ้มข้อมูลก่อนตามที่วิธีการที่กล่าวมาแล้ว จากนั้นจึงทำการกรอกข้อมูลเข้า (entering data) ดังต่อไปนี้

- เปิดแฟ้มข้อมูล (แฟ้มข้อมูลที่สร้างไว้และยังไม่มีข้อมูล)
- คลิกเมาส์ที่แถบ **Data View** ด้านล่างซ้าย
- พิมพ์ค่าตัวแปร (Value) ใต้ช่องชื่อตัวแปร (Variable Name)
- คลิกเมาส์ที่ช่องว่างของตัวแปรถัดไป หรือใช้คีย์บอร์ดปุ่มลูกศรซ้าย เพื่อเลื่อน Cursor ไปยังช่องข้อมูลของตัวแปรถัดไป



ระหว่างกรอกข้อมูล ควรบันทึกข้อมูลเป็นระยะ โดยไปที่เมนู **File -> Save... (Ctrl + S)** หรือ **Save As... (Shift + Ctrl + S)** หรือคลิกเมาส์ที่ ไอคอน

2. การนำเข้าข้อมูลจากแฟ้มข้อมูล

ข้อมูลที่นำมาใช้ในการวิเคราะห์ด้วยโปรแกรมคอมพิวเตอร์อาจได้มาจากหลายแหล่งข้อมูล เช่น การบันทึกข้อมูลด้วยโปรแกรมตารางคำนวณ (spreadsheet software) การรวบรวมข้อมูลด้วยแบบสอบถามแบบออนไลน์ (online questionnaire) ข้อมูลที่ส่งออก (export) ออกมาจากฐานข้อมูล (database) และข้อมูลเปิด (open data) ของหน่วยงานต่างๆที่เผยแพร่ไว้ในอินเทอร์เน็ต โดยทั่วไปข้อมูลที่เผยแพร่ข้อมูลสู่สาธารณะและการให้บริการแลกเปลี่ยนข้อมูล นิยมทำเป็นแฟ้มข้อมู 2 รูปแบบดังนี้

1. แฟ้มข้อมูลแบบตัวอักษร (Text Files) แฟ้มข้อมูลมีนามสกุลเป็น csv หรือ txt หรือ dat แบ่งข้อมูลด้วยตัวคั่น (delimiter-separated formats) เช่น เครื่องหมายจุลภาค (comma) หรือเครื่องหมายคำพูด (quote) เครื่องหมายเว้นระยะ (tab)

2. แฟ้มข้อมูลจากโปรแกรมตารางคำนวณ (Spreadsheet Files) ที่เป็นโปรแกรมเปิดรหัส (open-source) เช่น Gnumeric Spreadsheet (*.gnumeric) และ Open Document Spreadsheet (*.ods)

โปรแกรม PSPP เป็นโปรแกรมทางสถิติแบบรหัสเปิดหรือโอเพนซอร์ซ (Open Source Statistical Software: OSSS) ที่ออกแบบมาเพื่อเป็นตัวเลือกใช้งานแทนโปรแกรม SPSS นอกจากสามารถเปิดแฟ้มข้อมูลของโปรแกรม SPSS ได้แล้ว ยังสามารถใช้ข้อมูลจากแฟ้มข้อมูลของโปรแกรมตารางคำนวณได้ด้วย เพราะมีโครงสร้างรูปแบบการเก็บข้อมูลเป็นตาราง (table) มีสดมภ์ (column) และแถว (row) เหมือนกัน

อย่างไรก็ตามหากเป็นแฟ้มข้อมูลที่สร้างด้วยโปรแกรมตารางคำนวณแบบรหัสเปิด เช่น Gnumeric, LibreOffice or Openoffice แฟ้มข้อมูลมีนามสกุล *.gnumeric หรือ *.ods สามารถนำข้อมูลเข้าโปรแกรม (importing data) ได้ทันทีโดยไม่ต้องแปลงแฟ้มข้อมูล

ส่วนแฟ้มข้อมูลจากโปรแกรมตารางคำนวณเชิงพาณิชย์ที่นิยมใช้อย่างแพร่หลาย เช่น Microsoft Excel (*.xls หรือ *.xlsx) และ Google Sheet (*.gsheet) ไม่สามารถนำข้อมูลเข้าโปรแกรมทันที ต้องทำการแปลงแฟ้มข้อมูลที่โปรแกรม PSPP สามารถอ่านเข้าโปรแกรมได้ โดยต้องแปลงให้เป็นแฟ้มข้อมูลที่มีเฉพาะข้อมูลตัวอักษรเท่านั้น (plain text file) เช่น แฟ้มข้อมูลที่มีตัวคั่นเป็นเครื่องหมายจุลภาค (Comma Separated Values) มีนามสกุล CSV (*.csv) หรือแฟ้มข้อมูลที่มีเครื่องหมายเว้นระยะ (Tab Separated Values) มีนามสกุล TXT (*.txt)

Import Data		
Text Files *.csv, *.txt และ *.prn	Spreadsheet Files (Commercial Software) *.xls, *.xlsx และ *.gsheet	Spreadsheet Files (Open Source Software) *.gnumeric และ *.ods

2.1 การนำข้อมูลเข้าจากโปรแกรมตารางคำนวณ

โปรแกรมตารางคำนวณเชิงพาณิชย์ที่นิยมใช้กันอย่างแพร่หลายมี 2 โปรแกรม คือ Microsoft Excel และ Google Sheet การทดลองแปลงแฟ้มข้อมูลทั้ง 2 โปรแกรมเป็นแฟ้มข้อมูลแบบตัวอักษรนามสกุล CSV และนำข้อมูลเข้าโปรแกรมไม่พบปัญหาใดๆ ส่วนการแปลงแฟ้มข้อมูลนามสกุล xls , xlsx และ gsheet เป็นแฟ้มข้อมูลนามสกุล ODS พบปัญหาการเข้ากันไม่ได้ (incompatibility) ของข้อมูลที่ได้จากแฟ้มข้อมูลนามสกุล xls และ xlsx ดังนั้น เพื่อให้ผู้เรียนได้ประสบการณ์การนำข้อมูลแบบตารางคำนวณเข้าโปรแกรมจึงแบ่งเป็น 2 วิธี โดยมีขั้นตอนดังนี้

ดาวโหลดแฟ้มข้อมูล “**รายงานด้านสังคม**” จากศูนย์กลางข้อมูลเปิดภาครัฐ (Open Government Data) ที่ **https://data.go.th** มา 1 ปี (ตัวอย่างปี พ.ศ. 2558) และ**บันทึก (Save)** เป็นแฟ้มข้อมูลชื่อ **ASPSS_Data_Excel.xls**

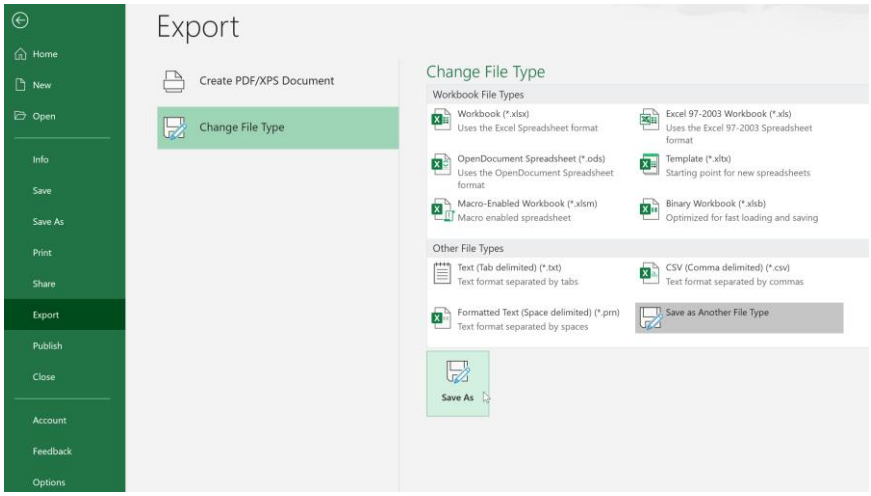
- เปิดแฟ้มข้อมูล **ASPSS_Data_Excel.xls** ด้วยโปรแกรม **Microsoft Excel**

	A	B	C	D	E	F	G	H
1	จำนวนชุมชนประเภทต่าง ๆ ในกรุงเทพมหานคร ณ เดือนมิถุนายน 2558							
2	สำนักงานเขต	เคหะชุมชน	หมู่บ้านจัดสรร	ชานเมือง	เมือง	อาคารสูง	แออัด	รวม
3	ดินแดง	2	-	8	2	3	8	23
4	ดุสิต	-	-	28	-	9	9	46
5	ป้อมปราบศัตรูพ่าย	-	-	4	-	-	11	15
6	พญาไท	-	-	1	2	12	14	29
7	พระนคร	-	-	15	-	-	5	20
8	ราชเทวี	-	-	-	-	6	19	25
9	วังทองหลาง	-	-	-	3	-	16	19
10	สัมพันธวงศ์	-	-	18	-	-	-	18

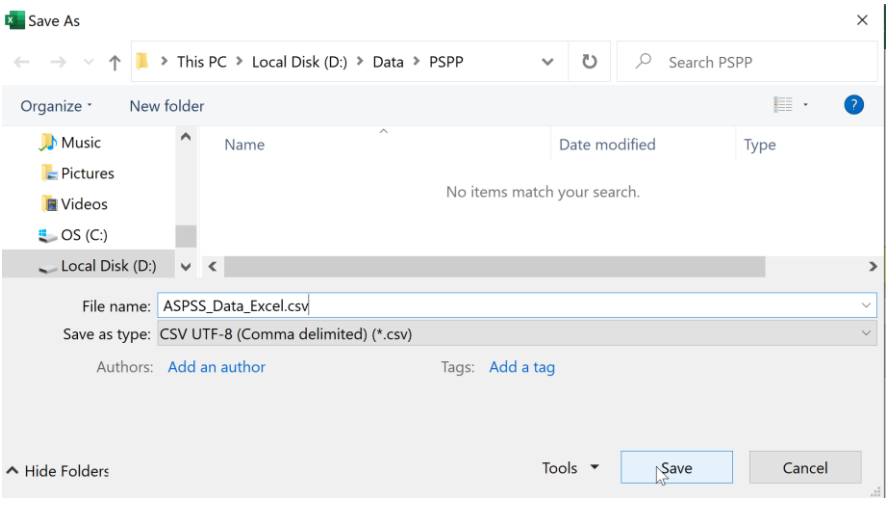
สมรส หย่า ฯ | หลังคาเรือนของชุมชน | ชุมชน | โครงการ อบรม สัมมนา | โรงเรียนฝึกอาชีพ

คลิกเมาส์ที่แถบข้อมูล ชุมชน ตัดข้อมูลแถวที่ 1 (จำนวนชุมชนประเภทต่าง ๆ ในกรุงเทพมหานคร ณ เดือนมิถุนายน 2558) และข้อมูลแถวที่ 53 (แหล่งข้อมูล : กองการพัฒนาชุมชน สำนักพัฒนาสังคม กรุงเทพมหานคร)

- เลือกเมนู **File**
- โปรแกรมแสดงหน้าต่าง Home
- คลิกเมาส์ที่คำสั่ง **Export -> Change File Type -> Save as Another File Type** (หรือเลือกใช้คำสั่ง Save As แทน Export)
- กดปุ่ม **Save As**

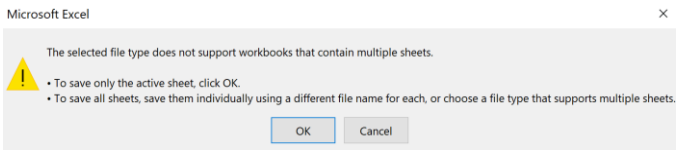


- เลือกพื้นที่เก็บข้อมูล ->พิมพ์ชื่อแฟ้มข้อมูลในช่อง **File name: ->**
- เลือกประเภทแฟ้มข้อมูลใน **Save as type:** เป็น **CSV UTF-8 (Comma delimited) (*.csv)**
- กดปุ่ม **Save**

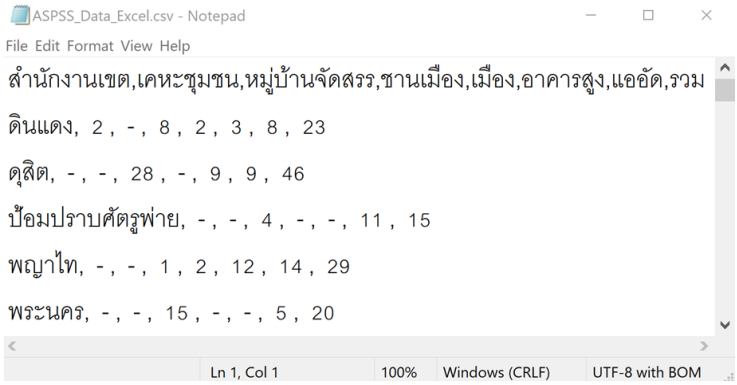


การเลือก**บันทึกแฟ้มข้อมูลเป็นแบบ UTF-8** (Universal Coded Character Set – 8 Bit) หากแฟ้มข้อมูลมีหลายภาษา เช่น ตัวเลข ตัวอักษรภาษาอังกฤษ ภาษาไทย และภาษาอื่นๆ จะทำให้การแสดงผลร่วมกันได้หรือแสดงภาษาผสมกัน ไม่เป็นข้อมูลที่อ่านไม่ออก

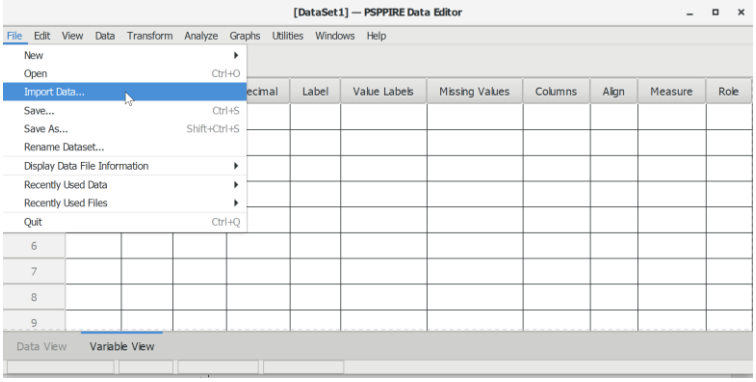
- โปรแกรมแสดงข้อความเตือนว่า “ประเภทแฟ้มข้อมูลที่เลือกไม่สนับสนุนสมุดงาน (workbooks) ที่มีอยู่ด้วยกันหลายแผ่นงาน (sheets) ” หากต้องการบันทึกเฉพาะแผ่นงานที่กำลังเปิดใช้งาน (active) ให้เลือกคลิก **OK**
- กดปุ่ม **OK**



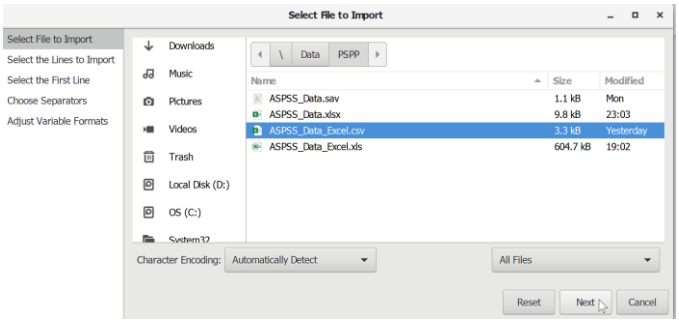
- ทดลอง**เปิดแฟ้ม**ข้อมูล **ASPSS_Data_Excel.csv** ด้วยโปรแกรม **Notepad** (ดูว่ามีตัวคั่นหรือแบ่งข้อมูลแบบใดบ้าง เพราะบางแฟ้มข้อมูลอาจมีเครื่องหมายแบ่งข้อมูลมากกว่า 1 แบบ)



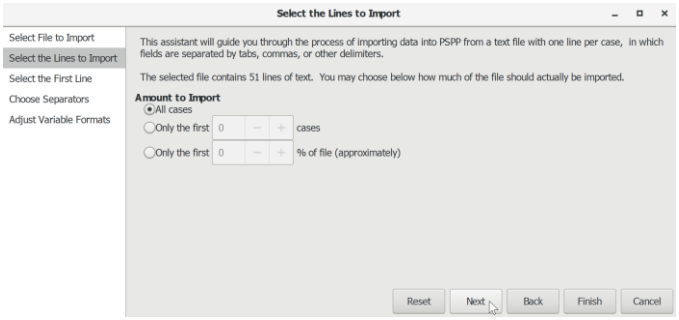
- เปิดโปรแกรม PSPP
- เลือกเมนู **File -> Import Data ...**



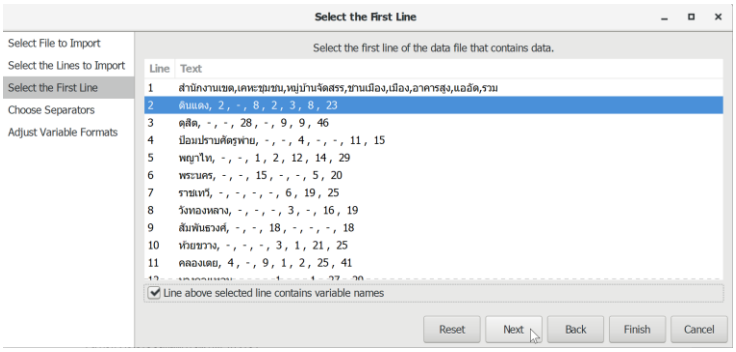
- เลือกแฟ้มข้อมูล **ASPSS_Data_Excel.csv**
- เมนูด้านซ้ายแสดงขั้นตอนการนำแฟ้มข้อมูลเข้า 5 ขั้นตอน
- กดปุ่ม **Next**



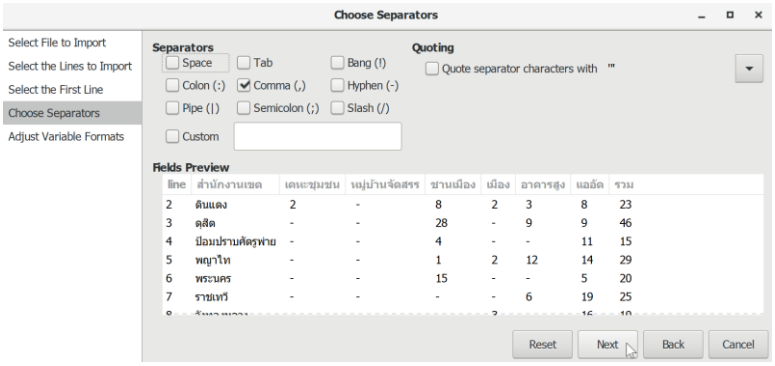
- กำหนดจำนวนข้อมูลที่ต้องการนำข้อมูลเข้า
  - **All cases** (ทั้งหมด)
  - Only the first … cases (จำนวน)
  - Only the first … % of file (approximately) (ร้อยละ)
- กดปุ่ม **Next**



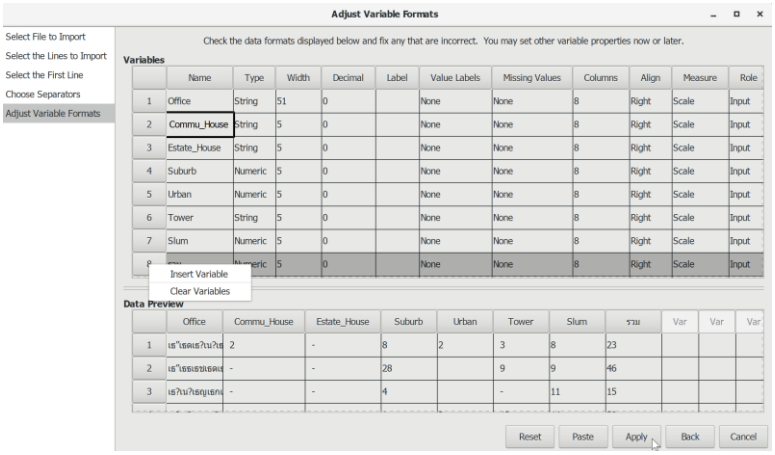
- คลิกเมาส์หรือใช้คีย์บอร์ดลูกศรลง -> เลื่อนแถบสีไป **Line ที่ 2**
- **Line above selected line contains variable names** (บรรทัดบนที่เลือกเป็นบรรทัดของชื่อตัวแปร) เลือก /
- กดปุ่ม **Next**



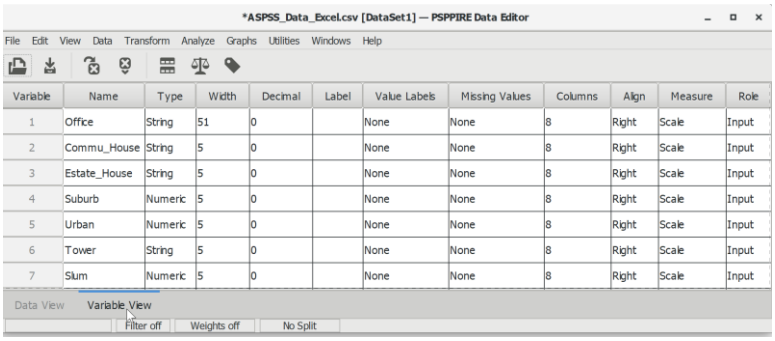
- **Separators** (ก่อนนำแฟ้มข้อมูลเข้าต้องตรวจดูว่ามีตัวคั่นอื่นหรือไม่ หากมีเครื่องหมายคั่นแบบอื่น ให้เลือกด้วย) คลิกเมาส์ / ที่ **Comma (,)**
- กดปุ่ม **Next**



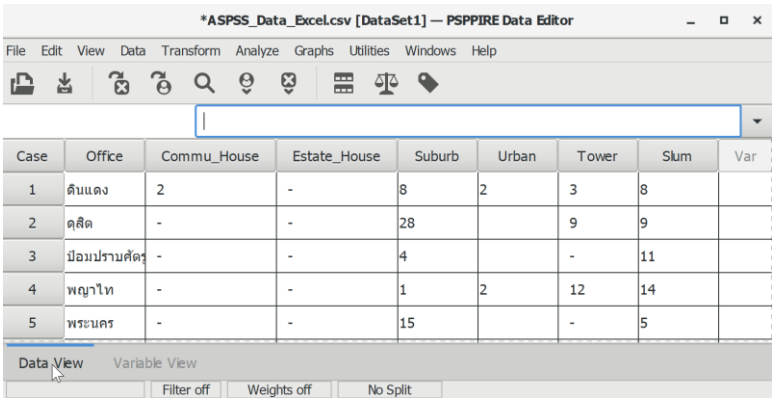
- กำหนดคุณสมบัติตัวแปร (ไม่กำหนดก็ได้ ไปกำหนดภายหลังใน Data Editor ก็ได้)
- คลิกเมาส์ปุ่มขวาที่ตัวแปรรวมและเลือกคำสั่ง **Clear Variables**
- กดปุ่ม **Apply**



- โปรแกรมแสดงหน้าต่าง **Data Editor** พร้อม **Output Viewer**



- คลิกเมาส์ที่แถบ **Data View** ด้านล่างซ้ายของโปรแกรมเพื่อดูข้อมูลที่นำเข้ามา



บันทึกแฟ้มข้อมูลเป็นแฟ้มข้อมูลนามสกุลเป็น sav เพื่อความสะดวกเปิดใช้ในครั้งต่อไป

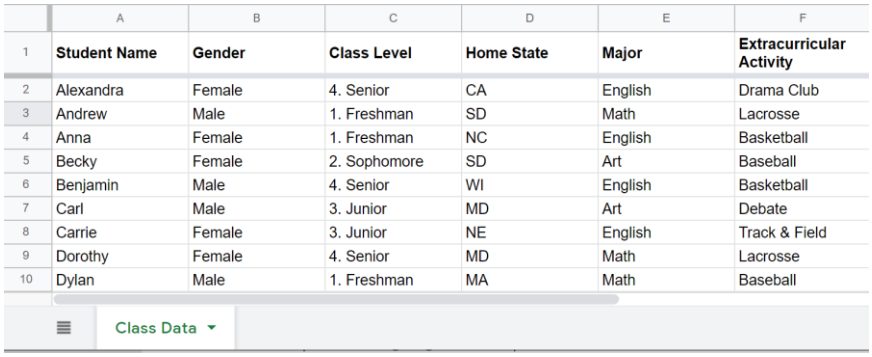
- เลือกเมนู **File -> Save As... (Shift + Ctrl + S)**

หากข้อมูลที่เป็นภาษาไทยแสดงผลเป็นภาษาต่างด้าว ให้แก้ไข Regional format ใน Settings ของระบบปฏิบัติการ Windows เป็น Thai (Thailand)

2.2 การนำข้อมูลเข้าจากแบบสอบถามออนไลน์

ปัจจุบันนักวิจัยนิยมใช้ Google Form สร้างแบบสอบถามออนไลน์ (online/digital questionnaire) มากขึ้น เพราะสามารถส่งแบบสอบถามผ่านอินเทอร์เน็ตไปถึงผู้ตอบได้สะดวดและรวดเร็ว ผลลัพธ์ที่ได้จากการสำรวจสามารถดูได้จากระบบ View Response ได้ทันที แต่หากต้องการนำมาวิเคราะห์ด้วยโปรแกรม PSPP ต้องบันทึกข้อมูลเข้าโปรแกรม Google Sheet เป็นแฟ้มข้อมูลนามสกุล *.gsheet และแปลงแฟ้มข้อมูลให้เป็นแฟ้มข้อมูลแบบตัวอักษรหรือเป็นแฟ้มข้อมูลของโปรแกรมตารางคำนวณแบบรหัสเปิด สำหรับในที่นี้จะแปลงเป็น *.ods เพื่อให้ผู้เรียนได้ทดลองนำเข้าข้อมูลจากแฟ้มข้อมูลแบบตารางคำนวณ

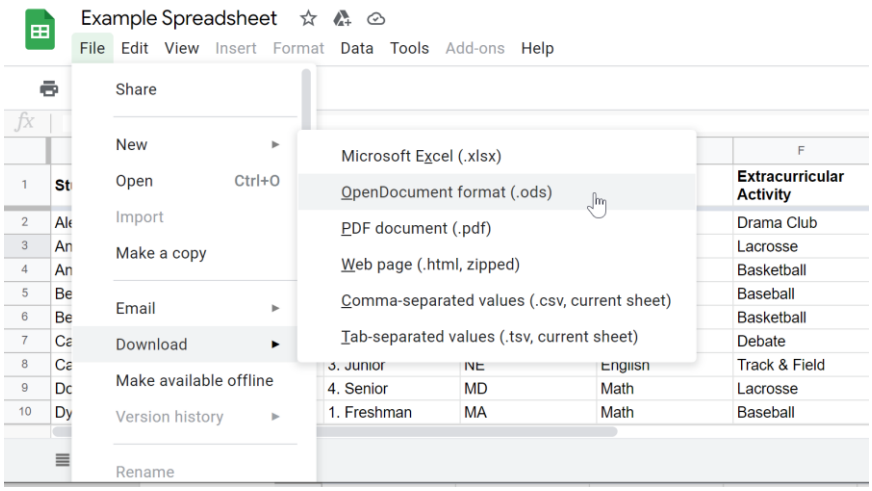
เปิดแฟ้มข้อมูล “Example Spreadsheet” ของ Google Sheet จากอินเทอร์เน็ต (**เปิด https://www.google.com** และ**ค้นหา**คำว่า “**example spreadsheet google sheet**”) ที่ลิงค์ https://docs.google.com/spreadsheets/d/1BxiMVs0XRA5nFMdKvBdBZjgmUUqptlbs74OgvE2upms/edit



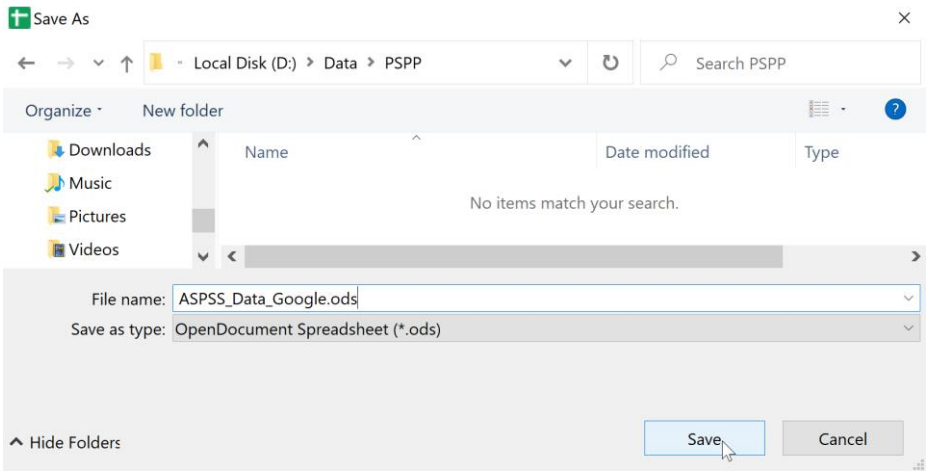
	A	B	C	D	E	F
1	**Student Name**	**Gender**	**Class Level**	**Home State**	**Major**	**Extracurricular Activity**
2	Alexandra	Female	4. Senior	CA	English	Drama Club
3	Andrew	Male	1. Freshman	SD	Math	Lacrosse
4	Anna	Female	1. Freshman	NC	English	Basketball
5	Becky	Female	2. Sophomore	SD	Art	Baseball
6	Benjamin	Male	4. Senior	WI	English	Basketball
7	Carl	Male	3. Junior	MD	Art	Debate
8	Carrie	Female	3. Junior	NE	English	Track & Field
9	Dorothy	Female	4. Senior	MD	Math	Lacrosse
10	Dylan	Male	1. Freshman	MA	Math	Baseball

Class Data

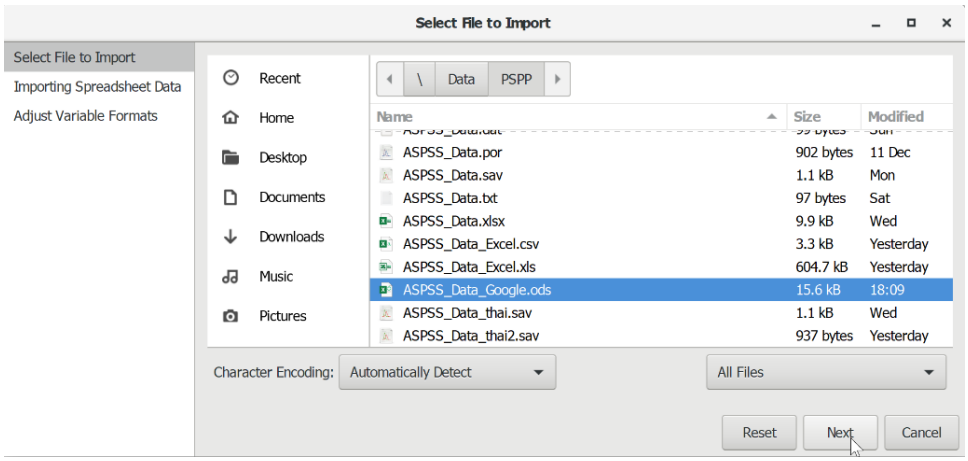
- เลือกเมนู **File -> Download -> OpenDocument format (.ods)**



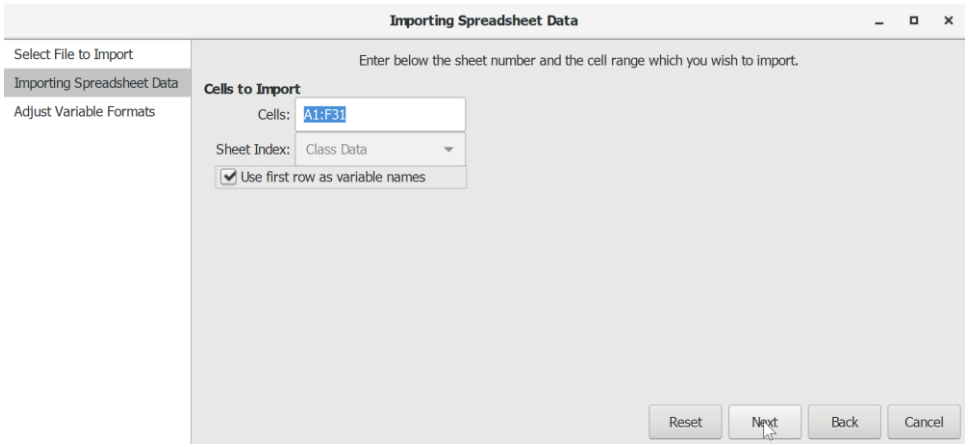
- **บันทึกแฟ้ม**ข้อมูลเป็นชื่อ **ASPSS_Data_Google.ods**



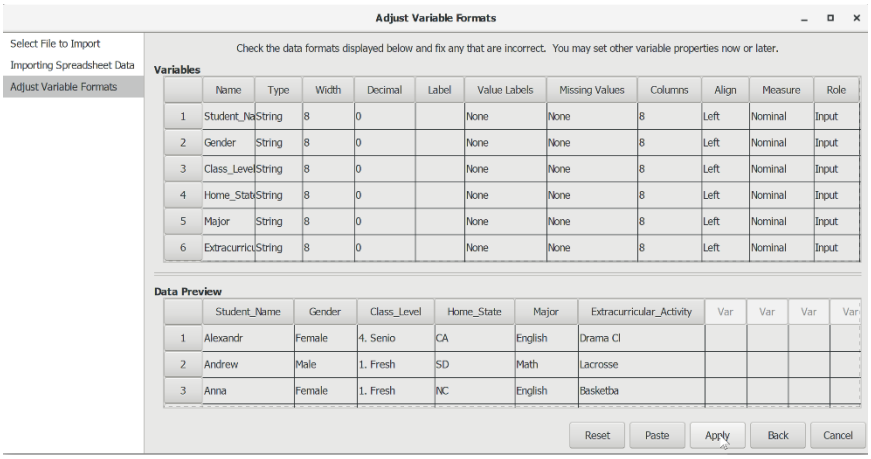
- เปิดโปรแกรม PSPP
- เมนู **File -> Import Data ...**
- เลือกแฟ้มข้อมูล **ASPSS_Data_Google.ods**
- เมนูด้านซ้ายแสดงขั้นตอนการนำแฟ้มข้อมูลเข้า 3 ขั้นตอน
- กดปุ่ม **Next**



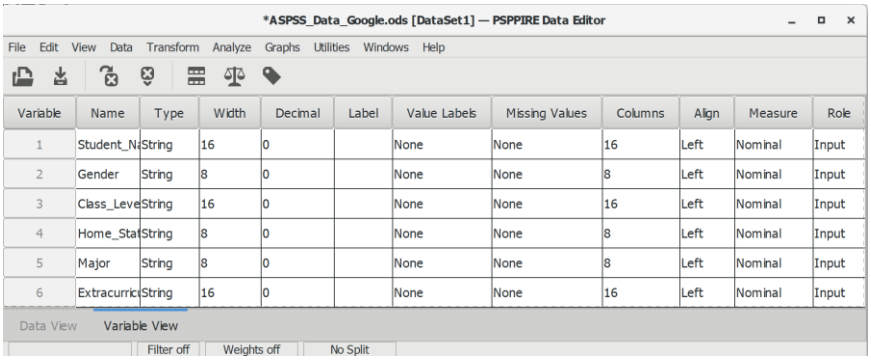
- ในช่อง **Cells:** กำหนดช่วงข้อมูล (ตรวจสอบจากแฟ้มข้อมูลให้ตรงกับที่โปรแกรมกำหนด)
- ใน **Sheet Index:** เลือกตารางข้อมูล (หากมีหลายตารางข้อมูลจะแสดงชื่อให้เลือก)
- **Use first row as variable names** (กำหนดให้แถวแรกเป็นชื่อตัวแปร) เลือก /
- กดปุ่ม **Next**



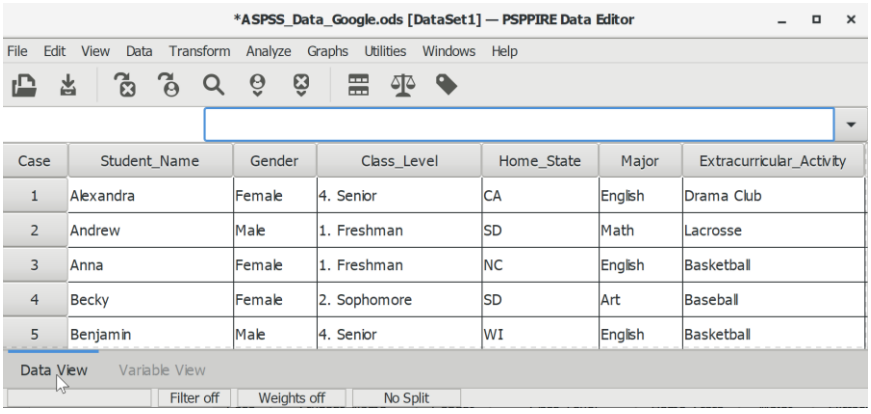
- กำหนดคุณสมบัติตัวแปร (ไม่กำหนดก็ได้ ไปกำหนดภายหลังใน Data Editor ก็ได้)
- กดปุ่ม **Apply**



- โปรแกรมแสดงหน้าต่าง Data Editor พร้อม Output Viewer



- คลิกเมาส์ที่แถบ Data View ด้านล่างซ้ายของโปรแกรมเพื่อดูข้อมูลที่นำเข้ามา



- บันทึกแฟ้มข้อมูลเป็นแฟ้มข้อมูลนามสกุลเป็น sav เพื่อความสะดวกเปิดใช้ในครั้งต่อไป
- เลือกเมนู **File -> Save As... (Shift + Ctrl + S)**

สรุป

การนำข้อมูลเข้าโปรแกรม คือ การเตรียมข้อมูลเพื่อนำข้อมูลจากแบบสอบถามหรือจากเอกสารที่บันทึกข้อมูลเข้าโปรแกรมวิเคราะห์ข้อมูลและบันทึกไว้เป็นแฟ้มข้อมูล วิธีการนำข้อมูลเข้าโปรแกรมมีทั้งวิธีการกรอกข้อมูลด้วยตนเอง และดึงข้อมูลจากแฟ้มข้อมูลที่มีอยู่แล้ว เช่น แฟ้มข้อมูลจากโปรแกรมตารางคำนวณ และแฟ้มข้อมูลแบบตัวอักษร

แบบฝึกหัดท้ายบท

ให้ผู้เรียนฝึกสร้างแฟ้มข้อมูลจากแบบสอบถาม และนำข้อมูลจากโปรแกรมตารางคำนวณเข้าโปรแกรมดังต่อไปนี้

1. ให้ผู้เรียนใช้แบบสอบถามที่สร้างไว้สำรวจข้อมูลภายในชั้นเรียนคนละ 5 ชุด และกรอกข้อมูลเข้าโปรแกรม
2. ให้ผู้เรียนสร้างแบบสอบถามออนไลน์ด้วย Google Form และสำรวจข้อมูลภายในชั้นเรียนคนละ 5 ชุด
3. ให้ผู้เรียนแปลงแฟ้มข้อมูลที่ได้จากแบบสอบถามออนไลน์และนำข้อมูลเข้าโปรแกรม
4. ให้ผู้เรียนค้นหาแฟ้มข้อมูลแบบตารางคำนวณ (*.xls และ *.xlsx) และแฟ้มข้อมูลแบบตัวอักษร (*.csv) และทดลองนำข้อมูลเข้าจากแฟ้มข้อมูล

บทที่ 7
การจัดการข้อมูล

ความมุ่งหมายของบทเรียน

ผู้เรียนสามารถเลือกใช้ข้อมูลและเปลี่ยนรูปแบบข้อมูลให้ตรงกับวัตถุประสงค์ของการวิเคราะห์ข้อมูล

เนื้อหา

1. การปรับข้อมูล
2. การเปลี่ยนรูปแบบข้อมูล

วิธีการสอนและกิจกรรมการเรียนการสอน

บรรยาย อธิบาย แสดงตัวอย่าง และทำแบบฝึกหัดท้ายบท

สื่อและอุปกรณ์การสอน

เอกสารคำสอน ภาพเลื่อน คอมพิวเตอร์ โปรแกรม PSPP โปรแกรมสอนออนไลน์ (สอนแบบออนไลน์) และเครื่องฉาย (สอนในชั้นเรียน)

งานที่มอบหมาย

ทำแบบฝึกหัดท้ายบท

การวัดผลและประเมินผล

การเข้าชั้นเรียน สังเกตความสนใจจากการเรียน สอบถามความเข้าใจจากการบรรยาย และการทำแบบฝึกหัดท้ายบท

เนื้อหา

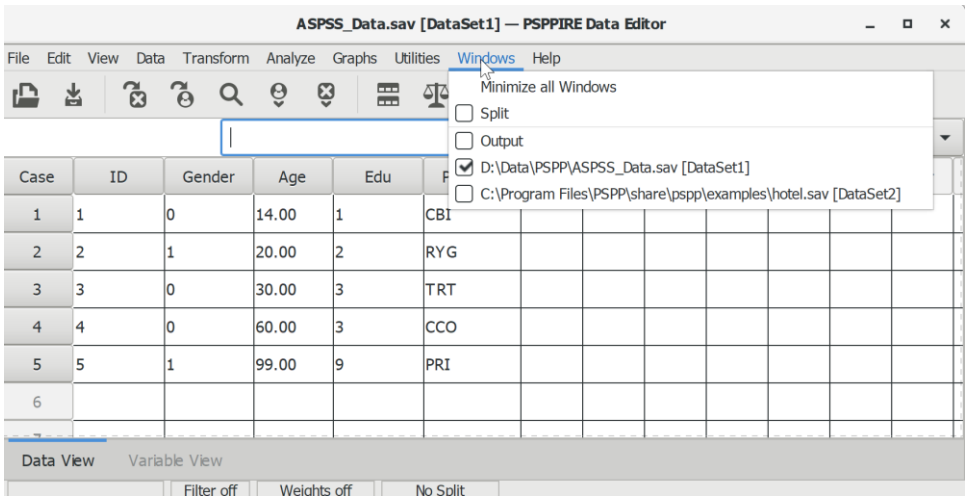
การจัดการข้อมูล (Data Management) คือ การปรับข้อมูล (data manipulation) ที่เป็นตัวแปรหรือค่าของตัวแปรในแฟ้มข้อมูล เช่น เลือกข้อมูล (selecting data) บางส่วนมาวิเคราะห์ หรือเปลี่ยนรูปแบบข้อมูล (data transformation) ให้เป็นตัวแปรใหม่หรือมีค่าของตัวแปรเป็นค่าใหม่

คำสั่งที่ใช้ในการจัดการข้อมูลมี 2 กลุ่มคำสั่ง คือ Data และ Transform ทั้ง 2 คำสั่งจะมีผลทำให้แฟ้มข้อมูลมีการเปลี่ยนแปลงแบบชั่วคราว (temporarily) แต่จะมีผลทำให้แฟ้มข้อมูลมีการเปลี่ยนแปลงแบบถาวร (permanently) เมื่อทำการบันทึกแฟ้มข้อมูล (save) บางคำสั่งโปรแกรมจะแสดงสถานะของแต่ละคำสั่งเป็น off หรือ by ในแถบสถานะภาพการทำงานด้านล่าง โดยมีรายละเอียดดังต่อไปนี้

Data	Transform
- Sort Cases... (การจัดเรียงตัวอย่าง)	- Compute... (การสร้างตัวแปรแบบค่าตัวเลข)
- Transpose... (การไขว้ตัวแปรและตัวอย่าง)	- Count... (การนับจำนวนตัวอย่าง)
- Aggregate... (การรวมค่าของกลุ่มตัวอย่าง)	- Rank Cases... (การจัดลำดับตัวอย่าง)
- Split File... (การแบ่งกลุ่มตัวอย่าง)*	- Automatic Recode...(การเปลี่ยนค่าตัวแปรเป็นตัวแปรใหม่แบบอัตโนมัติ)
- Select Cases... (การเลือกตัวอย่าง)*	- Recode into Same Variables... (การเปลี่ยนค่าตัวแปรในตัวแปรเดิม)
- Weight Cases... (การกำหนดน้ำหนักตัวอย่าง) *	- Recode into Different Variables... (การเปลี่ยนค่าตัวแปรเป็นตัวแปรใหม่)
	- Run Pending Transforms (การสั่งประมวลผลต่อ)

*คำสั่งแบบชั่วคราวจนกว่าจะมีการยกเลิกใช้คำสั่ง (turn off) หรือกำหนดให้เป็นค่าปกติ (default value)

บทนี้จะใช้แฟ้มข้อมูล ASPSS_Data.sav ที่เคยสร้างไว้ในบทที่ 5 การนำข้อมูลเข้าโปรแกรม และ hotel.sav ในโฟลเดอร์ examples (C:\Program Files\PSPP\share\pspp\examples) โดยสามารถเปิดใช้งานพร้อมกัน และสลับเลือกใช้งานจากเมนู Windows ดังนี้



การเลือกตัวแปรในรายการตัวแปรด้านซ้ายในหน้าต่างของแต่ละคำสั่ง มี 2 วิธี คือ ใช้เมาส์ชี้ที่ชื่อตัวแปรที่ต้องการเลือกแล้วกดปุ่มแบบดับเบิลคลิกที่ปุ่มซ้ายของเมาส์ หรือใช้เมาส์ชี้ที่ชื่อตัวแปรที่ต้องการเลือก ▸ แล้วกดปุ่มลูกศร

1. การปรับข้อมูล

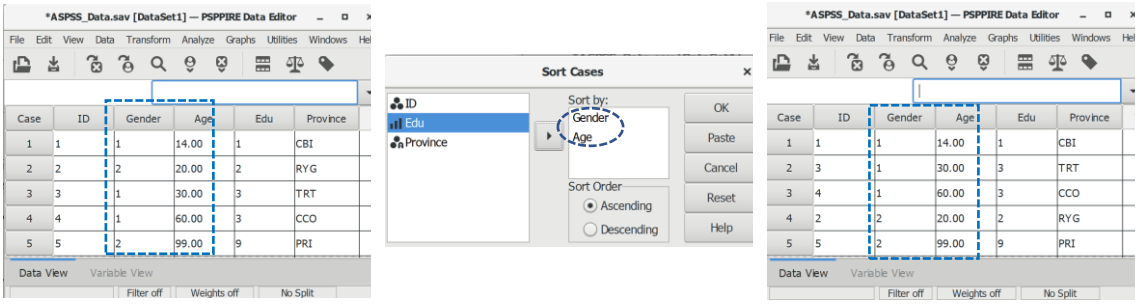
การปรับข้อมูล คือ การเปลี่ยนแปลงรูปแบบของข้อมูลทั้งหมด (global changes) ในชุดข้อมูล เช่น การจัดเรียงข้อมูล การไขว้ตัวแปร การรวมข้อมูล การแบ่งข้อมูล การเลือกข้อมูล และการกำหนดน้ำหนักข้อมูล โดยมีรายละเอียดดังต่อไปนี้

1.1 การจัดเรียงตัวอย่าง

การจัดเรียงตัวอย่าง (Sort Cases) คือ การเรียงตัวอย่างตามค่าของตัวแปร 1 ตัวแปร หรือหลายตัวแปร และสามารถเลือกเรียงลำดับจากค่าน้อยไปมาก (ascending order: 1-9 or A-Z) หรือจากค่ามากไปน้อย (ascending order: 9-1 or Z-A) โดยมีวิธีการดังนี้

- เลือกเมนู **Data -> Sort Cases...**
- **Sort by:** เลือกตัวแปรค่าจัดประเภท/ค่าต่อเนื่อง 1 ตัวแปรหรือมากกว่า จากรายการตัวแปรด้านซ้าย
- ในกลุ่มคำสั่ง **Sort Order** เลือกวิธีการจัดเรียง
  - Ascending (เรียงจากค่าน้อยไปมาก เช่น 1->9 หรือ A->Z)
  - Descending (เรียงจากค่ามากไปน้อย เช่น 9->1 หรือ Z->A)
- กดปุ่ม **OK**

ตัวอย่าง: การจัดเรียงตัวอย่างด้วยตัวแปร Gender และ Age จากค่าน้อยไปมาก



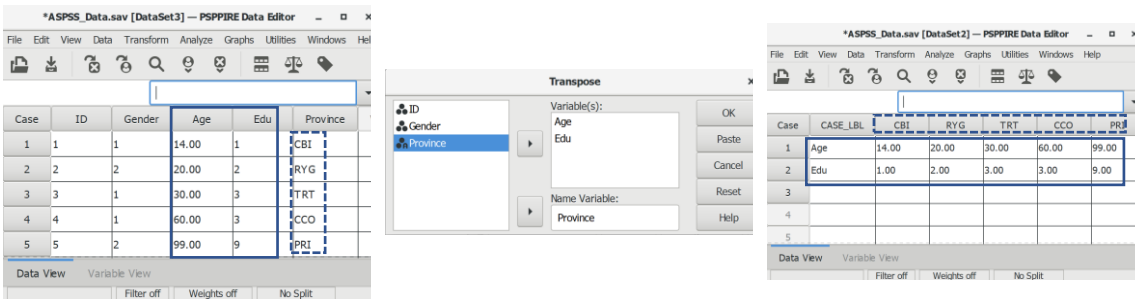
1.2 การไขว้ตัวแปรและตัวอย่าง

การไขว้ตัวแปรและตัวอย่าง (Transpose) คือ การสลับตัวแปร (variables) ที่อยู่ในรูปของสดมภ์ (columns) ไปเป็นตัวอย่าง (cases) อยู่ในรูปของแถว (rows) และสลับตัวอย่าง (cases) ที่อยู่ในรูปของแถว (rows) ไปเป็นตัวแปร (variables) อยู่ในรูปของสดมภ์ (columns)

ตัวแปรที่ต้องการไขว้ไปเป็นตัวอย่าง ต้องเป็นตัวแปรแบบตัวเลข (numeric variables) และตัวแปรที่ต้องการไขว้ค่าของตัวแปรไปเป็นตัวแปรใหม่ควรเป็นตัวแปรแบบตัวอักษร (string variables) เพราะค่าตัวแปรที่เป็นตัวเลข ตัวเลขแรกจะถูกเป็นตัวอักษร V ทั้งหมด โดยมีวิธีการดังนี้

- เมนู **Data -> Transpose...**
- **Variables:** เลือกตัวแปรค่าต่อเนื่อง 1 ตัวแปรหรือมากกว่า (ตัวแปร -> ตัวอย่าง) จากรายการตัวแปรด้านซ้าย
- **Name Variables:** เลือกตัวแปรค่าจัดกลุ่มในช่องด้านซ้าย 1 ตัวแปร (ตัวอย่าง -> ตัวแปร)
- กดปุ่ม **OK**

ตัวอย่าง: การไขว้ตัวแปร Age และ Edu เป็นตัวอย่าง และไขว้ค่าใน Province เป็นตัวแปร



1.3 การรวมกลุ่มตัวอย่าง

การรวมกลุ่มตัวอย่าง (Aggregate) คือ การรวมกลุ่ม (summarizes groups) ของตัวอย่าง (cases) ที่มีค่าเหมือนกันจากหลายตัวแปร (break variables) เช่น รวมรายได้ อายุ และประสบการณ์ทำงานของตัวอย่างที่มีระบบการศึกษาเหมือนกันให้เป็นกลุ่มการศึกษา โดยเลือกสร้างเป็นชุดข้อมูลใหม่ (data set) หรือสร้างเป็นแฟ้มข้อมูล (file) ใหม่ก็ได้ โดยมีวิธีการดังนี้

- เลือกเมนู **Data -> Aggregate...**
- **Break variable(s):** เลือกตัวแปรค่าจัดกลุ่ม 1 ตัวแปร จากรายการตัวแปรด้านซ้าย
- ในกลุ่มคำสั่ง **Aggregated variables:** กำหนดวิธีการรวมค่าของกลุ่มตัวอย่าง
  - **Variable Name:** พิมพ์ชื่อตัวแปรที่เป็นค่ารวมกลุ่มข้อมูล
  - **Variable Label:** พิมพ์คำอธิบายตัวแปร
  - ในกลุ่มคำสั่ง **Functions:** เลือกวิธีการรวมกลุ่มข้อมูล
    - Sum of values (ค่ารวม)
    - Mean average (ค่าเฉลี่ย)
    - Median average (ค่าเฉลี่ยมัธยฐาน)
    - Standard deviation (ค่าเบี่ยงเบนมาตรฐาน)
    - Maximum value (ค่าสูงสุด)
    - Minimum value (ค่าต่ำสุด)
    - Percentage greater than (ค่าร้อยละที่มากกว่า)
    - Percentage less than (ค่าร้อยละที่น้อยกว่า)
    - Percentage included in range (ค่าร้อยละที่อยู่ในช่วง)
    - Percentage excluded from range (ค่าร้อยละที่ไม่อยู่ในช่วง)
    - Fraction greater than (เศษส่วนที่มากกว่า)
    - Fraction less than (เศษส่วนที่น้อยกว่า)
    - Fraction included in range (เศษส่วนที่อยู่ในช่วง)
    - Fraction excluded from range (เศษส่วนที่ไม่อยู่ในช่วง)
    - Number of cases (จำนวนของตัวอย่าง)
    - Number of cases (unweighted) (จำนวนของตัวอย่างที่ไม่ถ่วงน้ำหนัก)

- Number of missing values (จำนวนของค่าสูญหาย)
- Number of missing values (unweighted) (จำนวนของค่าสูญหายที่ไม่ถ่วงน้ำหนัก)
- First non-missing value (ค่าแรกของค่าปกติ)
- Last non-missing value (ค่าสุดท้ายของค่าปกติ)

- **เลือกตัวแปรที่ต้องการรวมกลุ่มข้อมูล** 1 ตัวแปร จากรายการตัวแปรด้านซ้ายใส่ในช่องว่างหลังปุ่มลูกศร
- **Argument 1:** และ **Argument 2:** กำหนดเงื่อนไขการรวบกลุ่มข้อมูล (บาง Functions เช่น Percentage included in range -> Argument1: ค่าเริ่มต้น และ Argument 2: ค่าสิ้นสุด)

◎ กดปุ่ม **Add** เพื่อใส่คำสั่งการรวมกลุ่มตัวอย่างเข้าโปรแกรม

◎ ในกลุ่มคำสั่ง **Save** กำหนดวิธีการสร้างผลลัพธ์

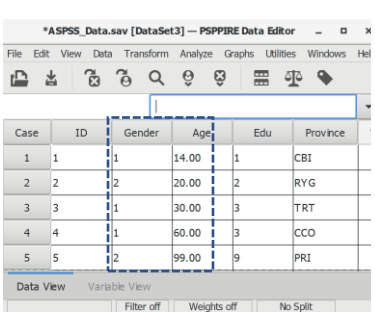
- Add aggregated variables to the active dataset (เพิ่มตัวแปรการรวมกลุ่มตัวอย่างลงในชุดข้อมูลที่เปิดใช้งาน)
- Replace the current dataset with the aggregated variables (สร้างตัวแปรการรวมกลุ่มตัวอย่างแทนที่ชุดข้อมูลที่เปิดใช้งาน)
- Write a new data file containing only the aggregated variables (สร้างแฟ้มข้อมูลใหม่ มีเฉพาะตัวแปรการรวมกลุ่มตัวอย่าง)

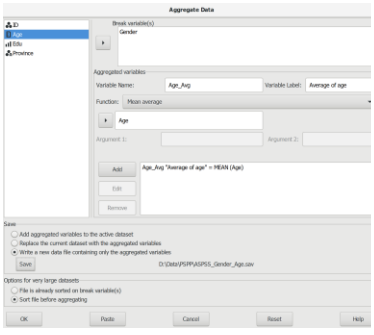
◎ ในกลุ่มคำสั่ง **Options for very large datasets** กำหนดวิธีการจัดเรียงตัวอย่าง (กรณีเป็นชุดข้อมูลขนาดใหญ่มาก)

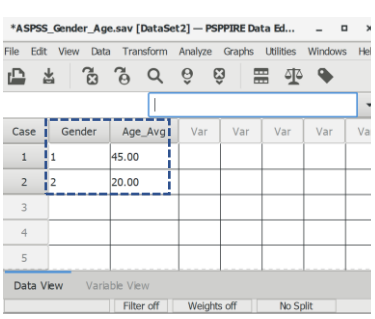
- File is already sorted on break variables(s). (แฟ้มข้อมูลทำการจัดเรียงตัวอย่างแล้ว)
- Sort file before aggregating. (จัดเรียงตัวอย่างก่อนทำการรวมกลุ่มตัวอย่าง)

◎ กดปุ่ม **OK**

ตัวอย่าง: การรวมค่าของตัวแปร Age เป็นค่าเฉลี่ยในแต่ละกลุ่มของตัวแปร Gender







1.4 การแบ่งกลุ่มตัวอย่าง

การแบ่งกลุ่มตัวอย่าง (Split File) คือ การแบ่งกลุ่มตัวอย่างในแฟ้มข้อมูลตามค่าของตัวแปรแบบจัดประเภท (categorical variables) เพื่อวิเคราะห์ตัวอย่างเป็นกลุ่มย่อยแบบชั่วคราว (temporary) จนกว่าจะยกเลิกการแบ่งกลุ่ม โดยมีวิธีการดังนี้

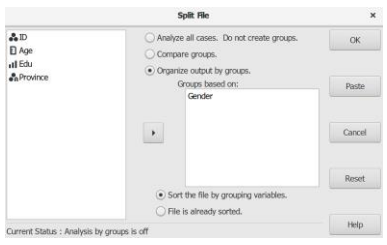
- ทดลองวิเคราะห์ค่าเฉลี่ยของอายุ
  - เลือกเมนู **Analyze -> Descriptive Statistics -> Descriptives...**
  - Variables: เลือกตัวแปร Age ในรายการตัวแปรด้านซ้าย
- เลือกเมนู **Data -> Split File...**
- ในกลุ่มคำสั่งด้านขวามือ เลือกวิธีการแบ่งกลุ่มตัวอย่าง
  - Analyze all cases. Do not create groups. (ไม่แบ่งกลุ่ม หรือ Split File Off)
  - Compare groups. (เปรียบเทียบค่าเฉลี่ยแต่ละกลุ่ม)
  - Organize output by groups. (แบ่งแสดงผลแต่ละกลุ่ม)
- **Groups based on:** เลือกตัวแปรค่าจัดกลุ่ม 1 ตัวแปรหรือมากกว่า จากรายการตัวแปรด้านซ้าย
- เลือกวิธีการจัดเรียงกลุ่ม
  - Sort the file by grouping variables. (เรียงกลุ่มตามตัวแปร)
  - File is already sorted. (แฟ้มข้อมูลนี้จัดเรียงกลุ่มแล้ว)
- กดปุ่ม **OK**
- ทดลองวิเคราะห์ค่าเฉลี่ยของอายุอีกครั้งและดูเปรียบเทียบผลลัพธ์

ตัวอย่าง: ผลการวิเคราะห์ค่าเฉลี่ยของตัวแปร Age ด้วยการแบ่งกลุ่มด้วยตัวแปร Gender

```
DESCRIPTIVES
  /VARIABLES= Age.
```

Descriptive Statistics

	N	Mean	Std Dev	Minimum	Maximum
Age of Responder	3	36.67	20.82	20.00	60.00
Valid N (listwise)	5				
Missing N (listwise)	2				




```
SORT CASES BY Gender.

SPLIT FILE SEPARATE BY  Gender.

DESCRIPTIVES
  /VARIABLES= Age.
```

Split Values

Variable	Value
Gender of Responder	Male

Descriptive Statistics

	N	Mean	Std Dev	Minimum	Maximum
Age of Responder	2	45.00	21.21	30.00	60.00
Valid N (listwise)	3				
Missing N (listwise)	1				

Split Values

Variable	Value
Gender of Responder	Female

Descriptive Statistics

	N	Mean	Std Dev	Minimum	Maximum
Age of Responder	1	20.00	.	20.00	20.00
Valid N (listwise)	2				
Missing N (listwise)	1				

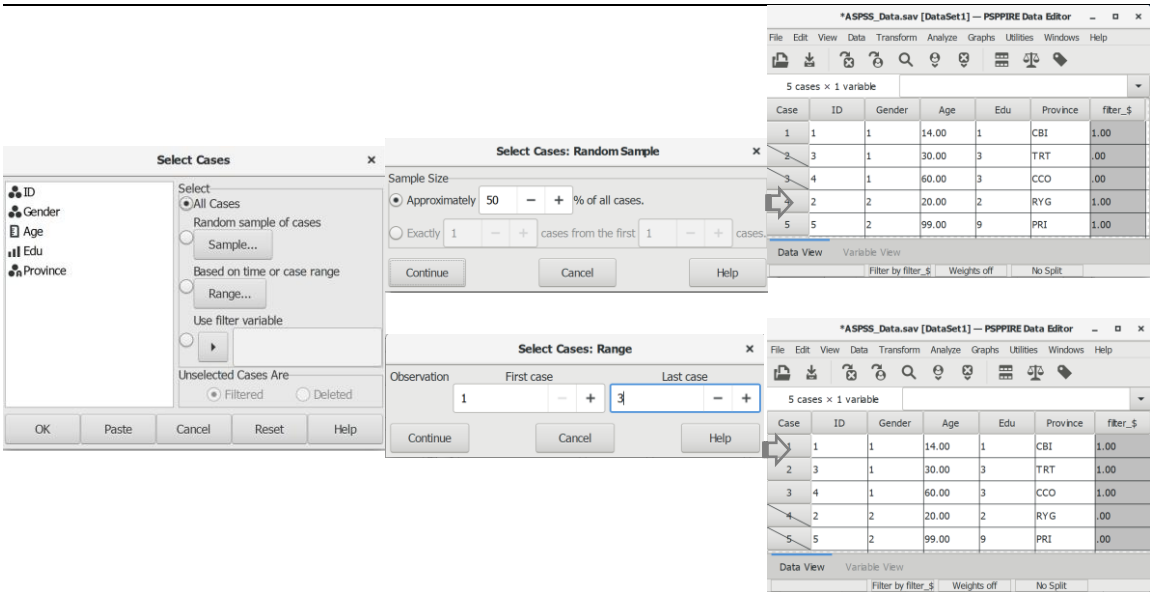
1.5 การเลือกตัวอย่าง

การเลือกตัวอย่าง (Select Cases) คือ การกำหนดเงื่อนไขเพื่อเลือกตัวอย่างสำหรับใช้วิเคราะห์ ข้อมูลที่ไม่ถูกเลือกตามเงื่อนไขจะถูกลบออกจากชุดข้อมูลที่ใช้วิเคราะห์ (data stream) แบบชั่วคราว (temporary) จนกว่าเงื่อนไขที่กำหนดไว้จะถูกยกเลิก โดยมีวิธีการดังนี้

เงื่อนไขที่เป็นค่าปกติ (default) คือ เลือกทั้งหมด (all cases) และมีตัวเลือก คือ เลือกสุ่มเป็นจำนวนร้อยละ (random) เลือกเฉพาะช่วงข้อมูล (range) และเลือกเฉพาะตัวอย่างที่ไม่มีค่าขาดหาย (use filter variable) ข้อมูลที่ไม่ถูกเลือก โดยมีวิธีการดังนี้

- เลือกเมนู **Data -> Select Cases...**
- ในกลุ่มคำสั่ง **Select** เลือกวิธีการเลือกตัวอย่าง
  - All Cases (เลือกทั้งหมด หรือ Filter off)
  - Random sample of cases (สุ่มขนาดตัวอย่างเป็นจำนวนร้อยละ)
    - กำหนดจำนวนร้อยละ
    - กดปุ่ม Continue
  - Based on time or case range (เลือกตัวอย่างตามช่วงที่กำหนด)
    - กำหนดช่วงตัวอย่าง
    - กดปุ่ม Continue
  - Use filter variable (ใช้ตัวแปรเลือกตัวอย่างที่ไม่มีค่าเป็น 0 หรือค่าสูญหาย)
    - เลือกตัวแปรจากช่องด้านซ้าย -> ใส่ในช่องว่าง
    - กดปุ่ม Continue
- ในกลุ่มคำสั่ง **Unselected Cases Are** กำหนดวิธีการเลือกตัวอย่าง
  - Filtered (กรองออกด้วยการทำสัญลักษณ์เป็นเครื่องหมาย \ ที่หน้าข้อมูลของแต่ละตัวอย่าง)
  - Deleted (ลบออกจากชุดข้อมูล)
- กดปุ่ม **OK**

ตัวอย่าง: การเลือกตัวอย่างด้วยวิธีสุ่มขนาดตัวอย่างเป็นจำนวนร้อยละและเลือกตัวอย่างตามช่วงที่กำหนด



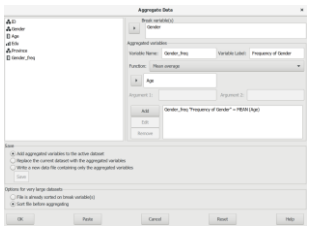
1.6 การกำหนดน้ำหนักตัวอย่าง

การกำหนดน้ำหนักตัวอย่าง (Weight Cases) คือ การกำหนดจำนวนตัวอย่างในการวิเคราะห์ให้มีจำนวนที่แตกต่างไปจากเดิมแบบชั่วคราว (temporary) ตามสัดส่วนของค่าในตัวแปรดัชนีที่มีอยู่ในแฟ้มข้อมูล จนกว่าจะยกเลิกการกำหนดน้ำหนักตัวอย่าง โดยมีวิธีการดังนี้

- ทดลองวิเคราะห์เปรียบเทียบค่าเฉลี่ยอายุระหว่างกลุ่มเพศ
- เลือกเมนู **Analyze -> Compare Means -> Means...**
  - **Dependent List:** เลือกตัวแปร Age จากรายการตัวแปรด้านซ้าย
  - **Independent List:** เลือกตัวแปร Gender จากรายการตัวแปรด้านซ้าย
  - กดปุ่ม **OK**
- ใช้คำสั่ง **Aggregate สร้างตัวแปร Gender_freq** เพื่อรวมอายุของกลุ่มตัวอย่าง
- เลือกเมนู **Data -> Weight Cases...**
- ในกลุ่มคำสั่งด้านขวา เลือกวิธีการ**กำหนดน้ำหนักตัวอย่าง**
  - Do not weight cases (ไม่กำหนดน้ำหนักตัวอย่าง หรือ Weight Off)
  - Weight cases by (กำหนดน้ำหนักจากตัวแปรที่เลือก)
    - เลือกตัวแปรความถี่จากรายการตัวแปรในช่องด้านซ้าย -> Frequency Variable
- กดปุ่ม **OK**

◎ ทดลองวิเคราะห์เปรียบเทียบค่าเฉลี่ยอายุระหว่างกลุ่มเพศอีกครั้งและดูเปรียบเทียบผลลัพธ์

ตัวอย่าง: ผลการวิเคราะห์ค่าเฉลี่ยตัวแปร Age ก่อนและหลังการกำหนดน้ำหนักตัวอย่างด้วยตัวแปร Gender



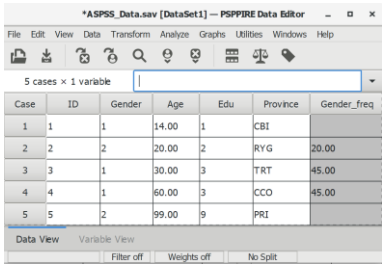
MEANS TABLES = Age
BY Gender.

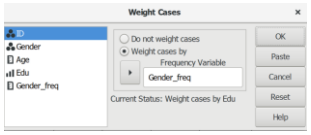
Case Processing Summary

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
Age * Gender	8	88.9%	1	11.1%	9	100.0%

Report

Gender of Responder	Mean	N	Std. Deviation
Male	45.00	6	16.43
Female	20.00	2	.00
Total	38.75	8	18.08





WEIGHT BY Gender_freq.

MEANS TABLES = Age
BY Gender.

Case Processing Summary

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
Age * Gender	110	100.0%	0	.0%	110	100.0%

Report

Gender of Responder	Mean	N	Std. Deviation
Male	45.00	90	15.08
Female	20.00	20	.00
Total	40.45	110	16.72

2. การแปลงข้อมูล

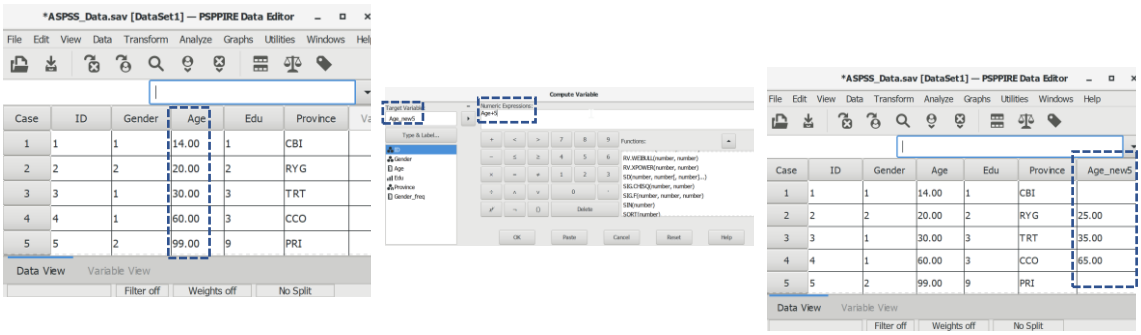
การแปลงข้อมูล คือ การเปลี่ยนแปลงข้อมูลเฉพาะตัวแปรที่เลือก เช่น การเปลี่ยนค่าของตัวแปร การสร้างตัวแปรใหม่

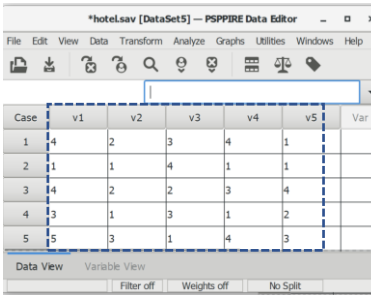
2.1 การสร้างตัวแปรแบบค่าต่อเนื่อง

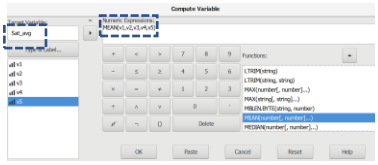
การสร้างตัวแปรแบบค่าต่อเนื่อง (Compute) คือ การสร้างหรือเปลี่ยนค่าตัวแปรที่มีค่าเป็นตัวเลขแบบต่อเนื่อง (numeric) ให้เป็นตัวแปรใหม่ โดยใช้ฟังก์ชั่น (function) ทางคณิตศาสตร์ เช่น + (บวก) – (ลบ) * (คูณ) / (หาร) Sum (ผลรวม) Mean (ค่าเฉลี่ย) SD (ค่าเบี่ยงเบนมาตรฐาน) Median (ค่ามัธยฐาน) Mod (ฐานนิยม) เขียนเป็นนิพจน์เชิงตัวเลข (numerical expression) โดยมีวิธีการดังนี้

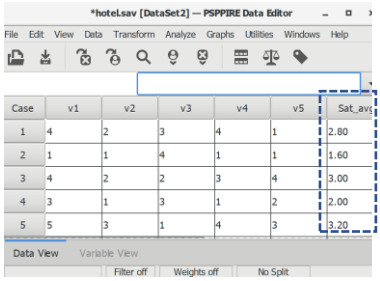
- เลือกเมนู **Transform -> Compute...**
- **Target Variable:** พิมพ์ชื่อตัวแปรใหม่
- กดปุ่ม **Type & Label...** -> กำหนดรายละเอียดเกี่ยวกับตัวแปร เช่น คำอธิบายชื่อตัวแปร และประเภทตัวแปร (ไม่กำหนดก็ได้)
- ในช่อง **Numeric Expression** สร้างนิพจน์เชิงตัวเลข มี 2 วิธีการดังนี้
  - สร้างนิพจน์เชิงตัวเลขเอง เช่น Age+5 หรือ V1+V2+V3 หรือ (V1+V2+V3)/3
  - เลือกใช้โปรแกรมฟังก์ชั่นในกลุ่มคำสั่ง Functions: เช่น SUM(V1, V2, V3) หรือ MEAN (V1, V2, V3)

ตัวอย่าง: การสร้างตัวแปร Age_new5 โดยเพิ่มค่า 5 ปี ในตัวแปร Age และการสร้างตัวแปรคะแนนเฉลี่ยความพึงพอใจ Sat_avg จากตัวแปรคะแนนความพึงพอใจ 5 ตัวแปร (V1, V2, V3, V4, V5)







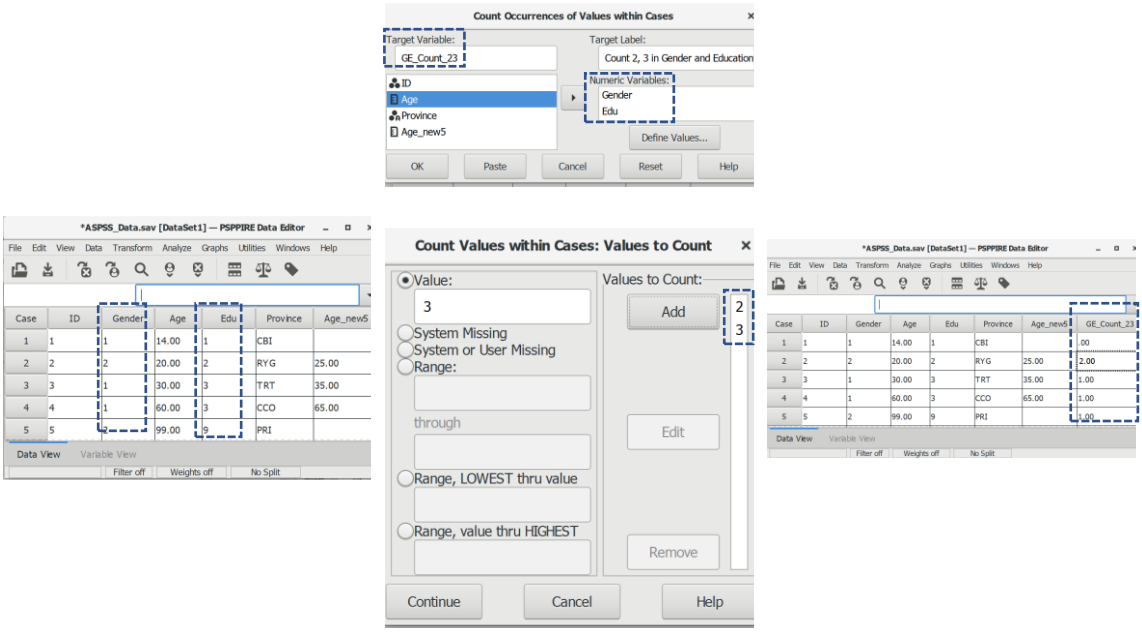


2.2 การนับจำนวนข้อมูล

การนับจำนวนข้อมูล (Count) คือ การสร้างตัวแปรใหม่จากการนับค่าตามเงื่อนไข (criterion value) ที่เกิดขึ้น (occurrences) ภายในรายการตัวแปร (variable) ของแต่ละตัวอย่าง (cases) โดยมีวิธีการดังนี้

- เลือกเมนู **Transform -> Count...**
- **Target Variable:** พิมพ์ชื่อตัวแปรใหม่
- **Numeric Variables:** เลือกตัวแปรในช่องรายการตัวแปรที่ต้องการนับค่า
- กดปุ่ม **Define Values...**
- **เลือกวิธีการนับค่า**
  - Value: (ค่าใดค่าหนึ่งของตัวแปร)
  - System Missing (ค่าสูญหายของระบบ)
  - System or User Missing (ค่าสูญหายของระบบหรือผู้ใช้)
  - Range: (ช่วงค่าของตัวแปร)
  - Rang, LOWEST thru value (ช่วงค่าของตัวแปรที่กำหนดถึงค่าต่ำสุด)
  - Rang, value thru HIGHEST (ช่วงค่าของตัวแปรที่กำหนดถึงค่าสูงสุด)
- ในกลุ่มคำสั่ง **Values to Count:** กดปุ่ม **Add**
- กดปุ่ม **Continue**
- กดปุ่ม **OK**

ตัวอย่าง: การนับจำนวนค่า 2 และ 3 ในตัวแปร Gender และ Edu และสร้างเป็นตัวแปร GE_Count_23



2.3 การจัดลำดับตัวอย่าง

การจัดลำดับตัวอย่าง (Rank Cases) คือ การสร้างตัวแปรใหม่จากการจัดลำดับกลุ่มตัวอย่างด้วยการจัดเรียงค่า (order) ของตัวแปรประเภทตัวเลข (numeric variables) ตามคำสั่งที่กำหนด (mandatory) เช่น ค่าที่เท่ากัน (tied) ค่าลำดับที่ (rank) ค่าเปอร์เซ็นต์ไทล์ (percentile) ค่าสัดส่วน (proportion) ฯลฯ โดยมีวิธีการดังนี้

- เลือกเมนู **Transform -> Rank Cases...**
- **Variable(s):** เลือกตัวแปรแบบค่าต่อเนื่อง 1 ตัวแปรหรือมากกว่า จากรายการตัวแปรด้านซ้าย
- **By:** เลือกหรือไม่เลือกตัวแปรก็ได้
- ในกลุ่มคำสั่ง **Assign rank 1 to:** เลือกวิธีจัดลำดับ
  - Smallest Value (ค่าน้อยสุด)
  - Largest Value (ค่ามากสุด)
- กดปุ่ม **Rank Types** เพื่อเลือกประเภทการจัดลำดับ
  - Rank (จัดลำดับ)

- Fractional rank as % (ค่าร้อยละของลำดับค่าเศษส่วน)
- Savage score (คะแนนดิบ)
- Sum of case weights (ค่ารวมของการกำหนดน้ำหนักของตัวอย่าง)
- Fractional rank (ลำดับของค่าเศษส่วน)
- Ntiles (ค่าเปอร์เซ็นไทล์)
- Proportion Estimates (ประมาณการของค่าสัดส่วน)
- Normal Scores (คะแนนมาตรฐาน)
- ในกลุ่มคำสั่ง Proportion Estimation Formula เลือกสูตรการประมาณการค่าสัดส่วน
  - Blom (บลอม)
  - Tukey (ทูกีย์)
  - Rankit (รันชิต)
  - Van der Waerden (แวน เดอ แวร์เดน)

◎ กดปุ่ม **Continues**

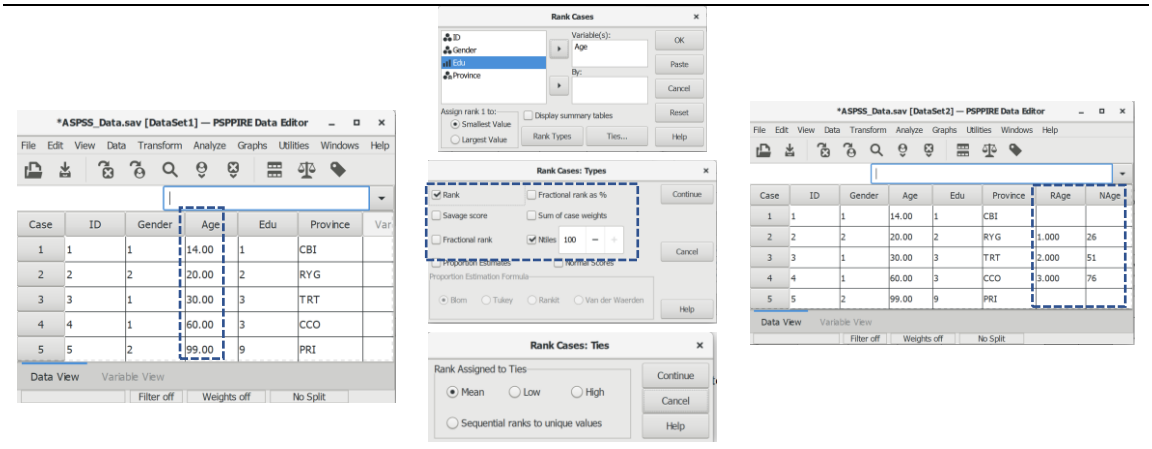
◎ กดปุ่ม **Ties...** เพื่อกำหนดวิธีจัดอันดับหากมีค่าเท่ากัน

- Mean (ค่าเฉลี่ย)
- Low (ค่าน้อย)
- Hight (ค่ามาก)
- Sequential ranks to unique value (เรียงอันดับต่อไปจนถึงค่าที่ไม่เท่ากัน)

◎ กดปุ่ม **Continuous**

◎ กดปุ่ม **OK**

ตัวอย่าง: การจัดลำดับตัวอย่างด้วยตัวแปร Age โดยใช้ Rank และ Ntiles

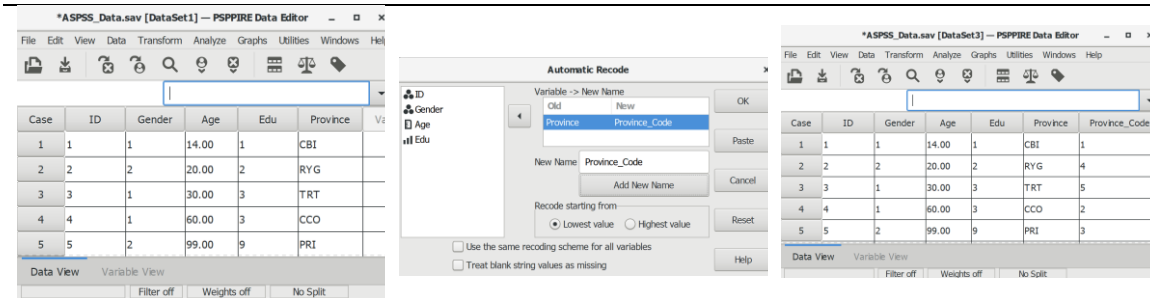


2.4 การสร้างตัวแปรแบบอัตโนมัติ

การสร้างตัวแปรแบบอัตโนมัติ (Automatic Recode) คือ การสร้างตัวแปรใหม่โดยให้โปรแกรมเปลี่ยนค่าตัวแปรที่เป็นตัวอักษร (string) ให้เป็นตัวเลข (numeric) แบบอัตโนมัติ ค่าใหม่ที่เป็นตัวเลขเกิดจากการใช้ตัวอักษรไปเทียบค่ากับรหัสอักขระ (character code) แล้วแปลงเป็นตัวเลขตามลำดับ เช่น ค่าที่เป็นตัวอักษรในตัวแปรเพศที่มีค่าเป็น “ชาย” ถูกเปลี่ยนเป็น 1 และ “หญิง” ถูกเปลี่ยนเป็น 2 แต่ถ้าเป็น “male” ถูกเปลี่ยนเป็น 2 และ “female” ถูกเปลี่ยนเป็น 1

- เลือกเมนู **Transform -> Automatic Recode...**
- **Old:** เลือกตัวแปรแบบตัวอักษรในช่องรายการตัวแปร 1 ตัวแปรหรือมากกว่า
- **New Name**: คลิกเมาส์ที่ชื่อตัวแปรเก่า -> พิมพ์ชื่อตัวแปรใหม่
- กดปุ่ม **Add New Name** (ใส่ชื่อตัวแปรใหม่ในช่อง New)
- **เลือกวิธีการสร้างตัวแปรแบบอัตโนมัติ**
  - Recode starting form (สร้างค่าตัวแปรจาก)
    - Lowest value (ค่าน้อยสุด)
    - Highest value (ค่ามากสุด)
  - Use the same recode scheme for all variables (ใช้วิธีการแปลงค่าเหมือนกันทุกตัวแปร)
  - Treat blank string values as missing (ตัวแปรที่มีค่าว่างให้ทำเป็นค่าสูญหาย)
- กดปุ่ม **OK**

ตัวอย่าง: การสร้างรหัสหรือค่าตัวแปรใหม่แบบอัตโนมัติจากตัวแปรที่มีค่าเป็นตัวอักษร

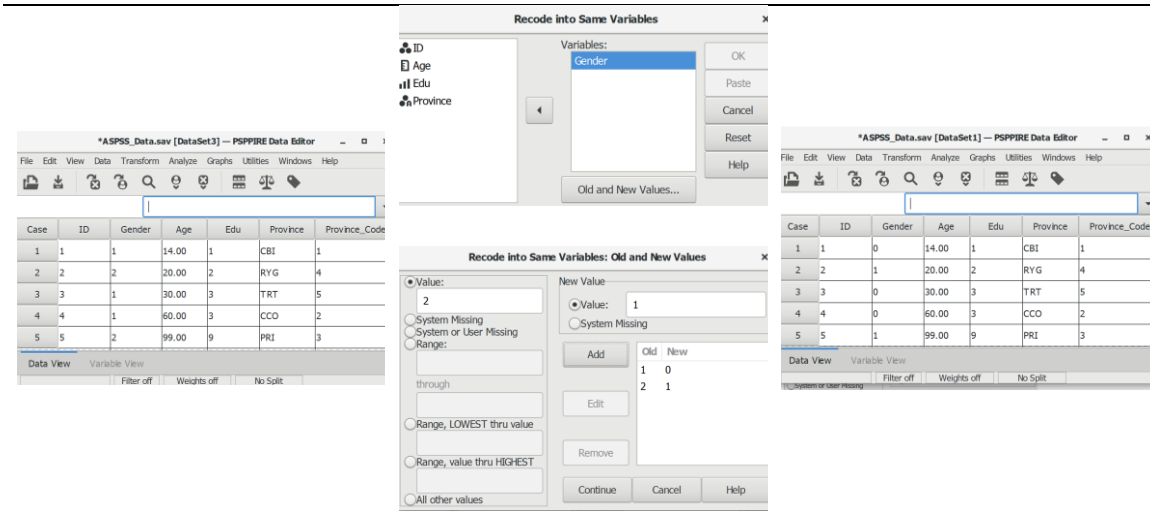


2.5 การเปลี่ยนค่าตัวแปรในตัวแปรเดิม

การเปลี่ยนค่าตัวแปรในตัวแปรเดิม (Recode into Same Variables) คือ การแทนค่าหรือเปลี่ยนค่าเก่าให้เป็นค่าใหม่ในตัวแปรเดิม การเปลี่ยนค่าตัวแปรแบบนี้ต้องมั่นใจว่าจะไม่ใช้ค่าเดิมแล้ว โดยมีวิธีการดังนี้

- เลือกเมนู **Transform -> Recode into Same Variables...**
- **Variables:** เลือกตัวแปรแบบตัวเลข/ตัวอักษร 1 ตัวแปรหรือมากกว่า จากรายการตัวแปรด้านซ้าย -> คลิกเมาส์ที่ชื่อตัวแปรในช่อง Variables
- กดปุ่ม **Old and New Values...**
  - ในกลุ่มคำสั่งด้านซ้าย **กำหนดค่าเก่าของตัวแปร**
    - Value: (ค่าใดค่าหนึ่งของตัวแปร)
    - System Missing (ค่าสูญหายของระบบ)
    - System or User Missing (ค่าสูญหายของระบบหรือผู้ใช้)
    - Rang: through (ช่วงจากค่า...ถึงค่า...)
    - Range, LOWEST thru value (ช่วงจากค่าต่ำสุดถึงค่า...)
    - Range, value thru HIGHEST (ช่วงจากค่าสูงสุดถึงค่า...)
    - All other values (ค่าอื่นๆทั้งหมด)
  - ในกลุ่มคำสั่ง **New Value** กำหนดค่าใหม่
    - Value (ค่าใหม่)
    - System Missing (กำหนดให้เป็นค่าสูญหายของระบบ)
    - กดปุ่ม **Add** (เพิ่มค่า) หรือ Edit (แก้ไข) หรือ Remove (ลบทิ้ง)
  - กดปุ่ม **Continue**
- กดปุ่ม **OK**

ตัวอย่าง: การแปลงตัวแปร Gender ให้เป็นตัวแปรหุ่น (dummy variables) โดยการแทนค่าเดิม 1 = Male และ 2 = Female ให้เป็นค่าใหม่ 1 = Female และ 0 = Non-femle



2.6 การสร้างตัวแปรแบบค่าจัดกลุ่ม

การสร้างตัวแปรแบบค่าจัดกลุ่ม (Recode into Different Variables) คือ การสร้างตัวแปรจากค่าเก่าที่เป็นค่าต่อเนื่องหรือค่าจัดกลุ่มให้เป็นตัวแปรใหม่มีค่าเป็นค่าจัดกลุ่ม โดยยังคงเก็บชื่อตัวแปรและค่าของตัวแปรไว้ตามเดิม โดยมีวิธีการดังนี้

- เลือกเมนู **Transform -> Recode into Different Variables...**
- **Variables:** เลือกตัวแปรแบบตัวเลข/ตัวอักษรในช่องรายการตัวแปร 1 ตัวแปรหรือมากกว่า
- ใต้คำว่า **Old** ในช่อง **Variables:** คลิกเมาส์ที่ชื่อตัวแปรเก่า
- ในกลุ่มคำสั่ง **Output Variable** ช่อง Name: พิมพ์ชื่อตัวแปรใหม่ และช่อง Label: คำอธิบายของตัวแปรใหม่ -> กดปุ่ม **Change** (หากมีหลายตัวแปรให้ทำวิธีเดียวกัน)
- กดปุ่ม **Old and New Values...**
  - ในกลุ่มคำสั่งด้านซ้าย **กำหนดค่าเก่าของตัวแปร**
    - Value: (ค่า)
    - System Missing (ค่าสูญหายของระบบ)
    - System or User Missing (ค่าสูญหายของระบบหรือผู้ใช้)
    - Rang: through (ช่วงจากค่า...ถึงค่า...)
    - Range, LOWEST thru value (ช่วงจากค่าต่ำสุดถึงค่า...)
    - Range, value thru HIGHEST (ช่วงจากค่าสูงสุดถึงค่า...)

- All other values (ค่าอื่นๆทั้งหมด)

o กำหนดค่าใหม่ในกลุ่มคำสั่ง **New Value**

- Value (ค่าใหม่)

- System Missing (กำหนดให้เป็นค่าสูญหายของระบบ)

- Copy old values (สำเนาหรือใช้ค่าเก่า)

- กดปุ่ม **Add** (เพิ่มค่า) หรือ Edit (แก้ไข) หรือ Remove (ลบทิ้ง)

o กำหนดคุณลักษณะของตัวแปรดังนี้ (หากต้องการแปลงตัวแปรให้เป็นตัวแปรแบบตัวอักษรหรือตัวแปรแบบตัวเลข)

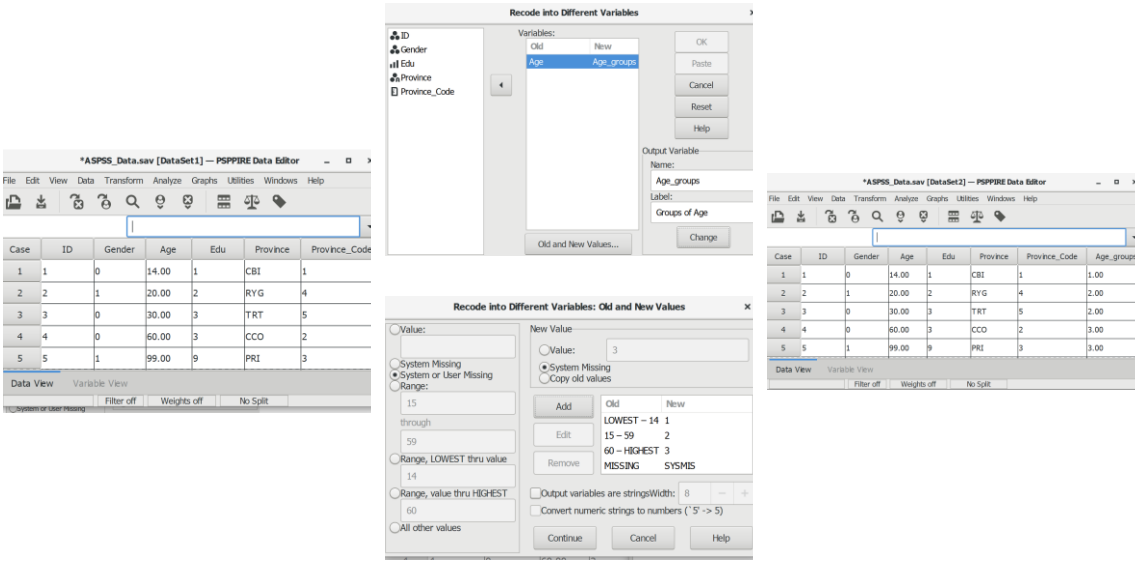
- Output variables are stringsWidth: (กำหนดเป็นตัวแปรแบบตัวอักษร)

- Covert numeric strings to numbers ('5' -> 5) (เปลี่ยนตัวเลขแบบตัวอักษรเป็นค่าแบบจำนวน)

o กดปุ่ม **Continue**

◎ กดปุ่ม **OK**

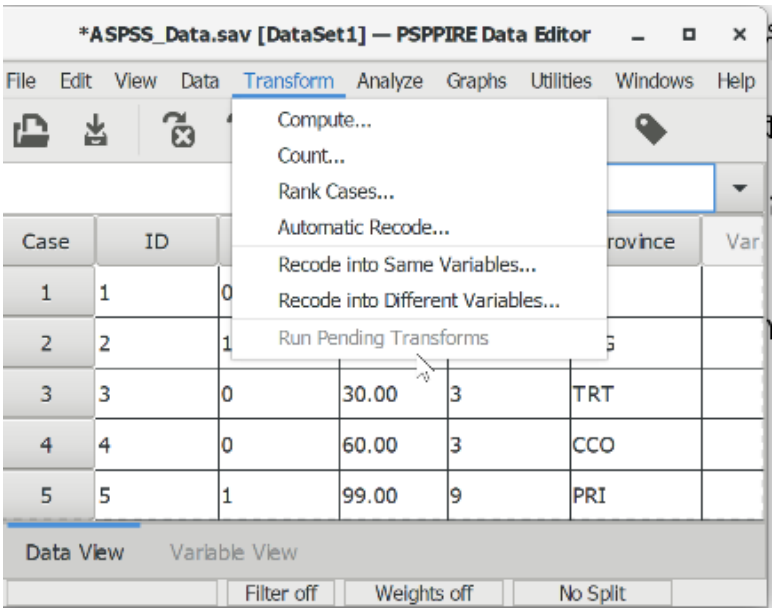
ตัวอย่าง: การจัดกลุ่มตัวอย่างจากตัวแปร Age ที่เป็นค่าต่อเนื่องให้เป็นตัวแปรมีค่าเป็นกลุ่มอายุ (1 = ต่ำกว่า 15 ปี 2 = 15 ถึง 59 ปี และ 3 = 60 ขึ้นไป)



2.7 การสั่งประมวลผลต่อ

การสั่งประมวลผลต่อ (Run Pending Transforms) คือ การสั่งให้โปรแกรมทำงานต่อระหว่างโปรแกรมหยุดการประมวลผลเพื่อเตือนให้ตรวจสอบและแก้ไข (ในแถบสถานการณ์ทำงานด้านล่างจะแสดงข้อความว่า Transformation Pending) เมื่อตรวจสอบหรือแก้ไขแล้วจึงสั่งให้ประมวลผลต่อ โดยมีวิธีการดังนี้

- เลือกเมนู **Transform -> Run Pending Transforms** (เมื่อโปรแกรมเตือนให้ตรวจสอบหรือแก้ไข ข้อความ Run Pending Transforms จะเปลี่ยนเป็นสีเข้ม)



สรุป

การจัดการข้อมูล คือ การปรับข้อมูลหรือค่าของตัวแปรในแฟ้มข้อมูลเพื่อเลือกข้อมูลบางส่วนมาวิเคราะห์ หรือเปลี่ยนข้อมูลให้เป็นตัวแปรใหม่หรือมีค่าของตัวแปรเป็นค่าใหม่

คำสั่งที่ใช้ในการจัดการข้อมูลมีทั้งคำสั่งแบบชั่วคราวที่มีผลต่อคำสั่งต่อไปที่ใช้เท่านั้น และคำสั่งแบบถาวรที่มีผลต่อการเปลี่ยนแปลงของชุดข้อมูลตลอดไป คำสั่งที่ใช้ในการจัดการข้อมูลแบ่งได้เป็น 2 กลุ่ม ดังนี้

1. คำสั่งสำหรับการปรับข้อมูล เป็นคำสั่งที่ใช้ในการเปลี่ยนแปลงรูปแบบของข้อมูลทั้งหมดในชุดข้อมูล เช่น การจัดเรียงข้อมูล การไขว้ตัวแปร การรวมข้อมูล การแบ่งข้อมูล การเลือกข้อมูล และการกำหนดน้ำหนักข้อมูล

2. คำสั่งสำหรับการแปลงข้อมูล เป็นคำสั่งที่ใช้ในการเปลี่ยนแปลงข้อมูลเฉพาะตัวแปรที่เลือก เช่น การเปลี่ยนค่าของตัวแปร การสร้างตัวแปรใหม่

แบบฝึกหัดท้ายบท

ให้ผู้เรียนฝึกจัดการข้อมูลและแปลงข้อมูลดังต่อไปนี้

1. เปิดแฟ้มข้อมูล ASPSS_Data.sav และจัดการข้อมูลด้วยคำสั่งดังต่อไปนี้

- Transpose -> ไขว้ข้อมูลในแฟ้มข้อมูลที่บันทึกเก็บไว้ให้เหมือนกับข้อมูลใน ASPSS_Data.sav
- Aggregate -> การรวมค่าของตัวแปร Age เป็นค่าเฉลี่ยในแต่ละกลุ่มของตัวแปร Edu
- Split File -> วิเคราะห์ค่าเฉลี่ยของตัวแปร Age ด้วยการแบ่งกลุ่มด้วยตัวแปร Edu
- Select Cases -> เลือกตัวอย่างด้วยวิธีสุ่มขนาดตัวอย่างเป็นจำนวนร้อยละ 75
- Count -> นับค่า 0 ในตัวแปร Gender
- Recode into Same Variables -> เปลี่ยนค่าของตัวแปร Gender ให้เป็นตามเดิม
- Recode into Difference Variables -> สร้างตัวแปร Edu_level มี 2 กลุ่ม คือ ระดับต่ำกว่าอุดมศึกษา และระดับอุดมศึกษา

2. เปิดแฟ้มข้อมูล hotel.sav และจัดการข้อมูลด้วยคำสั่งดังต่อไปนี้

- Compute -> สร้างตัวแปรใหม่ชื่อ Sat_sum รวมคะแนนความพึงพอใจจากตัวแปร V1 ถึง V5

ส่วนที่ 3

การวิเคราะห์เครื่องมือการวิจัย

และตรวจสอบข้อมูล

บทที่ 8
การวิเคราะห์เครื่องมือการวิจัย

ความมุ่งหมายของบทเรียน

ผู้เรียนสามารถวิเคราะห์คุณภาพของเครื่องมือการวิจัยได้

เนื้อหา

1. การวิเคราะห์ความถูกต้อง
2. การวิเคราะห์ความน่าเชื่อถือ

วิธีการสอนและกิจกรรมการเรียนการสอน

บรรยาย อธิบาย แสดงตัวอย่าง และทำแบบฝึกหัดท้ายบท

สื่อและอุปกรณ์การสอน

เอกสารคำสอน ภาพเลื่อน คอมพิวเตอร์ โปรแกรม PSPP โปรแกรมสอนออนไลน์ (สอนแบบออนไลน์) และเครื่องฉาย (สอนในชั้นเรียน)

งานที่มอบหมาย

ทำแบบฝึกหัดท้ายบท

การวัดผลและประเมินผล

การเข้าชั้นเรียน สังเกตความสนใจจากการเรียน สอบถามความเข้าใจจากการบรรยาย และประเมินจากการทำแบบฝึกหัดท้ายบท

เนื้อหา

หัวใจที่สำคัญประการหนึ่งของการวัด (measurement) ในการวิจัยเชิงปริมาณด้านสังคมศาสตร์ คือ การใช้เครื่องมือวัดกำหนดตัวเลข (numbers) ให้กับปรากฏการณ์ทางสังคมหรือพฤติกรรมของมนุษย์ เพื่อให้เป็นปริมาณที่สามารถนับได้ (Drost, 2011, Kimberlin & Winterstein, 2008) และเป็นเรื่องที่ยอมรับได้ในมุมมองของนักปฏิฐานนิยม (positivist) หรือวิธีการวิเคราะห์แบบประจักษ์นิยม (empirical analytic approach) (Drost, 2011)

แบบสอบถามเป็นเครื่องมืออันหนึ่งที่ได้รับความนิยมอย่างแพร่หลายในการสำรวจและรวบรวมข้อมูลในการวิจัยเชิงปริมาณด้านสังคมศาสตร์ แม้ว่าเครื่องมือที่ใช้ในการวิจัยด้านสังคมศาสตร์จะไม่สามารถวัดได้ค่าความถูกต้อง (accurate) 100 เปอร์เซ็นต์ แต่สิ่งที่สำคัญ คือ ทำอย่างไรให้ผลการวิจัยได้รับการยอมรับว่ามีความถูกต้อง (Kumar, 2014: 212-219)

ในกระบวนการพัฒนาและการตรวจสอบเครื่องมือต้องทำให้เครื่องมือมีคุณภาพ (quality) คือ มีทั้งความเที่ยงตรงและความน่าเชื่อถือ เพื่อให้ได้ข้อมูลที่ตรงกับความเป็นจริงและลดความคลาดเคลื่อนจากการวัด (Kimberlin & Winterstein, 2008; Taherdoost, 2016)

1. การวิเคราะห์ความเที่ยงตรง

ความเที่ยงตรง (Validity) คือ ความสัมพันธ์ระหว่างคำถามการวิจัยหรือค่าที่ได้จากการสังเกต (observed value) กับความจริง (objectives) หรือค่าที่เป็นจริง (true value) ที่ทำการศึกษา หรือวัดแล้วได้สิ่งที่ต้องการจะวัด (measure what is intended to be measured) หรือวัดสิ่งใดก็วัดได้สิ่งนั้นจริง

การสร้างเครื่องมือวิจัยในการวิจัยเชิงปริมาณมีความสำคัญเป็นอย่างมาก เพราะความจริง (objective) คำถามการวิจัย (research question) และสมมติฐาน (hypothesis) มีความเชื่อมโยงกัน ดังนั้นเครื่องมือที่ใช้วัดต้องมีความเที่ยงตรงหรือความถูกต้องดังต่อไปนี้

1. ความเที่ยงตรงด้านข้อเท็จจริง (Face validity) หมายถึง ความถูกต้องที่ตรงกับความเข้าใจทั่วไปของแนวคิด (common understandings of the concept)

2. ความเที่ยงตรงในเนื้อหา (Content validity) หมายถึง ความถูกต้องที่ข้อความทั้งหมดสามารถครอบคลุมขอบข่ายของความหมายของตัวแปรที่เราต้องการจะวัด

3. ความเที่ยงตรงที่สัมพันธ์กับขอบข่าย (Criteria-related validity) หมายถึง ความถูกต้องในการทำนาย และในการพิสูจน์ตามแนวคิด

4. ความเที่ยงตรงในแนวความคิด (Construct validity) หมายถึง ความถูกต้องเชิงทฤษฎีในการใช้เหตุผลแบบนิรนัย (deductive) และอุปนัย (inductive)

1.1 การวิเคราะห์ความสอดคล้องระหว่างข้อคำถาม-วัตถุประสงค์ของการวิจัย

วิธีการวิเคราะห์ความเที่ยงตรงเครื่องมือวิจัยเชิงปริมาณในการวิจัยด้านสังคมศาสตร์วิธีการหนึ่งที่ได้รับความนิยมอย่างแพร่หลายโดยมีจุดกำเนิดมาจากสาขาวิชาด้านการศึกษาที่พัฒนาโดยริชาร์ด เจ. โรวิเนลลี (Richard J. Rovinelli) และโรแนล เค. แฮเบิลตัน (Ronald K. Hambleton) คือ ดัชนีความสอดคล้องระหว่างข้อคำถาม-วัตถุประสงค์ของการวิจัย (Index of Item-Objective Congruence: IOC)

โรวิเนลลีและแฮเบิลตัน เชื่อว่าวิธีการตัดสินโดยผู้เชี่ยวชาญมีระดับความเชื่อมั่นสูง (high degree of confidence) เพราะเป็นการทดสอบความตรงของดัชนีตามวัตถุประสงค์ของการวัดและการออกแบบเครื่องมือของนักวิจัย และเป็นการวัดแบบอ้างอิงตามเกณฑ์ (criterion-referenced test) โดยให้ผู้เชี่ยวชาญในเนื้อหาของแต่ละเรื่องเป็นผู้ตัดสิน (judgements of content specialists) ประมาณ 3-5 คน และประเมินตามวัตถุประสงค์ในการสร้างเครื่องมือ (instructional objectives) อีกทั้งไม่ต้องใช้วิธีการทางสถิติที่ซับซ้อน (Rovinelli and Hambletion, 1967)

วิธีการตัดสินจากผู้เชี่ยวชาญ คือ ให้คะแนนข้อคำถามที่สอดคล้องกับวัตถุประสงค์ = 1 ไม่สอดคล้อง = -1 และไม่สามารถตัดสินใจได้ = 0 นำคะแนนที่สอดคล้องกับวัตถุประสงค์ (1) มาคำนวณแบ่งข้อคำถามที่ ดี (good) และไม่ดี (bad) ด้วยคะแนนครึ่งหนึ่งหรือร้อยละ 50 (0.5) โดยมีสูตรดังนี้

$$IOC = \frac{\sum R}{N}$$

IOC แทน ดัชนีความสอดคล้อง

$\sum R$ แทน ผลรวมคะแนนผู้เชี่ยวชาญที่มีความคิดเห็นว่าวัดได้ตรง

N แทน จำนวนผู้เชี่ยวชาญ

ตัวอย่างแบบประเมินความคิดเห็นของผู้เชี่ยวชาญเกี่ยวกับความเที่ยงตรงของแบบสอบถาม

ข้อ	ข้อคำถาม	ความเที่ยงตรง			ข้อเสนอแนะ
		1	0	-1	
1	เว็บไซต์นี้ของหน่วยงานนี้ให้บริการได้ตรงกับความต้องการ				
2	เว็บไซต์นี้มีความรวดเร็วในการแก้ไขปัญหาที่เกิดขึ้น				
3	เว็บไซต์นี้ทำงานได้รวดเร็ว				
4	เว็บไซต์นี้ใช้งานได้ง่าย				
5	เว็บไซต์นี้ออกแบบได้สวยงาม				
6	เว็บไซต์นี้มีการบูรณาการข้อมูลกับหน่วยงานที่เกี่ยวข้องได้ดี				
7	เว็บไซต์นี้สามารถใช้งานได้แบบที่เดียวเบ็ดเสร็จ (One-stop-services)				
8	เว็บไซต์นี้ให้ข้อมูลที่เชื่อถือได้				
9	เว็บไซต์นี้ให้บริการข้อมูลที่เป็นปัจจุบัน (Up-to-date)				
10	เว็บไซต์นี้เชื่อมั่นได้ว่ามีการรักษาข้อมูลส่วนตัวของผู้ใช้บริการได้ดี				

เกณฑ์การให้คะแนน

+1 ข้อคำถามตรงกับวัตถุประสงค์การวิจัย

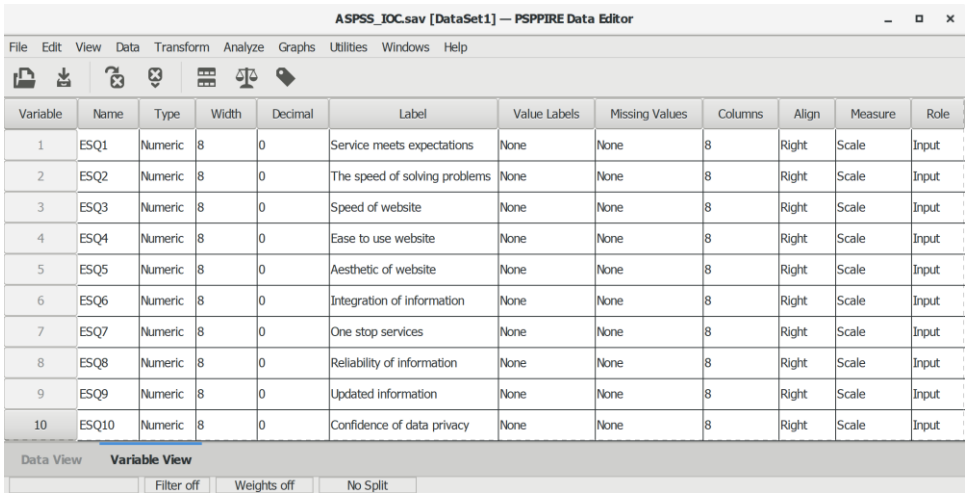
0 ข้อคำถามไม่แน่ใจว่าตรงวัตถุประสงค์การวิจัยหรือไม่

- 1 ข้อคำถามไม่ตรงกับวัตถุประสงค์การวิจัย

การวิเคราะห์ IOC อาจไม่จำเป็นต้องใช้โปรแกรมวิเคราะห์ข้อมูลด้านสถิติ เพราะยุ่งยากกว่าการวิเคราะห์ด้วยเครื่องคิดเลขหรือโปรแกรมตารางการคำนวณ (spreadsheet) แต่ในที่นี้จะแสดงเป็นตัวอย่างไว้หากต้องการใช้งาน

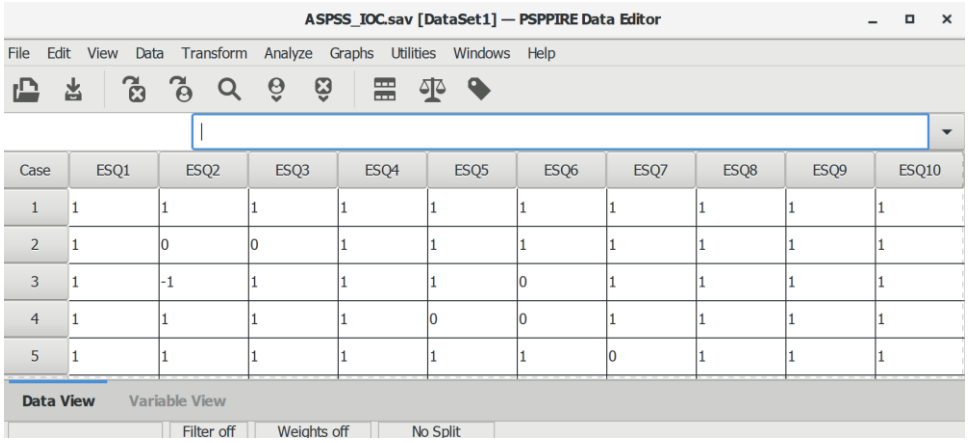
การใช้โปรแกรมวิเคราะห์ข้อมูล

- เปิดโปรแกรมและสร้างแฟ้มข้อมูล ASPSS_IOC.sav
- สร้างตัวแปร 10 ตัวแปร คือ ESQ1 ถึง ESQ10 โดยมีรายละเอียดดังนี้
  - ใส่หรือไม่ต้องใส่ Label, Value Labels แต่ต้องกำหนด Measure = Scale



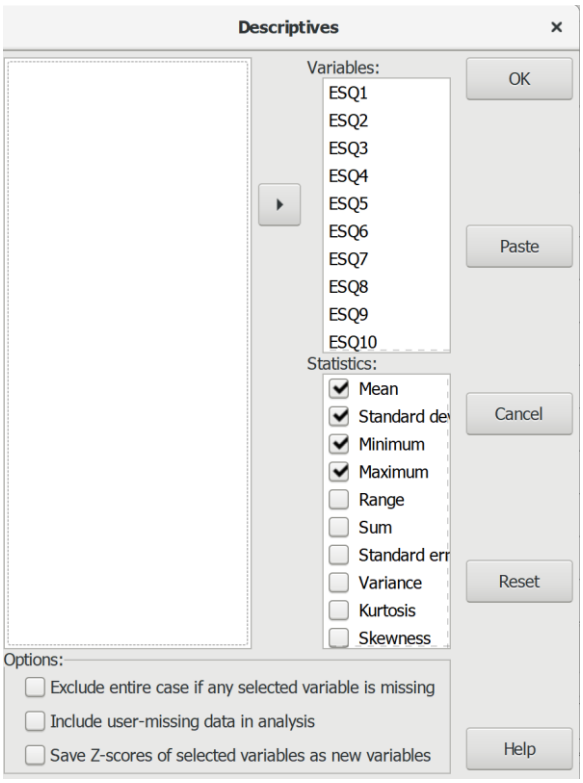
Variable	Name	Type	Width	Decimal	Label	Value Labels	Missing Values	Columns	Align	Measure	Role
1	ESQ1	Numeric	8	0	Service meets expectations	None	None	8	Right	Scale	Input
2	ESQ2	Numeric	8	0	The speed of solving problems	None	None	8	Right	Scale	Input
3	ESQ3	Numeric	8	0	Speed of website	None	None	8	Right	Scale	Input
4	ESQ4	Numeric	8	0	Ease to use website	None	None	8	Right	Scale	Input
5	ESQ5	Numeric	8	0	Aesthetic of website	None	None	8	Right	Scale	Input
6	ESQ6	Numeric	8	0	Integration of information	None	None	8	Right	Scale	Input
7	ESQ7	Numeric	8	0	One stop services	None	None	8	Right	Scale	Input
8	ESQ8	Numeric	8	0	Reliability of information	None	None	8	Right	Scale	Input
9	ESQ9	Numeric	8	0	Updated information	None	None	8	Right	Scale	Input
10	ESQ10	Numeric	8	0	Confidence of data privacy	None	None	8	Right	Scale	Input

- ใส่ค่าตัวแปรหรือข้อมูลโดยมีรายละเอียดดังนี้



Case	ESQ1	ESQ2	ESQ3	ESQ4	ESQ5	ESQ6	ESQ7	ESQ8	ESQ9	ESQ10
1	1	1	1	1	1	1	1	1	1	1
2	1	0	0	1	1	1	1	1	1	1
3	1	-1	1	1	1	0	1	1	1	1
4	1	1	1	1	0	0	1	1	1	1
5	1	1	1	1	1	1	0	1	1	1

- เลือกเมนู **Analyze -> Descriptive Statistics -> Descriptives...**
- ในหน้าต่าง Descriptives
  - **Variables:** เลือกตัวแปร ESQ1 ถึง ESQ10 จากรายการตัวแปรด้านซ้าย
  - **Statistics:** ใช้ค่าที่โปรแกรมกำหนด (default)
  - กดปุ่ม **OK**



ผลการวิเคราะห์

Descriptive Statistics

	N	Mean	Std Dev	Minimum	Maximum
Service meets expectations	5	1.00	.00	1	1
The speed of solving problems	5	.40	.89	-1	1
Speed of website	5	.80	.45	0	1
Ease to use website	5	1.00	.00	1	1
Aesthetic of website	5	.80	.45	0	1
Integration of information	5	.60	.55	0	1
One stop services	5	.80	.45	0	1
Reliability of information	5	1.00	.00	1	1
Updated information	5	1.00	.00	1	1
Confidence of data privacy	5	1.00	.00	1	1
Valid N (listwise)	5				
Missing N (listwise)	0				

การแปลผลการวิเคราะห์ค่า IOC จากค่า Mean

จากการประเมินเครื่องมือวิจัยโดยให้ผู้เชี่ยวชาญตัดสินพบว่า มีข้อคำถามที่มีค่า IOC น้อยกว่า 0.5 จำนวน 1 ข้อ คือ ข้อ 1 (ESQ2) และมีข้อคำถามที่มีค่า IOC มากกว่า 0.5 จำนวน 9 ข้อ คือ ข้อ 1, 3-10 (ESQ1, ESQ3 ถึง ESQ10)

สรุป มีข้อคำถามที่มีความเที่ยงตรงและสามารถนำไปใช้เก็บรวบรวมข้อมูล 9 ข้อ และมีข้อคำถามที่ไม่มีความเที่ยงตรงต้องทำการปรับปรุงหรือตัดออกจากแบบสอบถามจำนวน 1 ข้อ

ข้อ	ค่า IOC	การตัดสินใจ
1	1.00	ใช้ได้
2	0.40	ปรับปรุงหรือตัดทิ้ง
3	0.80	ใช้ได้
4	1.00	ใช้ได้
5	0.80	ใช้ได้
6	0.60	ใช้ได้
7	0.80	ใช้ได้
8	1.00	ใช้ได้
9	1.00	ใช้ได้
10	1.00	ใช้ได้

1.2 การวิเคราะห์ความเห็นสอดคล้องกัน

ในการวิเคราะห์ความเที่ยงตรง อาจทำการประเมินแบบสอบถามด้วยวิธีการวิเคราะห์ความเที่ยงตรงภายนอก (face validity) เพื่อทดสอบความเป็นไปได้ (feasibility) ความเข้าใจจากการอ่าน (readability) การใช้สำนวนและรูปแบบของภาษาที่มีความกลมกลืนกัน (consistency of style and formatting) และความชัดเจนของภาษา (clarity of the language) จากการให้คะแนนโดยผู้ประเมิน (inter-rater) ด้วยมาตรวัดแบบสองค่า (dichotomous scale) คือ ใช่ (Yes) = เหมาะสม และไม่ (No) = ไม่เหมาะสม

จากนั้นวิเคราะห์หาความเที่ยงตรงระหว่างคะแนนของผู้ประเมิน (inter-ratter) ด้วยค่าดัชนีแคปปาของโคเฮน (Cohen's Kappa Index: CKI) และใช้ค่า Kappa ตั้งแต่ 0.60 ขึ้นไป เป็นเกณฑ์ในการตัดสินใจยอมรับได้ว่า ผู้ประเมินมีความเห็นสอดคล้องกัน (Taherdoost, 2016) และแปลระดับความสอดคล้องจากค่า Kappa ได้จากตารางต่อไปนี้ (McHugh, 2012)

ตารางที่ 8.1 การแปลค่าแคปปาของโคเฮน

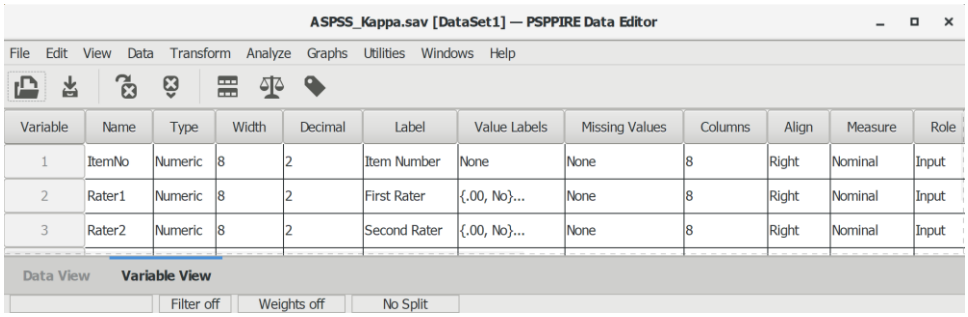
ค่าแคปปา (Value of Kappa)	**ระดับความสอดคล้อง (Level of Agreement)**	**จำนวนร้อยละของข้อมูลที่เชื่อถือได้ (% of Data that are Reliable)**
มากกว่า 0.90	มากที่สุด	82–100%
0.90-0.80	มาก	64–81%
0.79-0.60	ปานกลาง	35–63%
0.59-0.40	น้อย	15–35%
0.39-0.21	น้อยที่สุด	4–15%
0.20-0.00	ไม่สอดคล้อง	0–4%

ตัวอย่างแบบประเมินความเห็นเกี่ยวกับความเหมาะสมของแบบสอบถาม

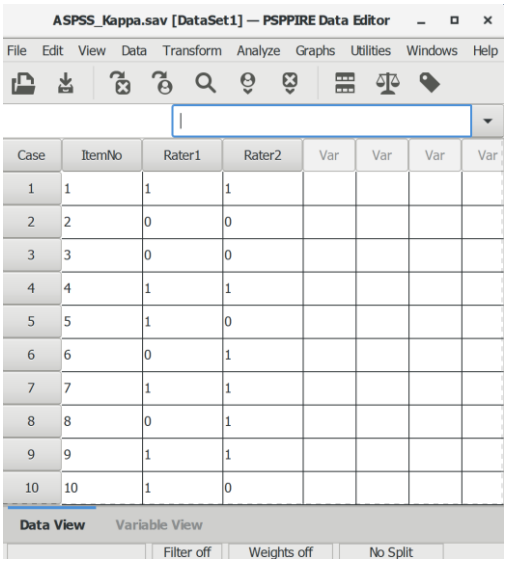
ข้อ	ข้อคำถาม	ความเหมาะสม		ข้อเสนอแนะ
		ใช่	ไม่	
1	เว็บไซต์นี้ของหน่วยงานนี้ให้บริการได้ตรงกับความต้องการ			
2	เว็บไซต์นี้มีความรวดเร็วในการแก้ไขปัญหาที่เกิดขึ้น			
3	เว็บไซต์นี้ทำงานได้รวดเร็ว			
4	เว็บไซต์นี้ใช้งานได้ง่าย			
5	เว็บไซต์นี้ออกแบบได้สวยงาม			
6	เว็บไซต์นี้มีการบูรณาการข้อมูลกับหน่วยงานที่เกี่ยวข้องได้ดี			
7	เว็บไซต์นี้สามารถใช้งานได้แบบที่เดียวเบ็ดเสร็จ (One-stop-services)			
8	เว็บไซต์นี้ให้ข้อมูลที่เชื่อถือได้			
9	เว็บไซต์นี้ให้บริการข้อมูลที่เป็นปัจจุบัน (Up-to-date)			
10	เว็บไซต์นี้เชื่อมั่นได้ว่ามีการรักษาข้อมูลส่วนตัวของผู้ใช้บริการได้ดี			

การใช้โปรแกรมวิเคราะห์ข้อมูล

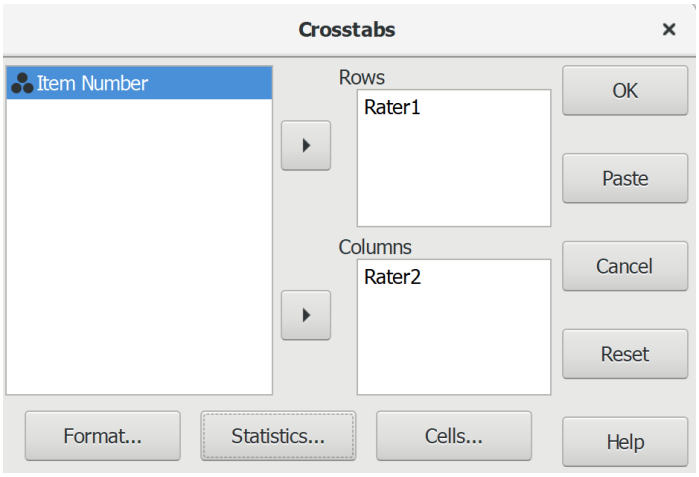
- เปิดโปรแกรมและสร้างแฟ้มข้อมูล ASPSS_Kappa.sav
- สร้างตัวแปร 3 ตัวแปร ItemNo, Rater1 และ Rater2 โดยมีรายละเอียดดังนี้
  - ตัวแปร ItemNo (Label = Item Number และ Measure = Nominal)
  - ตัวแปร Rater1 (Label = First Rater, Value Labels 0 = No และ 1 = Yes และ Measure = Nominal)
  - ตัวแปร Rater2 (Label = Second Rater, Value Labels 0 = No และ 1 = Yes และ Measure = Nominal)



- ใส่ค่าตัวแปรหรือข้อมูลโดยมีรายละเอียดดังนี้



- เลือกเมนู **Analyze -> Descriptive Statistics -> Crosstabs...**
- ในหน้าต่าง **Crosstabs**
  - **Rows:** เลือกตัวแปร Ratter1 -> **Columns:** และเลือกตัวแปร Ratter2



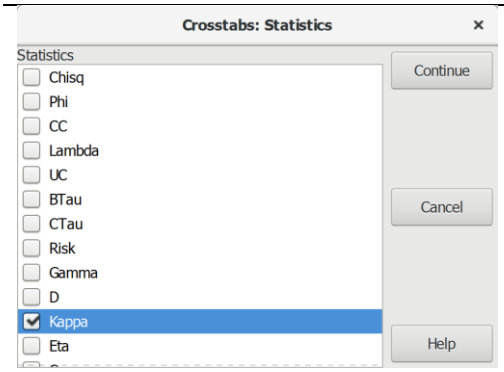
- กดปุ่ม **Statistics**

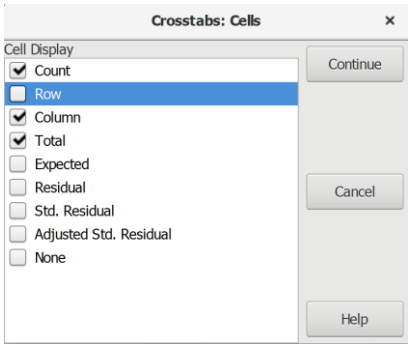
- Statistics เลือกเฉพาะ Kappa
- กดปุ่ม Continue

◎ กดปุ่ม Cells

- Cell Display เลือกเฉพาะ Count, Column และ Total
- กดปุ่ม Continue

◎ กดปุ่ม OK





ตารางผลการวิเคราะห์

Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
First Rater × Second Rater	10	100.0%	0	.0%	10	100.0%

First Rater × Second Rater

			Second Rater		
			No	Yes	Total
First Rater	No	Count	2	2	4
		Column %	50.0%	33.3%	40.0%
		Total %	20.0%	20.0%	40.0%
	Yes	Count	2	4	6
		Column %	50.0%	66.7%	60.0%
		Total %	20.0%	40.0%	60.0%
Total		Count	4	6	10
		Column %	100.0%	100.0%	100.0%
		Total %	40.0%	60.0%	100.0%

Symmetric Measures

		Value	Asymp. Std. Error	Approx. T
Measure of Agreement	Kappa	.17	.32	.53
N of Valid Cases		10		

การแปลผลการวิเคราะห์

ผลการวิเคราะห์ความเหมาะสมของแบบสอบถามจำนวน 10 ข้อ ผู้ประเมินให้คะแนนเหมาะสม (1) สอดคล้องกันน้อยกว่าไม่สอดคล้อง (Yes x Yes = 40.0%) และได้ค่า Kapp = 0.17 (Sig. =0.53) แสดงว่าแบบสอบถามฉบับนี้มีความไม่เหมาะสม

2. การวิเคราะห์ความน่าเชื่อถือ

ความน่าเชื่อถือ (Reliability) คือ ความสามารถในการวัดอย่างคงเส้นคงวา (consistent) หรือความแน่นอน (stability) ในการวัดแต่ละครั้ง เมื่อนำเครื่องมือไปใช้ภายใต้สถานการณ์ที่เหมือนกัน หรือประชากรที่เหมือนกัน ต้องได้ผลลัพธ์ที่เหมือนกัน (Drost, 2011)

การวิเคราะห์ความน่าเชื่อถือมีหลายวิธี แต่ละวิธีมีข้อดีและข้อเสียแตกต่างกัน เช่น ประหยัดเวลาและค่าใช้จ่ายในการเก็บข้อมูลครั้งเดียว หรือต้องใช้เวลาและเสียเวลามากในการเก็บข้อมูลหลายครั้ง แต่วิธีการทั้งหมดสามารถแบ่งได้เป็น 2 กลุ่ม ดังนี้

1. การทดสอบความน่าเชื่อถือภายใน (Internal reliability) เป็นการทดสอบความสอดคล้องภายในระหว่างข้อคำถามภายในแบบสอบถามแต่ละชุดที่เก็บข้อมูลเพียงครั้งเดียว เช่น การทดสอบแบบแบ่งครึ่ง (split-half) และการทดสอบความสอดคล้องภายในด้วยค่าสัมประสิทธิ์แอลฟา (alpha coefficient)

2. การทดสอบความน่าเชื่อถือภายนอก (External reliability) เป็นการทดสอบความสอดคล้องเชิงเปรียบเทียบระหว่างสองช่วงเวลา ระหว่างผู้ประเมิน และระหว่างชุดแบบสอบถาม เช่น การทดสอบ-ทดสอบซ้ำ (test-retest) การให้คะแนนระหว่างผู้ประเมิน (inter-rater) และการทดสอบหลากรูปแบบหรือรูปแบบเผื่อเลือก (multiple/alternative forms)

วิธีการวิเคราะห์ความน่าเชื่อถือที่นิยมใช้กันทั่วไป คือ การวัดความสอดคล้องภายในด้วยค่าสัมประสิทธิ์แอลฟาของครอนบัก (Cronbach Alpha Coefficient) ซึ่งเหมาะกับเครื่องมือที่ใช้ข้อคำถามแบบให้คะแนน (scored items) แบบสองค่า (dichotomous) เช่น มาตรวัดแบบให้คะแนน 0 และ 1 และแบบหลายค่า (polychotomous) เช่น มาตรวัดแบบลิเคิร์ต (Likert scale) (Ritter, 2010) ใช้วิธีการเก็บข้อมูลเพียงครั้งเดียว และข้อคำถามไม่จำเป็นต้องมีจำนวนเป็นเลขคู่เหมือนกับวิธีการทดสอบแบบแบ่งครึ่ง (split-half

แม้ว่าจะไม่มีกฎที่แน่นอน แต่มีหลักทั่วไป (rule of thumb) ที่ยอมรับกันอย่างแพร่หลาย คือ ค่าความสอดคล้องภายในต้องไม่ต่ำกว่า 0.70 (Taherdoost, 2016) และแปลระดับความน่าเชื่อถือจากค่า Cronbach ’s Alpha สำหรับคำถามแบบสองค่า (dichotomous questions) หรือคำถามแบบลิเคิร์ต (Likert scale questions) ได้จากตารางต่อไปนี้ (Glen, 2014)

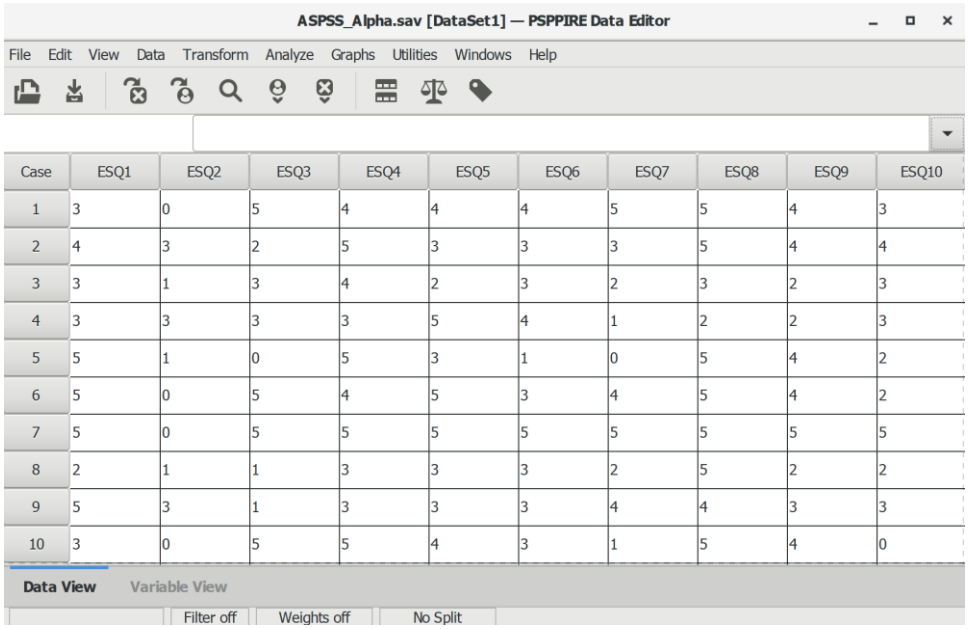
ตารางที่ 8.2 การแปลค่าสัมประสิทธิ์แอลฟาของครอนบัก

ค่าแอลฟาของครอนบัก (Value of Cronbach' s Alpha)	**ความสอดคล้องภายใน (Internal Consistency)**
α >= 0.9	ดีมาก (Excellent)
0.9 > α >= 0.9	ดี (Good)
0.8 > α >= 0.7	ยอมรับได้ (Acceptable)
0.7 > α >= 0.6	ไม่แน่นอน (Questionable)
0.6 > α >= 0.5	ต่ำ (Poor)
0.5 > 0.7 > α	ไม่ยอมรับ (Unacceptable)

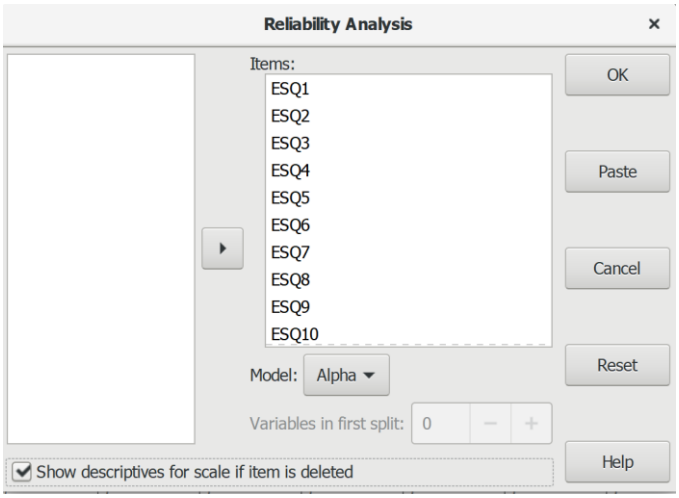
ตัวอย่างที่นำมาใช้ในการวิเคราะห์ความน่าเชื่อถือ คือ แฟ้มข้อมูลตัวอย่าง ASPSS_Alpha.sav โดยมีวิธีการดังต่อไปนี้

การใช้โปรแกรมวิเคราะห์ข้อมูล

◎ เปิดโปรแกรมและเปิดแฟ้มข้อมูล ASPSS_Alpha.sav



- เลือกเมนู **Analyze -> Reliability...**
- ในหน้าต่าง Reliability Analysis
  - **Items:** เลือกตัวแปร ESQ1 ถึง ESQ10 จากรายการตัวแปรด้านซ้ายมาใส่
  - **Model:** เลือก **Alpha**
  - เลือก **Show descriptives for scale if item is deleted** (แสดงค่าสถิติเพื่อปรับค่า Cronbach 's Alpha)
  - กดปุ่ม **OK**



ผลการวิเคราะห์

Case Processing Summary

Cases	N	Percent
Valid	10	100.0%
Excluded	0	.0%
Total	10	100.0%

Reliability Statistics

Cronbach's Alpha	N of Items
.67	10

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Service meets expectations	28.40	36.27	.37	.64
The speed of solving problems	31.00	48.67	-.42	.77
Speed of website	29.20	30.40	.39	.64
Ease to use website	28.10	38.99	.26	.66
Aesthetic of website	28.50	35.39	.48	.62
Integration of information	29.00	34.89	.55	.61
One stop services	29.50	27.17	.67	.55
Reliability of information	27.80	38.84	.20	.67
Updated information	28.80	33.07	.68	.59
Confidence of data privacy	29.50	34.28	.42	.63

การแปลผลการวิเคราะห์

ผลการวิเคราะห์ความน่าเชื่อถือข้อคำถามจำนวน 10 ข้อ จากกลุ่มตัวอย่าง 10 คน ได้ค่า Cronbach's Alpha = 0.67 โดยค่าสถิติ Cronbach Alpha if Item Deleted ในตาราง Item-Total Statistics แนะนำให้ลบคำถามข้อ ESQ2 (The speed of solving problems) ออกจากการวิเคราะห์ จะได้ค่า Alpha = 0.77

ทำการวิเคราะห์ความน่าเชื่อถือใหม่ด้วยวิธีการเดิมโดยเอาตัวแปร ESQ2 ออกจากรายการ Items: เหลือข้อคำถามที่ใช้ในการวิเคราะห์จำนวน 9 ข้อ

ผลการวิเคราะห์หลังลบตัวแปร ESQ2

Case Processing Summary

Cases	N	Percent
Valid	10	100.0%
Excluded	0	.0%
Total	10	100.0%

Reliability Statistics

Cronbach's Alpha	N of Items
.77	9

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Service meets expectations	27.20	42.40	.34	.77
Speed of website	28.00	33.33	.51	.75
Ease to use website	26.90	44.10	.33	.77
Aesthetic of website	27.30	40.68	.51	.75
Integration of information	27.80	40.62	.53	.74
One stop services	28.30	32.23	.66	.71
Reliability of information	26.60	43.38	.29	.77
Updated information	27.60	37.82	.73	.72
Confidence of data privacy	28.30	41.34	.32	.77

ผลการวิเคราะห์ความน่าเชื่อถือจากการลบตัวแปร ESQ2 ออกจากการวิเคราะห์ ได้ค่า Cronbach's Alpha = 0.77 มากกว่าค่าความสอดคล้องภายใน 0.70 ตามที่กล่าวมา และเป็นเครื่องมือวัดมีความน่าเชื่อถือหรือคงเส้นคงวาในการวัดแต่ละครั้ง

สำหรับข้อคำถาม ESQ2 หากมีความสำคัญและต้องใช้เป็นดัชนีในการวัด ต้องทำการปรับปรุงและนำไปทดสอบกับกลุ่มตัวอย่างใหม่ แล้วจึงนำมาทำการวิเคราะห์หาความน่าเชื่อถือใหม่อีกครั้ง

สรุป

การวิเคราะห์เครื่องมือการวิจัย คือ การตรวจสอบความมีคุณภาพของเครื่องมือทั้งด้านความเที่ยงตรงและความน่าเชื่อถือเพื่อลดความคลาดเคลื่อนจากการวัด และได้ข้อมูลที่ตรงกับความเป็นจริงอย่างคงเส้นคงวา

วิธีการวิเคราะห์ความเที่ยงตรงที่ได้รับความนิยมอย่างแพร่หลาย คือ การทดสอบความตรงของดัชนีตามวัตถุประสงค์ของการวัดและการออกแบบเครื่องมือของนักวิจัยด้วยค่าความสอดคล้องระหว่างข้อคำถาม-วัตถุประสงค์ของการวิจัย (Index of Item-Objective Congruence: IOC) และดัชนีความสอดคล้องระหว่างคะแนนของผู้ประเมินด้วยค่าแคปปาของโคเฮน (Cohen's Kappa Index: CKI)

วิธีการวิเคราะห์ความน่าเชื่อถือที่นิยมใช้กันทั่วไป คือ การวัดความสอดคล้องภายในด้วยค่าสัมประสิทธิ์แอลฟาของครอนบัก (Cronbach Alpha Coefficient)

แบบฝึกหัดท้ายบท

ให้ผู้เรียนฝึกวิเคราะห์ความเที่ยงตรงและความน่าเชื่อถือดังต่อไปนี้

1. เปิดแฟ้มข้อมูล ASPSS_IOC.sav
   - เพิ่มค่าตัวแปร Item4-Item10 และใส่ค่าตัวแปรให้ครบ
   - วิเคราะห์ค่า IOC และตัดสินใจว่าข้อคำถามใดใช้ได้หรือใช้ไม่ได้
2. เปิดแฟ้มข้อมูล ASPSS_Kappa.sav
   - เพิ่มค่าตัวแปรใน ItemNo, First Rater และ Second Rater ตัวต่อจากข้อที่ 3 ให้ครบ 10 รายการ
   - วิเคราะห์ค่า Kappa และประเมินว่าแบบสอบถามมีความเหมาะสมหรือไม่
3. เปิดแฟ้มข้อมูล hotel.sav แล้วดำเนินการดังนี้
   - เพิ่มตัวแปร V6-V10 หรือข้อคำถามอีก 5 ข้อ ต่อจากตัวแปร V5 และใส่ค่าตัวแปรให้ครบ
   - วิเคราะห์ความน่าชื่อถือ หากค่า Cronbach’s Alpha ต่ำกว่า 0.70 ให้ตัดข้อคำถามตามที่โปรแกรมแนะนำเพื่อปรับค่า Cronbach’s Alpha ให้อยู่ในระดับที่ยอมรับได้

บทที่ 9

การวิเคราะห์ข้อมูลเชิงสำรวจ

ความมุ่งหมายของบทเรียน

ผู้เรียนสามารถตรวจสอบความถูกต้องและความสมบูรณ์ของข้อมูล รวมถึงตรวจสอบลักษณะของข้อมูลตามข้อตกลงเบื้องต้นของสถิติ

เนื้อหา

1. การตรวจสอบความถูกต้องและความสมบูรณ์ของข้อมูล
2. การตรวจสอบข้อมูลตามข้อตกลงเบื้องต้น

วิธีการสอนและกิจกรรมการเรียนการสอน

บรรยาย อธิบาย แสดงตัวอย่าง และทำแบบฝึกหัดท้ายบท

สื่อและอุปกรณ์การสอน

เอกสารคำสอน ภาพเลื่อน คอมพิวเตอร์ โปรแกรม PSPP โปรแกรมสอนออนไลน์ (สอนแบบออนไลน์) และเครื่องฉาย (สอนในชั้นเรียน)

งานที่มอบหมาย

ทำแบบฝึกหัดท้ายบท

การวัดผลและประเมินผล

การเข้าชั้นเรียน สังเกตความสนใจจากการเรียน สอบถามความเข้าใจจากการบรรยาย และประเมินจากการทำแบบฝึกหัดท้ายบท

เนื้อหา

ความถูกต้องและไม่ถูกต้องของข้อมูลเกิดขึ้นได้ตั้งแต่การออกแบบแฟ้มข้อมูลและการนำข้อมูลเข้าแฟ้มข้อมูล ก่อนการวิเคราะห์ข้อมูลควรทำความสะอาดข้อมูล (data cleansing) เพื่อตรวจสอบให้มั่นใจว่าข้อมูลในแฟ้มข้อมูลมีความถูกต้องและตรงกับแบบสอบถาม แม้ว่าข้อมูลในแฟ้มข้อมูลจะมีความถูกต้องและตรงกับข้อมูลที่ได้จากการสำรวจแล้ว ก่อนการใช้สถิติทดสอบต้องตรวจสอบข้อมูล (data screening) ให้มั่นใจว่าลักษณะของข้อมูลเป็นไปตามข้อตกลงเบื้องต้นของสถิติแต่ละประเภท

1. การตรวจสอบความถูกต้องและความสมบูรณ์ของข้อมูล

การตรวจสอบความถูกต้องและความสมบูรณ์ของข้อมูล คือ การตรวจสอบการออกแบบข้อมูล (data design) เพื่อทบทวนการสร้างตัวแปรให้ถูกต้องและตรงกับแบบสอบถาม การตรวจสอบการนำเข้าข้อมูล (data entry) เพื่อตรวจสอบความถูกต้องของข้อมูลระหว่างแบบสอบถามกับแฟ้มข้อมูล และการตรวจความสมบูรณ์ของข้อมูล เพื่อตรวจสอบข้อมูลมีค่าสูญหายมากน้อยเพียงใด

1.1 การตรวจสอบการออกแบบข้อมูล

การตรวจสอบการออกแบบข้อมูล คือ การทบทวนการออกแบบแฟ้มข้อมูลที่ออกแบบไว้ก่อนการนำเข้าข้อมูลว่ามีความถูกต้องตรงกับแบบสอบถามหรือไม่ โดยการเปิดแฟ้มข้อมูลและเข้าไปที่หน้าต่างแสดงตัวแปร (Variable View) แล้วทำการตรวจสอบดังนี้

1. การกำหนดชื่อตัวแปร ควรใช้คำหรือชื่อย่อที่สื่อความหมาย หากเป็นกลุ่มตัวแปรดัชนี ควรใช้ชื่อเดียวกันแล้วตามด้วยตัวเลข เพื่อความสะดวกในการใช้คำสั่งที่ต้องใช้ตัวแปรหลายตัวแปร

2. การกำหนดความหมายตัวแปร ควรสั้นและเข้าใจง่าย เพื่อให้การแสดงผลในผลการวิเคราะห์ (output) อยู่ภายในบรรทัดเดียวกัน

3. การกำหนดค่าตัวแปร ให้ความหมายตรงกับคำตอบในแบบสอบถามหรือไม่ โดยเฉพาะตัวแปรแบบสองค่า (dichotomous variable) ควรให้ค่าแบบ 0 และ 1 หรือ 1 และ 2

4. การกำหนดค่าสูญหาย มีการกำหนดค่าที่ไม่ต้องการนำมาใช้ในการคำนวณหรือไม่ เช่น กำหนดให้เลข 8 = ไม่ตอบ หรือ 9 = ไม่ทราบ เพราะหากไม่กำหนดเป็นค่าสูญหายจะทำให้ค่าสถิติเบี่ยงเบนไปจากค่าที่เป็นจริง

5. การกำหนดระดับการวัดตัวแปร กำหนดตรงกับลักษณะข้อมูลในแบบสอบถามและระดับการวัดหรือไม่

การตรวจสอบความถูกต้องในขั้นตอนนี้แม้ว่าจะเป็นการทำงานที่ซ้ำซ้อน แต่ก็ช่วยสร้างความมั่นใจของความถูกต้องของข้อมูล และอาจพบข้อผิดพลาดที่ควรทำการแก้ไขก็ได้

วิธีที่ 1 การตรวจสอบพร้อมกันทุกตัวแปร

การใช้โปรแกรมวิเคราะห์ข้อมูล

- เปิดแฟ้มข้อมูลที่ต้องการตรวจสอบ (ตัวอย่างแฟ้มข้อมูล ASPSS_Data.sav)
- เลือกเมนู **File -> Display Data File Information -> Working File**
- ดูผลลัพธ์ในหน้าต่าง **Output Viewer**

ตัวอย่าง: การตรวจสอบลักษณะตัวแปรทั้งแฟ้มข้อมูล

```
DISPLAY DICTIONARY.
```

Variables

Name	Position	Label	Measurement Level	Role	Width	Alignment	Print Format	Write Format	Missing Values
ID	1	Questionnaire Identification	Nominal	None	8	Right	F5.0	F5.0	
Gender	2	Gender of Responder	Nominal	Input	7	Right	F1.0	F1.0	0
Age	3	Age of Responder	Scale	Both	7	Right	F8.2	F8.2	0 THRU 14; 99
Edu	4	Education of Responder	Ordinal	Both	8	Right	F2.0	F2.0	9
Province	5	Residence of Responder	Nominal	Input	7	Right	A3	A3	

Value Labels

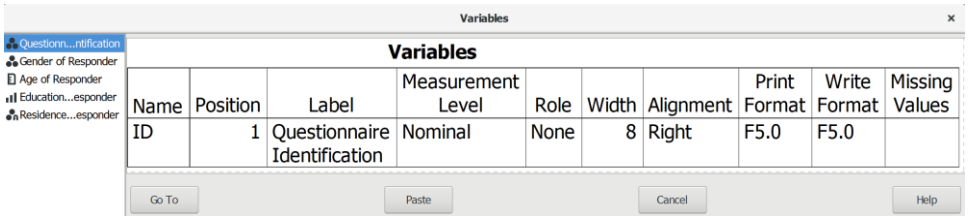
Variable Value		Label
Gender of Responder	1	Male
	2	Female
Education of Responder	1	Primary Education
	2	Secondary Education
	3	Higher Education

วิธีที่ 2 การตรวจสอบทีละตัวแปร

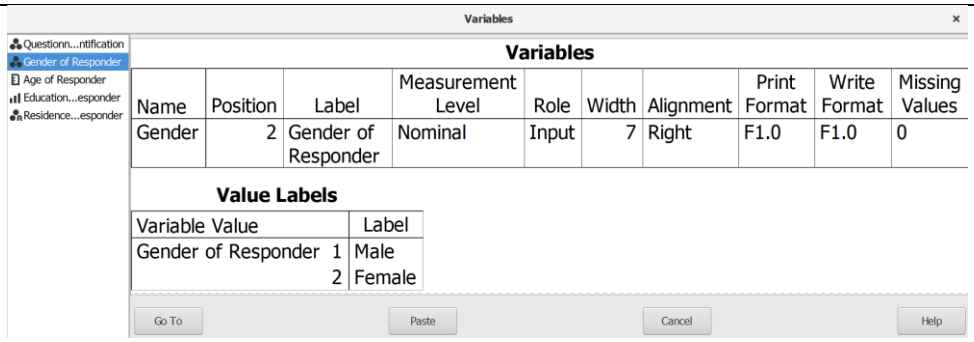
การใช้โปรแกรมวิเคราะห์ข้อมูล

- เปิดแฟ้มข้อมูลที่ต้องการตรวจสอบ
- เลือกเมนู **Utilities -> Variables...**
- คลิกเมาส์เลือกดุตัวแปรในเมนูด้านซ้ายตรวจสอบทีละตัวแปร

ตัวอย่าง: การตรวจสอบลักษณะตัวแปรทีละตัวแปร (ตัวแปร ID)



ตัวอย่าง: การตรวจสอบลักษณะตัวแปรทีละตัวแปร (ตัวแปร Gender)



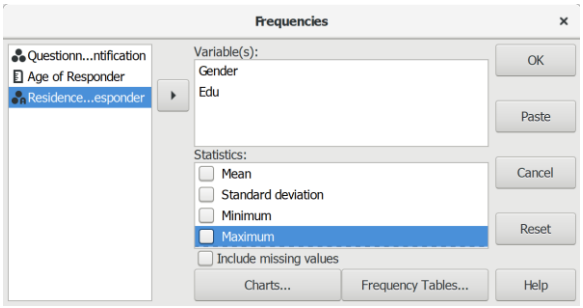
1.2 การตรวจสอบความถูกต้องของข้อมูล

การตรวจสอบความถูกต้องของข้อมูล คือ การตรวจสอบความผิดพลาดในการบันทึกข้อมูล (coding error และความผิดพลาดจากความไม่ตั้งใจในการตอบแบบสำรวจหรือบันทึกข้อมูลระหว่างการสัมภาษณ์ (logical error) วิธีการตรวจสอบความถูกต้องของข้อมูลที่กล่าวมามี 2 วิธี ดังนี้ (SPSS, 2005: 3)

1. การตรวจสอบแบบตัวแปรเดียว (Single-variable rule) เป็นการค้นหาความผิดพลาดในการบันทึกข้อมูล และเป็นการค้นหาค่าที่ไม่อยู่ในช่วงของค่าตัวแปร (out-of-range values) ในแฟ้มข้อมูล ซึ่งมีสาเหตุมาจากการพิมพ์ค่าตัวแปรผิดระหว่างการบันทึกข้อมูล หรือมีกลุ่มตัวอย่างที่มีคุณลักษณะของข้อมูลอยู่นอกช่วงของค่าตัวแปรที่กำหนดไว้ในการออกแบบสอบถาม

การใช้โปรแกรมวิเคราะห์ข้อมูล

- เลือกเมนู **Analyze -> Descriptive Statistics -> Frequencies...**
  - **Variable (s):** ใส่ตัวแปร Gender และ Edu
  - **Statistics: ยกเลิก**การเลือกค่า Mean, Standard deviation, Minimum และ Maximum



ตัวอย่าง: ผลลัพธ์การตรวจสอบข้อมูลตัวแปรค่าแบ่งกลุ่มแบบตัวแปรเดียว

Gender of Responder

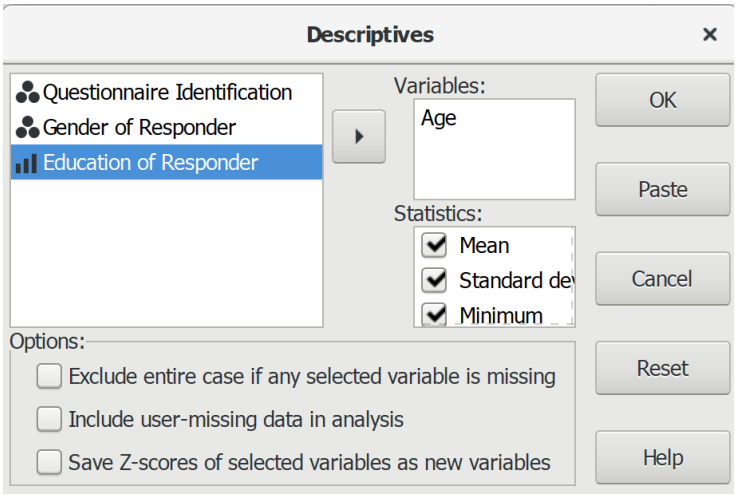
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	3	60.0%	60.0%	60.0%
	Female	2	40.0%	40.0%	100.0%
Total		5	100.0%		

Education of Responder

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Primary Education	1	20.0%	25.0%	25.0%
	Secondary Education	1	20.0%	25.0%	50.0%
	Higher Education	2	40.0%	50.0%	100.0%
Missing	9	1	20.0%		
Total		5	100.0%		

การตรวจสอบ: มีค่าของตัวแปรอยู่นอกเหนือจากค่า (value) ที่กำหนดไว้หรือไม่

- เลือกเมนู **Analyze -> Descriptive Statistics -> Descriptives...**
  - **Variable (s):** ใส่ตัวแปร Age
  - **Statistics: เลือก**ค่า Mean, Standard deviation, Minimum และ Maximum



ตัวอย่าง: ผลลัพธ์การตรวจสอบข้อมูลตัวแปรค่าต่อเนื่องแบบตัวแปรเดียว

Descriptive Statistics

	N	Mean	Std Dev	Minimum	Maximum
Age of Responder	3	36.67	20.82	20.00	60.00
Valid N (listwise)	5				
Missing N (listwise)	2				

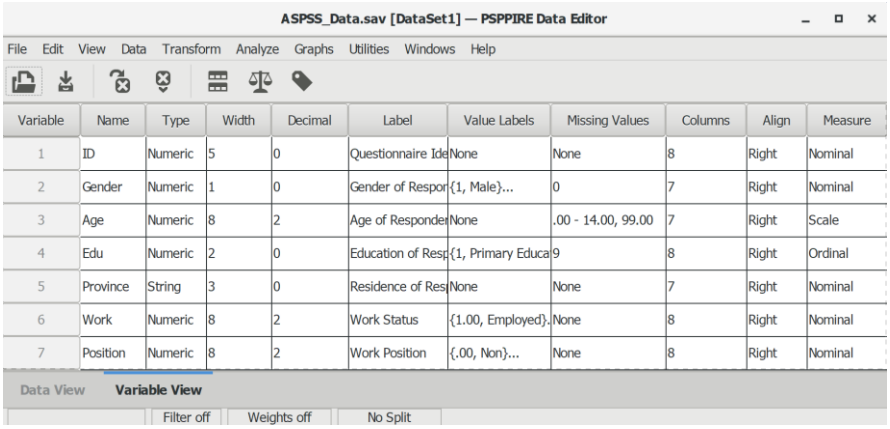
การตรวจสอบ: มีข้อมูลอยู่นอกขอบเขตหรือไม่

2. การตรวจสอบแบบไขว้ตัวแปร (Cross-variable rule) เป็นการค้นหาความผิดพลาดจากความไม่ตั้งใจในการตอบแบบสำรวจหรือผู้สัมภาษณ์บันทึกข้อมูลในแบบสอบถามผิด กล่าวคือ เป็นการค้นหาค่าตัวแปรที่ไม่สอดคล้องกับความเป็นจริงหรือความขัดแย้งของข้อมูลระหว่างตัวแปร เช่น ครัวเรือนที่มีสมาชิก 1 คน เป็นเพศชาย เป็นโสด แต่มีบุตรจำนวน 1 คนหรือมากกว่า เป็นต้น

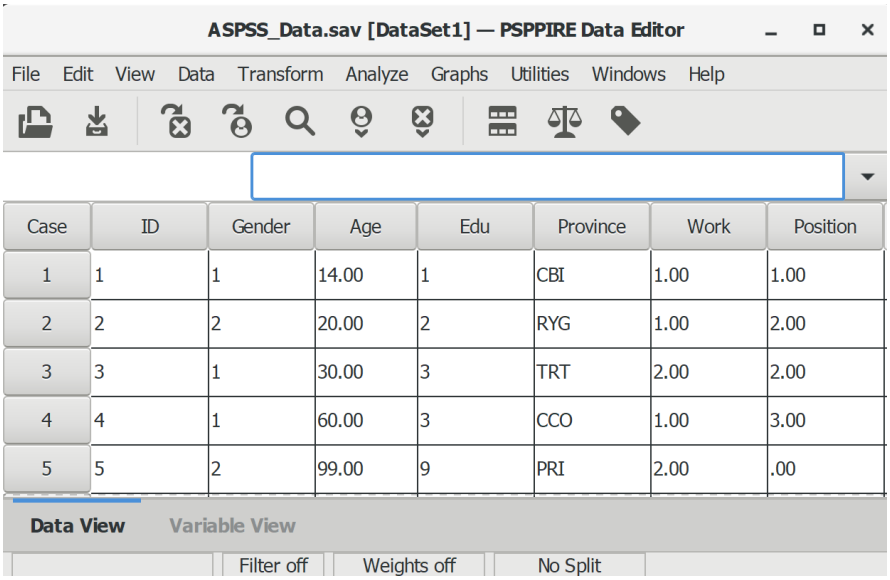
กรณีที่ 1 ตัวแปรที่ไม่น่าจะมีความสัมพันธ์กัน เช่น สถานภาพการทำงาน กับ ตำแหน่งงาน

การใช้โปรแกรมวิเคราะห์ข้อมูล

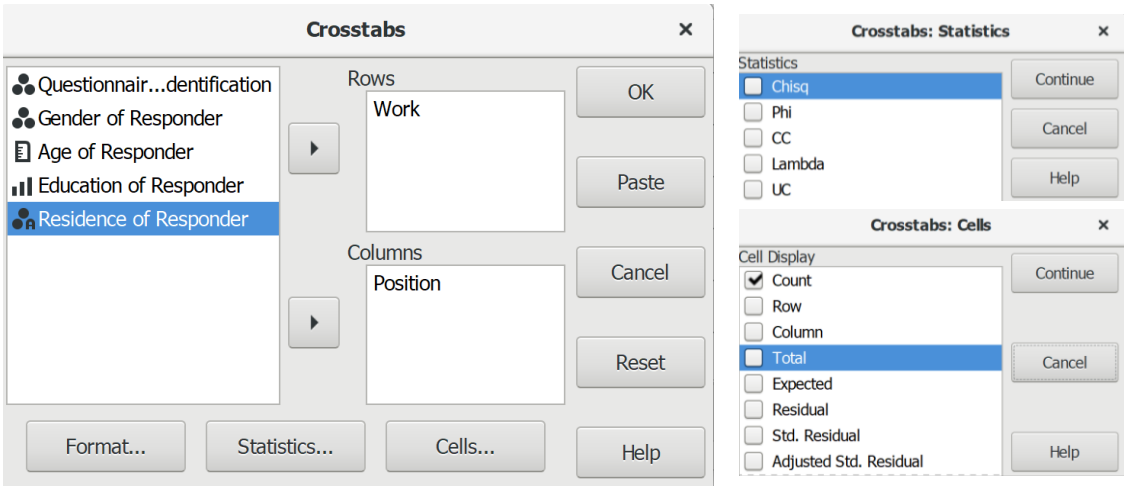
- เปิดแฟ้มข้อมูล **File -> Open** -> ASPSS_Data.sav
- **สร้างตัวแปร 2 ตัวแปร** คือ Work และ Position โดยมีรายละเอียดดังนี้
  - ตัวแปร **Work (Label = Work Status, Value 1 = Employed 2 = Unemployed และ Measure = Nominal)**
  - ตัวแปร **Position (Label = Work Position, Value Labels 0 = Non 1 = Manager 2 = Employee 3 = Other และ Measure = Nominal)**



- ใส่ค่าตัวแปรหรือข้อมูลโดยมีรายละเอียดดังนี้



- เลือกเมนู **Analyze -> Descriptive Statistics -> Crosstab...**
  - **Rows** ใส่ตัวแปร Work
  - **Columns** ใส่ตัวแปร Position
  - กดปุ่ม **Statistics...** ในกลุ่มคำสั่ง **Statistics ยกเลิก**การเลือก Chisq -> กดปุ่ม **Continues**
  - กดปุ่ม **Cells...** ในกลุ่มคำสั่ง **Cell Display** เลือกเฉพาะ **Count** -> กดปุ่ม **Continues**
  - กดปุ่ม **OK**



ตัวอย่าง: ผลลัพธ์การตรวจสอบข้อมูลแบบไขว้ระหว่างตัวแปรค่าแบ่งกลุ่ม X ตัวแปรค่าแบ่งกลุ่ม

Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Work Status × Work Position	5	100.0%	0	.0%	5	100.0%

Work Status × Work Position

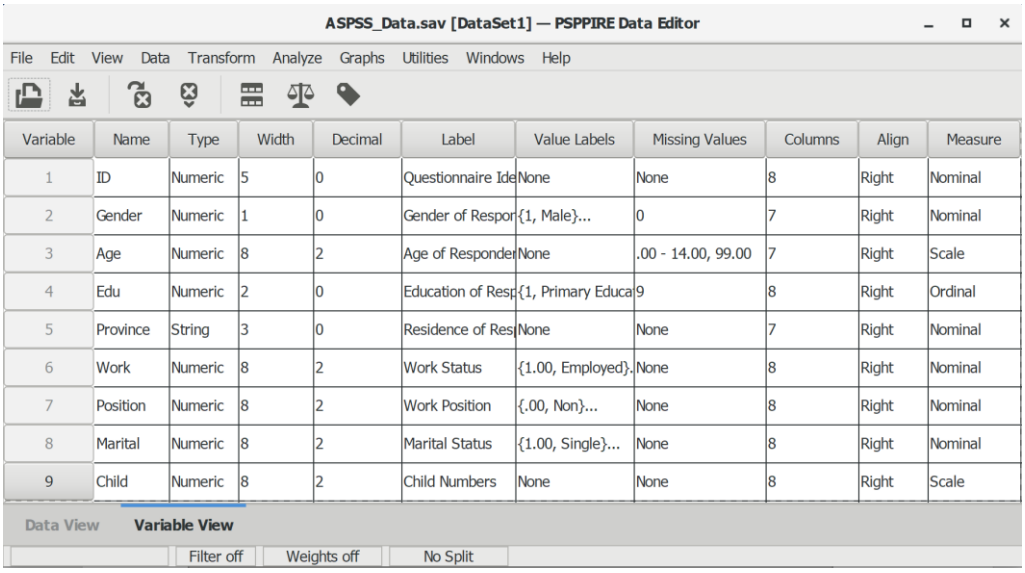
			Work Position				
			Non	Manager	Employee	Other	Total
Work Status	Employed	Count	0	1	1	1	3
	Unemployed		1	0	1	0	2
Total			1	1	2	1	5

การตรวจสอบ: ค่าตัวแปรที่ไม่สอดคล้องกัน เช่น Work Status = Unemployed แต่ Work Position ใน Employee = 1

กรณีที่ 2 ตัวแปรที่ไม่น่าจะมีความสัมพันธ์กัน เช่น สถานภาพสมรส กับ จำนวนบุตร

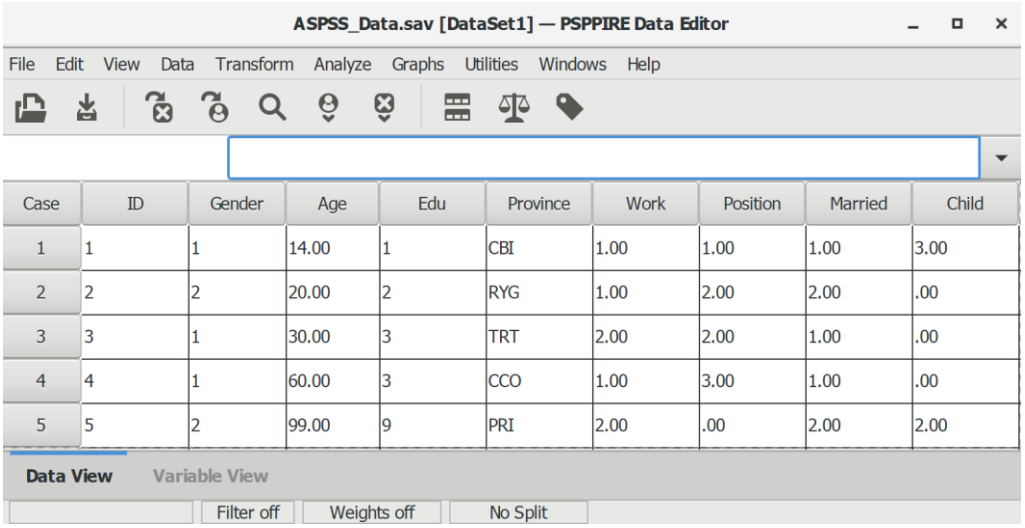
การใช้โปรแกรมวิเคราะห์ข้อมูล

- เปิดแฟ้มข้อมูล **File -> Open** -> ASPSS_Data.sav
- **สร้างตัวแปร 2 ตัวแปร** คือ Married และ Child โดยมีรายละเอียดดังนี้
  - ตัวแปร **Marital (Label = Marital Status, Value Labels 1 = Single 2 = Married 3 = Other และ Measure = Nominal)**
  - ตัวแปร **Child (Label = Child Numbers และ Measure = Scale)**



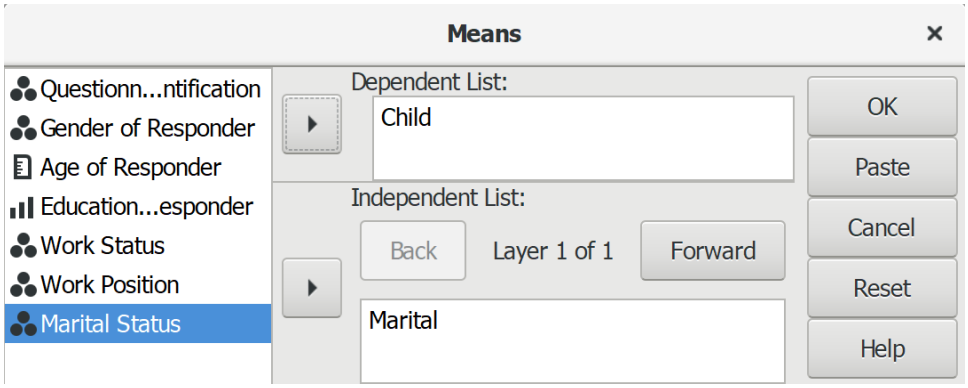
Variable	Name	Type	Width	Decimal	Label	Value Labels	Missing Values	Columns	Align	Measure
1	ID	Numeric	5	0	Questionnaire Ide	None	None	8	Right	Nominal
2	Gender	Numeric	1	0	Gender of Respor	{1, Male}...	0	7	Right	Nominal
3	Age	Numeric	8	2	Age of Responder	None	.00 - 14.00, 99.00	7	Right	Scale
4	Edu	Numeric	2	0	Education of Resp	{1, Primary Educa	9	8	Right	Ordinal
5	Province	String	3	0	Residence of Res	None	None	7	Right	Nominal
6	Work	Numeric	8	2	Work Status	{1.00, Employed}.	None	8	Right	Nominal
7	Position	Numeric	8	2	Work Position	{.00, Non}...	None	8	Right	Nominal
8	Marital	Numeric	8	2	Marital Status	{1.00, Single}...	None	8	Right	Nominal
9	Child	Numeric	8	2	Child Numbers	None	None	8	Right	Scale

- ใส่ค่าตัวแปรหรือข้อมูลโดยมีรายละเอียดดังนี้



Case	ID	Gender	Age	Edu	Province	Work	Position	Married	Child
1	1	1	14.00	1	CBI	1.00	1.00	1.00	3.00
2	2	2	20.00	2	RYG	1.00	2.00	2.00	.00
3	3	1	30.00	3	TRT	2.00	2.00	1.00	.00
4	4	1	60.00	3	CCO	1.00	3.00	1.00	.00
5	5	2	99.00	9	PRI	2.00	.00	2.00	2.00

- เลือกเมนู **Analyze -> Compare Mean -> Means...**
  เลือกตัวแปรที่คิดว่าไม่น่าจะมีความสัมพันธ์กัน เช่น สถานภาพสมรส กับ จำนวนบุตร
  - **Dependent List:** เลือกตัวแปร Child จากรายการตัวแปรด้านซ้าย
  - **Independent List:** เลือกตัวแปร Married จากรายการตัวแปรด้านซ้าย
  - กดปุ่ม **OK**



ตัวอย่าง: ผลลัพธ์การตรวจสอบข้อมูลแบบไขว้ระหว่างตัวแปค่าแบ่งกลุ่ม X ตัวแปรค่าต่อเนื่อง

Case Processing Summary

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
Child * Marital	5	100.0%	0	.0%	5	100.0%

Report

Marital Status	Mean	N	Std. Deviation
Single	1.00	3	1.73
Married	1.00	2	1.41
Total	1.00	5	1.41

การตรวจสอบ: ค่าตัวแปรที่ไม่สอดคล้องกัน เช่น Mairtal Status = Single แต่ค่าเฉลี่ย Child = 1

1.3 การตรวจสอบค่าสูญหาย

ค่าสูญหาย (Missing values) คือ ค่าตัวแปรที่ไม่ทราบค่าภายในแฟ้มข้อมูลมีมากน้อยเพียงใด และมีรูปแบบปกติหรือไม่ปกติ เช่น ค่าว่างเปล่า (null) ที่โปรแกรมกำหนดให้เป็นค่าสูญหาย (system missing) และค่าที่ผู้ใช้กำหนดให้เป็นค่าสูญหาย (user missing) เช่น ไม่ทราบ ไม่ตอบ ผลกระทบของค่าสูญหายมีผลทำให้ผลการวิเคราะห์ผิดพลาด (bias) และค่าสถิติไม่มีอำนาจในการทำนายที่ดี (statistical power)

การวิเคราะห์หาข้อมูลสูญหายภายในแฟ้มข้อมูลว่ามีลักษณะดังต่อไปนี้หรือไม่ (Hair and Others, 2006: 55-56)

1. จำนวนค่าสูญหาย มีจำนวนค่าสูญหายของแต่ละชุดข้อมูล (case) และแต่ละตัวแปร (variable) ว่ามีค่าสูญหายจำนวนมากน้อยเพียงใด ถ้ากำหนดไม่ให้นำเอาข้อมูลที่มีค่าสูญหายมาใช้ในการคำนวณ จำนวนตัวอย่างที่ถูกนำไปใช้ในการคำนวณก็จะมีขนาดลดลง หากมีข้อมูลสูญหายจำนวนมากอาจมีผลทำให้มีข้อมูลไม่เพียงพอในการวิเคราะห์แบบหลายตัวแปร (multivariate analysis) ตามมา

2. รูปแบบค่าสูญหาย มีรูปแบบการสูญหายอย่างมีระบบหรือสูญหายลักษณะคล้ายกัน (nonrandom) ที่เกิดจากผู้ตอบไม่ต้องการเปิดเผยข้อมูลของตน เช่น คนรวยไม่ต้องการบอกรายได้ของตน เป็นต้น รูปแบบการสูญหายอย่างมีระบบที่พบกันอย่างแพร่หลาย เช่น การสูญหายของข้อมูลเนื่องจากผู้ตอบแบบสอบถามจำนวนหนึ่งไม่ตอบคำถามข้อใดข้อหนึ่ง หรือส่วนใดส่วนหนึ่งในแบบสอบถาม หรือการสูญหายของข้อมูลเนื่องจากการไม่ตอบที่เพิ่มมากขึ้นในตอนท้ายของแบบสอบถาม เป็นต้น ถ้าค่าสูญหายมีอยู่ในตัวแปรตาม ค่าสูญหายนั้นจะมีผลทำให้ค่าของตัวแปรตามเปลี่ยนตามไปด้วย และอาจมีความสัมพันธ์กับค่าของตัวแปรที่สังเกตได้

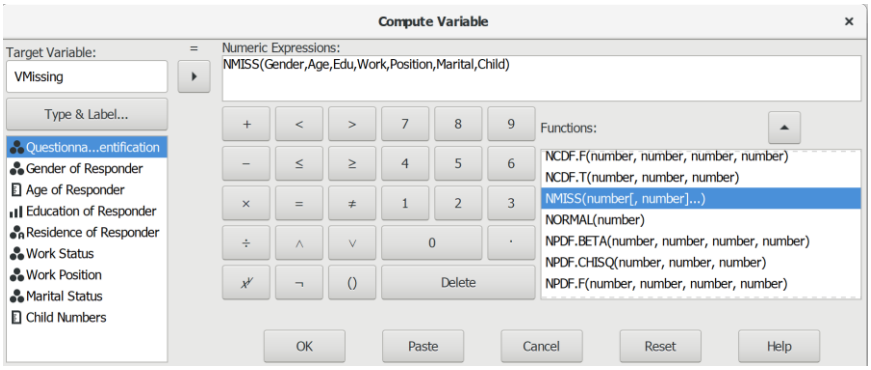
การไม่ให้ความสนใจ (ignore) และการแก้ไข (remedy) การสูญหายของข้อมูล มีผลกระทบต่อผลการวิเคราะห์ทั้งสองกรณี สิ่งสำคัญที่นักวิจัยควรทำก่อนการวิเคราะห์ข้อมูล คือ ทำความเข้าใจการสูญหายของข้อมูลว่า การสูญหายของข้อมูลมีสาเหตุมาจากอะไร และมีผลกระทบอย่างไร เพราะจะทำให้การตัดสินใจแก้ไขการสูญหายของข้อมูลมีความถูกต้องมากกว่า และป้องกันปัญหาความผิดพลาดของผลการวิเคราะห์ได้เป็นอย่างดี โดยทำการตรวจสอบจำนวนค่าสูญหายต่อตัวแปร (missing per variable) และจำนวนค่าสูญหายต่อชุดข้อมูล (missing per case) มีจำนวนมากน้อยเพียงใด และเกิดการสูญหายอย่าง

มีระบบหรือสูญหายลักษณะคล้ายกัน (nonrandom) หรือไม่ ตามที่กล่าวมาแล้ว แต่สำหรับในที่นี้จะกล่าวถึงเฉพาะการตรวจสอบจำนวนค่าสูญหายเท่านั้น โดยมีรายละเอียดดังนี้

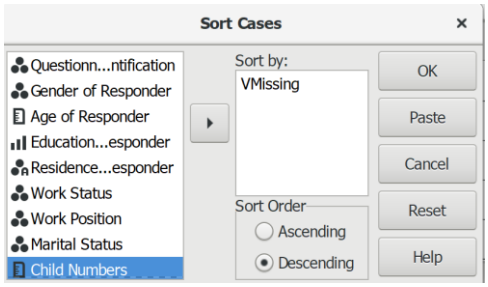
1. จำนวนค่าสูญหายของแต่ละตัวอย่าง (Missing values by case) ตัวอย่างที่มีค่าสูญหายต่ำกว่าร้อยละ 10 ไม่ต้องทำการปรับแก้ แต่หากมากกว่านั้นควรทำการแก้ไข และหากมีจำนวนมากร้อยละ 50 ควรตัดออกจากกระบวนการวิเคราะห์ข้อมูล

การใช้โปรแกรมวิเคราะห์ข้อมูล

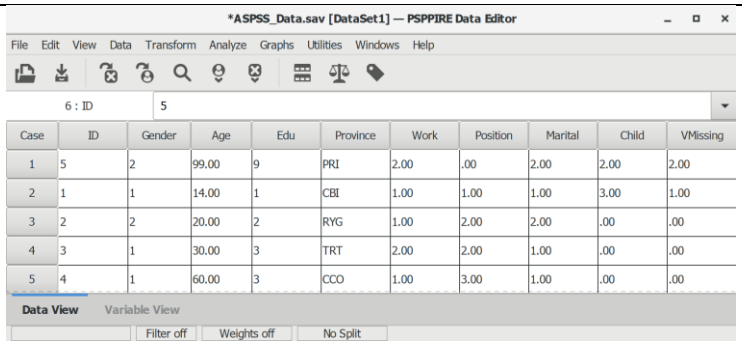
- เปิดแฟ้มข้อมูล **File -> Open** -> ASPSS_Data.sav
- เลือกเมนู **Transform -> Compute Variable**
  - **Target Variable:** พิมพ์ชื่อแปร VMissing
  - **Function** เลือก **NMIS** (นับเฉพาะค่า System Missing หรือค่าว่างเปล่า และ User Missing)
  - เลือกตัวแปรทั้งหมดที่ต้องการตรวจสอบใส่ในวงเล็บ (เฉพาะตัวแปรแบบตัวเลข)
  - กดปุ่ม **OK**



- เลือกเมนู **Data -> Sort Cases...**
  - **Sort by:** ใส่ตัวแปร VMissing
  - **Sort Order** เลือก Descending
  - กดปุ่ม **OK**



ตัวอย่าง: ผลลัพธ์การตรวจสอบจำนวนค่าสูญหายของแต่ละชุดข้อมูล



Case	ID	Gender	Age	Edu	Province	Work	Position	Marital	Child	VMissing
1	5	2	99.00	9	PRI	2.00	.00	2.00	2.00	2.00
2	1	1	14.00	1	CBI	1.00	1.00	1.00	3.00	1.00
3	2	2	20.00	2	RYG	1.00	2.00	2.00	.00	.00
4	3	1	30.00	3	TRT	2.00	2.00	1.00	.00	.00
5	4	1	60.00	3	CCO	1.00	3.00	1.00	.00	.00

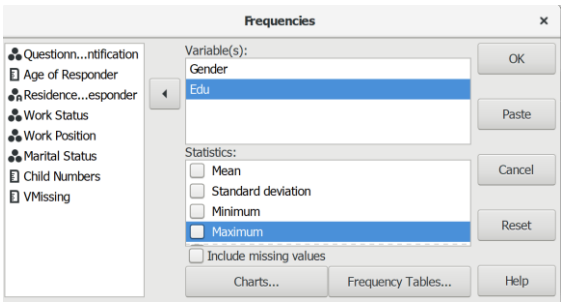
การตรวจสอบ: ดูจำนวนค่าสูญหายในตัวแปร VMissing ตัวอย่างที่ 5 มีค่าสูญหายจำนวน 2 ตัวแปร (ร้อยละ 25) และตัวอย่างที่ 1 มีค่าสูญจำนวน 1 ตัวแปร (ร้อยละ 12.5)

2. จำนวนค่าสูญหายของแต่ละตัวแปร (Missing values by variable) ตัวแปรที่มีค่าสูญหายเล็กน้อยประมาณร้อยละ 15 ไม่ต้องทำการปรับแก้ แต่หากมากกว่านั้นควรทำการแก้ไข และหากมีจำนวนมากร้อยละ 50 ควรตัดออกจากกระบวนการวิเคราะห์ข้อมูล

วิธีที่ 1 การตรวจสอบตัวแปรค่าแบ่งกลุ่ม

การใช้โปรแกรมวิเคราะห์ข้อมูล

- เปิดแฟ้มข้อมูล **File -> Open** -> ASPSS_Data.sav
- เลือกเมนู **Analyze -> Descriptive Statistics -> Frequencies**
  - **Variable (s):** ใส่ตัวแปร Gender และ Edu
  - ในกลุ่มคำสั่ง **Statistics: ยกเลิก**การเลือกค่า Mean, Standard deviation, Minimum และ Maximum
  - กดปุ่ม **OK**



ตัวอย่าง: ผลลัพธ์การตรวจสอบค่าสูญหายของตัวแปรค่าแบ่งกลุ่ม

Gender of Responder

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	3	60.0%	60.0%	60.0%
	Female	2	40.0%	40.0%	100.0%
Total		5	100.0%		

Education of Responder

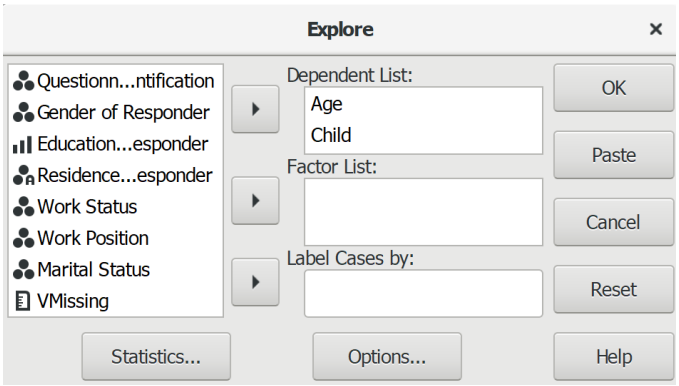
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Primary Education	1	20.0%	25.0%	25.0%
	Secondary Education	1	20.0%	25.0%	50.0%
	Higher Education	2	40.0%	50.0%	100.0%
Missing	9	1	20.0%		
Total		5	100.0%		

การตรวจสอบ: ถ้ามีค่าสูญหายน้อยกว่า 5% เลือกใช้การวิเคราะห์แบบ Listwise และถ้ามีค่าสูญหายมากกว่าหรือเท่ากับ 5% กำหนดค่าสูญหายเป็นกลุ่มใหม่ (new category)

วิธีที่ 2 การตรวจสอบตัวแปรค่าต่อเนื่อง

การใช้โปรแกรมวิเคราะห์ข้อมูล

- เปิดแฟ้มข้อมูล **File -> Open** -> ASPSS_Data.sav
- เลือกเมนู **Analyze -> Descriptive Statistics -> Explore...**
  - **Dependent List:** ใส่ตัวแปร Age และ Child



  - กดปุ่ม **Options...**
    - **Missing Values** เลือก **Report values**
    - กดปุ่ม **Continue**
  - กดปุ่ม **OK**

ตัวอย่าง: ผลลัพธ์การตรวจสอบค่าสูญหายของตัวแปรค่าต่อเนื่อง

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Age of Responder	3	60.0%	2	40.0%	5	100.0%
Child Numbers	3	100.0%	0	.0%	3	100.0%

การตรวจสอบ: ถ้ามีค่าสูญหายน้อยกว่า 5% เลือกใช้การวิเคราะห์แบบ Listwise ถ้ามีค่าสูญหายระหว่าง 5% ถึง 15% การแทนค่าน้อยกว่า 15% ของข้อมูลมีผลกระทบต่อผลลัพธ์เล็กน้อย ให้คำนวณหาค่าเฉลี่ยและแทนค่าโดยใช้คำสั่ง Recode into Difference Variables... แต่ถ้าต้องแทนค่ามากกว่า 15% ให้พิจารณาตัดตัวแปรออกจากการวิเคราะห์

เนื่องจากโปรแกรมทางสถิติส่วนใหญ่จะไม่นำเอาข้อมูล (exclude) ที่มีค่าสูญหาย (missing value) มาทำการวิเคราะห์ ซึ่งไม่ตรงกับความสนใจด้านประชากรที่เป็นเป้าหมายในการศึกษา และมีผลทำให้ผลการวิเคราะห์มีความคลาดเคลื่อน ดังนั้นจึงวิธีการแก้ไขมีดังนี้

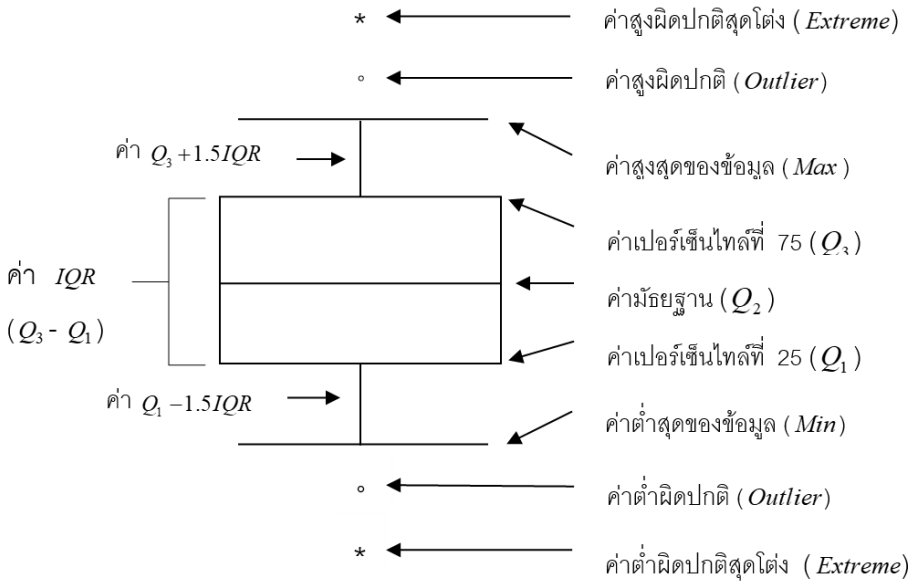
1. วิเคราะห์เฉพาะชุดข้อมูลที่ทุกตัวแปรไม่มีค่าสูญหาย (Listwise)
2. วิเคราะห์เฉพาะชุดข้อมูลที่คู่ตัวแปรไม่มีค่าสูญหาย (Pairwise)
3. ประมาณค่าที่สูญหายด้วยวิธีความเป็นไปได้สูงสุด (Maximum Likelihood Estimation: MLE)
4. แทนค่าที่สูญหาย (Imputation) ดังนี้
   - แทนค่าด้วยค่าเฉลี่ยของทั้งชุดข้อมูล (Series mean)
   - แทนค่าด้วยค่าเฉลี่ยของค่าที่อยู่ข้างบนและข้างล่างของค่าสูญหาย (Mean of nearby points)
   - แทนค่าด้วยค่ามัธยฐานของค่าที่อยู่ข้างบนและข้างล่างของค่าสูญหาย (Median of nearby points)
   - แทนค่าด้วยค่าที่ได้จากการลากเส้นตรงจากค่าสุดท้ายก่อนค่าสูญหายไปยังค่าแรกหลังค่าสูญหาย (Linear interpolation)
   - แทนค่าด้วยค่าทำนายที่ได้จากสมการถดถอย (Linear trend at point)

1.4 การตรวจสอบค่าผิดปกติและค่าต่างสุด

ค่าผิดปกติและค่าต่างสุดอาจรวมอยู่ในตัวแปรเดียวกัน แต่ค่าผิดปกติ (outliers) และค่าต่างสุด (extremes) ในแต่ละตำรามีนิยามและหลักเกณฑ์ไม่เหมือนกัน ดังนี้

1. ค่าผิดปกติ บางตำรากล่าวว่าเป็นค่าตัวแปรที่ไม่ถูกต้องที่ปรากฏอยู่ในชุดข้อมูล หรือเป็นค่าที่ไม่เป็นจริงจากประชากรหรือกลุ่มตัวอย่าง เช่น เงินเดือน 15 บาท/เดือน แต่อีกตำราหมายถึงค่าตัวแปรที่น้อยกว่า Q_1-1.5(IQR) และค่าตัวแปรที่มากกว่า Q_3+1.5(IQR)

2. ค่าต่างสุด บางตำรากล่าวว่าเป็นค่าที่เป็นจริงแต่เบี่ยงเบน (deviates) ออกไปจากค่าปกติที่ได้จากประชากรหรือกลุ่มตัวอย่าง อาจเป็นค่าที่มากกว่าหรือน้อยกว่าค่าปกติในชุดตัวแปรเดียวกัน เช่น เงินเดือน 100,000 บาท/เดือน บางตำรากำหนดให้ค่าต่างสุดอยู่ที่น้อยกว่า Q_1-1.5(IQR) และค่าต่างสุดอยู่ที่มากกว่า Q_3+1.5(IQR) แต่บางตำรากำหนดให้ค่าต่างสุดอยู่ที่น้อยกว่า Q_1-3(IQR) และค่าต่างสุดอยู่ที่มากกว่า Q_3+3(IQR)



รูปที่ 9.1 ค่าผิดปกติและค่าต่างสุด

แม้ว่าค่าผิดปกติและค่าต่างสุดจะมีนิยามและเกณฑ์ในการกำหนดค่าผิดปกติและค่าต่างสุดไม่เหมือนกัน แต่ทั้งสองค่าที่กล่าวมามีสาเหตุการเกิดดังนี้

1. การกรอกข้อมูล คือ กรอกข้อมูลผิดพลาด (wrong entry) เป็นค่าของตัวแปรที่ไม่อยู่ในช่วงของข้อมูลตามแบบสอบถาม

2. ข้อมูลจากกลุ่มตัวอย่าง มีกลุ่มตัวอย่างบางกลุ่มตัวอย่างมีคุณลักษณะที่แตกต่างไปจากกลุ่มตัวอย่างทั่วไป เช่น สูงมาก มีบุตรมาก มีรายได้สูงมาก

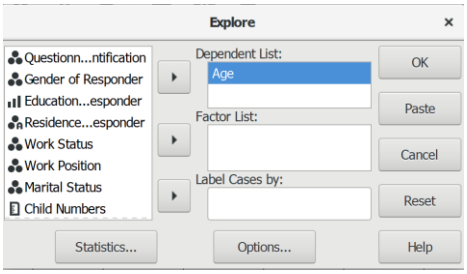
เนื่องจากในหัวข้อการการตรวจสอบความถูกต้องของข้อมูลได้กล่าวถึงการตรวจสอบค่าตัวแปรที่ไม่ปกติแล้ว ประกอบกับคำสั่งในโปรแกรมใช้คำว่า Extremes ดังนั้นเพื่อไม่ให้เกิดความสับสนในที่นี้จะใช้คำว่า ค่าต่างสุด (extremes) ซึ่งหมายถึง ค่าของตัวแปรที่มีค่าต่ำกว่าปกติหรือสูงกว่าปกติเมื่อเทียบกับค่าปกติที่ได้มาจากกลุ่มตัวอย่าง

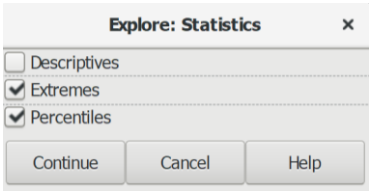
ค่าต่างสุดมีผลทำให้เกิดความคาดเคลื่อนของค่าสถิติ เช่น ค่าเฉลี่ยต่ำหรือสูงเกินจริง ค่าเบี่ยงเบนมาตรฐานสูงกว่าปกติ การแจกแจงข้อมูลไม่ปกติ ค่าความเบ้และความโด่งไม่ปกติ ค่าสัมประสิทธิ์สหสัมพันธ์ (correlation coefficient) เพิ่มขึ้นหรือลดลง ค่าความชัน (slope) ในการพยากรณ์ความถดถอย (regression) เบี่ยงเบนไป เป็นต้น นำไปสู่การ การตัดสินใจผิดพลาด โดยปฏิเสธและยอมรับสมมติฐานไม่ถูกต้อง เช่น ปฏิเสธสมมติฐาน H_0 ทั้งที่เป็นสมติฐานที่ถูกต้อง (Type I error) หรือยอมรับสมมติฐาน H_0 ทั้งที่เป็นสมติฐานที่ไม่ถูกต้อง (Type II error) (Struble, 2002; Verma, & Abdel-Salam, 2019: p. 76-79)

วิธีการตรวจสอบค่าต่างสุดในเบื้องต้นมี 2 วิธี คือ การตรวจสอบเชิงข้อเท็จจริงจากลักษณะทางประชากร และการตรวจสอบเชิงสถิติจากค่า Tukey’s Hinges ค่าต่างสุด คือ ค่าที่อยู่ต่ำกว่าค่า Lower Hinge (H_L) และสูงกว่าค่า Upper Hinge (H_U) (Hoaglin, 2003) โดยมีวิธีการดังต่อไปนี้

การใช้โปรแกรมวิเคราะห์ข้อมูล

- เปิดแฟ้มข้อมูล **File -> Open** -> ASPSS_Data.sav
- เลือกเมนู **Analyze -> Descriptive Statistics -> Explore...**
  - **Dependent List:** ใส่ตัวแปร Age
  - กดปุ่ม **Statistics... ->** เลือกคำสั่ง **Extremes** และ **Percentiles ->** กดปุ่ม **Continue**





ตัวอย่าง: ผลลัพธ์การตรวจสอบค่าต่างสุด

Extreme Values

		Case Number	Value
Age of Responder	Highest 1	5	60.00
	2	4	30.00
	3	3	20.00
	4	0	.00[a]
	5	0	.00[a]
	Lowest 1	3	20.00
	2	4	30.00
	3	5	60.00
	4	5	60.00
	5	5	60.00

a. User-missing value.

Percentiles

		Percentiles						
		5	10	25	50	75	90	95
Age of Responder	Weighted Average	4.00	8.00	20.00	30.00	60.00	60.00	60.00
	Tukey's Hinges			25.00	30.00	45.00		

การตรวจสอบ: 1. ตรวจสอบค่าต่างสุดในตาราง Extreme Value โดยดูค่าสูงสุดและต่ำสุดว่ามีความเป็นไปได้หรือสอดคล้องกับลักษณะทางประชากรของตัวอย่างหรือไม่ เช่น กลุ่มตัวอย่างเป็นวัยผู้ใหญ่ มีความเป็นไปได้หรือไม่ที่จะมีอายุสูงสุด = 60 ปี หรือถ้ากลุ่มตัวอย่างเป็นวัยหนุ่มสาว มีความเป็นไปได้หรือไม่ที่จะมีอายุสูงสุด = 60 ปี 2. ตรวจสอบจากค่าต่างสุดในตาราง Percentiles โดยพิจารณาจากค่า Tukey's Hinges ค่าต่างสุด มีค่าที่ต่ำกว่า H_L = 25 และสูงกว่า H_U = 45 หรือไม่ จากตัวอย่าง มีเฉพาะอายุ 60 ปี เท่านั้นที่เป็นค่าต่างสุด แม้ว่าในแฟ้มข้อมูลมีอายุ 14 ปี แต่ถูกกำหนดให้เป็นค่าสูญหายแบบ user missing ไม่ให้เอามาคำนวณไว้แล้ว

ค่าต่างสุดอาจไม่ต้องกำจัดทิ้ง เพราะเป็นค่าที่แท้จริงของกลุ่มตัวอย่าง แต่หากเกิดจากความผิดพลาดในการนำเข้าข้อมูล หรือเกิดจากความผิดพลาดในการเก็บข้อมูลที่ได้ข้อมูลมาจากประชากรที่อยู่นอกกลุ่มตัวอย่างต้องกำจัดทิ้ง

2. การตรวจสอบข้อมูลตามข้อตกลงเบื้องต้น

การตรวจสอบข้อมูลตามข้อตกลงเบื้องต้น คือ การตรวจสอบหรือประเมินข้อมูลตามเงื่อนไขเบื้องต้นของสถิติแต่ละประเภท (basic assumptions) โดยเฉพาะวิเคราะห์ข้อมูล 1 ตัวแปร (univariate analysis) และ 2 ตัวแปร (bivariate analysis) ด้วยสถิติเชิงอ้างอิง (inferential statistics) มีข้อตกลงทั่วไปที่ต้องตรวจสอบ คือ การแจกแจงปกติ (normality) ความเท่ากันของความแปรปรวน (equality of variance) และความสัมพันธ์เชิงเส้นตรง (linearity)

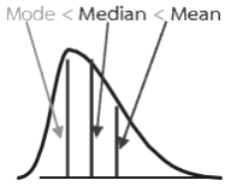
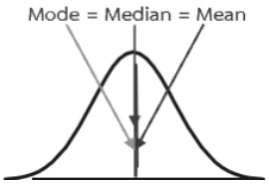
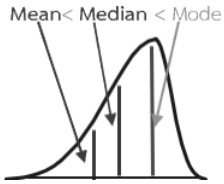
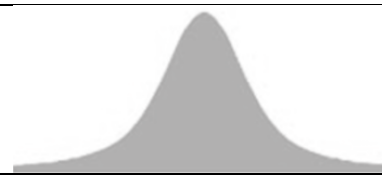
2.1 การแจกแจงแบบปกติ

การแจกแจงแบบปกติ (Normality/Normal distribution) หรืออาจเรียกว่า การแจกแจงแบบเกาส์ (Gaussian distribution) คือ ลักษณะของข้อมูลที่การแจกแจงแบบความน่าจะเป็น (probability distribution) หรือมีค่าแบบสุ่ม (random) ที่แสดงว่า ข้อมูลส่วนใหญ่มีความถี่หรือปรากฏอยู่ใกล้ๆกับค่าเฉลี่ย (mean)

เงื่อนไขเบื้องต้นของสถิติเชิงอ้างอิงในการทดสอบแบบอิงค่าพารามิเตอร์ (parametric test) ต้องทราบลักษณะข้อมูลว่า มีการแจกแจงปกติหรือไม่ ถ้าข้อมูลมีการแจกแจงปกติมีโอกาสน้อยที่จะเกิดความคลาดเคลื่อนเชิงสถิติ แต่หากข้อมูลมีการแจกแจงไม่ปกติการทดสอบสมมติฐานและการสรุปก็จะไม่ถูกต้อง (Verma, & Abdel-Salam, 2019: p. 66-67) ควรตัดสินใจเลือกการทดสอบแบบไม่อิงค่าพารามิเตอร์ (non-parametric test)

ผลกระทบที่เกิดจากข้อมูลที่มีความเบี่ยงเบนไปจากการแจกแจงปกติ คือ ถ้าเป็นข้อมูลของกลุ่มตัวอย่างขนาดใหญ่ อาจทำให้เกิดความสัมพันธ์ระหว่างตัวแปรอันอย่างมีนัยสำคัญทางสถิติ ทั้งที่ในความเป็นจริงแล้วอาจไม่ความสัมพันธ์ระหว่างเหตุการณ์ที่เกิดขึ้นหรือเป็นประเด็นปัญหาที่ไม่มีความสำคัญแต่อย่างใด และถ้าข้อมูลของกลุ่มตัวอย่างมีขนาดเล็ก อาจไม่แสดงความสัมพันธ์ระหว่างตัวแปร

ข้อมูลที่มีลักษณะแจกแจงปกติ (normal distribution) คือ ข้อมูลที่มีรูปทรงแบบระฆังคว่ำที่สมมาตร (symmetric bell-shaped curve) ข้อมูลที่มีการแจกแจงปกติตามทฤษฎีมีค่าความเบ้ (skewness) และความโด่ง (kurtosis) เท่ากับ 0

เบ้ขวา (Right-Skewed)	ปกติ (Normal or Symmetric)	เบ้ซ้าย (Left-Skewed)
		
ข้อมูลแจกแจงแบบเบ้ขวา ค่า Skewness เป็น + แสดงว่าข้อมูลส่วนใหญ่มีค่าต่ำกว่าค่ามัธยฐาน	ข้อมูลแจกแจงแบบปกติ ค่า Skewness เป็น 0 แสดงว่าข้อมูลส่วนใหญ่มีค่าใกล้เคียงกับค่ามัธยฐาน	ข้อมูลแจกแจงแบบเบ้ซ้าย ค่า Skewness เป็น – แสดงว่าข้อมูลส่วนใหญ่มีค่าสูงกว่าค่ามัธยฐาน
ยอดแบน (Platykurtic)	**ยอดปานกลาง (Mesokurtic)**	**ยอดสูง (Leptokurtic)**
		
ข้อมูลแจกแจงแบบยอดแบน ค่า Kurtosis เป็น – แสดงว่าข้อมูลกระจายมากหรือแปรปรวนมาก	ข้อมูลแจกแจงแบบยอดปานกลาง ค่า Kurtosis เป็น 0 แสดงว่าข้อมูลกระจายปานกลางหรือแปรปรวนปานกลาง	ข้อมูลแจกแจงแบบยอดสูง ค่า Kurtosis เป็น + แสดงว่าข้อมูลกระจายน้อยหรือแปรปรวนน้อย

รูปที่ 9. 2 ความเบ้และความโด่ง

ในความเป็นจริงข้อมูลที่ได้จากการสังเกตและเก็บรวบรวมมาจากปรากฏการทางสังคมหรือคนในสังคมไม่น่าจะมีความเป็นไปได้หรือเป็นไปได้น้อยมากที่จะมีการแจกแจงปกติ ขนาดของค่าความเบ้และความโด่งที่ใช้ในการตัดสินว่า ข้อมูลมีการแจกแจงใกล้เคียงโค้งปกติหรือไม่นั้น มีการกำหนดขนาดของค่าที่แตกต่างกัน แต่ส่วนใหญ่ใช้ ± 1 หรือ ± 2 หรือ ± 3 (Garson, Online: 2008; Hair and Others, 2006: 40; Hardy and Bryman, 2004: 44)

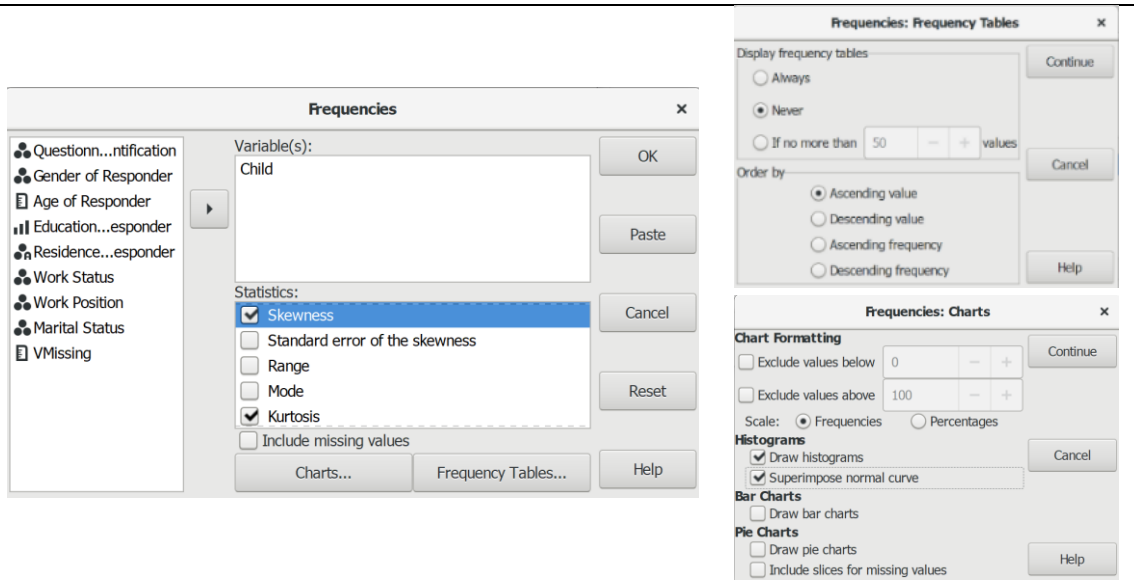
การตรวจสอบการแจกแจงปกติอาจใช้วิทัศน์หรือกราฟ (visual/graphical methods) เช่น กราฟแจกแจงความถี่ (histogram) กราฟลำต้นและใบ (stem-and-leaf diagram) กราฟกล่องและหนวด (boxplot and whisker plot) กราฟการกระจาย (scatter plot) แต่การใช้สายตาตรวจสอบจากภาพของกราฟอาจไม่รับประกันได้ว่าข้อมูลมีการกระจายปกติหรือไม่ ดังนั้น การใช้วิธีการทางสถิติ น่าจะ

ได้รับการยอมรับและมีความน่าเชื่อถือมากกว่า โดยมีวิธีการทดสอบที่สำคัญๆ เช่น ค่าความเบ้ (skewness) และค่าความโด่ง (kurtosis) การทดสอบด้วยวิธีการทางสถิติ เช่น กลุ่มตัวอย่างขนาดใหญ่ใช้การทดสอบแบบคอลมอโกรอฟ-สเมอร์นอฟ (Kolmogorov-Smirnov test: K-S test) และกลุ่มตัวอย่างขนาดเล็กถึงขนาดกลางใช้การทดสอบแบบแชพิโล-วิลค์ (Shapiro-Wilk test: S-W test) สำหรับในที่นี้จะทำเสนอวิธีการตรวจสอบโดยใช้ค่าความเบ้และความโด่งประกอบการใช้กราฟแจกแจงความถี่ และการทดสอบด้วยวิธีการทางสถิติแบบแบบคอลมอโกรอฟ-สเมอร์นอฟเท่านั้น โดยมีรายละเอียดดังนี้

2.1.2 การตรวจสอบความเบ้และความโด่ง

2.1.2.1 การใช้โปรแกรมวิเคราะห์ข้อมูล

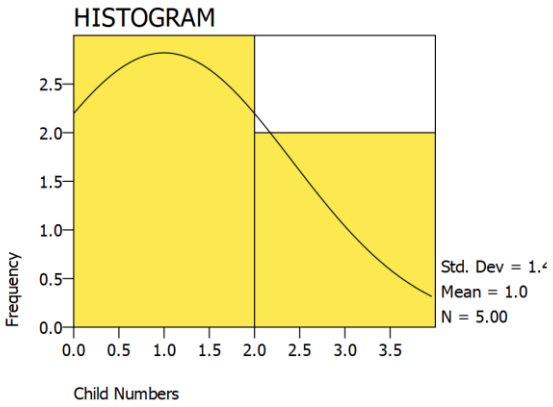
- เปิดแฟ้มข้อมูล **File -> Open** -> ASPSS_Data.sav
- เลือกเมนู **Analyze -> Descriptive Statistics -> Frequencies...**
  - **Frequencies** ใส่ตัวแปร (ค่าต่อเนื่อง) Child
  - กดปุ่ม **Chart**
    - **Histograms** เลือก **Draw histograms** และ **Superimpose normal curve**
    - กดปุ่ม **Continue**
  - กดปุ่ม **Chart**
    - **Display frequency tables** เลือก **Never**



ตัวอย่าง: ผลลัพธ์การตรวจสอบค่าความเบ้และความโด่ง

Statistics

	Child Numbers
N Valid	5
Missing	0
Kurtosis	-1.75
Skewness	.88



การตรวจสอบ: ข้อมูลตัวแปร Child (จำนวนบุตร) มีลักษณะเบ้ขวาและยอดแบน แสดงว่า ข้อมูลส่วนใหญ่มีค่าต่ำกว่าค่ามัธยฐาน และมีการกระจายหรือแปรปรวนมาก แต่ค่าความเบ้และความโด่งไม่เกิน ±2 ยอมรับได้ว่า ข้อมูลมีการแจกแจงใกล้เคียงโค้งปกติ

2. การทดสอบด้วยวิธีการทางสถิติ

2.1 กำหนดสมมติฐาน

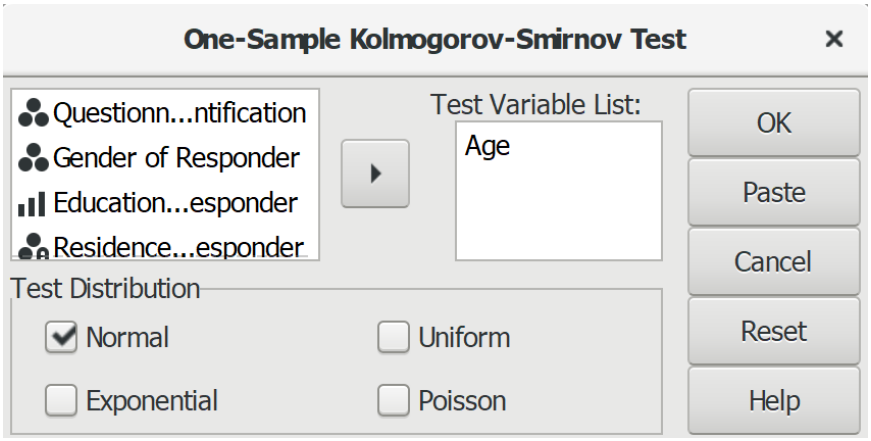
H_0: ข้อมูลมีการแจกแจงปกติ
: ตัวแปรอายุมีการแจกแจงปกติ

H_1: ข้อมูลมีการแจกแจงไม่ปกติ
: ตัวแปรอายุมีการแจกแจงไม่ปกติ

เกณฑ์ในการทดสอบสมมุติฐาน คือ หากการทดสอบการแจกแจงของข้อมูลแบบคอลมอโกรอฟ-สเมอร์นอฟ (Kolmogorov-Smirnov test: K-S test) ได้ค่าความน่าจะเป็น (p-value) มากกว่าค่านัยสำคัญทางสถิติ (α = 0.05) จะยอมรับสมมติฐาน H_o แต่หากได้ค่าความน่าจะเป็น (p-value) เท่ากับหรือน้อยกว่าค่านัยสำคัญทางสถิติ (α = 0.05) จะปฏิเสธสมมติฐาน H_o และยอมรับสมมติฐาน H_a

2.2 การใช้โปรแกรมวิเคราะห์ข้อมูล

- เปิดแฟ้มข้อมูล **File -> Open** -> ASPSS_Data.sav
- เลือกเมนู **Analyze -> Non-Parametric Statistics -> 1-Sample K-S...**
  - **Test Variable List:** ใส่ตัวแปร (ตัวแปรค่าต่อเนื่อง) Age
  - **Test Distribution** เลือกคำสั่งสถิติ **Normal**
  - กดปุ่ม **OK**



2.3 ผลการวิเคราะห์

One-Sample Kolmogorov-Smirnov Test

		Age of Responder
N		3
Normal Parameters	Mean	36.67
	Std. Deviation	20.82
Most Extreme Differences	Absolute	.29
	Positive	.29
	Negative	-.21
Kolmogorov-Smirnov Z		.51
Asymp. Sig. (2-tailed)		.960

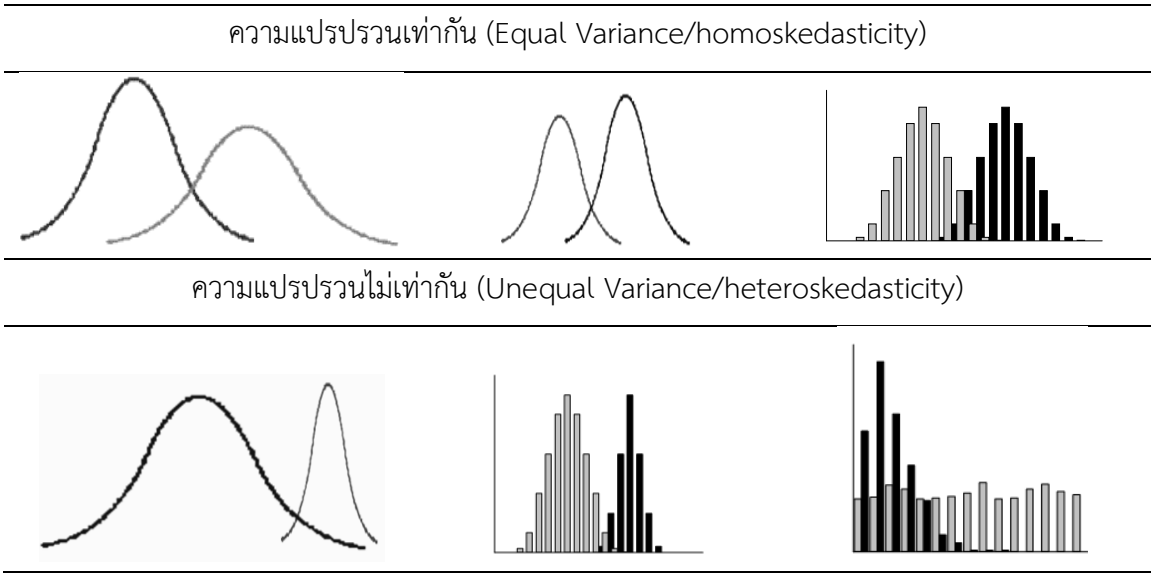
2.4 สรุป

จากผลการทดสอบแสดงให้เห็นว่า ค่าความน่าจะเป็น (Sig.) = 0.960 มีค่ามากกว่าค่านัยสำคัญทางสถิติที่กำหนดไว้ (α = 0.05) จึงยอมรับสมมติฐาน H_0: ตัวแปรอายุมีการแจกแจงปกติ

2.2 ความเท่ากันของความแปรปรวน

ความเท่ากันของความแปรปรวน (Equality of variance) หรือเรียกว่า ความเป็นแบบเดียวกันของความแปรปรวน (Homogeneity of variance) คือ ลักษณะของข้อมูลที่อยู่รอบค่าเฉลี่ยของ 2 กลุ่มตัวอย่าง หรือมากกว่า 2 กลุ่มตัวอย่าง มีลักษณะการแจกแจงหรือการกระจายเท่ากัน

ความไม่เท่ากันของแปรปรวนมีผลทำให้บางค่าหรือบางระดับของตัวแปรอิสระมีความสามารถในการทำนายได้ดีกว่าค่าอื่นของตัวแปรอิสระ รวมถึงมีผลต่อความคาดเคลื่อนมาตรฐาน และการยอมรับ/ไม่ยอมรับผลของการทดสอบสมมติฐาน ดังนั้นจึงต้องทำการทดสอบความแปรปรวนหรือการกระจายของข้อมูลระหว่างกลุ่มในแต่ละตัวแปรตามมีความแปรปรวนเท่ากันหรือไม่ เช่น การตรวจสอบความเท่ากันของความแปรปรวนระหว่างกลุ่มของแต่ละตัวแปร (homogeneity of variance) การตรวจสอบความเท่ากันของความแปรปรวน-ความแปรปรวนร่วมพหุ (homogeneity of variance-covariance matrices) และการตรวจสอบความเท่ากันของความแปรปรวนของความคลาดเคลื่อน (homoscedasticity)



รูปที่ 9.3 ความแปรปรวน

การทดสอบความเท่ากันของความแปรปรวนมีทั้งการทดสอบจากข้อมูลระหว่างกลุ่มของตัวแปรตามแต่ละตัวหรือตัวแปรเดียว และการทดสอบร่วมของตัวแปรตหลายตัวแปรร่วมกัน สำหรับในที่นี้จะ

กล่าวถึงเฉพาะการตรวจสอบความเท่ากันของความแปรปรวนระหว่างกลุ่มของแต่ละตัวแปรตาม (homogeneity of variance) เท่านั้น โดยมีวิธีการดังนี้

1. กำหนดสมมติฐาน

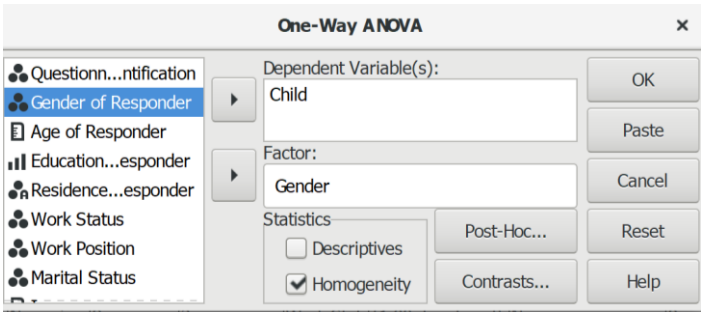
H_0: ข้อมูลระหว่างกลุ่มตัวอย่างมีความแปรปรวนหรือการกระจายเท่ากัน
: ค่าเฉลี่ยจำนวนบุตรระหว่างกลุ่มเพศมีความแปรปรวนเท่ากัน

H_1: ข้อมูลระหว่างกลุ่มตัวอย่างอย่างน้อย 2 กลุ่มประชากรมีความแปรปรวนไม่เท่ากัน
: ค่าเฉลี่ยจำนวนบุตรระหว่างกลุ่มเพศมีความแปรปรวนไม่เท่ากัน

เกณฑ์ในการทดสอบสมมุติฐาน คือ หากทดสอบความเท่ากันของแปรปรวนระหว่างกลุ่มตัวอย่าง (Test of Homogeneity of Variances) ด้วยวิธีการทางสถิติของเลวีน (Levene Statistic) ด้วยค่าทดสอบ 0.05 ได้ค่าความน่าจะเป็น (p-value) มากกว่าค่านัยสำคัญทางสถิติ (α = 0.05) จะยอมรับสมมติฐาน H_0 แต่หากได้ค่าความน่าจะเป็น (p-value) เท่ากับหรือน้อยกว่าค่านัยสำคัญทางสถิติ (α = 0.05) จะปฏิเสธสมมติฐาน H_0 และยอมรับสมมติฐาน H_1

2. การใช้โปรแกรมวิเคราะห์ข้อมูล

- เปิดแฟ้มข้อมูล ASPSS_Data.sav
- **Analyze -> Compare Means -> One Way ANOVA...**
  - **Test Variables(s):** ใส่ตัวแปร (ตัวแปรค่าต่อเนื่อง) Age
  - **Grouping Variable:** ใส่ตัวแปร (ตัวแปรค่าแบ่งกลุ่ม) Gender
  - **Statistics** เลือก **Homomeneity**



3. ผลการวิเคราะห์

Test of Homogeneity of Variances

	Levene Statistic	df1	df2	Sig.
Child Numbers	.60	1	3	.495

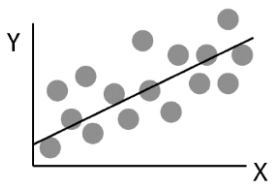
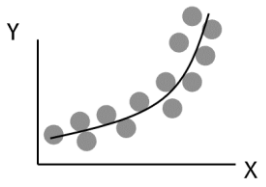
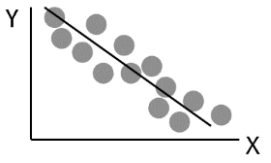
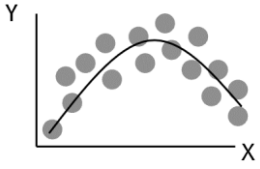
4. สรุป

จากผลการทดสอบแสดงให้เห็นว่า ค่าความน่าจะเป็น (Sig.) = 0.495 มีค่ามากกว่าค่านัยสำคัญทางสถิติที่กำหนดไว้ (α = 0.05) จึงยอมรับสมมติฐาน H_0: ค่าเฉลี่ยจำนวนบุตรระหว่างกลุ่มเพศมีความแปรปรวนเท่ากัน

2.3 ความเป็นเส้นตรง

ความเป็นเส้นตรง (Linearity) คือ ลักษณะความสัมพันธ์ของข้อมูลระหว่าง 2 ตัวแปรที่มีการเปลี่ยนแปลงไปในทิศทางเดียวกันในเชิงเส้นตรง เมื่อค่าของตัวแปรหนึ่งมีการเปลี่ยนแปลงเพิ่มขึ้นค่าของอีกตัวแปรหนึ่งมีการเปลี่ยนแปลงเพิ่มขึ้นด้วย และเมื่อตัวแปรหนึ่งมีค่าเปลี่ยนแปลงลดลงค่าของอีกตัวแปรหนึ่งก็มีการเปลี่ยนแปลงลดลงด้วยเช่นกัน

การวิเคราะห์ข้อมูลแบบสองตัวแปร (bivariate analysis) เช่น การวิเคราะห์การถดถอยแบบง่าย (simple regression) และการวิเคราะห์ข้อมูลแบบหลายตัวแปร (multivariate analysis) เช่น การวิเคราะห์การถดถอยแบบพหุหรือหลายตัวแปร (multiple regression) การวิเคราะห์การถดถอยแบบลอจิสติก (logistic regression) การวิเคราะห์ปัจจัย (factor analysis) และการวิเคราะห์ตัวแบบสมการเชิงโครงสร้าง (structural equation model) ข้อมูลที่ใช้ในการวิเคราะห์ต้องมีความสัมพันธ์ระหว่างกันเชิงเส้นตรง ไม่เป็นความสัมพันธ์เชิงเส้นโค้ง

ความสัมพันธ์เชิงเส้นตรง (Linear relationships)	**ความสัมพันธ์เชิงเส้นโค้ง (Curvilinear relationships)**
	
	

รูปที่ 9.4 ความสัมพันธ์เชิงเส้นตรงและเส้นโค้ง

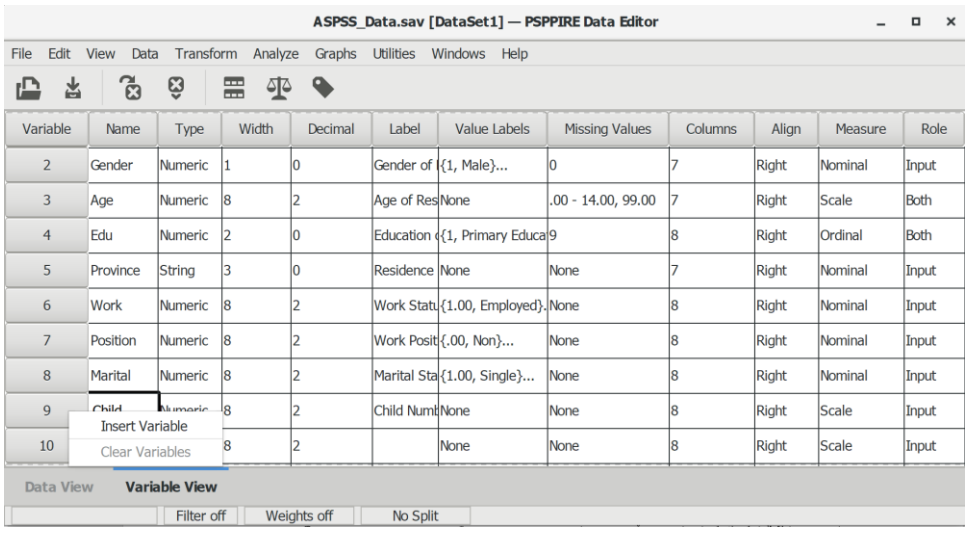
การตรวจสอบความเป็นเส้นตรงระหว่างตัวแปรมีหลายวิธีขึ้นอยู่กับจำนวนตัวแปร แต่โดยทั่วไปมี 2 วิธี ดังนี้

1. การตรวจสอบด้วยกราฟหรือแผนภาพ เช่น แผนภาพการกระจาย (scatterplot) โดยกำหนดให้ตัวแปรอิสระ (X) อยู่ในแกนนอน (X Axis) และตัวแปรตาม (Y) อยู่ในแกนตั้ง (Y Axis) ถ้าตัวแปรมีความสัมพันธ์เชิงเส้นตรง การกระจายของจุดข้อมูลจะมีเป็นรูปทรงรี (oval shape) หรือคล้ายซิการ์ (cigar shape)

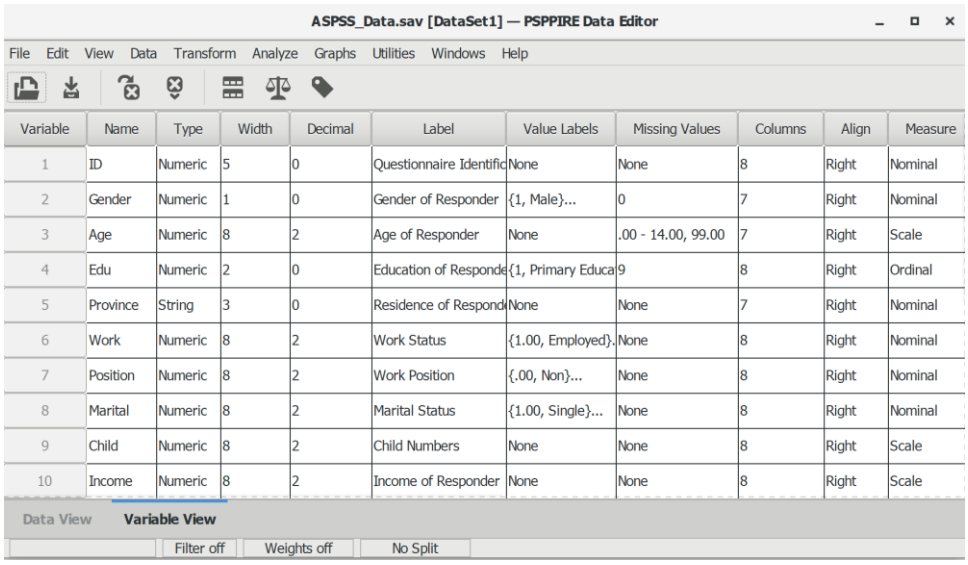
2. การตรวจสอบด้วยค่าสถิติ เช่น การตรวจสอบด้วยค่าสัมประสิทธิ์สหสัมพันธ์ (Correlation coefficient) ตัวแปรที่มีความสัมพันธ์ระหว่างกัน แสดงว่าเป็นตัวแปรที่มีความสัมพันธ์เชิงเส้น (Hair and Others, 2006: 205) โดยค่าสัมประสิทธิ์สหสัมพันธ์ (r) เป็นดัชนีบอกระดับความสัมพันธ์ (degree of association) หากมีค่าใกล้ 0 แสดงว่า มีความสัมพันธ์เชิงเส้นระหว่างกันน้อย แต่หากมีค่าใกล้ 1 แสดงว่า มีความสัมพันธ์เชิงเส้นระหว่างกันมาก

สำหรับในที่นี้จะนำเสนอเฉพาะการตรวจสอบความสัมพันธ์เชิงตรงระหว่าง 2 ตัวแปรด้วยค่าสัมประสิทธิ์สหสัมพันธ์เท่านั้น โดยมีวิธีการดังนี้

- เปิดแฟ้มข้อมูล **File -> Open** -> ASPSS_Data.sav
- **สร้างตัวแปร 1 ตัวแปร** คือ Income ต่อจากตัวแปร Child โดยมีรายละเอียดดังนี้
  - คลิกเมาส์ปุ่มขวาที่เลข 9 หน้าตัวแปร Child -> เลือกคำสั่ง **Insert Variable**



- เพิ่มตัวแปร Income (Label = Income of Responder และ Measure = Scale)



ASPSS_Data.sav [DataSet1] — PSPPIRE Data Editor

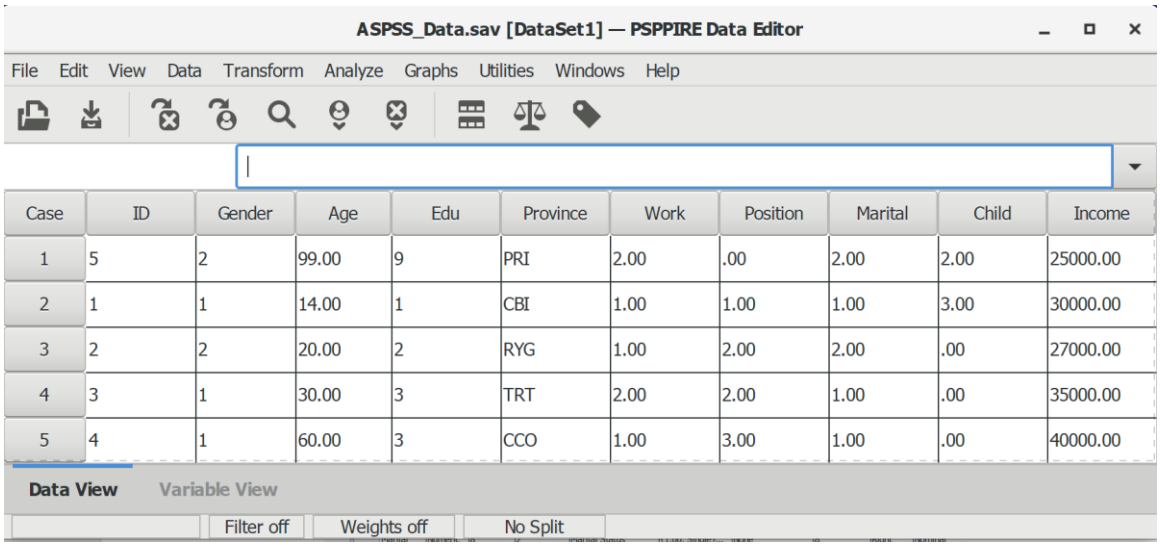
File Edit View Data Transform Analyze Graphs Utilities Windows Help

Variable	Name	Type	Width	Decimal	Label	Value Labels	Missing Values	Columns	Align	Measure
1	ID	Numeric	5	0	Questionnaire Identific	None	None	8	Right	Nominal
2	Gender	Numeric	1	0	Gender of Responder	{1, Male}...	0	7	Right	Nominal
3	Age	Numeric	8	2	Age of Responder	None	.00 - 14.00, 99.00	7	Right	Scale
4	Edu	Numeric	2	0	Education of Responde	{1, Primary Educa	9	8	Right	Ordinal
5	Province	String	3	0	Residence of Respond	None	None	7	Right	Nominal
6	Work	Numeric	8	2	Work Status	{1.00, Employed}...	None	8	Right	Nominal
7	Position	Numeric	8	2	Work Position	{.00, Non}...	None	8	Right	Nominal
8	Marital	Numeric	8	2	Marital Status	{1.00, Single}...	None	8	Right	Nominal
9	Child	Numeric	8	2	Child Numbers	None	None	8	Right	Scale
10	Income	Numeric	8	2	Income of Responder	None	None	8	Right	Scale

Data View | Variable View

Filter off | Weights off | No Split

- ใส่ค่าตัวแปรหรือข้อมูลโดยมีรายละเอียดดังนี้



ASPSS_Data.sav [DataSet1] — PSPPIRE Data Editor

File Edit View Data Transform Analyze Graphs Utilities Windows Help

Case	ID	Gender	Age	Edu	Province	Work	Position	Marital	Child	Income
1	5	2	99.00	9	PRI	2.00	.00	2.00	2.00	25000.00
2	1	1	14.00	1	CBI	1.00	1.00	1.00	3.00	30000.00
3	2	2	20.00	2	RYG	1.00	2.00	2.00	.00	27000.00
4	3	1	30.00	3	TRT	2.00	2.00	1.00	.00	35000.00
5	4	1	60.00	3	CCO	1.00	3.00	1.00	.00	40000.00

Data View | Variable View

Filter off | Weights off | No Split

1. กำหนดสมมติฐาน

H_0: ไม่มีความสัมพันธ์เชิงเส้นตรงระหว่างตัวแปร

: ไม่มีความสัมพันธ์เชิงเส้นตรงระหว่างตังแปรอายุและรายได้

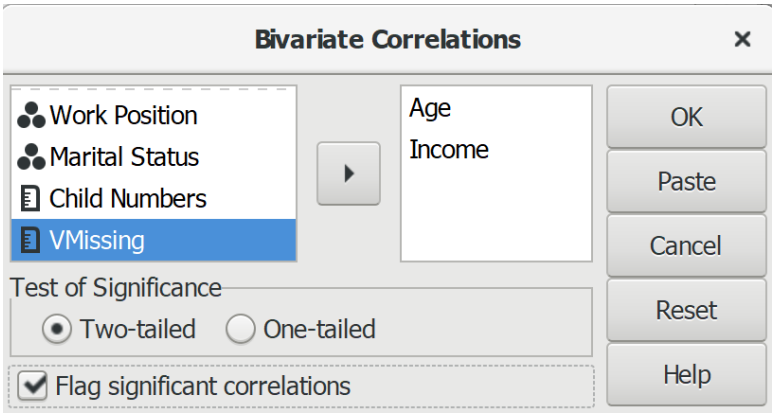
H_1: มีความสัมพันธ์เชิงเส้นตรงระหว่างตัวแปร

: มีความสัมพันธ์เชิงเส้นตรงระหว่างตังแปรอายุและรายได้

เกณฑ์ในการทดสอบสมมุติฐาน คือ หากการทดสอบสัมประสิทธิ์ความสัมพันธ์เพียร์สันได้ค่าความน่าจะเป็น (p-value) มากกว่าค่านัยสำคัญทางสถิติ (α = 0.05) จะยอมรับสมมติฐาน H_0 แต่หากได้ค่าความน่าจะเป็น (p-value) น้อยกว่าหรือเท่ากับค่านัยสำคัญทางสถิติ (α = 0.05) จะปฏิเสธสมมติฐาน H_0 และยอมรับสมมติฐาน H_1

2. การใช้โปรแกรมวิเคราะห์ข้อมูล

- เลือกเมนู **Analyze -> Bivariate Correlation...**
  - เลือกตัวแปร Age และ Income (ตัวแปรค่าต่อเนื่อง) จากรายการตัวแปรด้านซ้าย -> ใส่ช่องว่างด้านขวา
  - **Flag Significance Correlations** เลือก /
  - กดปุ่ม **OK**



3. ผลการวิเคราะห์

Correlations

		Age of Responder	Income of Responder
Age of Responder	Pearson Correlation	1.000	.916
	Sig. (2-tailed)		.263
	N	3	3
Income of Responder	Pearson Correlation	.916	1.000
	Sig. (2-tailed)	.263	
	N	3	5

4. สรุป

จากผลการทดสอบแสดงให้เห็นว่า ค่าความน่าจะเป็น (Sig.) = 0.263 มีค่ามากกว่าค่านัยสำคัญทางสถิติที่กำหนดไว้ (α = 0.05) จึงยอมรับสมมติฐาน H_0: ไม่มีความสัมพันธ์เชิงเส้นตรงระหว่างตังแปรอายุและรายได้

สรุป

การวิเคราะห์ข้อมูลเชิงสำรวจ คือ กระบวนการทำให้ข้อมูลมีความสมบูรณ์ (data cleansing) โดยการตรวจสอบข้อมูลให้มีความถูกต้องและตรงกับความเป็นจริง รวมถึงการตรวจหาค่าต่างสุด และการตรวจสอบค่าสูญหาย หากข้อมูลมีความไม่ถูกต้องและไม่สมบูรณ์จะมีผลทำให้การวิเคราะห์เบี่ยงเบนไปจากความเป็นจริง แม้ว่าข้อมูลจะมีความถูกต้องและสมบูรณ์แล้ว ก่อนการทดสอบแบบอิงค่าพารามิเตอร์ยังต้องมีการตรวจสอบ (data screening) ข้อตกลงเบื้องต้นเกี่ยวกับลักษณะของข้อมูล เช่น การแจกแจงปกติ ความเท่ากันของความแปรปรวน และความสัมพันธ์เชิงเส้นตรง เพราะการละเมิดข้อตกลงเบื้องต้นในการวิเคราะห์ข้อมูลด้วยสถิติแต่ละประเภทมีความเสี่ยงที่จะทำให้ผลวิเคราะห์เกิดความคลาดเคลื่อนและได้ข้อสรุปที่ไม่ถูกต้อง

แบบฝึกหัดท้ายบท

1. ให้ผู้เรียนตรวจสอบความถูกต้องของข้อมูลแบบไขว้ระหว่างตัวแปรอายุ (age) และระดับการศึกษา (edu)

2. ให้ผู้เรียนวิเคราะห์จำนวนค่าสูญหายของแต่ละตัวอย่างจากตัวแปรเพศ (gender) สถานการณ์ทำงาน (work) และตำแหน่งงาน (position)

3. ให้ผู้เรียนวิเคราะห์หาค่าต่างสุดของตัวแปรจำนวนบุตร (child)

4. ให้ผู้เรียนตรวจสอบการแจกแจงปกติแบบคอลมอโกรอฟ-สเมอร์นอฟ (Kolmogorov-Smirnov: K-S Test) ของตัวแปรจำนวนบุตร (child)

5. ให้ผู้เรียนตรวจสอบความเท่ากันของความแปรปรวนของตัวแปรอายุ (age) ระหว่างกลุ่มเพศ (gender)

6. ให้ผู้เรียนตรวจสอบความสัมพันธ์เชิงเส้นตรงระหว่างตัวแปรอายุ (age) กับจำนวนบุตร (child)

ส่วนที่ 4

การวิเคราะห์ข้อมูลด้วยสถิติเชิงพรรณนา

บทที่ 10
การวิเคราะห์ข้อมูลเชิงพรรณนา 1 ตัวแปร

ความมุ่งหมายของบทเรียน

ผู้เรียนสามารถใช้โปรแกรมประมวลผลการแจกแจงข้อมูลของตัวแปร 1 ตัวแปร ด้วยสถิติเชิงพรรณนา และเขียนบรรยายผลการวิเคราะห์ได้

เนื้อหา

1. การวิเคราะห์ตัวแปรแบบค่าแบ่งกลุ่ม
2. การวิเคราะห์ตัวแปรแบบค่าต่อเนื่อง

กิจกรรมและวิธีสอน

บรรยาย อธิบาย แสดงตัวอย่าง และทำแบบฝึกหัดท้ายบท

อุปกรณ์การสอน

เอกสารคำสอน ภาพเลื่อน คอมพิวเตอร์ โปรแกรม PSPP โปรแกรมสอนออนไลน์ (สอนแบบออนไลน์) และเครื่องฉาย (สอนในชั้นเรียน)

งานที่มอบหมาย

ทำแบบฝึกหัดท้ายบท

การวัดผลและประเมินผล

การเข้าชั้นเรียน สังเกตความสนใจจากการเรียน สอบถามความเข้าใจจากการบรรยาย และประเมินจากการทำแบบฝึกหัดท้ายบท

เนื้อหา

การนำเสนอข้อมูลทั่วไปเกี่ยวของประชากรและกลุ่มตัวอย่างส่วนใหญ่ใช้การวิเคราะห์ข้อมูล 1 ตัวแปร (univariate analysis) เพื่อนำเสนอลักษณะของข้อมูลและบรรยายลักษณะของข้อมูลของแต่ละตัวแปร และสถิติที่นิยมใช้คือสถิติเชิงพรรณนา ได้แก่ ค่าความถี่ (frequency) ค่าร้อยละ (percentage) ค่าฐานนิยม (mode) ค่าเฉลี่ย (mean) ค่าเบี่ยงเบนมาตรฐาน (standard deviation) และค่าสัมประสิทธิ์การกระจาย (coefficient of variation)

1. การวิเคราะห์ตัวแปรแบบค่าจัดประเภท

การวิเคราะห์ตัวแปรแบบค่าจัดประเภท 1 ตัวแปร เป็นการวิเคราะห์และอธิบายข้อมูลจากตัวแปรที่มีระดับการวัดเป็นค่าจัดประเภท (categorical variables) เช่น ตัวแปรแบบมาตราจัดกลุ่ม (nominal scale) และตัวแปรแบบมาตราจัดอันดับ (scale: ordinal) ด้วยสถิติเชิงพรรณนา (descriptive statistics) ดังนี้

1. การแจกแจงความถี่ (frequency distribution) ได้แก่ จำนวนนับ ความถี่ และค่าร้อยละ)

2. การวิเคราะห์ตำแหน่ง (location) หรือแนวโน้มเข้าศูนย์กลาง (central tendency) ได้แก่ ค่าฐานนิยม (mode)

1.1 การวิเคราะห์ข้อมูลด้วยค่าร้อยละ

ค่าร้อยละ (Percentage) คือ ส่วนหนึ่ง (a portion) ของผลรวมเมื่อเทียบกับส่วนร้อยที่แสดงเป็นตัวเลขระหว่าง 0 ถึง 100

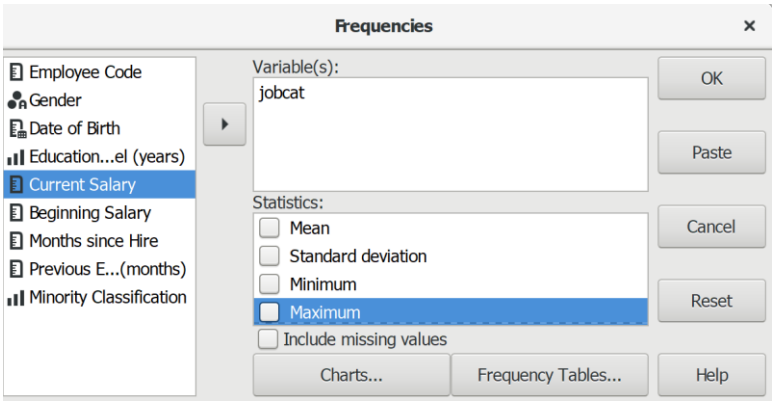
ตัวอย่างคำถามการวิจัย
จากการสำรวจข้อมูลของพนักงานในหน่วยงานต้องการทราบว่า พนักงานแต่ละตำแหน่งมีจำนวนมากน้อยเพียงใด

1.1.1 ตัวแปร

ชื่อตัวแปร	มาตรการวัด
ตำแหน่งงาน (jobcat)	จัดกลุ่ม

1.1.2 การใช้โปรแกรมวิเคราะห์ข้อมูล

- เปิดแฟ้มข้อมูล **File -> Open** -> Employee Data.sav
- เมนู **Analyze -> Descriptive Statistics -> Frequencies...**
  - **Variable (s):** เลือกตัวแปรแบบค่าแบ่งกลุ่ม 1 ตัวแปร หรือมากกว่า จากรายการตัวแปรด้านซ้าย
  - ในกลุ่มคำสั่ง **Statistic:** ไม่เลือกสถิติ Mean, Standard deviation, Minimum และ Maximum
  - กดปุ่ม **OK**



1.1.3 ผลการวิเคราะห์

Employment Category

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Clerical	363	76.6%	76.6%	76.6%
	Custodial	27	5.7%	5.7%	82.3%
	Manager	84	17.7%	17.7%	100.0%
Total		474	100.0%		

1.1.4 การสร้างตารางผลการวิเคราะห์

ตารางที่ (#) แสดงจำนวนและค่าร้อยละของกลุ่มตัวอย่างพนักงาน จำแนกตามตำแหน่งงาน

ตำแหน่งงาน	จำนวน	ร้อยละ
เจ้าหน้าที่ปฏิบัติงาน	363	76.6
หัวหน้างาน	27	5.7
ผู้จัดการ	84	17.7
รวม	474	100.0

1.1.5 การแปลผลและสรุป

จากตารางที่ (#) แสดงให้เห็นว่า กลุ่มตัวอย่างส่วนใหญ่เป็นเจ้าหน้าที่ปฏิบัติงาน รองลงมา คือ ผู้จัดการ และน้อยที่สุดคือ หัวหน้างาน

ตัวอย่างเพิ่มเติม: การแปลผลและสรุป

#1

จากตารางที่ (#) แสดงให้เห็นว่า กลุ่มตัวอย่างจำนวน 474 คน เป็นเจ้าหน้าที่ปฏิบัติงาน จำนวน 363 คน คิดเป็นร้อยละ 76.6 เป็นหัวหน้างาน จำนวน 27 คน คิดเป็นร้อยละ 5.7 และเป็นผู้จัดการ จำนวน 84 คน คิดเป็นร้อยละ 17.7

#2

จากตารางที่ (#) แสดงให้เห็นว่า กลุ่มตัวอย่างจำนวน 474 คน เป็นเจ้าหน้าที่ปฏิบัติงาน ร้อยละ 76.6 (363 คน) เป็นหัวหน้างาน ร้อยละ 5.7 (27 คน) และเป็นผู้จัดการ ร้อยละ 17.7 (84 คน)

#3

จากตารางที่ (#) แสดงให้เห็นว่า กลุ่มตัวอย่างจำนวน 474 คน ส่วนใหญ่เป็นเจ้าหน้าที่ปฏิบัติงาน ร้อยละ 76.6 (363 คน) รองลงมา คือ ผู้จัดการ ร้อยละ 17.7 (84 คน) และน้อยที่สุดเป็นหัวหน้างาน ร้อยละ 5.7 (27 คน)

1.2 การวิเคราะห์ข้อมูลด้วยค่าฐานนิยม

ค่าฐานนิยม (Mode) คือ ค่าที่แสดงความถี่มากที่สุด (most frequency) ของข้อมูล 1 ชุด หรือ 1 ตัวแปร ข้อมูลแต่ละชุดอาจมีค่าฐานนิยม 1 ค่า หรือมากกว่า 1 ค่า หรืออาจไม่มีค่าฐานนิยมเลยก็เป็นได้

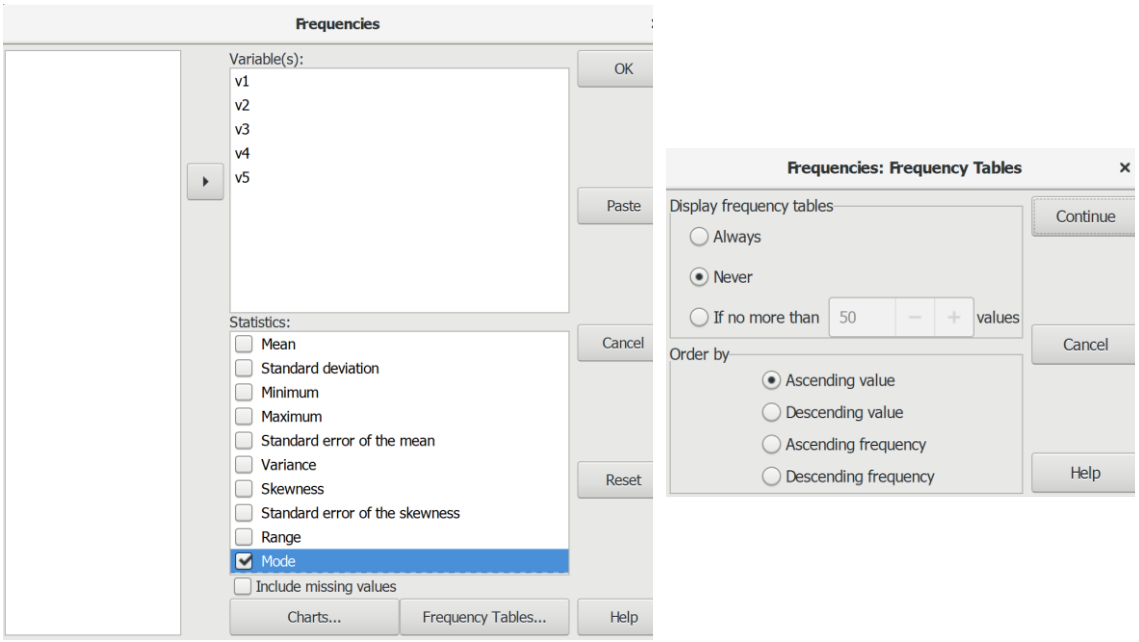
ตัวอย่างคำถามการวิจัย
จากข้อมูลการสำรวจผู้มาใช้บริการต้องการทราบว่า ผู้ใช้บริการให้คะแนนการให้บริการแต่ละด้านอยู่ในระดับมากที่สุด

1.2.1 ตัวแปร

ชื่อตัวแปร	**มาตรการวัด**
ความพึงพอใจกับการให้บริการ (V1)	จัดอันดับ
ความคุ้มค่าของเงินกับการให้บริการ (V2)	จัดอันดับ
ความล่าช้าของพนักงานในการให้บริการ (V3)	จัดอันดับ
ความมีประสิทธิภาพในการให้บริการ (V4)	จัดอันดับ
ความความดังของเสียงรบกวนภายในห้องพัก (V5)	จัดอันดับ

1.2.2 การใช้โปรแกรมวิเคราะห์ข้อมูล

- เปิดแฟ้มข้อมูล **File -> Open** -> hotel.sav
- เลือกเมนู **Analyze -> Descriptive Statistics -> Frequencies...**
  - **Variable (s):** เลือกตัวแปรแบบค่าแบ่งกลุ่ม 1 ตัวแปร หรือหลายตัวแปร จากรายการตัวแปรด้านซ้าย
  - ในกลุ่มคำสั่ง **Statistic:** ไม่เลือกสถิติ Mean, Standard deviation, Minimum, Maximum และเลือกสถิติ Mode
  - กดปุ่ม **Frequency Tables... -> Display frequency tables** เลือก **Never** -> กดปุ่ม **Continue**
  - กดปุ่ม **OK**



1.2.3 ผลการวิเคราะห์

Statistics

		I am satisfied with the level of service	The value for money was good	The staff were slow in responding	My concerns were dealt with in an efficient manner	There was too much noise in the rooms
N	Valid	17	17	17	17	17
	Missing	0	0	0	0	0
Mode		Strongly Disagree	Disagree	Strongly Agree	No Opinion	Strongly Disagree

1.2.4 การสร้างตารางผลการวิเคราะห์

ตารางที่ (#) แสดงค่าฐานนิยมของการประเมินการให้บริการของโรงแรม

ข้อ	การบริการ	จำนวน	ฐานนิยม
1.	ความพึงพอใจกับการให้บริการ	17	ไม่เห็นด้วยอย่างยิ่ง
2.	ความคุ้มค่าของเงินกับการให้บริการ	17	ไม่เห็นด้วย
3.	ความล่าช้าของพนักงานในการให้บริการ	17	เห็นด้วยอย่างยิ่ง
4.	ความมีประสิทธิภาพในการให้บริการ	17	ไม่มีความเห็น
5.	ความดังของเสียงรบกวนภายในห้องพัก	17	ไม่เห็นด้วยอย่างยิ่ง

1.2.5 การแปลผลและสรุป

จากตารางที่ (#) แสดงให้เห็นว่า กลุ่มตัวอย่างส่วนใหญ่ประเมินการให้บริการของโรงแรมเกือบทุกข้ออยู่ในระดับไม่พึงพอใจกับการให้บริการ ยกเว้นด้านเสียงภายในห้องพักที่ประเมินให้อยู่ในระดับพึงพอใจมาก

2. การวิเคราะห์ตัวแปรแบบค่าต่อเนื่อง

การวิเคราะห์ตัวแปรแบบค่าต่อเนื่อง 1 ตัวแปร เป็นการวิเคราะห์และอธิบายข้อมูลจากตัวแปรที่มีระดับการวัดเป็นค่าต่อเนื่อง (continuous variables) ด้วยสถิติเชิงพรรณนา เช่น การวิเคราะห์ตำแหน่ง (location) หรือแนวโน้มเข้าศูนย์กลาง (central tendency) ได้แก่ ค่าเฉลี่ย (mean) การวิเคราะห์การกระจาย (dispersion/variation) ได้แก่ ส่วนเบี่ยงเบนมาตรฐาน (standard deviation) และการวิเคราะห์รูปร่างของการแจกแจงข้อมูล (shape of the distributions) ได้แก่ ความเบ้ (skewness) และความโด่ง (kurtosis) ของข้อมูล

สถิติเชิงพรรณนาที่นิยมใช้วิเคราะห์ตัวแปรแบบค่าต่อเนื่อง คือ ค่าเฉลี่ย (Mean) ที่เป็นค่าการเฉลี่ย (average) หรือค่ากลาง (middle) ของข้อมูลประเภทหนึ่ง โดยการบวก (add) รวมค่าทั้งหมด (total) ในชุดข้อมูล และทำการหาร (divided) ด้วยจำนวนค่าของข้อมูลทั้งหมดในชุดข้อมูล

2.1 การวิเคราะห์ข้อมูลค่าเฉลี่ย (รายตัวแปร)

การวิเคราะห์รายตัวแปร คือ การวิเคราะห์ตัวแปรวัดที่ได้จากการวัดมิติเดียว เช่น อายุ เงินเดือน ระยะเวลาทำงาน น้ำหนัก ส่วนสูง

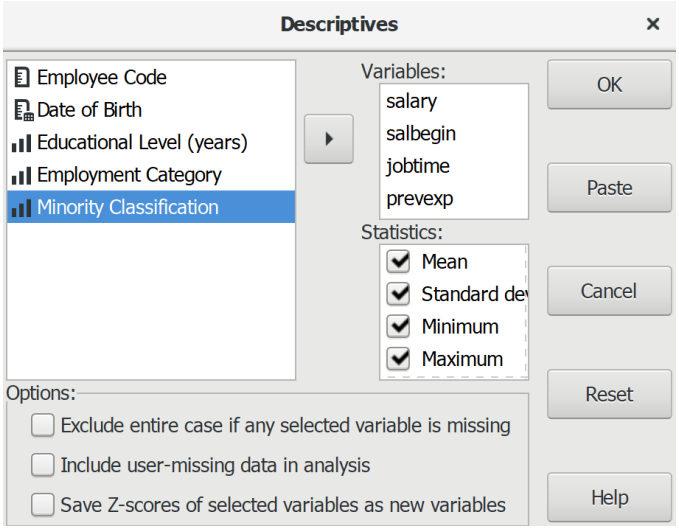
ตัวอย่างคำถามการวิจัย
จากการสำรวจข้อมูลของพนักงานในหน่วยงานต้องการทราบว่า พนักงานมีค่าเฉลี่ยของเงินเดือนปัจจุบัน เงินเดือนเริ่มต้น ระยะเวลาทำงาน และประสบการณ์ เป็นอย่างไร

2.1.1 ตัวแปร

ชื่อตัวแปร	มาตราการวัด
เงินเดือนปัจจุบัน (salary)	ค่าต่อเนื่อง/สเกล
เงินเดือนเริ่มต้น (salbegin)	ค่าต่อเนื่อง/สเกล
ระยะเวลาทำงาน (jobtime)	ค่าต่อเนื่อง/สเกล
ประสบการณ์ (prevexp)	ค่าต่อเนื่อง/สเกล

2.1.2 การใช้โปรแกรมวิเคราะห์ข้อมูล

- เปิดแฟ้มข้อมูล **File -> Open** -> Employee Data.sav
- เลือกเมนู **Analyze -> Descriptive Statistics -> Descriptives...**
  - **Variable (s):** เลือกตัวแปรแบบค่าต่อเนื่อง 1 ตัวแปร หรือมากกว่า จากรายการตัวแปรด้านซ้าย
  - ในกลุ่มคำสั่ง **Statistics:** เลือกสถิติ Mean, Standard deviation, Minimum และ Maximum
  - กดปุ่ม **OK**



2.1.3 ผลการวิเคราะห์

Descriptive Statistics

	N	Mean	Std Dev	Minimum	Maximum
Current Salary	474	34419.57	17075.66	$15,750	$135,000
Beginning Salary	474	17016.09	7870.64	$9,000	$79,980
Months since Hire	474	81.11	10.06	63	98
Previous Experience (months)	474	95.86	104.59	missing	476
Valid N (listwise)	474				
Missing N (listwise)	0				

2.1.4 การสร้างตารางผลการวิเคราะห์

ตารางที่ (#) แสดงค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐาน และค่าสัมประสิทธิ์การกระจายข้อมูลพื้นฐานของพนักงาน

ข้อมูลพื้นฐาน	จำนวน	ค่าต่ำสุด	ค่าสูงสุด	ค่าเฉลี่ย	S.D.	CV[16] (%)
เงินเดือนปัจจุบัน (ดอลล่าร์)	474	15,750	135,000	34,419.57	17,075.661	49.61
เงินเดือนเริ่มต้น (ดอลล่าร์)	474	9,000	79,980	17,016.09	7,870.638	46.25
ระยะเวลาทำงาน (เดือน)	474	63	98	81.11	10.061	12.40
ประสบการณ์ก่อนทำงาน (เดือน)	474	0	476	95.86	104.586	100.10

2.1.5 การแปลผลและสรุป

จากตารางที่ (#) แสดงให้เห็นว่า กลุ่มตัวอย่างจำนวน 474 คน มีเงินเดือนปัจจุบันเฉลี่ย 34,419.57 ดอลลาร์ มีเงินเดือนเริ่มต้นเฉลี่ย 17,016.09 ดอลลาร์ มีระยะเวลาทำงานเฉลี่ย 81.11 เดือน และมีประสบการณ์ก่อนทำงาน เฉลี่ย 95.86 เดือน

เมื่อพิจารณาจากค่าสัมประสิทธิ์การกระจาย (coefficient of variation: CV) แสดงให้เห็น กลุ่มตัวอย่างมีเงินเดือนปัจจุบันแตกต่างกันอยู่ในระดับน้อย มีเงินเดือนเริ่มต้นแตกต่างกันอยู่ในระดับน้อย มีระยะเวลาทำงานแตกต่างกันอยู่ในระดับน้อยมาก และมีประสบการณ์ก่อนทำงานแตกต่างกันอยู่ในระดับมาก

[16] สัมประสิทธิ์การกระจาย (Coefficient of Variation: C.V.) คือ ค่าวัดการกระจายของข้อมูลรอบค่าเฉลี่ยจากอัตราส่วน (ratio) ของค่าเบี่ยงเบนมาตรฐาน (standard deviation) ต่อค่าเฉลี่ย (mean) หรือคำนวณจากค่า Standard deviation หารด้วยค่า Mean และคูณด้วย 100 หากค่า C.V. น้อยกว่า 1 หรือ 100 แสดงว่าข้อมูลมีการกระจายน้อย แต่หากมากกว่า 1 หรือ 100 แสดงว่าข้อมูลมีการกระจายมาก สามารถนำไปใช้เปรียบเทียบการกระจายของข้อมูลระหว่างกลุ่มได้ง่ายกว่าค่าเบี่ยงเบนมาตรฐาน

ตัวอย่างเพิ่มเติม: การแปลผลและสรุป

#1

จากตารางที่ (#) แสดงให้เห็นว่า กลุ่มตัวอย่างจำนวน 474 คน มีเงินเดือนปัจจุบันเฉลี่ย 34,419.57 ดอลลาร์ โดยมีเงินเดือนปัจจุบันต่ำสุด 15,750 ดอลลาร์ และสูงสุด 135,000 ดอลลาร์ มีเงินเดือนเริ่มต้น เฉลี่ย 17,016.09 ดอลลาร์ โดยมีเงินเดือนเริ่มต้นต่ำสุด 9,000 ดอลลาร์ และสูงสุด 79,980 ดอลลาร์ มีระยะเวลาทำงาน เฉลี่ย 81.11 เดือน โดยมีระยะเวลาทำงานต่ำสุด 63 เดือน และสูงสุด 98 เดือน และมีประสบการณ์ก่อนทำงาน เฉลี่ย 95.86 เดือน โดยมีประสบการณ์ก่อนทำงานต่ำสุด 0 เดือน และสูงสุด 476 เดือน

เมื่อพิจารณาจากค่าสัมประสิทธิ์การกระจาย (coefficient of variation: CV) แสดงให้เห็น กลุ่มตัวอย่างมีเงินเดือนปัจจุบันแตกต่างกันน้อย (C.V. = 49.61) มีเงินเดือนเริ่มต้นแตกต่างกันน้อย (C.V. = 46.25) มีระยะเวลาทำงานแตกต่างกันน้อยมาก (C.V. = 12.40) และมีประสบการณ์ก่อนทำงานแตกต่างกันมาก (C.V. = 100.10)

#2

จากตารางที่ (#) แสดงให้เห็นว่า พนักงานของหน่วยงานนี้มีเงินเดือนเริ่มต้น เงินเดือนปัจจุบัน และระยะเวลาทำงานแตกต่างกันอยู่ในระดับน้อย แต่มีประสบการณ์ก่อนทำงานแตกต่างกันอยู่ในระดับมาก

2.2 การวิเคราะห์ข้อมูลด้วยค่าเฉลี่ย (ตัวแปรดัชนี)

การวิเคราะห์ตัวแปรดัชนี คือ การวิเคราะห์ข้อคำตอบหรือดัชนีของตัวแปรเชิงโครงสร้าง (construct variables) หรือตัวแปรแฝง (latent variables) เช่น คุณภาพการให้บริการ ความสุขในการทำงาน ความผูกพันต่อองค์การ

ตัวอย่างคำถามการวิจัย
จากการสำรวจการให้บริการผู้มาใช้บริการต้องการทราบว่า ผู้ให้บริการให้คะแนนการให้บริการรายข้ออยู่ในระดับใด

2.2.1 ตัวแปร

ตัวแปรเชิงโครงสร้างหรือตัวแปรแฝง (คุณภาพการให้บริการ)

ดัชนี	ประเภทคำถาม	การให้คะแนน
ความพึงพอใจกับการให้บริการ (V1)	+	ปกติ (5= มาก <--> 1 = น้อย)
ความคุ้มค่าของเงินกับการให้บริการ (V2)	+	ปกติ (5= มาก <--> 1 = น้อย)
ความล่าช้าของพนักงานในการให้บริการ (V3)	-	กลับค่า (5= น้อย <--> 1 = มาก)
ความมีประสิทธิภาพในการให้บริการ (V4)	+	ปกติ (5= มาก <--> 1 = น้อย)
ความความดังของเสียงรบกวนภายในห้องพัก (V5)	-	กลับค่า (5= น้อย <--> 1 = มาก)

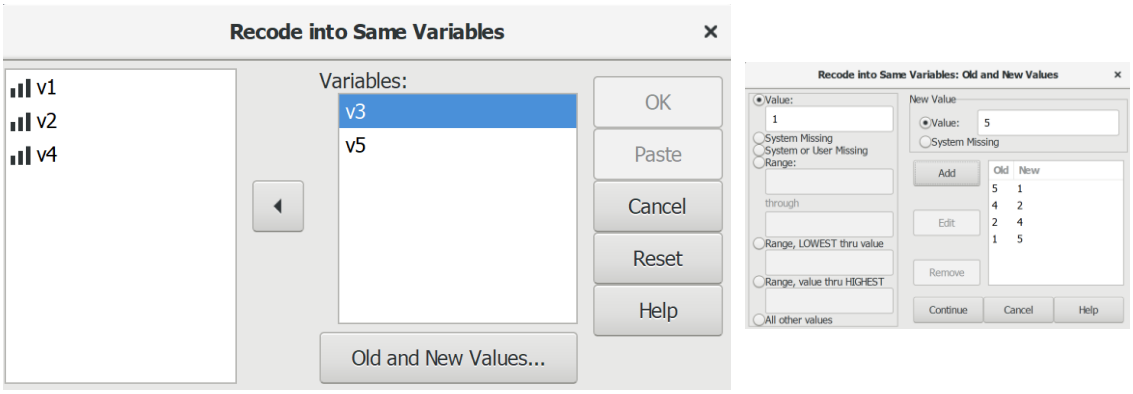
ผลการวิเคราะห์ก่อนกลับค่าของตัวแปร

Descriptive Statistics

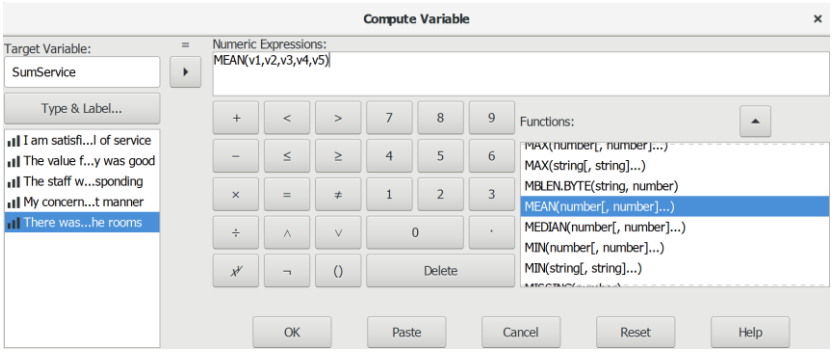
	N	Mean	Std Dev
I am satisfied with the level of service	17	2.53	1.33
The value for money was good	17	2.29	1.31
The staff were slow in responding	17	3.47	1.37
My concerns were dealt with in an efficient manner	17	2.41	1.12
There was too much noise in the rooms	17	2.00	1.17
Score of Services	17	2.54	.48
Valid N (listwise)	17		
Missing N (listwise)	0		

2.2.2 การใช้โปรแกรมวิเคราะห์ข้อมูล

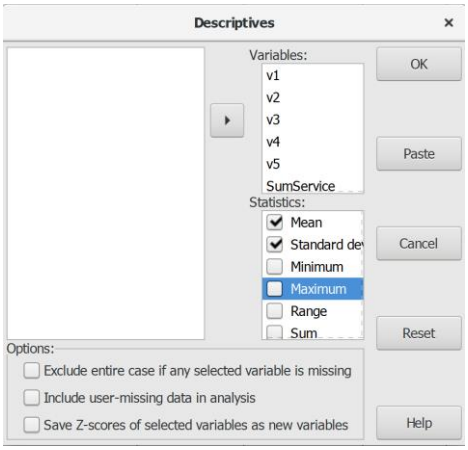
- เปิดแฟ้มข้อมูล **File -> Open** -> hotel.sav
- กลับค่าของตัวแปร V3 และ V5
  - เมนู **Transform -> Recode in Same Variables**
  - **Variables:** เลือกตัวแปร V3 และ V5 จากรายการตัวแปรด้านซ้าย
  - กดปุ่ม **Old and New Values...** -> เปลี่ยนค่า 5->1, 4 -> 2, 2 -> 4 และ 1 -> 5
  - กดปุ่ม **Continue**
  - กดปุ่ม **OK**



- สร้างตัวแปร SumService (รวมคะแนนประเมินการให้บริการด้วยค่าเฉลี่ย)



- เลือกเมนู **Analyze -> Descriptive Statistics -> Descriptives...**
  - **Variable (s):** เลือกตัวแปร V1, V2, V3, V4, V5 และ SumService จากรายการตัวแปรด้านซ้าย
  - ในกลุ่มคำสั่ง **Statistics:** เลือกสถิติ Mean และ Standard deviation
  - กดปุ่ม **OK**



2.2.3 ผลการวิเคราะห์

Descriptive Statistics

	N	Mean	Std Dev
I am satisfied with the level of service	17	2.53	1.33
The value for money was good	17	2.29	1.31
The staff were slow in responding	17	2.53	1.37
My concerns were dealt with in an efficient manner	17	2.41	1.12
There was too much noise in the rooms	17	4.00	1.17
SumService	17	2.75	.59
Valid N (listwise)	17		
Missing N (listwise)	0		

2.2.4 การสร้างตารางผลการวิเคราะห์

ตารางที่ (#) แสดงค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐาน และค่าสัมประสิทธิ์การกระจายคะแนนการให้บริการ

การประเมินของผู้ใช้บริการ	จำนวน	ค่าเฉลี่ย	S.D.	C.V. (%)
ความพึงพอใจในการให้บริการ	17	2.53	1.33	58.08
ความคุ้มค่าของเงินกับการให้บริการ	17	2.29	1.31	57.21
ความรวดเร็วของพนักงานในการให้บริการ	17	2.53	1.37	54.15
ความมีประสิทธิภาพในการให้บริการ	17	2.41	1.12	46.47
ความเงียบของเสียงภายในห้องพัก	17	4.00	1.17	29.25
รวม	17	2.75	0.59	21.45

2.2.5 การแปลผลและสรุป

จากตารางที่ (#) แสดงให้เห็นว่า กลุ่มตัวอย่างให้คะแนนเฉลี่ยรวมการให้บริการอยู่ที่ 2.75 จากคะแนนเต็ม 5 คะแนน โดยให้คะแนนด้านความเงียบของเสียงภายในห้องพักมากที่สุด รองลงมา คือ ความพึงพอใจในการให้บริการและความรวดเร็วของพนักงาในการให้บริการ ความมีประสิทธิภาพในการให้บริการ และน้อยที่สุด คือ ความคุ้มค่าของเงินกับการให้บริการ

เมื่อพิจารณาจากค่าสัมประสิทธิ์การกระจาย (C.V.) แสดงให้เห็นว่า กลุ่มตัวอย่างให้คะแนนการประเมินโดยรวมและรายข้อทุกข้อแตกต่างกันอยู่ในระดับน้อย

สรุป

การวิเคราะห์ข้อมูล 1 ตัวแปร เป็นการแจกแจงให้เห็นลักษณะของข้อมูลของแต่ละตัวแปร ในการวิเคราะห์ต้องเลือกใช้สถิติให้สอดคล้องกับลักษณะของข้อมูลของแต่ละตัวแปร คือ ตัวแปรแบบค่าแบ่งกลุ่ม ใช้ค่าความถี่ ค่าร้อยละ และค่าฐานนิยม ส่วนตัวแปรแบบค่าต่อเนื่อง ใช้ค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐาน และค่าสัมประสิทธิ์การกระจาย

แบบฝึกหัดท้ายบท

ให้ผู้เรียนเปิดแฟ้มข้อมูล 1991 US. General Social Survey.sav แล้วประมวลผลข้อมูล เพื่อวิเคราะห์และอธิบายผล ดังนี้

1. วิเคราะห์และอธิบายผลตัวแปรภูมิภาค (region) อายุ (age) และการศึกษา (educ) ด้วยคำสั่ง Frequencies หรือ Descriptive

ตารางที่ (#) จำนวนและค่าร้อยละของ..... จำแนกตาม.....

.....	จำนวน	ร้อยละ
รวม		

ตารางที่ (#) แสดงค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐาน และค่าสัมประสิทธิ์การกระจายของ.....

ตัวแปร	จำนวน	ค่าต่ำสุด	ค่าสูงสุด	ค่าเฉลี่ย	S.D.	CV[17] (%)

2. วิเคราะห์และอธิบายผลตัวแปรพี่น้อง (sibs) และบุตร (childs) ด้วยคำสั่ง Explore แล้วอธิบายลักษณะข้อมูลด้วยค่าเฉลี่ย (mean) ค่ามัธยฐาน (median) ค่าต่ำสุด (minimum) ค่าสูงสุด (maximum) ค่าพิสัย (range) ค่าความเบ้ (skewness) ค่าความโด่ง (kurtosis) และค่าผิดปกติ (outlier)

ตารางที่ (#) แสดง.....

ตัวแปร	จำนวน	ค่าต่ำสุด	ค่าสูงสุด	ค่าเฉลี่ย	ค่ามัธยฐาน	ค่าพิสัย	ค่าความเบ้	ค่าความโด่ง	ค่าผิดปกติ

[17] สัมประสิทธิ์การกระจาย (Coefficient of Variation)

บทที่ 11

การวิเคราะห์ข้อมูลเชิงพรรณนา 2 ตัวแปร

ความมุ่งหมายของบทเรียน

ผู้เรียนสามารถใช้โปรแกรมประมวลผลการเปรียบเทียบข้อมูลของตัวแปร 2 ตัวแปร ด้วยสถิติเชิงพรรณนา และเขียนบรรยายผลการวิเคราะห์ได้

เนื้อหา

1. การวิเคราะห์ข้อมูลระหว่างตัวแปรแบบค่าแบ่งประเภท
2. การวิเคราะห์ข้อมูลระหว่างตัวแปรแบบค่าแบ่งประเภทกับตัวแปรแบบค่าต่อเนื่อง

กิจกรรมและวิธีสอน

บรรยาย อธิบาย แสดงตัวอย่าง และทำแบบฝึกหัดท้ายบท

อุปกรณ์การสอน

เอกสารคำสอน ภาพเลื่อน คอมพิวเตอร์ โปรแกรม PSPP โปรแกรมสอนออนไลน์ (สอนแบบออนไลน์) และเครื่องฉาย (สอนในชั้นเรียน)

งานที่มอบหมาย

ทำแบบฝึกหัดท้ายบท

การวัดผลและประเมินผล

การเข้าชั้นเรียน สังเกตความสนใจจากการเรียน สอบถามความเข้าใจจากการบรรยาย และประเมินจากการทำแบบฝึกหัดท้ายบท

เนื้อหา

การวิเคราะห์ข้อมูล 2 ตัวแปร (bivariate analysis) คือ วิธีการนำเอาตัวแปร 2 ตัวแปร มาไขว้หรือเปรียบเทียบกันเพื่อหารูปแบบของความสัมพันธ์ระหว่างตัวแปรด้วยสถิติเชิงพรรณนา เช่น ค่าความถี่ (frequency) ค่าร้อยละ (percentage) ค่าเฉลี่ย (mean) ค่าเบี่ยงเบนมาตรฐาน (standard deviation) และค่าสัมประสิทธิ์การกระจาย (coefficient of variation)

1. การวิเคราะห์ข้อมูลระหว่างตัวแปรแบบค่าจัดประเภท

การวิเคราะห์ข้อมูลระหว่างตัวแปรแบบค่าจัดประเภท เป็นการการนำเสนอข้อมูลในรูปของตารางเพื่อสรุปข้อมูลที่เกิดจากความสัมพันธ์ของตัวแปร 2 ตัวแปร (2x2) หรือมากกว่า (NxN)

1.1 การวิเคราะห์แบบตารางไขว้ 2 ตัวแปร

การวิเคราะห์แบบตารางไขว้ 2 ตัวแปร เป็นการวิเคราะห์ข้อมูลระหว่างตัวแปรแบบค่าจัดประเภท 2 ตัวแปร เพื่อสรุปข้อมูลและเปรียบเทียบความสัมพันธ์ระหว่าง 2 ตัวแปร โดยใช้สถิติเชิงพรรณนา ได้แก่ ค่าความถี่และค่าร้อยละ

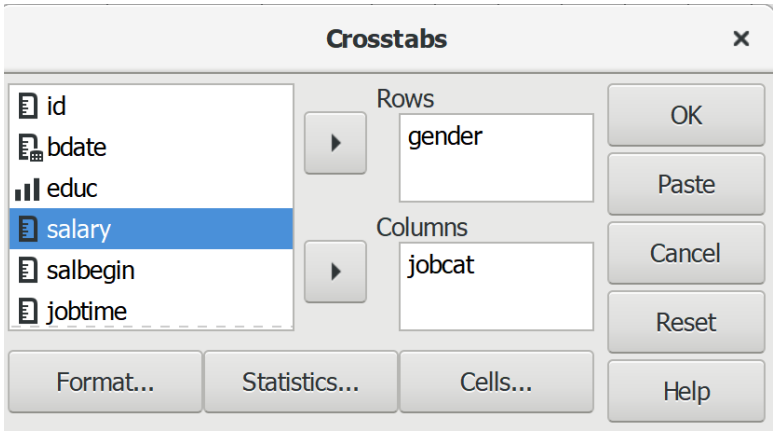
ตัวอย่างคำถามการวิจัย
จากการสำรวจข้อมูลของพนักงานในหน่วยงานต้องการทราบว่า พนักงานแต่ละเพศมีตำแหน่งงานแตกต่างกันมากน้อยเพียงใด

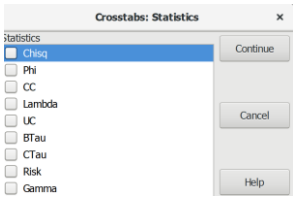
1.1.1 ตัวแปร

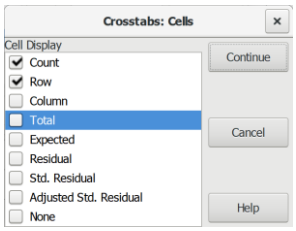
ตัวแปรอิสระ		**ตัวแปรตาม**
ค่าจัดกลุ่ม		ค่าจัดกลุ่ม
เพศ (gender) - ชาย - หญิง	->	ตำแหน่งงาน (jobcat) - เจ้าหน้าที่ปฏิบัติงาน - หัวหน้างาน - ผู้จัดการ

1.1.2 การใช้โปรแกรมวิเคราะห์ข้อมูล

- เปิดแฟ้มข้อมูล Employee Data.sav
- เลือกเมนู **Analyze -> Descriptive Statistics -> Crosstabs ...**
  - **Rows:** เลือกตัวแปรแบบค่าจัดประเภท 1 ตัวแปร หรือหลายตัวแปร (ตัวแปรที่ต้องการให้เป็นตัวแปรอิสระ) จากรายการตัวแปรด้านซ้าย
  - **Column:** เลือกตัวแปรแบบค่าจัดประเภท 1 ตัวแปร หรือหลายตัวแปร (ตัวแปรที่ต้องการให้เป็นตัวแปรตาม) จากรายการตัวแปรด้านซ้าย
  - กดปุ่ม **Statistics...** ในกลุ่มคำสั่ง **Statistics** ไม่เลือกสถิติ Chisq -> กดปุ่ม **Continue**
  - กดปุ่ม **Cells...** ในกลุ่มคำสั่ง **Cell Display** เลือกค่าร้อยละเฉพาะ **Count** และ **Row** -> กดปุ่ม **Continue**
  - กดปุ่ม **OK**







1.1.3 ผลการวิเคราะห์

Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Gender × Employment Category	474	100.0%	0	.0%	474	100.0%

Gender × Employment Category

			Employment Category			
			Clerical	Custodial	Manager	Total
Gender	Female	Count	206	0	10	216
		Row %	95.4%	.0%	4.6%	100.0%
	Male	Count	157	27	74	258
		Row %	60.9%	10.5%	28.7%	100.0%
Total		Count	363	27	84	474
		Row %	76.6%	5.7%	17.7%	100.0%

1.1.4 การสร้างตารางแสดงผลการวิเคราะห์

ตารางที่ (#) แสดงค่าร้อยละของพนักงาน จำแนกตามเพศและตำแหน่งงาน

เพศ	ตำแหน่งงาน			รวม
	เจ้าหน้าที่ปฏิบัติงาน	หัวหน้างาน	ผู้จัดการ	
หญิง	95.4	0.0	4.6	100.0 (216)
ชาย	60.9	10.5	28.7	100.0 (258)
รวม	76.6	5.7	17.7	100.0 (474)

1.15 การแปลผลและสรุป

จากตารางที่ (#) แสดงให้เห็นว่า กลุ่มตัวอย่างทั้งเพศหญิงและเพศชายส่วนใหญ่เป็นเจ้าหน้าที่ปฏิบัติงาน เพศหญิงเป็นผู้จัดการน้อยที่สุด ส่วนเพศชายเป็นหัวหน้างานน้อยที่สุด

เมื่อเปรียบเทียบการดำรงตำแหน่งงานของเพศหญิงและเพศชาย แสดงให้เห็นว่า เพศหญิงเป็นเจ้าหน้าที่ปฏิบัติงานมากกว่าเพศชาย เพศชายเป็นหัวหน้างานและผู้จัดการมากกว่าเพศหญิง

สรุป เพศชายมีตำแหน่งงานสูงกว่าเพศหญิง

ตัวอย่างเพิ่มเติม: การสร้างตารางแสดงผลการวิเคราะห์

#1

เพศ	ตำแหน่งงาน			รวม
	เจ้าหน้าที่ปฏิบัติงาน	หัวหน้างาน	ผู้จัดการ	
หญิง	205 (95.4)	0 (0.0)	10 (4.6)	216 (100.0)
ชาย	157 (60.9)	27 (10.5)	74 (28.7)	258 (100.0)
รวม	363 (76.6)	27 (5.7)	84 (17.7)	474 (100.0)

#2

เพศ	ตำแหน่งงาน			รวม
	เจ้าหน้าที่ปฏิบัติงาน	หัวหน้างาน	ผู้จัดการ	
หญิง	95.4 (205)	0.0 (0)	4.6 (10)	100.0 (216)
ชาย	60.9 (157)	10.5 (27)	28.7 (74)	100.0 (258)
รวม	76.6 (363)	5.7 (27)	17.7 (84)	100.0 (474)

ตัวอย่างเพิ่มเติม: การแปลผลและสรุป

#1

จากตารางที่ (#) แสดงให้เห็นว่า กลุ่มตัวอย่างที่เป็นพนักงานทั้งหมดจำนวน 474 คน แบ่งเป็นเพศหญิง จำนวน 216 คน คิดเป็นจำนวนร้อยละ 100.0 เป็นเจ้าหน้าที่ปฏิบัติงาน จำนวน 205 คน คิดเป็นจำนวนร้อยละ 95.4 เป็นหัวหน้างาน จำนวน 0 คน คิดเป็นจำนวนร้อยละ 0.0 และเป็นผู้จัดการ จำนวน 10 คน คิดเป็นจำนวนร้อยละ 4.6 และแบ่งเป็นเพศชาย จำนวน 258 คน คิดเป็นจำนวนร้อยละ 100.0 เป็นเจ้าหน้าที่ปฏิบัติงาน จำนวน 157 คน คิดเป็นจำนวนร้อยละ 60.9 เป็นหัวหน้างาน จำนวน 27 คน คิดเป็นจำนวนร้อยละ 10.5 และเป็นผู้จัดการ จำนวน 74 คน คิดเป็นจำนวนร้อยละ 28.7

#2

จากตารางที่ (#) แสดงให้เห็นว่า กลุ่มตัวอย่างที่เป็นพนักงานทั้งหมด จำนวน 474 คน แบ่งเป็นเพศหญิง จำนวน 216 คน และเพศชาย จำนวน 258 คน เพศหญิงส่วนใหญ่เป็นเจ้าหน้าที่ปฏิบัติงาน จำนวนร้อยละ 95.4 รองลงมา คือ ผู้จัดการ จำนวนร้อยละ 4.6 และไม่มีเพศหญิงในตำแหน่งหัวหน้างาน

ส่วนกลุ่มตัวอย่างเพศชายส่วนใหญ่เป็นเจ้าหน้าที่ปฏิบัติงาน จำนวนร้อยละ 60.9 รองลงมา คือ ผู้จัดการ จำนวนร้อยละ 28.7 และน้อยที่สุด คือ หัวหน้างาน จำนวนร้อยละ 10.5

เมื่อเปรียบเทียบการดำรงตำแหน่งงานระหว่างเพศหญิงและเพศชาย แสดงให้เห็นว่า เพศหญิงเป็นเจ้าหน้าที่ปฏิบัติงานมากกว่าเพศชาย เพศชายเป็นหัวหน้างานและผู้จัดการมากกว่าเพศหญิง

โดยสรุปแสดงให้เห็นว่า เพศชายมีตำแหน่งงานสูงกว่าเพศหญิง

#3

จากตารางที่ (#) แสดงให้เห็นว่า กลุ่มตัวอย่างที่เป็นเพศหญิงส่วนใหญ่เป็นเจ้าหน้าที่ปฏิบัติงาน รองลงมา คือ ผู้จัดการ และไม่มีเพศหญิงในตำแหน่งหัวหน้างาน ส่วนกลุ่มตัวอย่างที่เป็นเพศชายส่วนใหญ่เป็นเจ้าหน้าที่ปฏิบัติงาน รองลงมา คือ ผู้จัดการ และน้อยที่สุด คือ หัวหน้างาน

เมื่อเปรียบเทียบการดำรงตำแหน่งงานระหว่างเพศหญิงและเพศชาย แสดงให้เห็นว่า เพศหญิงเป็นเจ้าหน้าที่ปฏิบัติงานมากกว่าเพศชาย เพศชายเป็นหัวหน้างานและผู้จัดการมากกว่าเพศหญิง

สรุป เพศชายมีตำแหน่งงานสูงกว่าเพศหญิง

1.2 การวิเคราะห์แบบตารางไขว้ 2 ตัวแปร โดยมีตัวแปรคุม

การวิเคราะห์แบบตารางไขว้ 2 ตัวแปร โดยมีตัวแปรคุม เป็นการวิเคราะห์ข้อมูลระหว่างตัวแปรแบบค่าจัดกลุ่ม 2 ตัวแปรเหมือนกันแต่มีตัวแปรคุมเป็นตัวแปรค่าจัดกลุ่ม เพื่อสรุปข้อมูลและเปรียบเทียบความสัมพันธ์ระหว่างตัวแปรมากกว่า 2 ตัวแปร ในรูปของตาราง 2x3 หรือ 3x2 หรือมากกว่า (NxN) ด้วยสถิติเชิงพรรณนา ได้แก่ ค่าความถี่และค่าร้อยละ

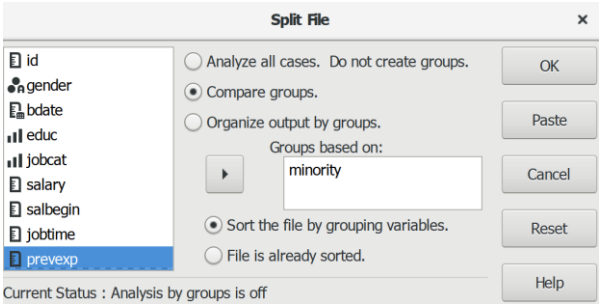
ตัวอย่างคำถามการวิจัย
จากการสำรวจข้อมูลของพนักงานในหน่วยงานต้องการทราบว่า พนักงานแต่ละเพศในแต่ละกลุ่มชนมีตำแหน่งงานแตกต่างกันมากน้อยเพียงใด

1.2.1 ตัวแปร

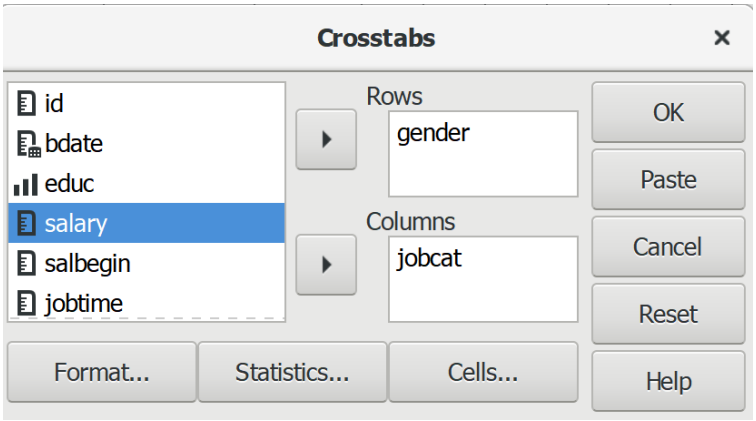
ตัวแปรอิสระ		**ตัวแปรตาม**
ค่าจัดกลุ่ม	**ตัวแปรคุม**	ค่าจัดกลุ่ม
เพศ (gender) - ชาย - หญิง	ค่าจัดกลุ่ม กลุ่มชน (minority) ->	ตำแหน่งงาน (jobcat) - เจ้าหน้าที่ปฏิบัติงาน - หัวหน้างาน - ผู้จัดการ

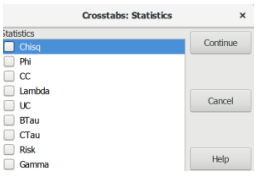
1.2.2 การใช้โปรแกรมวิเคราะห์ข้อมูล

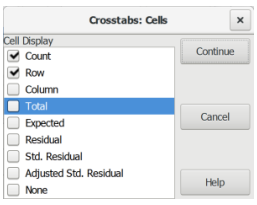
- เปิดแฟ้มข้อมูล Employee Data.sav
- เลือกเมนู **Data -> Split File...**
  - เลือกคำสั่ง **Compare groups.**
  - ใน **Group base on** เลือกตัวแปรแบบค่าจัดประเภทจากรายการเมนูด้านซ้าย 1 ตัวแปร หรือมากกว่า
  - กดปุ่ม **OK**



- เมนู **Analyze -> Descriptive Statistics -> Crosstabs ...**
  - **Rows:** เลือกตัวแปรแบบค่าจัดประเภท 1 ตัวแปร หรือหลายตัวแปร (ตัวแปรที่ต้องการให้เป็นตัวแปรอิสระ) จากรายการตัวแปรด้านซ้าย
  - **Column:** เลือกตัวแปรแบบค่าจัดประเภท 1 ตัวแปร หรือหลายตัวแปร (ตัวแปรที่ต้องการให้เป็นตัวแปรตาม) จากรายการตัวแปรด้านซ้าย
  - กดปุ่ม **Statistics...** ในกลุ่มคำสั่ง **Statistics** ไม่เลือกสถิติ Chisq -> กดปุ่ม **Continue**
  - กดปุ่ม **Cells...** ในกลุ่มคำสั่ง **Cell Display** เลือกค่าร้อยละเฉพาะ **Count** และ **Row** -> กดปุ่ม **Continue**
  - กดปุ่ม **OK**







1.2.3 ผลการวิเคราะห์

Split Values

Variable	Value
Minority Classification	No

Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Gender × Employment Category	370	100.0%	0	.0%	370	100.0%

Gender × Employment Category

			Employment Category			
			Clerical	Custodial	Manager	Total
Gender	Female	Count	166	0	10	176
		Row %	94.3%	.0%	5.7%	100.0%
	Male	Count	110	14	70	194
		Row %	56.7%	7.2%	36.1%	100.0%
Total		Count	276	14	80	370
		Row %	74.6%	3.8%	21.6%	100.0%

Split Values

Variable	Value
Minority Classification	Yes

Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Gender × Employment Category	104	100.0%	0	.0%	104	100.0%

Gender × Employment Category

			Employment Category			
			Clerical	Custodial	Manager	Total
Gender	Female	Count	40	0	0	40
		Row %	100.0%	.0%	.0%	100.0%
	Male	Count	47	13	4	64
		Row %	73.4%	20.3%	6.3%	100.0%
Total		Count	87	13	4	104
		Row %	83.7%	12.5%	3.8%	100.0%

1.1.4 การสร้างตารางผลการวิเคราะห์

ตารางที่ (#) แสดงค่าร้อยละของตำแหน่งงาน จำแนกตามกลุ่มชนและเพศ

กลุ่มชน	เพศ	ตำแหน่งงาน			รวม
		เจ้าหน้าที่ปฏิบัติงาน	หัวหน้างาน	ผู้จัดการ	
ชนกลุ่มหลัก	หญิง	94.3	0.0	5.7	100.0 (176)
	ชาย	56.7	7.2	36.1	100.0 (194)
	รวม	74.6	3.8	21.6	100.0 (370)
ชนกลุ่มน้อย	หญิง	100.0	0.0	0.0	100.0 (40)
	ชาย	73.4	20.3	6.3	100.0 (64)
	รวม	83.7	12.5	3.8	100.0 (104)

1.1.5 การแปลผลและสรุป

จากตารางที่ (#) สามารถอธิบายตำแหน่งงานของกลุ่มตัวอย่างจำแนกตามกลุ่มชนและเพศ ได้ดังนี้

1. เปรียบเทียบการดำรงตำแหน่งของชนกลุ่มหลักกับชนกลุ่มน้อย แสดงให้เห็นว่า ชนกลุ่มหลักที่เป็นเพศหญิง ดำรงตำแหน่งงานเป็นเจ้าหน้าที่ปฏิบัติงานน้อยกว่าชนกลุ่มน้อยที่เป็นเพศหญิง แต่ชนกลุ่มหลักที่เป็นเพศหญิง ดำรงตำแหน่งงานเป็นผู้จัดการมากกว่าชนกลุ่มน้อยที่เป็นเพศหญิง

2. เปรียบเทียบการดำรงตำแหน่งของชนกลุ่มหลักกับชนกลุ่มน้อย แสดงให้เห็นว่า ชนกลุ่มหลักที่เป็นเพศชาย ดำรงตำแหน่งงานเป็นเจ้าหน้าที่ปฏิบัติงานและหัวหน้างานน้อยกว่าชนกลุ่มน้อยที่เป็นเพศชาย แต่ชนกลุ่มหลักที่เป็นเพศชาย ดำรงตำแหน่งงานเป็นผู้จัดการมากกว่าชนกลุ่มน้อยที่เป็นเพศชาย

สรุป ชนกลุ่มหลักทั้งเพศชายและหญิงมีตำแหน่งงานสูงกว่าชนกลุ่มน้อยทั้งเพศชายและหญิง

ตัวอย่างเพิ่มเติม: การแปลผลและสรุป

#1

จากตารางที่ (#) สามารถอธิบายตำแหน่งงานของกลุ่มตัวอย่างจำแนกตามกลุ่มชนและเพศ ได้ดังนี้

1. กลุ่มตัวอย่างที่เป็นชนกลุ่มหลักจำนวน 370 คน เป็นพนักงานเพศหญิง จำนวน 179 คน คิดเป็นร้อยละ 100.0 ดำรงตำแหน่งเป็นเจ้าหน้าที่ปฏิบัติงาน ร้อยละ 94.3 เป็นหัวหน้างาน ร้อยละ 0.0 และเป็นผู้จัดการร้อยละ 5.7 ส่วนพนักงานเพศชาย จำนวน 194 คน คิดเป็นร้อยละ 100.0 ดำรงตำแหน่งเป็นเจ้าหน้าที่ปฏิบัติงาน ร้อยละ 56.7 เป็นหัวหน้างานร้อยละ 7.2 และเป็นผู้จัดการร้อยละ 36.1

2. กลุ่มตัวอย่างที่เป็นชนกลุ่มน้อยจำนวน 104 คน เป็นพนักงานเพศหญิง จำนวน 40 คน คิดเป็นจำนวนร้อยละ 100.0 ดำรงตำแหน่งเป็นเจ้าหน้าที่ปฏิบัติงานร้อยละ 100.0 ดำรงตำแหน่งหัวหน้างานร้อยละ 0.0 และดำรงตำแหน่งเป็นผู้จัดการร้อยละ 0.0 ส่วนพนักงานเพศชาย จำนวน 64 คน คิดเป็นจำนวนร้อยละ 100.0 ดำรงตำแหน่งเป็นเจ้าหน้าที่ปฏิบัติงาน ร้อยละ 73.4 เป็นหัวหน้างานร้อยละ 20.3 และเป็นผู้จัดการร้อยละ 6.3

สรุป กลุ่มตัวอย่างเพศหญิงทั้งชนกลุ่มหลักและชนกลุ่มน้อยส่วนใหญ่มีตำแหน่งงานต่ำกว่าพนักงานที่เป็นเพศชายทั้งชนกลุ่มหลักและชนกลุ่มน้อย

#2

จากตารางที่ (#) สามารถอธิบายตำแหน่งงานของกลุ่มตัวอย่างจำแนกตามกลุ่มชนและเพศ ได้ดังนี้

1. กลุ่มตัวอย่างที่เป็นชนกลุ่มหลักและเป็นเพศหญิง จำนวน 176 คน คิดเป็นจำนวนร้อยละ 100.0 ดำรงตำแหน่งเป็นเจ้าหน้าที่ปฏิบัติงาน ร้อยละ 94.3 เป็นหัวหน้างาน ร้อยละ 0.0 และเป็นผู้จัดการ ร้อยละ 5.7 ส่วนพนักงานที่เป็นชนกลุ่มน้อยและเป็นเพศหญิง จำนวน 40 คน คิดเป็นร้อยละ 100.0 ดำรงตำแหน่งเป็นเจ้าหน้าที่ปฏิบัติงาน ร้อยละ 100.0 เป็นหัวหน้างาน ร้อยละ 0.0 และเป็นผู้จัดการ ร้อยละ 0.0

2. กลุ่มตัวอย่างที่เป็นชนกลุ่มหลักและเป็นเพศชาย จำนวน 194 คน คิดเป็นร้อยละ 100.0 ดำรงตำแหน่งเป็นเจ้าหน้าที่ปฏิบัติงาน ร้อยละ 56.7 เป็นหัวหน้างาน ร้อยละ 7.2 และเป็นผู้จัดการ ร้อยละ

36.1 ส่วนพนักงานที่เป็นชนกลุ่มน้อยและเป็นเพศชาย 64 คน คิดเป็นร้อยละ 100.0 ดำรงตำแหน่งเป็นเจ้าหน้าที่ปฏิบัติงาน ร้อยละ 73.4 เป็นหัวหน้างาน ร้อยละ 20.3 และเป็นผู้จัดการร้อยละ 6.3

สรุป กลุ่มชนมีอิทธิพลต่อการดำรงตำแหน่งของพนักงาน กล่าวคือ พนักงานที่เป็นชนกลุ่มหลักมีแนวโน้มที่จะดำรงตำแหน่งงานในระดับบังคับบัญชามากกว่าพนักงานที่เป็นชนกลุ่มน้อยทั้งเพศชายและเพศหญิง

#3

จากตารางที่ (#) เปรียบเทียบการดำรงตำแหน่งของกลุ่มตัวอย่างที่เป็นชนกลุ่มหลักและชนกลุ่มน้อยทั้งเพศชายและเพศหญิง แสดงให้เห็นว่า ชนกลุ่มหลักทั้งเพศชายและเพศหญิงดำรงตำแหน่งเป็นเจ้าหน้าที่ปฏิบัติงานน้อยกว่าชนกลุ่มน้อยทั้งเพศชายและเพศหญิง ชนกลุ่มหลักและชนกลุ่มน้อยที่เป็นเพศหญิงไม่มีตำแหน่งเป็นหัวหน้างานเลย ส่วนชนกลุ่มหลักที่เป็นเพศชายดำรงตำแหน่งหัวหน้างานน้อยกว่าชนกลุ่มน้อยที่เป็นเพศชาย แต่ชนกลุ่มหลักที่เป็นเพศชายดำรงตำแหน่งเป็นผู้จัดการมากกว่าชนกลุ่มน้อยที่เป็นเพศชาย

สรุป ชนกลุ่มหลักทั้งเพศชายและเพศหญิงดำรงตำแหน่งงานเป็นผู้บริหารระดับสูงมากกว่าชนกลุ่มน้อยทั้งเพศชายและเพศหญิง ส่วนชนกลุ่มน้อยทั้งเพศชายและเพศหญิงดำรงตำแหน่งเป็นเจ้าหน้าที่ปฏิบัติงานและหัวหน้างานมากกว่าชนกลุ่มน้อยทั้งเพศชายและเพศหญิง

---

2. การวิเคราะห์ข้อมูลระหว่างตัวแปรแบบค่าแบ่งประเภทกับตัวแปรแบบค่าต่อเนื่อง

การวิเคราะห์ข้อมูลระหว่างตัวแปรแบบค่าแบ่งประเภท (ตัวแปรเชิงคุณภาพ) กับตัวแปรแบบค่าต่อเนื่อง (ตัวแปรเชิงปริมาณ) เป็นการเปรียบเทียบข้อมูลของตัวแปรแบบค่าแบ่งประเภทด้วยค่าเฉลี่ย (compare means) การวิเคราะห์แบบนี้มีข้อมูลมีตั้งแต่ 2 กลุ่ม และมากกว่า 2 กลุ่ม อาจเรียกว่า การวิเคราะห์เปรียบเทียบค่าเฉลี่ย 2 กลุ่ม และ การวิเคราะห์เปรียบเทียบค่าเฉลี่ยมากกว่า 2 กลุ่ม โดยมีรายละเอียดดังต่อไปนี้

2.1 การเปรียบเทียบค่าเฉลี่ย 2 กลุ่ม

การเปรียบเทียบค่าเฉลี่ย 2 กลุ่ม เป็นวิธีการวิเคราะห์แบบไขว้ตัวแปรแบบหนึ่ง แต่ตัวแปรอิสระเป็นตัวแปรแบบค่าจัดกลุ่มส่วนตัวแปรตามเป็นตัวแปรแบบค่าต่อเนื่อง โดยใช้สถิติเชิงพรรณนา ได้แก่ ค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐาน และค่าสัมประสิทธิ์การกระจาย

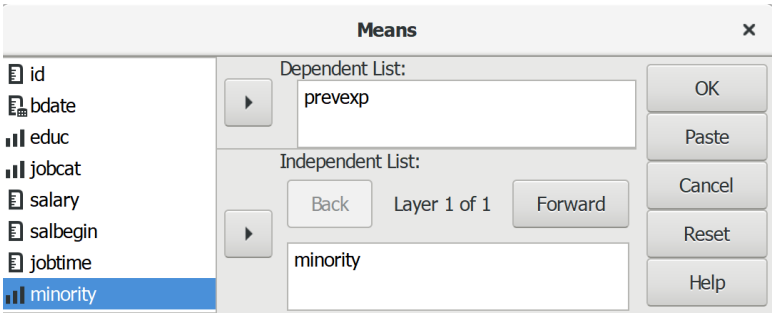
ตัวอย่างคำถามการวิจัย
จากการสำรวจข้อมูลของพนักงานในหน่วยงานต้องการทราบว่า พนักงานแต่ละกลุ่มชนมีประสบการณ์ก่อนทำงานแตกต่างกันมากน้อยเพียงใด

2.1.1 ตัวแปร

ตัวแปรอิสระ		ตัวแปรตาม
ค่าจัดกลุ่ม		ค่าต่อเนื่อง
กลุ่มชน (minority) - ชนกลุ่มหลัก - ชนกลุ่มน้อย	->	ประสบการณ์ก่อนทำงาน (prevexp)

2.1.2 การใช้โปรแกรมวิเคราะห์ข้อมูล

- เปิดแฟ้มข้อมูล **File -> Open** -> Employee data.sav
- เลือกเมนู **Analyze -> Compare Means -> Means...**
  - **Dependent List:** เลือกตัวแปรแบบค่าต่อเนื่อง 1 ตัวแปร (ตัวแปรที่ต้องการให้เป็นตัวแปรตาม) จากรายการตัวแปรด้านซ้าย
  - **Independent List:** เลือกตัวแปรแบบค่าจัดประเภท (2 กลุ่ม) 1 ตัวแปร หรือหลายตัวแปร (ตัวแปรที่ต้องการให้เป็นตัวแปรอิสระ) จากรายการตัวแปรด้านซ้าย
  - กดปุ่ม **OK**



2.1.3 ผลการวิเคราะห์

Case Processing Summary

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
prevexp * minority	474	100.0%	0	.0%	474	100.0%

Report

Minority Classification	Mean	N	Std. Deviation
No	88	370	105
Yes	124	104	100
Total	96	474	105

2.1.4 การสร้างตารางผลการวิเคราะห์

ตารางที่ (#) แสดงค่าเฉลี่ยประสบการณ์ก่อนทำงานของพนักงาน ระหว่างกลุ่มชน

ประสบการณ์ก่อนทำงาน (เดือน)	จำนวน	ค่าเฉลี่ย	S.D.	C.V. (%)
ชนกลุ่มหลัก	370	88	105	119
ชนกลุ่มน้อย	104	124	100	81
รวม	474	96	105	109

2.1.5 การแปลผลและสรุป

จากตารางที่ (#) แสดงให้เห็นว่า กลุ่มตัวอย่างที่เป็นชนกลุ่มน้อยมีประสบการณ์ก่อนทำงานโดยเฉลี่ยมากกว่าชนกลุ่มหลัก

เมื่อพิจารณาจากค่าสัมประสิทธิ์การกระจาย แสดงให้เห็นว่า ประสบการณ์ก่อนการทำงานของชนกลุ่มหลักมีความแตกต่างกันอยู่ในระดับมาก ส่วนชนกลุ่มน้อยมีความแตกต่างกันอยู่ในระดับน้อย

ตัวอย่างเพิ่มเติม: การแปลผลและสรุป

#1

จากตารางที่ (#) แสดงให้เห็นว่า กลุ่มตัวอย่างทั้งหมด จำนวน 474 คน มีประสบการณ์ก่อนทำงานต่ำสุด 0 เดือน สูงสุด 476 เดือน เฉลี่ย 96 เดือน และมีค่าสัมประสิทธิ์การกระจาย ร้อยละ 119 โดยชนกลุ่มหลัก จำนวน 370 คน มีประสบการณ์ก่อนทำงานต่ำสุด 0 เดือน สูงสุด 476 เดือน เฉลี่ย 88 เดือน และมีค่าสัมประสิทธิ์การกระจาย ร้อยละ 119 ส่วนชนกลุ่มน้อย จำนวน 104 คน มีประสบการณ์ก่อนทำงานต่ำสุด 0 เดือน สูงสุด 432 เดือน เฉลี่ย 124 เดือน และมีค่าสัมประสิทธิ์การกระจาย ร้อยละ 81

เมื่อพิจารณาจากค่าสัมประสิทธิ์การกระจาย แสดงให้เห็นว่า กลุ่มตัวอย่างทั้งหมดมีประสบการณ์ก่อนทำงานแตกต่างต่างกันมาก (C.V. = 109) โดยชนกลุ่มหลักมีประสบการณ์ก่อนการทำงานแตกต่างกันมาก (C.V. = 119) ส่วนชนกลุ่มน้อยมีประสบการณ์ก่อนการทำงานแตกต่างกันน้อย (C.V. = 81)

#2

จากตารางที่ (#) แสดงให้เห็นว่า กลุ่มตัวอย่างทั้งหมด จำนวน 474 คน มีประสบการณ์ก่อนทำงาน เฉลี่ย 96 เดือน โดยชนกลุ่มน้อยมีประสบการณ์ก่อนทำงานมากกว่าค่าเฉลี่ยรวม คือ 124 เดือน และชนกลุ่มหลักที่มีประสบการณ์ก่อนทำงานน้อยกว่าค่าเฉลี่ยรวม คือ 88 เดือน

เมื่อพิจารณาจากค่าสัมประสิทธิ์การกระจาย แสดงให้เห็นว่า กลุ่มตัวอย่างทั้งหมดมีประสบการณ์ก่อนการทำงานแตกต่างกันมาก (C.V. = 109) โดยชนกลุ่มหลักมีประสบการณ์ก่อนทำงานแตกต่างกันมากกว่า (C.V. = 119) ชนกลุ่มน้อย (C.V. = 109)

#3

จากตารางที่ (#) แสดงให้เห็นว่า กลุ่มตัวอย่างทั้งหมด จำนวน 474 คน มีประสบการณ์ก่อนทำงาน เฉลี่ย 96 เดือน โดยชนกลุ่มน้อยมีประสบการณ์ก่อนทำงาน เฉลี่ย 124 เดือน มากกว่าชนกลุ่มหลักที่มีประสบการณ์ก่อนทำงาน เฉลี่ย 88 เดือน

เมื่อพิจารณาจากค่าสัมประสิทธิ์การกระจาย แสดงให้เห็นว่า ประสบการณ์ก่อนการทำงานของกลุ่มตัวอย่างทั้งหมดมีความแตกต่างกันอยู่ในระดับมาก (C.V. = 109) และชนกลุ่มหลักมีความแตกต่างกันอยู่ในระดับมากเช่นกัน (C.V. = 119) ส่วนชนกลุ่มน้อยมีแตกต่างกันอยู่ในระดับน้อย (C.V. = 109)

2.2 การเปรียบเทียบค่าเฉลี่ยมากกว่า 2 กลุ่ม

การเปรียบเทียบค่าเฉลี่ยมากกว่า 2 กลุ่ม เป็นวิธีการวิเคราะห์แบบเดียวกับการเปรียบเทียบค่าเฉลี่ย 2 กลุ่ม แต่ตัวแปรอิสระเป็นค่าจัดกลุ่มมากกว่า 2 กลุ่มขึ้นไป และตัวแปรตามเป็นค่าต่อเนื่อง โดยใช้สถิติเชิงพรรณนา ได้แก่ ค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐาน และค่าสัมประสิทธิ์การกระจาย

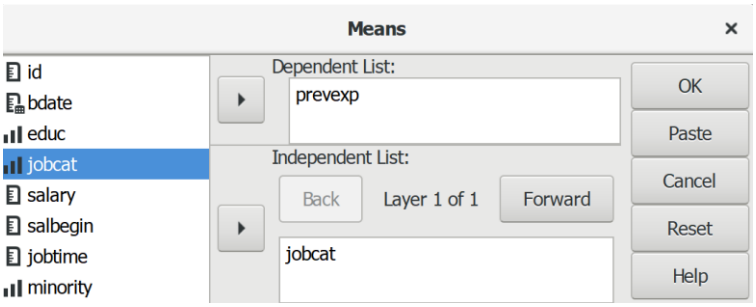
ตัวอย่างคำถามการวิจัย
จากการสำรวจข้อมูลของพนักงานในหน่วยงานต้องการทราบว่า พนักงานแต่ละตำแหน่งงานมีประสบการณ์ก่อนทำงานแตกต่างกันมากน้อยเพียงใด

2.2.1 ตัวแปร

ตัวแปรอิสระ		**ตัวแปรตาม**
ค่าจัดกลุ่ม		ค่าต่อเนื่อง
ตำแหน่งงาน (jobcat) - เจ้าหน้าที่ปฏิบัติงาน - หัวหน้างาน - ผู้จัดการ	->	ประสบการณ์ก่อนทำงาน (prevexp)

2.2.2 การใช้โปรแกรมวิเคราะห์ข้อมูล

- เปิดแฟ้มข้อมูล Employee data.sav
- เลือกเมนู **Analyze -> Compare Means -> Means...**
  - **Dependent List:** เลือกตัวแปรแบบค่าต่อเนื่อง 1 ตัวแปร (ตัวแปรที่ต้องการให้เป็นตัวแปรตาม) จากรายการตัวแปรด้านซ้าย
  - **Independent List:** เลือกตัวแปรแบบค่าจัดประเภท (มากกว่า 2 กลุ่ม) 1 ตัวแปร หรือหลายตัวแปร (ตัวแปรที่ต้องการให้เป็นตัวแปรอิสระ) จากรายการตัวแปรด้านซ้าย
  - กดปุ่ม **OK**



2.2.3 ผลการวิเคราะห์

Case Processing Summary

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
prevexp * jobcat	474	100.0%	0	.0%	474	100.0%

Report

Employment Category	Mean	N	Std. Deviation
Clerical	85	363	95
Custodial	298	27	101
Manager	78	84	73
Total	96	474	105

2.2.4 การสร้างตารางผลการวิเคราะห์

ตารางที่ (#) แสดงค่าเฉลี่ยประสบการณ์ก่อนทำงานของพนักงาน

ตำแหน่งงาน	จำนวน	ต่ำสุด	S.D.	C.V. (%)
เจ้าหน้าที่ปฏิบัติงาน	363	0	95	112
หัวหน้างาน	27	144	101	34
ผู้จัดการ	84	3	73	94
รวม	474	0	105	109

2.2.5 การแปลผลและสรุป

จากตารางที่ (#) แสดงให้เห็นว่า กลุ่มตัวอย่างที่มีประสบการณ์ก่อนทำงานโดยเฉลี่ยมากที่สุด คือ หัวหน้างาน รองลงมา คือ เจ้าหน้าที่ปฏิบัติงาน และน้อยที่สุด คือ ผู้จัดการ

เมื่อพิจารณาจากค่าสัมประสิทธิ์การกระจายแสดงให้เห็นว่า ประสบการณ์ก่อนทำงานของเจ้าหน้าที่ปฏิบัติงานมีความแตกต่างกันอยู่ในระดับมาก ส่วนหัวหน้างานและผู้จัดการมีความแตกต่างกันอยู่ในระดับน้อย

ตัวอย่างเพิ่มเติม: การแปลผลและสรุป

#1

จากตารางที่ (#) แสดงให้เห็นว่า กลุ่มตัวอย่างทั้งหมด 474 คน มีประสบการณ์ก่อนทำงาน เฉลี่ย 96 เดือน โดยตำแหน่งงานมีประสบการณ์ก่อนทำงานเฉลี่ยมากที่สุด คือ หัวหน้างาน เฉลี่ย 298 เดือน รองลงมา คือ เจ้าหน้าที่ปฏิบัติงาน เฉลี่ย 85 เดือน และน้อยที่สุด คือ ผู้จัดการ เฉลี่ย 78 เดือน

เมื่อพิจารณาจากค่าสัมประสิทธิ์การกระจาย แสดงให้เห็นว่า ประสบการณ์ก่อนทำงานของกลุ่มตัวอย่างทั้งหมดมีความแตกต่างกันอยู่ในระดับมาก (C.V. = 109) และเจ้าหน้าที่ปฏิบัติงานมีความแตกต่างกันอยู่ในระดับมากเช่นกัน (C.V. = 112) ส่วนหัวหน้างานและผู้จัดการมีความแตกต่างกันอยู่ในระดับน้อย (C.V. = 34 และ C.V. = 94)

2.3 การเปรียบเทียบค่าเฉลี่ยโดยมีตัวแปรคุม

การเปรียบเทียบค่าเฉลี่ยโดยมีตัวแปรคุม เป็นการวิเคราะห์เหมือนกับการเปรียบเทียบค่าเฉลี่ย 2 กลุ่ม และมากกว่า 2 กลุ่ม แต่มีตัวแปรคุมเป็นค่าจัดกลุ่ม 2 กลุ่มหรือมากกว่า แต่เพื่อไม่ให้ซับซ้อนควรเลือกตัวแปรแบบค่าจัดกลุ่มไม่เกิน 2 กลุ่ม โดยใช้สถิติเชิงพรรณนา ได้แก่ ค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐาน และค่าสัมประสิทธิ์การกระจาย

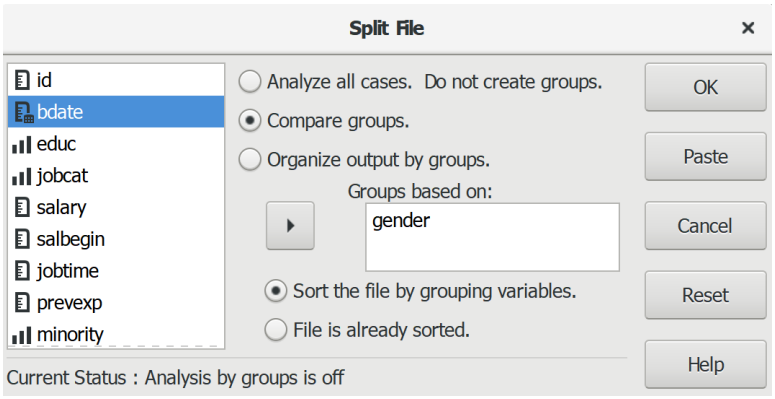
ตัวอย่างคำถามการวิจัย
จากการสำรวจข้อมูลของพนักงานในหน่วยงานต้องการทราบว่า พนักงานแต่กลุ่มชนที่มีเพศแตกต่างกันมีประสบการณ์ก่อนทำงานแตกต่างกันมากน้อยเพียงใด

2.3.1 ตัวแปร

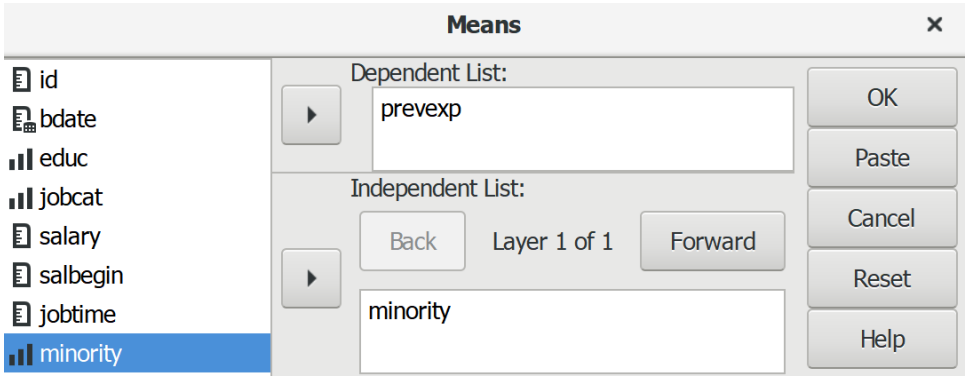
ตัวแปรอิสระ	**ตัวแปรคุม**	**ตัวแปรตาม**
ค่าจัดกลุ่ม	เพศ (gender)	ค่าต่อเนื่อง
กลุ่มชน (minority) - ชนกลุ่มหลัก - ชนกลุ่มน้อย	->	ประสบการณ์ก่อนทำงาน (prevexp)

2.3.2 การใช้โปรแกรมวิเคราะห์ข้อมูล

- เปิดแฟ้มข้อมูล Employee Data.sav
- เมนู **Data -> Split File...**
  - เลือกคำสั่ง **Compare groups.**
  - ใน **Group base on** เลือกตัวแปรแบบค่าจัดประเภทจากรายการเมนูด้านซ้าย 1 ตัวแปร หรือมากกว่า
  - กดปุ่ม **OK**



- เลือกเมนู **Analyze -> Compare Means -> Means...**
  - **Dependent List:** เลือกตัวแปรแบบค่าต่อเนื่อง 1 ตัวแปร (ตัวแปรที่ต้องการให้เป็นตัวแปรตาม) จากรายการตัวแปรด้านซ้าย
  - **Independent List:** เลือกตัวแปรแบบค่าจัดประเภท (มากกว่า 2 กลุ่ม) 1 ตัวแปร หรือหลายตัวแปร (ตัวแปรที่ต้องการให้เป็นตัวแปรอิสระ) จากรายการตัวแปรด้านซ้าย
  - กดปุ่ม **OK**



2.3.3 ผลการวิเคราะห์

Case Processing Summary

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
prevexp * minority	216	100.0%	0	.0%	216	100.0%

Report

Minority Classification	Mean	N	Std. Deviation
No	72	176	98
Yes	99	40	78
Total	77	216	95

Case Processing Summary

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
prevexp * minority	258	100.0%	0	.0%	258	100.0%

Report

Minority Classification	Mean	N	Std. Deviation
No	102	194	109
Yes	140	64	109
Total	112	258	110

2.3.4 การสร้างตารางผลการวิเคราะห์

ตารางที่ (#) แสดงค่าเฉลี่ยประสบการณ์ก่อนทำงานของพนักงาน ระหว่างเพศและกลุ่มชน

เพศ	กลุ่มชน	จำนวน	ค่าเฉลี่ย	S.D.	C.V. (%)
หญิง	ชนกลุ่มหลัก	176	72	98	136
	ชนกลุ่มน้อย	40	99	78	79
	รวม	216	77	95	123
ชาย	ชนกลุ่มหลัก	194	102	109	107
	ชนกลุ่มน้อย	64	140	109	78
	รวม	258	112	110	98

2.3.5 การแปลผลและสรุป

จากตารางที่ (#) แสดงให้เห็นว่า กลุ่มตัวอย่างเพศหญิงที่เป็นชนกลุ่มหลักมีประสบการณ์ก่อนทำงานโดยเฉลี่ย 72 เดือน น้อยกว่าเพศหญิงที่เป็นชนกลุ่มน้อยที่มีประสบการณ์ก่อนทำงานโดยเฉลี่ย 99 เดือน ส่วนกลุ่มตัวอย่างเพศชายที่เป็นชนกลุ่มหลักมีประสบการณ์ก่อนทำงานโดยเฉลี่ย 102 เดือน น้อยกว่าเพศชายที่เป็นชนกลุ่มน้อยที่มีประสบการณ์ก่อนทำงานโดยเฉลี่ย 140 เดือน

อาจกล่าวได้ว่า เพศหญิงและเพศชายที่เป็นชนกลุ่มหลักมีประสบการณ์ก่อนทำงานโดยเฉลี่ยน้อยกว่าเพศหญิงและเพศชายที่เป็นชนกลุ่มน้อย หรือชนกลุ่มหลักมีประสบการณ์ก่อนทำงานโดยเฉลี่ยน้อยกว่าชนกลุ่มน้อย

เมื่อพิจารณาจากค่าสัมประสิทธิ์การกระจายแสดงให้เห็นว่า ประสบการณ์ก่อนการทำงานของชนกลุ่มหลักทั้งเพศชายและเพศหญิงมีความแตกต่างกันอยู่ในระดับมาก ส่วนชนกลุ่มน้อยทั้งเพศชายและเพศหญิงมีความแตกต่างกันอยู่ในระดับน้อย

สรุป

การวิเคราะห์ข้อมูล 2 ตัวแปร หากเป็นตัวแปรแบบค่าแบ่งประเภทตั้งแต่ 2 กลุ่ม และมากกว่า ใช้การเปรียบเทียบความสัมพันธ์ระหว่างกลุ่มด้วยค่าความถี่ และค่าร้อยละ แต่ถ้าตัวแปรหนึ่งเป็นตัวแปรแบบค่าจัดประเภทและอีกตัวแปรหนึ่งเป็นตัวแปรแบบค่าต่อเนื่อง ใช้การเปรียบเทียบข้อมูลระหว่างกลุ่มด้วยค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐาน และค่าสัมประสิทธิ์การกระจาย

แบบฝึกหัดท้ายบท

ให้ผู้เรียนเปิดแฟ้มข้อมูล 1991 US. General Social Survey.sav แล้วประมวลผลข้อมูล เพื่อวิเคราะห์และอธิบายผล ดังนี้

1. วิเคราะห์และอธิบายผลตัวแปรอิสระคือ ภูมิภาค (region) กับตัวแปรตามคือ ประเภทของอาชีพ (occcat80) ด้วยคำสั่ง Crosstabs

2. วิเคราะห์และอธิบายผลตัวแปรอิสระคือ ภูมิภาค (region) กับตัวแปรตามคือ ความคิดเห็นเกี่ยวกับภาษี (tax) และคุมด้วยตัวแปรกลุ่มสีผิว (race) ด้วยคำสั่ง Crosstabs

3. วิเคราะห์และอธิบายผลตัวแปรภูมิภาค (region) กับจำนวนพี่น้อง (sibs) ด้วยคำสั่ง Means

4. วิเคราะห์และอธิบายผลตัวแปรภูมิภาค (region) กับจำนวนพี่น้อง (sibs) และคุมด้วยตัวแปรกลุ่มสีผิว (race) ด้วยคำสั่ง Means

ส่วนที่ 5

การวิเคราะห์ข้อมูลด้วยสถิติเชิงอ้างอิง

บทที่ 12
การทดสอบข้อมูลแบบอิงค่าพารามิเตอร์

ความมุ่งหมายของบทเรียน

ผู้เรียนสามารถใช้โปรแกรมประมวลผลทดสอบข้อมูลของตัวแปรด้วยสถิติเชิงอ้างอิงแบบอิงค่าพารามิเตอร์และเขียนบรรยายผลการทดสอบได้

เนื้อหา

1. การทดสอบค่าเฉลี่ยของกลุ่มตัวอย่าง 1 กลุ่ม
2. การทดสอบค่าเฉลี่ยของกลุ่มตัวอย่างที่ไม่เป็นอิสระจากกัน 2 กลุ่ม
3. การทดสอบค่าเฉลี่ยของกลุ่มตัวอย่างที่เป็นอิสระจากกัน 2 กลุ่ม
4. การทดสอบค่าเฉลี่ยของกลุ่มตัวอย่างที่เป็นอิสระจากกันมากกว่า 2 กลุ่ม
5. การทดสอบความสัมพันธ์รายคู่แบบเพียร์สัน

กิจกรรมและวิธีสอน

บรรยาย อธิบาย แสดงตัวอย่าง และทำแบบฝึกหัดท้ายบท

อุปกรณ์การสอน

เอกสารคำสอน ภาพเลื่อน คอมพิวเตอร์ โปรแกรม PSPP โปรแกรมสอนออนไลน์ (สอนแบบออนไลน์) และเครื่องฉาย (สอนในชั้นเรียน)

งานที่มอบหมาย

ทำแบบฝึกหัดท้ายบท

การวัดผลและประเมินผล

การเข้าชั้นเรียน สังเกตความสนใจจากการเรียน สอบถามความเข้าใจจากการบรรยาย และประเมินจากการทำแบบฝึกหัดท้ายบท

เนื้อหา

เมื่อข้อมูลเป็นไปตามข้อตกลงเบื้องต้น (meet basic assumptions) และใช้การทดสอบแบบอิงค่าพารามิเตอร์ (parametric test) ผลการทดสอบให้อำนาจจำแนกทางสถิติ (statistical power) สูงกว่าการทดสอบแบบไม่อิงค่าพารามิเตอร์

แม้ว่าการทดสอบแบบอิงค่าพารามิเตอร์ส่วนใหญ่วางอยู่บนฐานของข้อมูลที่มีระดับการวัดตั้งแต่มาตราแบบช่วงขึ้นไปหรือเป็นค่าต่อเนื่อง และข้อมูลต้องมีการแจกแจงแบบปกติ (normal distribution) แล้ว แต่การทดสอบแต่ละประเภทยังมีข้อตกลงเฉพาะเพิ่มเติมอีก เช่น การทดสอบค่าเฉลี่ยระหว่างกลุ่มตัวอย่าง ข้อมูลแต่ละกลุ่มตัวอย่างต้องมาจากประชากรที่มีความแปรปรวนเท่ากัน (homogeneity of variance) การทดสอบค่าเฉลี่ยแบบทดสอบซ้ำ ข้อมูลของแต่ละตัวอย่างต้องมีความเป็นอิสระจากกัน (independence)

1. การทดสอบค่าเฉลี่ยของกลุ่มตัวอย่าง 1 กลุ่ม

การทดสอบค่าเฉลี่ยของกลุ่มตัวอย่าง 1 กลุ่ม (One Sample t Test) คือ การเปรียบเทียบข้อมูลของตัวอย่าง 1 กลุ่ม ที่สุ่มมาจากประชากร กับค่าทดสอบ (value test) ของประชากรที่ทราบจำนวน (know populations) หรือเป็นค่าของสมมติฐาน (hypothesized) เพื่อตัดสินใจว่า ค่าเฉลี่ยของข้อมูลมีความแตกต่างกันหรือไม่แตกต่างกัน

การทดสอบแบบนี้อาจมีชื่อเรียกอีกชื่อหนึ่งว่า Single Sample t Test

ตัวอย่างคำถามการวิจัย
จากรายงานผลการสำรวจของสำนักงานสถิติด้านสังคมและเศรษฐกิจของประเทศพบว่า ผู้จบการศึกษาในทุกระดับการศึกษา มีการย้ายหรือเปลี่ยนงานหลังจากได้งานครั้งแรกแล้วโดยเฉลี่ยประมาณ 5 ปี (60 เดือน) ดังนั้นจากข้อมูลการสำรวจพนักงานในหน่วยงานต้องการทราบว่า พนักงานมีค่าเฉลี่ยของประสบการณ์ก่อนทำงานแตกต่างไปจากค่าเฉลี่ยตามรายงานผลการสำรวจของสำนักงานสถิติหรือไม่

1.1 ข้อตกลงเบื้องต้น

ขั้นตอน	ข้อตกลง
การออกแบบการวิจัย	1. กลุ่มตัวอย่างได้มาจากประชากรที่ทราบจำนวน (know population) โดยการสุ่มตัวอย่างแบบสุ่ม (independence random sampling) 2. มีกลุ่มตัวอย่างมากกว่าหรือน้อยกว่า 30 ตัวอย่าง 3. ตัวแปรทดสอบเป็นตัวแปรแบบค่าต่อเนื่อง (continuous variable) หรือคะแนน (scores)
การวิเคราะห์ข้อมูล	1. ข้อมูลมีการแจกแจงแบบปกติ (normal distribution) 2. ข้อมูลไม่มีค่าผิดปกติ (outliers) หรือค่าต่างสุด (extremes)

1.2 รูปแบบการทดสอบตัวแปร

ตัวแปร		ค่าทดสอบ
ค่าต่อเนื่อง	->	ค่าเฉลี่ย
ประสบการณ์ก่อนทำงาน (prevexp)		60 เดือน

1.3 การกำหนดสมมติฐานและเกณฑ์การทดสอบ

ตารางที่ (#) แสดงค่าเฉลี่ยประสบการณ์ก่อนทำงานของพนักงาน

ประสบการณ์ทำงาน	จำนวน	ค่าต่ำสุด	ค่าสูงสุด	ค่าเฉลี่ย	S.D.	C.V. (%)
ก่อนทำงาน (เดือน)	474	0	476	95.86	104.59	109.11

จากตารางที่ (#) แสดงให้เห็นว่า พนักงานจำนวน 474 คน มีประสบการณ์ก่อนทำงานเฉลี่ย 95.86 เดือน มีประสบการณ์ก่อนทำงานต่ำสุด 0 เดือน และมีประสบการณ์สูงสุด 476 เดือน เมื่อพิจารณาจากค่าสัมประสิทธิ์การกระจายแสดงให้เห็นว่า พนักงานมีประสบการณ์ก่อนทำงานแตกต่างกันมาก

จากข้อมูลค่าเฉลี่ยประสบการณ์ก่อนทำงานของพนักงาน ต้องการทดสอบว่า พนักงานของหน่วยงานมีค่าเฉลี่ยของประการณ์ก่อนทำงานเท่ากับหรือไม่เท่ากับค่าเฉลี่ยจากผลการสำรวจ จึงทำการทดสอบด้วยวิธีการทางสถิติ โดยมีสมมติฐานดังนี้

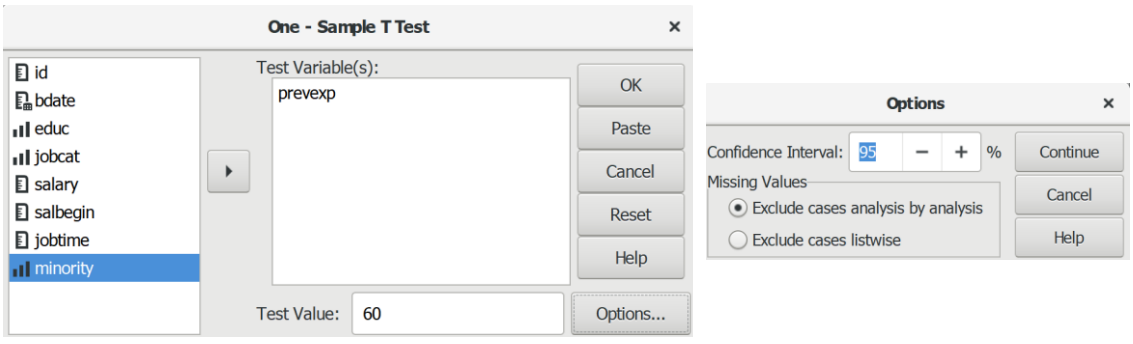
H_0: กลุ่มตัวอย่างที่เป็นพนักงานของหน่วยงานมีประสบการณ์ก่อนทำงานเท่ากับ 60 เดือน

H_1: กลุ่มตัวอย่างที่เป็นพนักงานของหน่วยงานมีประสบการณ์ก่อนทำงานไม่เท่ากับ 60 เดือน

เกณฑ์ในการทดสอบสมมติฐาน คือ หากการทดสอบความแตกต่างระหว่างค่าเฉลี่ยของกลุ่มตัวอย่างกับค่าทดสอบ (one-sample t-test) ได้ค่าความน่าจะเป็น (p-value) มากกว่าค่านัยสำคัญทางสถิติ (α = 0.05) จะยอมรับสมมติฐาน H_0 แต่หากได้ค่าความน่าจะเป็น (p-value) น้อยกว่าหรือเท่ากับค่านัยสำคัญทางสถิติ (α = 0.05) จะปฏิเสธสมมติฐาน H_0 และยอมรับสมมติฐาน H_1

1.4 การใช้โปรแกรมคอมพิวเตอร์วิเคราะห์ข้อมูล

- เปิดแฟ้มข้อมูล **File -> Open** -> Employee data.sav
- เลือกเมนู **Analyze -> Compare Means -> One Sample T Test...**
  - **Test Variable(s):** เลือกตัวแปรทดสอบ (ตัวแปรแบบค่าต่อเนื่อง) 1 ตัวแปร หรือมากกว่า 1 ตัวแปร จากรายการตัวแปรด้านซ้าย
  - **Test Value:** ใส่ค่าทดสอบในช่อง (ค่าทดสอบมาจากข้อมูลในอดีต ความรู้ที่มีอยู่ ทฤษฎี หรืองานวิจัย)
  - กดปุ่ม **Options**
    - **Confidence Interval Percentage:** กำหนดช่วงความเชื่อมั่นในช่อง (ค่าปกติ 0.95)
    - ในชุดคำสั่ง **Missing Values** เลือกวิธีการคำนวณค่าสูญหาย
      - Exclude cases analysis by analysis (วิเคราะห์ตัวอย่างที่ไม่มีข้อมูลสูญหายในแต่ละตัวแปรทดสอบ)
      - Exclude cases listwise (วิเคราะห์ตัวอย่างที่ไม่มีข้อมูลสูญหายในทุกตัวแปรทดสอบ)
  - กดปุ่ม **Continue**
  - กดปุ่ม **OK**



1.5 ผลการวิเคราะห์

One-Sample Statistics

	N	Mean	Std. Deviation	S.E. Mean
Previous Experience (months)	474	95.86	104.59	4.80

One-Sample Test

	Test Value = 60					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Previous Experience (months)	7.47	473	.000	35.86	26.42	45.30

1.6 การสร้างตารางแสดงผลการทดสอบ

ตารางที่ (#) แสดงการทดสอบการเปรียบเทียบค่าเฉลี่ยประสบการณ์ก่อนทำงานของพนักงาน

ค่าทดสอบ = 60 เดือน								
ประสบการณ์ทำงาน	**จำนวน (คน)**	**ค่าเฉลี่ย**	**t**	**df**	**Sig.**	Mean Difference	95% Confidence Level	
							Lower	Upper
ก่อนทำงาน	474	95.86	7.47	473	0.00	35.86	26.42	45.30

1.7 การแปลผลและสรุป

จากตารางที่ (#) ผลการทดสอบแสดงให้เห็นว่า ค่าความน่าจะเป็น (Sig.) = 0.00 มีค่าน้อยกว่าค่านัยสำคัญทางสถิติที่กำหนดไว้ (α = 0.05) จึงปฏิเสธสมมติฐาน H_0 และยอมรับสมมติฐาน H_a

สรุป พนักงานของหน่วยงานมีประสบการณ์ก่อนทำงานโดยเฉลี่ยไม่เท่ากับ 60 เดือน กล่าวคือ มีค่าเฉลี่ยแตกต่างจากค่าทดสอบประมาณ 35.86 เดือน โดยมีความเชื่อมั่นว่าความแตกต่างของค่าเฉลี่ยมีค่าต่ำสุดไม่เกิน 26.42 เดือน และสูงสุดไม่เกิน 45.30 เดือน

2. การทดสอบค่าเฉลี่ยของกลุ่มตัวอย่างที่ไม่เป็นอิสระจากกัน 2 กลุ่ม

การทดสอบค่าเฉลี่ยของกลุ่มตัวอย่างที่ไม่เป็นอิสระจากกัน 2 กลุ่ม (Paired Samples t Test) คือ การเปรียบเทียบข้อมูลของกลุ่มตัวอย่างที่สัมพันธ์กันหรือเกี่ยวข้องกัน (ได้มาจากสถานการณ์/บุคคล/สิ่งของเดียวกันแต่เป็นเรื่องที่มีความเกี่ยวข้องกันอาจเกิดขึ้นในช่วงเวลาหรือสถานที่ที่แตกต่างกันที่) เพื่อตัดสินใจว่า ค่าเฉลี่ยของข้อมูลมีความแตกต่างกันหรือไม่แตกต่างกัน เช่น รายได้เฉลี่ยของครัวเรือนในแต่ละชุมชนระหว่างปี พ.ศ. 2555 กับ ปี พ.ศ. 2560 จำนวนประชากรในแต่ละจังหวัดระหว่างปี พ.ศ. 2553 กับ ปี พ.ศ. 2563 เงินเดือนของพนักงานในหน่วยงานระหว่างปีปัจจุบันกับเมื่อ 5 ปีที่แล้ว

การทดสอบแบบนี้อาจมีชื่อเรียกเป็นอย่างอื่น ได้แก่ Dependent t Test, Related t Test และ Repeated Measures t Test

ตัวอย่างคำถามการวิจัย
จากการสำรวจข้อมูลของพนักงานในหน่วยงานต้องการทราบว่า พนักงานมีค่าเฉลี่ยของเงินเดือนเริ่มต้นแตกต่างกับเงินเดือนปัจจุบันหรือไม่

2.1 ข้อตกลงเบื้องต้น

ขั้นตอน	ข้อตกลง
การออกแบบการวิจัย	1. กลุ่มตัวอย่างได้มาจากประชากรที่ทราบจำนวน (know population) โดยการสุ่มตัวอย่างแบบสุ่มที่เป็นอิสระจากกัน (independence random sampling) 2. กลุ่มตัวอย่างอาจมีขนาดมากกว่าหรือน้อยกว่า 30 ตัวอย่าง 3. ข้อมูลได้มาจากกลุ่มตัวอย่างมีความสัมพันธ์หรือเกี่ยวข้องกัน (dependent/paired sample) 4. ข้อมูลของตัวแปรทดสอบทั้งคู่เป็นมีระดับการวัดแบบค่าต่อเนื่อง (continuous level)
การวิเคราะห์ข้อมูล	1. ข้อมูลของตัวแปรทดสอบทั้งคู่มีการแจกแจงปกติ (normal distribution) 2. ข้อมูลของตัวแปรทดสอบทั้งคู่ไม่มีค่าผิดปกติ (outliers) หรือค่าต่างสุด (extremes)

2.2 รูปแบบการทดสอบตัวแปร

ตัวแปรทดสอบ (การวัดครั้งที่ 1)	วัดซ้ำ	ตัวแปรทดสอบ (การวัดครั้งที่ 2)
ค่าต่อเนื่อง	->	ค่าต่อเนื่อง
เงินเดือนเริ่มต้น (salbegin)		เงินเดือนปัจจุบัน (salary)

2.3 การกำหนดสมมติฐานและเกณฑ์การทดสอบ

ตารางที่ (#) แสดงการเปรียบเทียบค่าเฉลี่ยเงินเดือนปัจจุบันกับเงินเดือนเริ่มต้น

เงินเดือน	จำนวน	ค่าเฉลี่ย	S.D.	C.V. (%)
เงินเดือนเริ่มต้น	474	17,016.09	7,870.64	46.25
เงินเดือนปัจจุบัน	474	34,419.57	17,075.66	49.61

จากตารางที่ (#) แสดงให้เห็นว่า กลุ่มตัวอย่างที่เป็นพนักงานจำนวน 474 คน มีเงินเดือนเริ่มต้นเฉลี่ย 17,016.09 ดอลลาร์สหรัฐ ส่วนเงินเดือนปัจจุบันเฉลี่ย 34,419.57 ดอลลาร์สหรัฐ กล่าวคือ ค่าเฉลี่ยของเงินเดือนปัจจุบันมากกว่าค่าเฉลี่ยของเงินเดือนเริ่มต้น

ดังนั้น เพื่อเป็นการยืนยันว่า ค่าเฉลี่ยของเงินเดือนปัจจุบันแตกต่างจากค่าเฉลี่ยของเงินเดือนเริ่มต้น จึงทำการทดสอบด้วยวิธีการทางสถิติ โดยมีสมมติฐาน ดังนี้

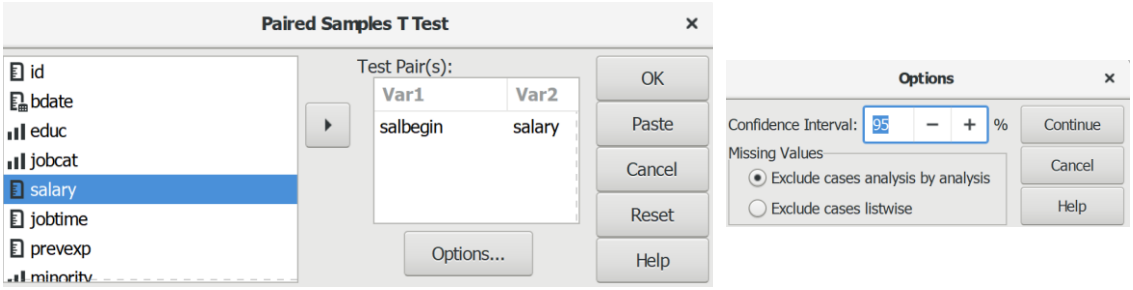
H_0: ค่าเฉลี่ยของเงินเดือนปัจจุบันกับค่าเฉลี่ยของเงินเดือนเริ่มต้นของพนักงานไม่แตกต่างกัน

H_1: ค่าเฉลี่ยของเงินเดือนปัจจุบันกับค่าเฉลี่ยของเงินเดือนเริ่มต้นของพนักงานแตกต่างกัน

เกณฑ์ในการทดสอบสมมติฐาน คือ หากการทดสอบความแตกต่างระหว่างค่าเฉลี่ยของกลุ่มตัวอย่างที่มีความสัมพันธ์กันหรือไม่เป็นอิสระจากกัน (paired sample t-test) ได้ค่าความน่าจะเป็น (p-value) มากกว่าค่านัยสำคัญทางสถิติ (α = 0.05) จะยอมรับสมมติฐาน H_0 แต่หากได้ค่าความน่าจะเป็น (p-value) น้อยกว่าหรือเท่ากับค่านัยสำคัญทางสถิติ (α = 0.05) จะปฏิเสธสมมติฐาน H_0 และยอมรับสมมติฐาน H_1

2.4 การใช้โปรแกรมคอมพิวเตอร์วิเคราะห์ข้อมูล

- เปิดแฟ้มข้อมูล **File -> Open** -> Employee data.sav
- เลือกเมนู **Analyze -> Compare Means -> Paired Samples T Test...**
  - เลือกตัวแปรทดสอบตัวที่ 1 (ตัวแปรแบบค่าต่อเนื่อง/เหตุการณ์ก่อน) จากช่องรายการตัวแปร -> ใส่ในช่อง **Var1** ของ **Test Pair(s):**
  - เลือกตัวแปรทดสอบตัวที่ 2 (ตัวแปรแบบค่าต่อเนื่อง/เหตุการณ์หลัง) จากช่องรายการตัวแปร -> ใส่ในช่อง **Var2** ของ **Test Pair(s):**
  - กดปุ่ม **Options**
    - กำหนดช่วงความเชื่อมั่นในช่อง **Confidence Interval Percentage:** (ค่าปกติ .95)
    - เลือกวิธีการคำนวณค่าสูญหายในชุดคำสั่ง **Missing Values**
      - Exclude cases analysis by analysis (วิเคราะห์ตัวอย่างที่ไม่มีข้อมูลสูญหายในแต่ละตัวแปรทดสอบ)
      - Exclude cases listwise (วิเคราะห์ตัวอย่างที่ไม่มีข้อมูลสูญหายในทุกตัวแปรทดสอบ)
  - กดปุ่ม **Continue**
  - กดปุ่ม **OK**



2.5 ผลการวิเคราะห์

Paired Sample Statistics

		N	Mean	Std. Deviation	S.E. Mean
Pair 1	Beginning Salary	474	17016.09	7870.64	361.51
	Current Salary	474	34419.57	17075.66	784.31

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Beginning Salary & Current Salary	474	.880	.000

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	S.E. Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Beginning Salary - Current Salary	-17403.5	10814.62	496.73	-18379.6	-16427.4	-35.04	473	.000

2.6 การสร้างตารางแสดงผลการทดสอบ

ตารางที่ (#) แสดงการทดสอบเปรียบเทียบค่าเฉลี่ยเงินเดือนปัจจุบันกับเงินเดือนเริ่มต้น

เงินเดือน	Mean Difference	Std. Error Difference	S.E. Mean	95% Confidence Level		t	df	Sig.
				Lower	Upper			
เริ่มต้น-ปัจจุบัน	-1,7403.48	10,814.62	496.73	-18,379.6	-16,427.4	-35.04	473	0.000

2.7 การแปลผลและสรุป

จากตารางที่ (#) ผลการทดสอบแสดงให้เห็นว่า ค่าความน่าจะเป็น (Sig.) = 0.00 มีค่าน้อยกว่าค่านัยสำคัญทางสถิติที่กำหนดไว้ (α = 0.05) จึงปฏิเสธสมมติฐาน H_0 และยอมรับสมมติฐาน H_1

สรุป ค่าเฉลี่ยของเงินเดือนปัจจุบันกับค่าเฉลี่ยของเงินเดือนเริ่มต้นของพนักงานแตกต่างกัน โดยค่าเฉลี่ยของเงินเดือนปัจจุบันแตกต่างจากค่าเฉลี่ยของเงินเดือนเริ่มต้นประมาณ 1,7403.48 ดอลล่าร์สหรัฐ โดยมีความเชื่อมั่นว่าความแตกต่างของค่าเฉลี่ยมีค่าต่ำสุดไม่เกิน 16427.41 ดอลลาร์สหรัฐ และสูงสุดไม่เกิน 18,379.59 ดอลลาร์สหรัฐ

ตัวอย่างเพิ่มเติม: การแปลผลและสรุป
#1 จากการทดสอบความแตกต่างระหว่างค่าเฉลี่ยของกลุ่มตัวอย่างที่มีความสัมพันธ์กัน (paired sample t-test) ได้ค่านัยสำคัญทางสถิติ (Sig.) = 0.000 ซึ่งน้อยกว่าเกณฑ์ที่กำหนดไว้ จึงปฏิเสธสมมติฐาน H_0 และยอมรับสมมติฐาน H_1 คือ ค่าเฉลี่ยของเงินเดือนปัจจุบันกับค่าเฉลี่ยของเงินเดือนเริ่มต้นของพนักงานแตกต่างกันประมาณ 1,7403.48 ดอลล่าร์สหรัฐ โดยมีความเชื่อมั่นว่าความแตกต่างของค่าเฉลี่ยมีค่าต่ำสุดไม่เกิน 16427.41 ดอลลาร์สหรัฐ และสูงสุดไม่เกิน 18,379.59 ดอลลาร์สหรัฐ
#2 จากการทดสอบความแตกต่างระหว่างค่าเฉลี่ยของเงินเดือนปัจจุบันกับค่าเฉลี่ยของเงินเดือนเริ่มต้นของพนักงานแสดงให้เห็นว่า มีแตกต่างกันอย่างมีนัยสำคัญทางสถิติ

3. การทดสอบค่าเฉลี่ยของกลุ่มตัวอย่างที่เป็นอิสระจากกัน 2 กลุ่ม

การทดสอบค่าเฉลี่ยของกลุ่มตัวอย่างที่เป็นอิสระจากกัน 2 กลุ่ม (2 Samples t Tests) คือ การเปรียบเทียบข้อมูลของกลุ่มตัวอย่างที่ไม่มีความสัมพันธ์กัน 2 กลุ่ม เพื่อตัดสินใจว่า ค่าเฉลี่ยของข้อมูลมีความแตกต่างกันหรือไม่แตกต่างกัน

การทดสอบแบบนี้อาจมีชื่อเรียกเป็นอย่างอื่น ได้แก่ Independent t Test, Independent Measures t Test, Independent Two-sample t Test, Student t Test, Uncorrelated Scores t Test, Unpaired t Test และ Unrelated t Test

ตัวอย่างคำถามการวิจัย
จากข้อมูลการสำรวจพนักงานในหน่วยงานต้องการทราบว่า พนักงานระหว่างชนกลุ่มหลักกับชนกลุ่มน้อยมีค่าเฉลี่ยของประสบการณ์ก่อนทำงานแตกต่างกันหรือไม่

3.1 ข้อตกลงเบื้องต้น

ขั้นตอน	**ข้อตกลง**
การออกแบบการวิจัย	1. กลุ่มตัวอย่างได้มาจากประชากรที่ทราบจำนวน (know population) โดยการสุ่มตัวอย่างแบบสุ่มที่เป็นอิสระจากกัน (independence random sampling) 2. กลุ่มตัวอย่างมีขนาดใหญ่ (มากกว่า 30) 3. ตัวแปรอิสระเป็นตัวแปรแบบค่าจัดประเภท (categorical variable) มีอย่างน้อย 2 กลุ่ม 4. ตัวแปรตามเป็นตัวแปรแบบค่าต่อเนื่อง (continuous variable)
การวิเคราะห์ข้อมูล	1. ข้อมูลทั้งสองกลุ่มมีการแจกแจงแบบปกติ (normal distribution) 2. ข้อมูลระหว่างกลุ่มมีความแปรปรวนเท่ากัน (homogeneity of Variance) 3. ข้อมูลทั้งสองกลุ่มไม่มีค่าผิดปกติ (outliers) หรือค่าต่างสุด (extremes)

3.2 รูปแบบการทดสอบตัวแปร

ตัวแปรอิสระ		ตัวแปรตาม
ค่าจัดประเภท		ค่าต่อเนื่อง
กลุ่มชน (minority) - ชนกลุ่มหลัก - ชนกลุ่มน้อย	->	ประสบการณ์ก่อนทำงาน (prevexp)

3.3 การกำหนดสมมติฐานและเกณฑ์การทดสอบ

ตารางที่ (#) แสดงการเปรียบเทียบค่าเฉลี่ยประสบการณ์ก่อนทำงานของพนักงานระหว่างกลุ่มชน

กลุ่มชน	จำนวน	ค่าเฉลี่ย	S.D.	C.V.
ชนกลุ่มหลัก	370	87.84	104.56	119.04
ชนกลุ่มน้อย	104	124.38	100.07	80.46
รวม	474	95.86	104.59	109.11

จากตารางที่ (#) แสดงให้เห็นว่า พนักงานที่เป็นกลุ่มตัวอย่างจำนวน 474 คน เป็นชนกลุ่มหลัก จำนวน 370 คน มีประสบการณ์ก่อนทำงานเฉลี่ย 87.84 เดือน ส่วนกลุ่มตัวอย่างที่เป็นกลุ่มน้อย จำนวน 104 คน มีประสบการณ์ก่อนทำงานเฉลี่ย 124.38 เดือน

จากการเปรียบเทียบพนักงานที่เป็นชนกลุ่มหลักมีประสบการณ์ก่อนการทำงานน้อยกว่าพนักงานที่เป็นชนกลุ่มน้อย

ดังนั้น เพื่อเป็นการยืนข้อสรุปที่ว่า กลุ่มตัวอย่างที่เป็นพนักงานชนกลุ่มหลักมีประสบการณ์ก่อนการทำงานแตกต่างกับกลุ่มตัวอย่างที่เป็นพนักงานชนกลุ่มน้อย จึงทำการทดสอบความแตกต่างระหว่างกลุ่มตัวอย่าง โดยกำหนดสมมติฐาน ดังนี้

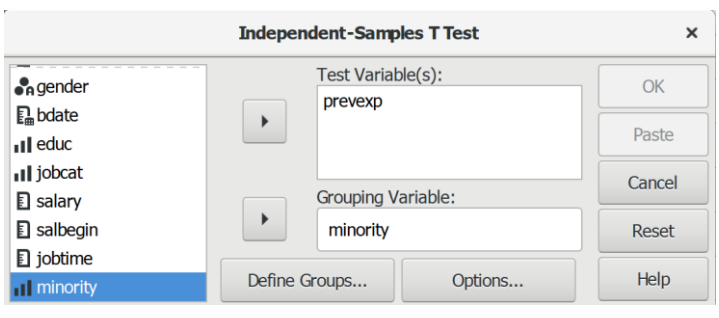
H_0: พนักงานที่เป็นชนกลุ่มหลักและชนกลุ่มน้อยมีค่าเฉลี่ยของประสบการณ์ก่อนทำงานเท่ากัน

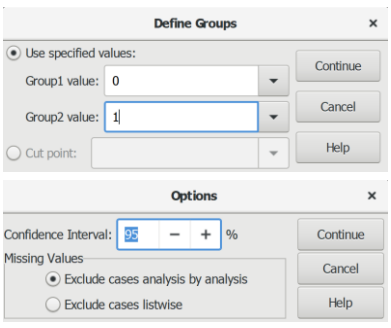
H_1: พนักงานที่เป็นชนกลุ่มหลักค่าเฉลี่ยของประสบการณ์ก่อนทำงานน้อยกว่าพนักงานที่เป็นชนกลุ่มน้อย

เกณฑ์ในการทดสอบสมมติฐาน คือ หากการทดสอบความแตกต่างของค่าเฉลี่ยระหว่างกลุ่มตัวอย่างที่เป็นอิสระจากกัน (Independent sample t-test) ได้ค่าความน่าจะเป็น (p-value) มากกว่าค่านัยสำคัญทางสถิติ (α = 0.05) จะยอมรับสมมติฐาน H_0 แต่หากได้ค่าความน่าจะเป็น (p-value) น้อยกว่าหรือเท่ากับค่านัยสำคัญทางสถิติ (α = 0.05) จะปฏิเสธสมมติฐาน H_0 และยอมรับสมมติฐาน H_1

3.4 การใช้โปรแกรมคอมพิวเตอร์วิเคราะห์ข้อมูล

- เปิดแฟ้มข้อมูล **File -> Open** -> Employee data.sav
- เลือกเมนู **Analyze -> Compare Means -> Independent Samples T Test...**
  - **Test Variable (s):** เลือกตัวแปรค่าต่อเนื่อง 1 ตัวแปร หรือมากกว่า 1 ตัวแปร จากรายการตัวแปรด้านซ้าย
  - **Grouping Variable:** เลือกตัวแปรค่าแบ่งกลุ่ม 1 ตัวแปร
  - กดปุ่ม **Define Groups...**
    - Use specified values: กำหนดค่าของตัวแปรค่าแบ่งกลุ่มใน Group 1: และ Group2: หรือ
    - Cut point: กำหนดจุดแบ่งกลุ่มใน (กรณีต้องการรวมกลุ่มที่มีลักษณะเป็นค่าต่อเนื่องเป็น 2 กลุ่ม)
  - กดปุ่ม **Continue**
  - กดปุ่ม **Options**
    - **Confidence Interval Percentage:** กำหนดช่วงความเชื่อมั่นในช่อง (ค่าปกติ .95)
    - **Missing Values** เลือกวิธีการคำนวณค่าสูญหายในชุดคำสั่ง
      - Exclude cases analysis by analysis (วิเคราะห์ตัวอย่างที่ไม่มีข้อมูลสูญหายในแต่ละตัวแปรทดสอบ)
      - Exclude cases listwise (วิเคราะห์ตัวอย่างที่ไม่มีข้อมูลสูญหายในทุกตัวแปรทดสอบ)
  - กดปุ่ม **Continue**
  - กดปุ่ม **OK**





3.5 ผลการวิเคราะห์

Group Statistics

	Group	N	Mean	Std. Deviation	S.E. Mean
Previous Experience (months)	No	370	87.84	104.56	5.44
	Yes	104	124.38	100.07	9.81

Independent Samples Test

		Levene's Test for Equality of Variances		T-Test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Previous Experience (months)	Equal variances assumed	.11	.744	-3.18	472.00	.002	-36.54	11.50	-59.13	-13.95
	Equal variances not assumed			-3.26	171.41	.001	-36.54	11.22	-58.68	-14.40

3.6 การสร้างตารางแสดงผลการทดสอบ

ตารางที่ (#) แสดงผลการทดสอบการเปรียบเทียบค่าเฉลี่ยประสบการณ์ก่อนทำงานของพนักงานระหว่างกลุ่มชน

ประสบการณ์ระหว่างกลุ่มชน	Leven's Test		t	df	Sig.	Mean Difference	95% Confidence Level	
	F	Sig.					Lower	Upper
ประสบการณ์ก่อนทำงาน	0.107	0.74	-3.18	472	0.002	-36.54	-59.13	-13.95

3.7 การแปลผลและสรุป

จากตารางที่ (#) แสดงให้เห็นว่า ค่าความน่าจะเป็น (Sig.) = 0.002 มีค่าน้อยกว่าค่านัยสำคัญทางสถิติที่กำหนดไว้ (α = 0.05) จึงปฏิเสธสมมติฐาน H_0 และยอมรับสมมติฐาน H_1

สรุป พนักงานที่เป็นชนกลุ่มหลักมีค่าเฉลี่ยของประสบการณ์ก่อนทำงานน้อยกว่าพนักงานที่เป็นชนกลุ่มน้อยประมาณ 36.54 เดือน โดยมีความเชื่อมั่นว่าความแตกต่างของค่าเฉลี่ยมีค่าต่ำสุดไม่เกิน 56.13 เดือน และสูงสุดไม่เกิน 13.95 เดือน

Tips: การแปลผลและสรุป
#1 จากตารางที่ (#) ผลการทดสอบแสดงให้เห็นว่า พนักงานที่เป็นชนกลุ่มหลักมีประสบการณ์ก่อนทำงานน้อยกว่าพนักงานที่เป็นชนกลุ่มน้อยโดยเฉลี่ย 36.54 เดือน อย่างมีนัยสำคัญทางสถิติ
#2 จากตารางที่ (#) ผลการทดสอบแสดงให้เห็นว่า พนักงานที่เป็นชนกลุ่มหลักค่าเฉลี่ยของประสบการณ์ก่อนทำงานน้อยกว่าพนักงานที่เป็นชนกลุ่มน้อย อย่างมีนัยสำคัญทางสถิติ
#3 จากตารางที่ (#) ผลการทดสอบแสดงให้เห็นว่า กลุ่มตัวอย่างที่เป็นพนักงานชนกลุ่มหลักมีประสบการณ์ก่อนทำงานแตกต่างกับกลุ่มตัวอย่างที่เป็นพนักงานชนกลุ่มน้อย อย่างมีนัยสำคัญทางสถิติ

4. การทดสอบค่าเฉลี่ยของกลุ่มตัวอย่างที่เป็นอิสระจากกันมากกว่า 2 กลุ่ม

การทดสอบค่าเฉลี่ยของกลุ่มตัวอย่างที่เป็นอิสระจากกันมากกว่า 2 กลุ่ม (One Way Analysis of Variance of Analysis: One-way ANOVA) คือ การเปรียบเทียบข้อมูลของกลุ่มตัวอย่างที่เป็นอิสระจากกันมากกว่า 2 กลุ่ม เพื่อตัดสินใจว่า ค่าเฉลี่ยของข้อมูลมีความแตกต่างกันหรือไม่แตกต่างกัน

การทดสอบแบบนี้อาจมีชื่อเรียกเป็นอย่างอื่น ได้แก่ One-Factor ANOVA, One-Way Analysis of Variance และ Between Subjects ANOVA

ตัวอย่างคำถามการวิจัย
จากข้อมูลการสำรวจพนักงานในหน่วยงานต้องการทราบว่า พนักงานที่มีตำแหน่งงานแตกต่างกันมีค่าเฉลี่ยของประสบการณ์ก่อนทำงานแตกต่างกันหรือไม่

4.1 ข้อตกลงเบื้องต้น

ขั้นตอน	ข้อตกลง
การออกแบบการวิจัย	1. กลุ่มตัวอย่างได้มาจากประชากรที่ทราบจำนวน (know population) โดยการสุ่มตัวอย่างแบบสุ่มที่เป็นอิสระจากกัน (independence random sampling) 2. กลุ่มตัวอย่างมีขนาดใหญ่ (มากกว่า 30) 3. ตัวแปรอิสระเป็นตัวแปรแบบค่าจัดประเภท (categorical variable) มีการแบ่งประเภทอย่างน้อย 3 กลุ่ม 4. ตัวแปรตามเป็นตัวแปรแบบค่าต่อเนื่อง (continuous variable)
การวิเคราะห์ข้อมูล	1. ข้อมูลของแต่ละกลุ่มมีการแจกแจงแบบปกติ (normal distribution) 2. ข้อมูลของแต่กลุ่มมีความแปรปรวนเท่ากัน (homogeneity of variance) 3. ข้อมูลของแต่ละกลุ่มไม่มีค่าผิดปกติ (outliers) หรือค่าต่างสุด (extremes)

4.2 รูปแบบการทดสอบตัวแปร

ตัวแปรอิสระ		ตัวแปรตาม
ค่าจัดประเภท (มาตราแบบจัดอันดับ 3 กลุ่มหรือมากกว่า)		ค่าต่อเนื่อง
ตำแหน่งงาน (jobcat) - เจ้าหน้าที่ปฏิบัติงาน - หัวหน้างาน - ผู้จัดการ	->	ค่าเฉลี่ยของประสบการณ์ก่อนทำงาน (prevexp)

4.3 การกำหนดสมมติฐานและเกณฑ์การทดสอบ

ตารางที่ (#) แสดงการเปรียบเทียบค่าเฉลี่ยประสบการณ์ก่อนทำงานของพนักงานระหว่างตำแหน่งงาน

ตำแหน่งงาน	จำนวน	ต่ำสุด	สูงสด	ค่าเฉลี่ย	S.D.
เจ้าหน้าที่ปฏิบัติงาน	363	0	476	85.04	95.27
หัวหน้างาน	27	144	460	298.11	101.43
ผู้จัดการ	84	3	285	77.62	73.26
รวม	474	0	476	95.86	104.59

จากตารางที่ (#) แสดงให้เห็นว่า กลุ่มตัวอย่างพนักงานที่เป็นหัวหน้างานมีประสบการณ์ก่อนทำงานมากที่สุด คือ 298.11 เดือน รองลงมาเป็นเจ้าหน้าที่ปฏิบัติงาน คือ 85.04 เดือน และผู้จัดการมีประสบการณ์ก่อนทำงานน้อยที่สุด คือ 77.62 เดือน หรือสรุปได้ว่า พนักงานที่มีตำแหน่งงานแตกต่างกันมีประสบการณ์การก่อนทำงานแตกต่างกัน

ดังนั้น เพื่อเป็นการยืนยันว่า พนักงานที่มีตำแหน่งงานแตกต่างกันมีประสบการณ์การก่อนทำงานแตกต่างกัน จึงทำการทดสอบด้วยวิธีการทางสถิติ โดยมีสมมติฐาน ดังนี้

H_0: กลุ่มตัวอย่างที่มีตำแหน่งงานแตกต่างกันมีค่าเฉลี่ยของประสบการณ์ก่อนทำงานไม่แตกต่างกัน

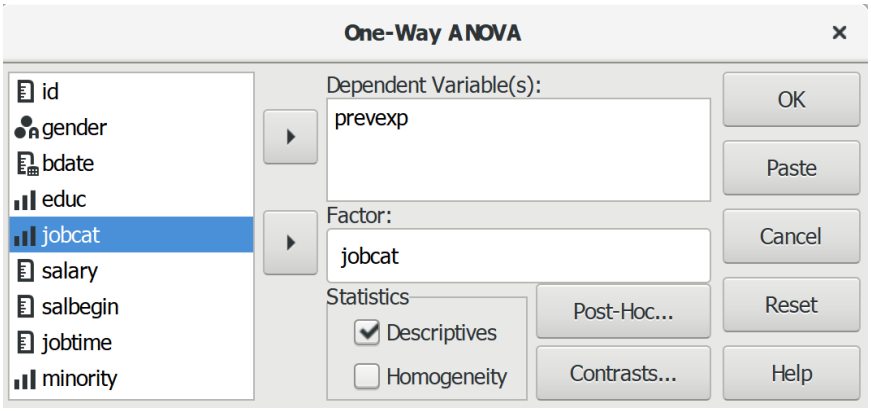
H_1: กลุ่มตัวอย่างที่มีตำแหน่งงานแตกต่างกันมีค่าเฉลี่ยของประสบการณ์ก่อนทำงานแตกต่างกัน

เกณฑ์ในการทดสอบสมมติฐาน คือ หากการทดสอบความแตกต่างของค่าเฉลี่ยระหว่างกลุ่มตัวอย่างมากกว่า 2 กลุ่ม (One-way ANOVA) ได้ค่าความน่าจะเป็น (p-value) มากกว่าค่านัยสำคัญทาง

สถิติ (α = 0.05) จะยอมรับสมมติฐาน H_0 แต่หากได้ค่าความน่าจะเป็น (p-value) น้อยกว่าหรือเท่ากับค่านัยสำคัญทางสถิติ (α = 0.05) จะปฏิเสธสมมติฐาน H_0 และยอมรับสมมติฐาน H_1

4.4 การใช้โปรแกรมคอมพิวเตอร์วิเคราะห์ข้อมูล

- เปิดแฟ้มข้อมูล **File -> Open** -> Employee data.sav
- เลือกเมนู **Analyze -> Compare Means -> One-Way ANOVA...**
  - **Dependent List:** เลือกตัวแปรแบบค่าต่อเนื่อง 1 ตัวแปร หรือมากกว่า 1 ตัวแปร จากรายการตัวแปรด้านซ้าย
  - **Factor:** เลือกตัวแปรแบบค่าจัดประเภท 1 ตัวแปร จากรายการตัวแปรด้านซ้าย
  - ในกลุ่มคำสั่ง **Statistics** เลือกค่าสถิติ **Descriptives**
  - กดปุ่ม **OK**



4.5 ผลการวิเคราะห์

Descriptives

	Employment Category	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
Previous Experience (months)	Clerical	363	85.04	95.27	5.00	75.20	94.87	.00	476.00
	Custodial	27	298.11	101.43	19.52	257.99	338.23	144.00	460.00
	Manager	84	77.62	73.26	7.99	61.72	93.52	3.00	285.00
	Total	474	95.86	104.59	4.80	86.42	105.30	.00	476.00

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Previous Experience (months)	Between Groups	1174907	2	587453.4	69.19	.000
	Within Groups	3998900	471	8490.23		
	Total	5173807	473			

4.6 การสร้างตารางแสดงผลการทดสอบ

ตารางที่ (#) แสดงผลการทดสอบการเปรียบเทียบค่าเฉลี่ยประสบการณ์ก่อนทำงานของพนักงานระหว่างตำแหน่งงาน

ตำแหน่งงาน	จำนวน	ค่าเฉลี่ย	S.D.	F	df	Sig.
เจ้าหน้าที่ปฏิบัติงาน	363	85.04	95.27	69.19	2	0.00
หัวหน้างาน	27	298.11	101.43			
ผู้จัดการ	84	77.62	73.26			
รวม	474	95.86	104.59			

4.7 การแปลผลและสรุป

จากตารางที่ (#) ผลการทดสอบแสดงให้เห็นว่า ค่านัยความน่าจะเป็น (Sig.) = 0.000 มีค่าน้อยกว่าค่านัยสำคัญทางสถิติที่กำหนดไว้ (α = 0.05) จึงปฏิเสธสมมติฐาน H_0 และยอมรับสมมติฐาน H_1

สรุป กลุ่มตัวอย่างที่มีตำแหน่งงานแตกต่างกันมีค่าเฉลี่ยของประสบการณ์ก่อนทำงานแตกต่างกัน โดยหัวหน้างานมีค่าเฉลี่ยของประสบการก่อนทำงานมากที่สุด รองลงมา คือ เจ้าหน้าที่ปฏิบัติงาน และผู้จัดการ ตามลำดับ

จากการทดสอบสมมติฐานพบว่า กลุ่มตัวอย่างที่มีตำแหน่งงานแตกต่างกันมีค่าเฉลี่ยของประสบการณ์ก่อนทำงานแตกต่างกัน นั่นแสดงว่า มีกลุ่มตัวอย่างที่มีตำแหน่งงานแตกต่างกันอย่างน้อย 1 คู่ ที่มีค่าเฉลี่ยของประสบการณ์ก่อนทำงานแตกต่างกัน

เนื่องจากกลุ่มตัวอย่างพนักงานที่มีตำแหน่งงานแตกต่างกันแต่ละกลุ่มมีจำนวนไม่เท่ากัน จึงทำการทดสอบความแปรปรวนระหว่างกลุ่มด้วยสถิติของเลวีน (Levene Statistic) เพื่อตัดสินใจเลือกการวิเคราะห์เปรียบเทียบความแตกต่างของค่าเฉลี่ย (Post Hoc Tests) ของประสบการณ์ก่อนการทำงานของพนักงานที่มีตำแหน่งงานแตกต่างกัน โดยมีสมมติฐานดังนี้

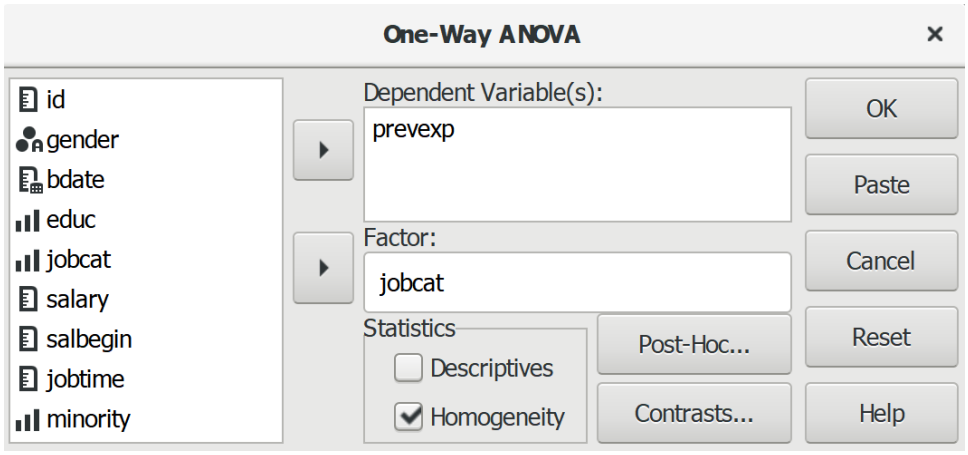
H_0: ความแปรปรวนของค่าเฉลี่ยประสบการณ์ก่อนการทำงานของพนักงานที่มีตำแหน่งงานแตกต่างกันเท่ากัน

H_1: ความแปรปรวนของค่าเฉลี่ยประสบการณ์ก่อนการทำงานของพนักงานที่มีตำแหน่งงานแตกต่างกันไม่เท่ากัน

เกณฑ์ในการทดสอบสมมติฐาน คือ หากการทดสอบความแปรปรวนด้วยวิธีของเลวีน (Levene's Test) ได้ค่าความน่าจะเป็น (p-value) มากกว่าค่านัยสำคัญทางสถิติ (α = 0.05) จะยอมรับสมมติฐาน H_0 แต่หากได้ค่าความน่าจะเป็น (p-value) น้อยกว่าหรือเท่ากับค่านัยสำคัญทางสถิติ (α = 0.05) จะปฏิเสธสมมติฐาน H_0 และยอมรับสมมติฐาน H_1

4.7.1 การทดสอบความแปรปรวน

- เลือกเมนู **Analyze -> Compare Means -> One-Way ANOVA...**
  - **Dependent List:** เลือกตัวแปรแบบค่าต่อเนื่อง 1 ตัวแปร หรือมากกว่า 1 ตัวแปร จากรายการตัวแปรด้านซ้าย
  - **Factor:** เลือกตัวแปรแบบค่าจัดประเภท 1 ตัวแปร ใส่ในช่อง
  - ในกลุ่มคำสั่ง **Statistics** เลือกค่าสถิติ **Homogeneity**
  - กดปุ่ม **OK**



4.7.1.1 ผลการทดสอบความแปรปรวน

Test of Homogeneity of Variances

	Levene Statistic	df1	df2	Sig.
Previous Experience (months)	2.54	2	471	.080

4.7.1.2 การสร้างตารางแสดงผลการทดสอบความแปรปรวน

ตารางที่ (#) แสดงผลการทดสอบความแปรปรวนของค่าเฉลี่ยประสบการณ์ก่อนการทำงานของพนักงานที่มีตำแหน่งงานแตกต่างกัน

Levene Statistic	df1	df2	Sig.
2.54	2	471	.080

จากตารางที่ (#) ผลการทดสอบแสดงให้เห็นว่า ค่าความน่าจะเป็น (Sig.) = 0.80 มีค่ามากกว่าค่านัยสำคัญทางสถิติที่กำหนดไว้ (α = 0.05) จึงยอมรับสมมติฐาน H_0

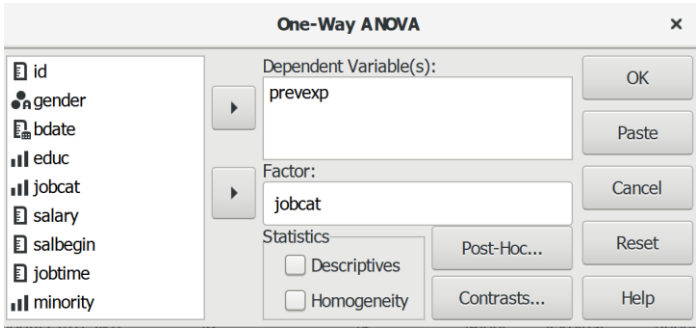
สรุป ความแปรปรวนของค่าเฉลี่ยประสบการณ์ก่อนการทำงานของพนักงานที่มีตำแหน่งงานแตกต่างกันเท่ากัน

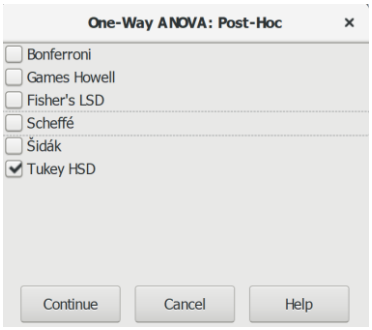
4.7.2 การทดสอบเปรียบเทียบค่าเฉลี่ยรายคู่หลังการทดสอบรวม

- เลือกเมนู **Analyze -> Compare Means -> One-Way ANOVA...**
  - **Dependent List:** เลือกตัวแปรแบบค่าต่อเนื่อง 1 ตัวแปร หรือมากกว่า 1 ตัวแปร จากรายการตัวแปรด้านซ้าย
  - **Factor:** เลือกตัวแปรแบบค่าจัดประเภท 1 ตัวแปร จากรายการตัวแปรด้านซ้าย
  - กดปุ่ม **Post-Hoc...** (ทดสอบค่าเฉลี่ยรายคู่)

    - เลือกค่าสถิติ Bonferroni หรือ Fisher's LSD หรือ Scheffé, Šídák หรือ Turkey HSD (ค่าความแปรปรวนของทั้งสองกลุ่มเท่ากัน)[18]

    - เลือกค่าสถิติ Games Howell (ค่าความแปรปรวนของทั้งสองกลุ่มไม่เท่ากัน)
  - กดปุ่ม **Continue**
  - กดปุ่ม **OK**





จากการทดสอบความแปรปรวนของค่าเฉลี่ยประสบการณ์ก่อนทำงานระหว่างกลุ่มพนักงานที่มีตำแหน่งแตกต่างกัน พบว่ามีความแปรวนเท่ากัน จึงทำการทดสอบความแตกต่างของค่าเฉลี่ยรายคู่ (Post Hoc Tests) ด้วยวิธีของทูกี้ (Tukey HSD: Tukey Honestly Significant Difference) โดยมีสมมติฐานดังนี้

H_0: ค่าเฉลี่ยของประสบการณ์ก่อนทำงานไม่แตกต่างกัน

H_1: มีค่าเฉลี่ยของประสบการณ์ก่อนทำงานแตกต่างกันอย่างน้อย 1 คู่

[18] ค่าสถิติที่นิยมใช้กันอย่างแพร่หลาย คือ Scheffé เพราะเป็นการทดสอบเชิงอนุรักษ์ (ค่านัยสำคัญทางสถิติของอัตราความคลาดเคลื่อนน้อยกว่า 0.05) มีโอกาสทำให้เกิดความผิดพลาดชนิดที่ 1 (Type I Error) น้อยกว่าสถิติเชิงก้าวหน้า (ค่านัยสำคัญทางสถิติของอัตราความคลาดเคลื่อนมากกว่า 0.05) เหมาะกับข้อมูลที่มีขนาดตัวอย่างเท่ากันและไม่เท่ากัน และมีการเปรียบเทียบรายคู่มากกว่า 3 กลุ่ม แต่หากทำการเปรียบเทียบรายคู่ 2-3 กลุ่ม มีขนาดกลุ่มตัวอย่างเท่ากันและไม่เท่ากัน นิยมใช้ Tukey HSD เพราะจะใช้การทดสอบเชิงอนุรักษ์

เกณฑ์ในการทดสอบสมมติฐาน คือ หากการทดสอบความแตกต่างของค่าเฉลี่ย (Mean difference) ได้ค่าความน่าจะเป็น (p-value) มากกว่าค่านัยสำคัญทางสถิติ (α = 0.05) จะยอมรับสมมติฐาน H_0 แต่หากได้ค่าความน่าจะเป็น (p-value) น้อยกว่าหรือเท่ากับค่านัยสำคัญทางสถิติ (α = 0.05) จะปฏิเสธสมมติฐาน H_0 และยอมรับสมมติฐาน H_1

4.7.2.1 ผลการทดสอบเปรียบเทียบค่าเฉลี่ยรายคู่

Multiple Comparisons (Previous Experience (months))

						95% Confidence Interval	
	(J) Family	(J) Family	Mean Difference (I - J)	Std. Error	Sig.	Lower Bound	Upper Bound
Tukey HSD	Clerical	Custodial	-213.07	18.38	.000	-256.29	-169.86
		Manager	7.42	11.16	.784	-18.81	33.65
	Custodial	Clerical	213.07	18.38	.000	169.86	256.29
		Manager	220.49	20.38	.000	172.57	268.42
	Manager	Clerical	-7.42	11.16	.784	-33.65	18.81
		Custodial	-220.49	20.38	.000	-268.42	-172.57

4.7.2.2 การสร้างตารางแสดงผลการทดสอบค่าเฉลี่ยรายคู่

ตารางที่ (#) แสดงผลการทดสอบการเปรียบเทียบค่าเฉลี่ยของประสบการณ์ก่อนทำงานระหว่างตำแหน่งงาน

ประสบการณ์ก่อนทำงาน(ค่าเฉลี่ย)	เจ้าหน้าที่ปฏิบัติงาน (85.04)	หัวหน้างาน (298.11)	ผู้จัดการ (77.62)
เจ้าหน้าที่ปฏิบัติงาน (85.04)	-	-213.073*	7.420
หัวหน้างาน (298.11)		-	220.492*
ผู้จัดการ (77.62)			-

* มีค่าความน่าจะเป็น (Sig.) = 0.000

จากตารางที่ (#) ผลการทดสอบแสดงให้เห็นว่า การเปรียบเทียบค่าเฉลี่ยประสบการณ์ก่อนทำงานรายคู่ระหว่างเจ้าหน้าที่ปฏิบัติงานกับหัวหน้างาน และหัวหน้างานกับผู้จัดการ มีค่าความน่าจะเป็น (Sig.) = 0.000 น้อยกว่าค่านัยสำคัญทางสถิติที่กำหนดไว้ (α = 0.05) จึงปฏิเสธสมมติฐาน H_0 และยอมรับสมมติฐาน H_1

สรุป มีกลุ่มตัวอย่างที่มีค่าเฉลี่ยของประสบการณ์ก่อนทำงานแตกต่างกันอย่างน้อย 1 คู่ ตำแหน่งงานที่มีค่าเฉลี่ยของประสบการณ์ก่อนทำงานแตกต่างกัน คือ เจ้าหน้าที่ปฏิบัติงานกับหัวหน้างาน และหัวหน้างานกับผู้จัดการ ส่วนเจ้าหน้าที่ปฏิบัติงานกับผู้จัดการ มีค่าเฉลี่ยของประสบการณ์ก่อนทำงานไม่แตกต่างกัน

ตัวอย่างเพิ่มเติม: การทดสอบ การแปลผลและสรุป

#1

ตารางที่ (#) แสดงผลการทดสอบเปรียบเทียบค่าเฉลี่ยประสบการณ์ก่อนทำงานของพนักงาน

ตำแหน่งงาน	จำนวน	ค่าเฉลี่ย	S.D.	F	Df	Sig.
เจ้าหน้าที่ปฏิบัติงาน	363	85.04	95.27	69.19	2	0.00
หัวหน้างาน	27	298.11	101.43			
ผู้จัดการ	84	77.62	73.26			
รวม	474	95.86	104.59			

Levene Statistic = 2.54, df1=2, df2=471, Sig.=0.08

จากตารางที่ (#) ผลการทดสอบแสดงให้เห็นว่า กลุ่มตัวอย่างที่เป็นพนักงานตำแหน่งหัวหน้างานมีประสบการณ์ก่อนทำงานมากที่สุด รองลงมา คือ เจ้าหน้าที่ปฏิบัติงาน และผู้จัดการ ตามลำดับ

สรุป พนักงานที่มีตำแหน่งงานแตกต่างกันมีประสบการณ์การก่อนทำงานแตกต่างกันอย่างมีนัยสำคัญทางสถิติ โดยมีความแปรปรวนของค่าเฉลี่ยของประสบการณ์ก่อนทำงานไม่แตกต่างกัน ดังนั้นจึงทำการทดสอบความแตกต่างของค่าเฉลี่ยรายคู่ (Post Hoc Tests) ด้วยวิธีของเชฟเฟ่ (Scheffe's Method)

ตารางที่ (#) แสดงผลการทดสอบเปรียบเทียบค่าเฉลี่ยประสบการณ์ก่อนทำงานของพนักงานที่มีตำแหน่งแตกต่างกัน

ประสบการณ์ก่อนทำงาน (ค่าเฉลี่ย)	เจ้าหน้าที่ปฏิบัติงาน (85.04)	หัวหน้างาน (298.11)	ผู้จัดการ (77.62)
เจ้าหน้าที่ปฏิบัติงาน (85.04)	-	-213.073*	7.420
หัวหน้างาน (298.11)		-	220.492*
ผู้จัดการ (77.62)			-

* มีค่าความน่าจะเป็น (Sig.) = 0.000

จากตารางที่ (#) ผลการทดสอบแสดงให้เห็นว่า กลุ่มตัวอย่างที่เป็นพนักงานตำแหน่งเจ้าหน้าที่ปฏิบัติงานมีประสบการณ์ก่อนทำงานโดยเฉลี่ยแตกต่าง (น้อยกว่า) กับตำแหน่งหัวหน้างาน และพนักงานตำแหน่งหัวหน้างานมีประสบการณ์ก่อนทำงานโดยเฉลี่ยแตกต่าง (มากกว่า) กับตำแหน่งผู้จัดการ อย่างมีนัยสำคัญทางสถิติ ส่วนกลุ่มตัวอย่างพนังงานตำแหน่งเจ้าหน้าที่ปฏิบัติงานมีประสบการณ์ก่อนทำงานไม่แตกต่างกับตำแหน่งผู้จัดการ หรือกล่าวได้ว่า กลุ่มตัวอย่างพนักงานตำแหน่งหัวหน้างานมีประสบการณ์ก่อนทำงานโดยเฉลี่ยมากกว่าตำแหน่งเจ้าหน้าที่ปฏิบัติงานและผู้จัดการ

#2

ตารางค่าเฉลี่ยเงินเดือนปัจจุบันของพนักงาน

ตำแหน่งงาน	จำนวน	ค่าเฉลี่ยของเงินเดือนปัจจุบัน (ดอลลาร์สหรัฐ)	S.D.
เจ้าหน้าที่ปฏิบัติงาน	363	27,838.54	7,567.995
หัวหน้างาน	27	30,938.89	2,114.616
ผู้จัดการ	84	63,977.80	18,244.776
รวม	474	34,419.57	17,075.661

F = 434.481, df = 2, Sig. = 0.000, Levene Statistic = 59.733, df1=2, df2=471, Sig.=0.000

จากตารางที่ (#) ผลการทดสอบแสดงให้เห็นว่า กลุ่มตัวอย่างที่เป็นพนักงานจำนวน 474 คน ตำแหน่งงานที่มีเงินเดือนปัจจุบันเฉลี่ยสูงสุดคือ ตำแหน่งผู้จัดการ รองลงมาเป็น หัวหน้างาน และตำแหน่งที่มีเงินเดือนปัจจุบันเฉลี่ยต่ำที่สุด คือ ตำแหน่งเจ้าหน้าที่ปฏิบัติงาน จากการทดสอบเปรียบค่าเฉลี่ยของเงินเดือนปัจจุบันด้วยวิธีการทางสถิติ พบว่า พนักงานที่มีตำแหน่งงานแตกต่างกันอย่างน้อย 1 คู่ จะมีเงินเดือนปัจจุบันเฉลี่ยแตกต่างกัน อย่างมีนัยสำคัญทางสถิติ และจากการทดสอบความแปรปรวนระหว่างกลุ่มพบว่า ค่าเฉลี่ยของเงินเดือนปัจจุบันมีความแปรปรวนแตกต่างกัน จึงทำการทดสอบความแตกต่างของค่าเฉลี่ยรายคู่ (Post Hoc Tests) ด้วยวิธีของทามเฮน (Tamhane's Method)

ตารางที่ (#) แสดงผลการทดสอบเปรียบเทียบค่าเฉลี่ยเงินเดือนปัจจุบันของพนักงานที่มีตำแหน่งแตกต่างกัน

เงินเดือนปัจจุบัน (ค่าเฉลี่ย)	เจ้าหน้าที่ปฏิบัติงาน ($27,838.54)	หัวหน้างาน ($30,938.89)	ผู้จัดการ ($63,977.80)
เจ้าหน้าที่ปฏิบัติงาน ($27,838.54)	-	-$3,100.349*	-$36,139.258*
หัวหน้างาน ($30,938.89)		-	-$33,038.909*
ผู้จัดการ ($63,977.80)			-

* มีค่าความน่าจะเป็น (Sig.) = 0.000

จากตารางที่ (#) ผลการทดสอบแสดงให้เห็นว่า กลุ่มตัวอย่างที่เป็นพนักงานตำแหน่งงานเป็นเจ้าหน้าที่ปฏิบัติงาน มีเงินเดือนปัจจุบันเฉลี่ยน้อยกว่า กลุ่มตัวอย่างพนักงานที่มีตำแหน่งงานเป็นหัวหน้างานและผู้จัดการ และกลุ่มตัวอย่างพนักงานที่มีตำแหน่งงานเป็นหัวหน้างานมีเงินเดือนปัจจุบันเฉลี่ยน้อยกว่า กลุ่มตัวอย่างพนักงานที่มีตำแหน่งงานเป็นผู้จัดการ อย่างมีนัยสำคัญทางสถิติ ดังนั้นจึงสรุปได้ว่า กลุ่มตัวอย่างพนักงานที่มีตำแหน่งงานแตกต่างกัน มีเงินเดือนปัจจุบันเฉลี่ยแตกต่างกันทุกตำแหน่ง

5. การทดสอบความสัมพันธ์รายคู่แบบเพียร์สัน

การทดสอบความสัมพันธ์รายคู่แบบเพียร์สัน (Pearson's Correlation) คือ การวัดความสัมพันธ์เชิงเส้นตรง (linear relationship) จากค่าสัมประสิทธิ์สหสัมพันธ์ (correlation coefficient) ระหว่างคู่ตัวแปร (2 ตัวแปร) ที่เป็นตัวแปรแบบค่าต่อเนื่อง (continuous variables) จะมีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติหรือไม่ เพื่ออธิบายความเข้ม/ขนาด (strength) และทิศทาง (direction)

ความสัมพันธ์เชิงเส้นตรง คือ ลักษณะของความสัมพันธ์ระหว่างตัวแปรเมื่อตัวแปรหนึ่งมีการเปลี่ยนแปลงจะมีความเกี่ยวข้องกับสัดส่วนการเปลี่ยนแปลงในอีกตัวแปรหนึ่ง

การทดสอบแบบนี้อาจมีชื่อเรียกในชื่ออื่น เช่น Pearson correlation coefficient, Pearson's correlation และ Pearson product-moment correlation (PPMC)

การแปรผลค่าความสัมพันธ์ขึ้นอยู่กับบริบท (context) และจุดประสงค์/ความต้องการ (purposes) ค่า r = 0.90 อาจสูงมากในสาขาวิชาด้านสังคมศาสตร์ แต่อาจมีค่าต่ำมากในสาขาวิชาฟิสิกส์ที่มีการใช้เครื่องมือวัดที่มีคุณภาพสูง (Araújo Jr, A.H.A. et al., 2011) ค่าที่คำนวณได้อาจมีค่าเป็น + หรือ – แสดงถึงความสัมพันธ์ที่แตกต่างกันดังนี้

ค่า r เป็น + แสดงว่า ตัวแปรทั้งสองตัวแปร (ตัวแปรอิสระและตัวแปรตาม) มีความสัมพันธ์ในทิศทางเดียวกัน

ค่า r เป็น – แสดงว่า ตัวแปรทั้งสองตัวแปร (ตัวแปรอิสระและตัวแปรตาม) มีความสัมพันธ์ในทิศทางตรงกันข้าม

ค่า r เป็น 0 แสดงว่า ตัวแปรทั้งสองตัวแปร (ตัวแปรอิสระและตัวแปรตาม) ไม่มีความสัมพันธ์กันเลย

ค่า r มีค่าเข้าใกล้ 1 แสดงว่า ตัวแปรทั้งสองตัวแปร (ตัวแปรอิสระและตัวแปรตาม) มีความสัมพันธ์กันมาก

ค่า r มีค่าเข้าใกล้ 0 แสดงว่า ตัวแปรทั้งสองตัวแปร (ตัวแปรอิสระและตัวแปรตาม) มีความสัมพันธ์กันน้อย

ตารางที่ 12.1 การแปลค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน

ค่าสัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficient)	
ค่าของความสัมพันธ์	**ความเข้มและทิศทางของความสัมพันธ์**
1.00 ถึง 0.80	มีความสัมพันธ์เชิงบวกระดับมากอย่างยิ่ง
0.79 ถึง 0.60	มีความสัมพันธ์เชิงบวกระดับมาก
0.59 ถึง 0.40	มีความสัมพันธ์เชิงบวกระดับปานกลาง
0.39 ถึง 0.20	มีความสัมพันธ์เชิงบวกระดับน้อย
0.19 ถึง 0.01	มีความสัมพันธ์เชิงบวกระดับน้อยมาก
0.00	ไม่มีความสัมพันธ์
-0.01 ถึง -0.19	มีความสัมพันธ์เชิงลบระดับน้อยมาก
-0.39 ถึง -0.20	มีความสัมพันธ์เชิงลบระดับน้อย
-0.59 ถึง -0.40	มีความสัมพันธ์เชิงลบระดับปานกลาง
-0.79 ถึง -0.60	มีความสัมพันธ์เชิงลบระดับมาก
-1.00 ถึง 0.-80	มีความสัมพันธ์เชิงลบระดับมากอย่างยิ่ง

ปรับปรุงมาจาก: Evans, J. D. (1996). Straightforward statistics for the behavioral sciences. Pacific Grove, CA: Brooks/Cole Publishing.

5.1 การทดสอบความสัมพันธ์แบบง่าย

การทดสอบความสัมพันธ์แบบง่าย (Simple correlation) คือ การทดสอบความสัมพันธ์ของข้อมูล 1 คู่ เช่น การวิเคราะห์หาความสัมพันธ์ระหว่างประสบการณ์ก่อนทำงานกับเงินเดือนเริ่มต้น

ตัวอย่างคำถามการวิจัย
จากการสำรวจข้อมูลของพนักงานในหน่วยงานต้องการทราบว่า ประสบการณ์ก่อนการทำงานกับเงินเดือนเริ่มต้นของพนักงานมีความสัมพันธ์กันหรือไม่

5.1.2 ข้อตกลงเบื้องต้น

ขั้นตอน	**ข้อตกลง**
การออกแบบการวิจัย	1. กลุ่มตัวอย่างได้มาจากประชากรที่ทราบจำนวน (know population) โดยการสุ่มตัวอย่างแบบสุ่มที่เป็นอิสระจากกัน (independence random sampling) 2. มีตัวแปรอิสระ (independent variable) 1 ตัว และตัวแปรตาม (dependent variable) 1 ตัว 3. ตัวแปรอิสระและตัวแปรตามเป็นตัวแปรแบบค่าต่อเนื่อง (continuous variable)
การวิเคราะห์ข้อมูล	1. ข้อมูลมีการแจกแจงแบบปกติทั้งสองตัวแปร (bivariate normal distribution) 2. ข้อมูลมีความคงที่ของความแปรปรวนของค่าความคาดเคลื่อน (homoscedasticity) ถ้า n = 30 หรือมากกว่า ไม่ต้องทดสอบ 3. ข้อมูลทั้ง 2 ตัวแปร ไม่มีค่าสูญหาย (missing values) 4. ข้อมูลมีความสัมพันธ์เชิงเส้นตรง (linear relationship)

5.1.3 รูปแบบการทดสอบตัวแปร

ตัวแปรอิสระ		**ตัวแปรตาม**
ค่าต่อเนื่อง	<->	ค่าต่อเนื่อง
ประสบการณ์ก่อนทำงาน (prevexp)		เงินเดือนเริ่มต้น (salbegin)

5.1.4 การกำหนดสมมติฐานและเกณฑ์การทดสอบ

ตารางที่ (#) แสดงค่าเฉลี่ยของประสบการณ์ก่อนทำงาน และเงินเดือนเริ่มต้น

ตัวแปร	จำนวน	ค่าเฉลี่ย	S.D.
ประสบการณ์ก่อนทำงาน (เดือน)	474	95.86	104.586
เงินเดือนเริ่มต้น (ดอลลาร์สหรัฐ)	474	17,016.09	7,870.638

จากตารางที่ (#) แสดงให้เห็นว่า พนักงาน 474 คน มีประสบการณ์ก่อนทำงานเฉลี่ย 95.86 เดือน และมีเงินเดือนเริ่มต้นเฉลี่ย 17,016.09 ดอลลาร์สหรัฐ

จากข้อมูลที่กล่าวมาน่าจะมีความสัมพันธ์ระหว่างตัวแปรประสบการณ์ก่อนทำงานและเงินเดือนเริ่มต้น ดังนั้นจึงทำการทดสอบความสัมพันธ์รายคู่โดยมีสมมติฐานดังนี้

H_0: ไม่มีความสัมพันธ์ระหว่างประสบการณ์ก่อนทำงานกับเงินเดือนเริ่มต้นของพนักงาน

H_1: มีความสัมพันธ์ระหว่างประสบการณ์ก่อนทำงานกับเงินเดือนเริ่มต้นของพนักงาน

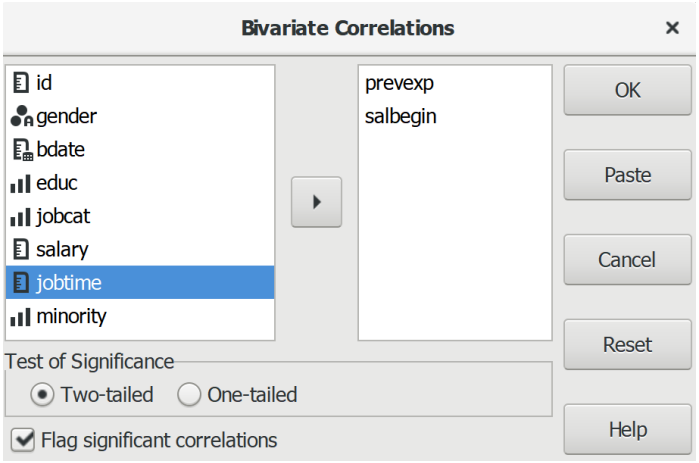
เกณฑ์ในการทดสอบสมมุติฐาน คือ หากการทดสอบความสัมพันธ์ด้วยวิธีของเพียร์สันได้ค่าความน่าจะเป็น (p-value) มากกว่าค่านัยสำคัญทางสถิติ (α = 0.05) จะยอมรับสมมติฐาน H_0 แต่หากได้ค่าความน่าจะเป็น (p-value) น้อยกว่าหรือเท่ากับค่านัยสำคัญทางสถิติ (α = 0.05) จะปฏิเสธสมมติฐาน H_0 และยอมรับสมมติฐาน H_1

5.1.5 การใช้โปรแกรมคอมพิวเตอร์วิเคราะห์ข้อมูล

- เปิดแฟ้มข้อมูล **File -> Open** -> Employee data.sav
- เลือกเมนู **Analyze -> Bivariate Correlation...**
  - เลือกตัวแปรค่าต่อเนื่อง (continuous variables) 2 ตัวแปร จากรายการตัวแปรในช่องด้านซ้าย -> ช่องด้านขวา
  - ในกลุ่มคำสั่ง **Test of Significance** เลือกประเภทการทดสอบนัยสำคัญ

    - Two-tailed (สมมติฐานแบบสองหาง $r = 0$)

    - One-tailed (สมมติฐานแบบหางเดียว $r > 0$ หรือ $r < 0$)
  - **Flag significant correlations** (แสดงเครื่องหมาย * หากมีความสัมพันธ์อย่างนัยสำคัญทางสถิติ) เลือก /
  - คลิกปุ่ม **OK**



5.1.6 ผลการวิเคราะห์

Correlations

		Previous Experience (months)	Beginning Salary
Previous Experience (months)	Pearson Correlation	1.000	.045
	Sig. (2-tailed)		.327
	N	474	474
Beginning Salary	Pearson Correlation	.045	1.000
	Sig. (2-tailed)	.327	
	N	474	474

5.1.7 การสร้างตารางแสดงผลการทดสอบ

ตารางที่ (#) แสดงผลการทดสอบความสัมพันธ์ระหว่างประสบการณ์ก่อนทำงานกับเงินเดือนเริ่มต้น

ตัวแปร	จำนวน	เงินเดือนเริ่มต้น (r)	Sig. (2-tailed)
ประสบการณ์ก่อนทำงาน (เดือน)	474	.05	0.327

5.1.8 การแปลผลและสรุป

จากการทดสอบสมมติฐานด้วยการวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สัน พบว่า ประสบการณ์ก่อนทำงานกับเงินเดือนเริ่มต้นของพนักงาน มีค่าความน่าจะเป็น (Sig.) = 0.327 มากกว่าค่านัยสำคัญทางสถิติที่กำหนดไว้ (α = 0.05) จึงยอมรับสมมติฐาน H_0

สรุป ไม่มีความสัมพันธ์ระหว่างประสบการณ์ก่อนทำงานกับเงินเดือนเริ่มต้นของพนักงาน

5.2 การทดสอบความสัมพันธ์แบบหพุ

การทดสอบความสัมพันธ์แบบพหุ (Multiple correlation) คือ การทดสอบความสัมพันธ์ของข้อมูลรายคู่มากกว่า 1 คู่ เช่น การวิเคราะห์หาความสัมพันธ์ระหว่างระดับการศึกษากับเงินเดือนเริ่มต้น ประสบการณ์ก่อนทำงานกับเงินเดือนเริ่มต้น

ตัวอย่างคำถามการวิจัย
จากข้อมูลการสำรวจพนักงานในหน่วยงานต้องการทราบว่า ระดับการศึกษาและประสบการณ์ก่อนทำงานของพนักงานมีความสัมพันธ์กับเงินเดือนเริ่มต้นหรือไม่ และเงินเดือนเริ่มต้นและระยะเวลาทำงานของพนักงานมีความสัมพันธ์กับเงินเดือนปัจจุบันหรือไม่

5.2.1 ข้อตกลงเบื้องต้น

ขั้นตอน	ข้อตกลง
การออกแบบการวิจัย	1. กลุ่มตัวอย่างได้มาจากประชากรที่ทราบจำนวน (know population) โดยการสุ่มตัวอย่างแบบสุ่มที่เป็นอิสระจากกัน (independence random sampling) 2. มีตัวแปรอิสระ (independent variable) อย่างน้อย 1 ตัว และตัวแปรตาม (dependent variable) อย่างน้อย 1 ตัว 3. ตัวแปรอิสระและตัวแปรตามเป็นตัวแปรแบบค่าต่อเนื่อง (continuous variable)
การวิเคราะห์ข้อมูล	1. ข้อมูลมีการแจกแจงแบบปกติทั้งสองตัวแปร (bivariate normal distribution) 2. ข้อมูลมีความคงที่ของความแปรปรวนของค่าความคาดเคลื่อน (homoscedasticity) ถ้า n = 30 หรือมากกว่า ไม่ต้องทดสอบ 3. ข้อมูลมีความสัมพันธ์เชิงเส้นตรง (linear relationship)

5.2.2 รูปแบบการทดสอบตัวแปร

ตัวแปรอิสระ		ตัวแปรตาม
ค่าต่อเนื่อง		ค่าต่อเนื่อง
ระดับการศึกษา: ปี (educ)	<->	เงินเดือนเริ่มต้น: ดอลลาร์สหรัฐ (salbegin)
ประสบการณ์ก่อนทำงาน: เดือน (prevexp)		เงินเดือนเริ่มต้น: ดอลลาร์สหรัฐ (salbegin)
เงินเดือนเริ่มต้น: ดอลลาร์สหรัฐ (salbegin)		เงินเดือนปัจจุบัน: ดอลลาร์สหรัฐ (salary)
ระยะเวลาทำงาน: เดือน (jobtime)		เงินเดือนปัจจุบัน: ดอลลาร์สหรัฐ (salary)

5.2.3 การกำหนดสมมติฐานและเกณฑ์การทดสอบ

ตารางที่ (#) แสดงค่าเฉลี่ยระดับการศึกษา เงินเดือนเริ่มต้น เงินเดือนปัจจุบัน ระยะเวลาทำงาน และประสบการณ์ก่อนทำงาน

ตัวแปร	จำนวน (คน)	ค่าเฉลี่ย	S.D.
ระดับการศึกษา (ปี)	474	13.49	2.89
เงินเดือนเริ่มต้น (ดอลลาร์สหรัฐ)	474	17,016.09	7,870.64
เงินเดือนปัจจุบัน (ดอลลาร์สหรัฐ)	474	34,419.57	17,075.66
ระยะเวลาทำงาน (เดือน)	474	81.11	10.06
ประสบการณ์ก่อนทำงาน (เดือน)	474	95.86	104.59

จากตารางที่ (#) แสดให้เห็นว่า พนักงาน 474 คน มีการศึกษาเฉลี่ย 13.49 ปี มีเงินเดือนเริ่มต้นเฉลี่ย 17,016.09 ดอลลาร์สหรัฐ มีเงินเดือนปัจจุบันเฉลี่ย 34,419.57 ดอลลาร์สหรัฐ มีระยะเวลาทำงานเฉลี่ย 81.11 เดือน และมีประสบการณ์ก่อนทำงานเฉลี่ย 65.86 เดือน

จากข้อมูลที่กล่าวมาน่าจะมีความสัมพันธ์ระหว่างตัวแปรบางคู่ ดังนั้นจึงทำการทดสอบความสัมพันธ์รายคู่โดยมีสมมติฐานดังนี้

H_1: มีความสัมพันธ์ระหว่างระดับการศึกษากับเงินเดือนเริ่มต้นของพนักงาน

H_2: มีความสัมพันธ์ระหว่างประสบการณ์ก่อนทำงานกับเงินเดือนเริ่มต้นของพนักงาน

H_3: มีความสัมพันธ์ระหว่างเงินเดือนเริ่มต้นกับเงินเดือนปัจจุบันของพนักงาน

H_4: มีความสัมพันธ์ระหว่างระยะเวลาทำงานกับเงินเดือนปัจจุบันของพนักงาน

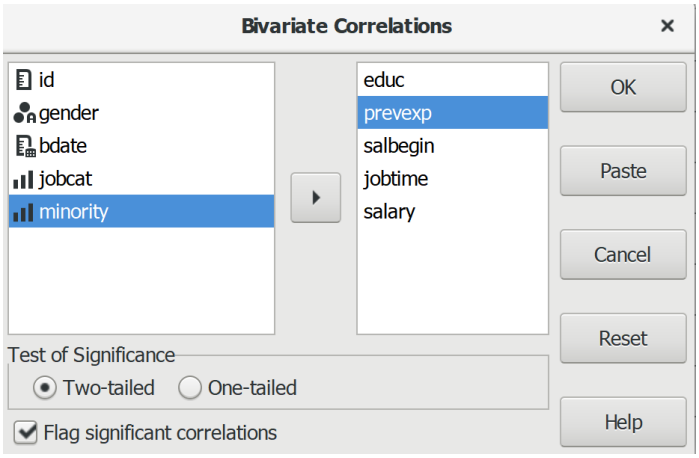
เกณฑ์ในการทดสอบสมมุติฐาน คือ หากการทดสอบความสัมพันธ์ด้วยวิธีของเพียร์สัน ได้ค่าความน่าจะเป็น (p-value) มากกว่าค่านัยสำคัญทางสถิติ (α = 0.05) จะปฏิเสธสมมติฐาน H_1 , H_2 , H_3 และ H_4 แต่หากได้ค่าความน่าจะเป็น (p-value) น้อยกว่าหรือเท่ากับค่านัยสำคัญทางสถิติ (α = 0.05) จะยอมรับสมมติฐาน H_1 , H_2 , H_3 และ H_4

5.2.4 การใช้โปรแกรมคอมพิวเตอร์วิเคราะห์ข้อมูล

- เปิดแฟ้มข้อมูล **File -> Open** -> Employee data.sav
- เมนู **Analyze -> Bivariate Correlation...**
  - เลือกตัวแปรค่าต่อเนื่อง (continuous variables) 3 ตัวแปร หรือมากกว่า จากรายการตัวแปรในช่องด้านซ้ายใส่ในช่องด้านขวา
  - ในกลุ่มคำสั่ง **Test of Significance** เลือกประเภทการทดสอบนัยสำคัญ

    - Two-tailed (สมมติฐานแบบสองหาง)

    - One-tailed (สมมติฐานแบบหางเดียว)
  - **Flag significant correlations** (แสดงเครื่องหมาย * หากมีความสัมพันธ์อย่างนัยสำคัญทางสถิติ) เลือก /
  - คลิกปุ่ม **OK**



5.2.5 ผลการวิเคราะห์

Correlations

		Educational Level (years)	Previous Experience (months)	Beginning Salary	Months since Hire	Current Salary
Educational Level (years)	Pearson Correlation	1.000	-.252[a]	.633[a]	.047	.661[a]
	Sig. (2-tailed)		.000	.000	.303	.000
	N	474	474	474	474	474
Previous Experience (months)	Pearson Correlation	-.252[a]	1.000	.045	.003	-.097[a]
	Sig. (2-tailed)	.000		.327	.948	.034
	N	474	474	474	474	474
Beginning Salary	Pearson Correlation	.633[a]	.045	1.000	-.020	.880[a]
	Sig. (2-tailed)	.000	.327		.668	.000
	N	474	474	474	474	474
Months since Hire	Pearson Correlation	.047	.003	-.020	1.000	.084
	Sig. (2-tailed)	.303	.948	.668		.067
	N	474	474	474	474	474
Current Salary	Pearson Correlation	.661[a]	-.097[a]	.880[a]	.084	1.000
	Sig. (2-tailed)	.000	.034	.000	.067	
	N	474	474	474	474	474

a. Significant at .05 level

5.2.6 การสร้างตารางแสดงผลการทดสอบ

ตารางที่ (#) แสดงผลการทดสอบความสัมพันธ์ระหว่างระดับการศึกษา ประสบการณ์ก่อนทำงาน เงินเดือนเริ่มต้น ระยะเวลาทำงาน และเงินเดือนปัจจุบัน

ความสัมพันธ์ระหว่างตัวแปร		จำนวน	r	Sig. (2-tailed)
ระดับการศึกษา	เงินเดือนเริ่มต้น	474	.63	0.000
ประสบการณ์ก่อนทำงาน	เงินเดือนเริ่มต้น	474	.05	0.327
เงินเดือนเริ่มต้น	เงินเดือนปัจจุบัน	474	.88	0.000
ระยะเวลาทำงาน	เงินเดือนปัจจุบัน	474	.08	0.067

5.2.7 การแปลผลและสรุป

จากตารางที่ (#) ผลการทดสอบแสดงให้เห็นว่า ระดับการศึกษากับเงินเดือนเริ่มต้นของพนักงานและเงินเดือนเริ่มต้นและเงินเดือนปัจจุบัน มีค่าความน่าจะเป็น (Sig.) = 0.00 น้อยกว่าค่านัยสำคัญทางสถิติที่กำหนดไว้ (α = 0.05) ส่วนประสบการณ์ก่อนทำงานและเงินเดือนเริ่มต้น มีค่านัยสำคัญทางสถิติ (Sig.) = 0.327 มากกว่าค่านัยสำคัญทางสถิติที่กำหนดไว้ (α = 0.05) และระยะเวลาทำงานกับเงินเดือนปัจจุบัน มีค่านัยสำคัญทางสถิติ (Sig.) = 0.067 มากกว่าค่านัยสำคัญทางสถิติที่กำหนดไว้ (α = 0.05) จึงยอมรับสมมติฐาน H_1 และ H_3 แต่ปฏิเสธสมมติฐาน H_2 และ H_4

สรุป การศึกษาของพนักงานมีความสัมพันธ์เชิงบวกระดับปานกลางกับเงินเดือนเริ่มต้น เงินเดือนเริ่มต้นของพนักงานมีความสัมพันธ์เชิงบวกระดับปานกลางกับเงินเดือนปัจจุบันของพนักงาน (พนักงานที่มีระดับการศึกษาสูงจะได้รับเงินเดือนเริ่มต้นสูงกว่าพนักงานที่มีการศึกษาต่ำกว่า และพนักงานที่ได้รับเงินเดือนเริ่มต้นสูงจะได้รับเงินเดือนปัจจุบันสูงกว่าพนักงานที่ได้รับเงินเดือนเริ่มต้นต่ำกว่า)

สรุป

การทดสอบแบบอิงค่าพารามิเตอร์มีค่าตัวเลขทางประชากรที่ใช้ในการอ้างอิงไปสู่คุณลักษณะ (characteristics) ของประชากร ก่อนใช้การทดสอบแบบอิงค่าพารามิเตอร์ต้องทำการตรวจสอบข้อตกลงเบื้องต้นของการทดสอบแต่ละวิธี การละเมิดข้อตกลงบางครั้งอาจไม่มีผลต่อข้อสรุปหรือผลการวิเคราะห์แต่อย่างไร แต่บางครั้งอาจมีผลต่อการแปลความหมายเป็นอย่างมากและอาจทำให้ได้ข้อสรุปที่ไม่ถูกต้อง

แบบฝึกหัดท้ายบท

ให้ผู้เรียนเปิดแฟ้มข้อมูล 1991 US. General Social Survey.sav แล้วประมวลผลข้อมูล เพื่อทดสอบและแปลผล ดังนี้

1. การทดสอบค่าเฉลี่ย 1 กลุ่ม และ 2 กลุ่ม

1.1 วิเคราะห์และอธิบายผล โดยการทดสอบเปรียบเทียบว่า กลุ่มตัวอย่างมีพี่น้อง (sib) มากกว่าหรือน้อยกว่า 4 คน

1.2 วิเคราะห์และอธิบายผล โดยการทดสอบเปรียบเทียบว่า กลุ่มตัวอย่างที่อยู่ภาค (region) ตะวันออกกับตะวันตก (สร้างตัวแปรใหม่เป็นตัวแปร region_group) มีพี่น้อง (sibs) แตกต่างกันหรือไม่

1.3 วิเคราะห์และอธิบายผล โดยการทดสอบเปรียบเทียบว่า การศึกษาของบิดา (paeduc) และมารดา (maeduc) แตกต่างกันหรือไม่

2. การทดสอบค่าเฉลี่ยของกลุ่มตัวอย่างมากกว่า 2 กลุ่ม

วิเคราะห์เปรียบค่าเฉลี่ย ทดสอบสมมติฐาน และทดสอบเปรียบเทียบค่าเฉลี่ยรายคู่คะแนนความมีเกียรติในอาชีพ (prestg80) ของคนในแต่ละภูมิภาค (region)

3. สหสัมพันธ์แบบเพียร์สัน

3.1 วิเคราะห์และอธิบายผลความสัมพันธ์แบบง่ายระหว่างการศึกษาของผู้ตอบ (edu) และจำนวนพี่น้อง (sibs)

3.2 วิเคราะห์และอธิบายผลความสัมพันธ์แบบพหุระหว่างการศึกษาของผู้ตอบ (edu) กับจำนวนพี่น้อง (sibs) และจำนวนบุตร (childs)

บทที่ 13
การทดสอบข้อมูลแบบไม่อิงค่าพารามิเตอร์ (1 ตัวแปร)

ความมุ่งหมายของบทเรียน

ผู้เรียนสามารถใช้โปรแกรมประมวลผลทดสอบข้อมูลของตัวแปร 1 ตัวแปร ด้วยสถิติเชิงอ้างอิงแบบไม่อิงค่าพารามิเตอร์ และเขียนบรรยายผลการทดสอบได้

เนื้อหา

1. การทดสอบการแจกแจงของตัวอย่าง
2. การทดสอบสัดส่วนของตัวอย่าง

กิจกรรมและวิธีสอน

บรรยาย อธิบาย แสดงตัวอย่าง และทำแบบฝึกหัดท้ายบท

อุปกรณ์การสอน

เอกสารคำสอน ภาพเลื่อน คอมพิวเตอร์ โปรแกรม PSPP โปรแกรมสอนออนไลน์ (สอนแบบออนไลน์) และเครื่องฉาย (สอนในชั้นเรียน)

งานที่มอบหมาย

ทำแบบฝึกหัดท้ายบท

การวัดผลและประเมินผล

การเข้าชั้นเรียน สังเกตความสนใจจากการเรียน สอบถามความเข้าใจจากการบรรยาย และประเมินจากการทำแบบฝึกหัดท้ายบท

เนื้อหา

การทดสอบข้อมูลด้วยสถิติเชิงอ้างอิงหากข้อมูลมีลักษณะหนึ่งดังต่อไปนี้ คือ มีจำนวนตัวอย่างน้อยกว่า 30 ตัวอย่าง หรือเป็นกลุ่มตัวอย่างขนาดเล็ก (small sample size) ข้อมูลมีการแจกแจงไม่ปกติ ข้อมูลเป็นมาตราแบบจัดกลุ่มหรือจัดอันดับ ข้อมูลมีค่าผิดปกติ (outlier) ที่ไม่สามารถกำจัดออกไปได้ ควรตัดสินใจเลือกการทดสอบแบบไม่อิงค่าพารามิเตอร์ (non-parametric test) หรือการทดสอบแบบแจกแจงอิสระ (distribution-free test) การฝืนใช้การทดสอบแบบอิงค่าพารามิเตอร์โดยละเมิดข้อตกลงเบื้องต้นมีความเสี่ยงทำให้ผลลัพธ์และข้อสรุปที่ไม่ถูกต้องตามมา

การทดสอบข้อมูลแบบไม่อิงพารามิเตอร์ สมมติฐานไม่มีความเกี่ยวข้องกับค่าพารามิเตอร์ของประชากร (population parameters) เช่น สมมติฐานหลัก μ=50 เป็นการเปรียบเทียบข้อมูลของตัวอย่างที่ได้มาจากประชากรกับค่าอ้างอิงที่ใช้ในการทดสอบว่าเท่ากันหรือไม่เท่ากัน หรือ μ_1=μ_2 เป็นการเปรียบเทียบตำแหน่งของข้อมูลระหว่าง 2 กลุ่มตัวอย่างมีแนวโน้มเข้าสู่ส่วนกลาง (central tendency) เหมือนกันหรือแตกต่างกัน

แม้ว่าการทดสอบด้วยสถิติแบบไม่อิงค่าพารามิเตอร์จะเป็นอิสระจากข้อตกลง (assumption-free test) แต่มีข้อตกลงทั่วไป (common assumptions) ดังนี้ (Verma, & Abdel-Salam, 2019: 141-142)

1. ความเป็นสุ่มหรือความไม่เป็นแบบแผน (Randomness) คือ กลุ่มตัวอย่างได้มาจากประชากรแบบสุ่ม (random sampling) หรือเลือกแบบปราศจากความลำเอียง (bias)

2. ความเป็นอิสระ (Independence) คือ ค่าสังเกตแต่ละค่ามีความเป็นอิสระจากกัน (independence of observations) หรือค่าสังเกตได้มาจากแต่ละตัวอย่างที่มีความเป็นอิสระจากกัน ไม่ได้มาจากบุคคลที่เป็นคนเดียวกัน (independent sample)

การทดสอบแบบอิงค่าพารามิเตอร์มีวิธีการทดสอบข้อมูล 1 ตัวแปร คือ การทดสอบค่าเฉลี่ยของกลุ่มตัวอย่าง 1 กลุ่ม (One-sample t-test) ใช้สำหรับทดสอบกับค่าทดสอบ (value test) ที่เป็นค่าเฉลี่ยจากประชากรที่ทราบจำนวน (know populations) ส่วนการทดสอบแบบไม่อิงค่าพารามิเตอร์มีการทดสอบข้อมูล 1 ตัวแปรเช่นกัน แต่เป็นการทดสอบการแจกแจงของข้อมูล (มีการแจกแจงปกติหรือไม่ มี

การแจกแจงเป็นไปอย่างสุ่มหรือไม่) และการทดสอบสัดส่วน (เปรียบเทียบสัดส่วนระหว่างกลุ่มตัวอย่าง เช่น กลุ่มตัวอย่างมีสัดส่วนระหว่างเพศชายกับเพศหญิงเช่นเดียวกับสัดส่วนประชากรของประเทศหรือไม่)

1. การทดสอบการแจกแจงของข้อมูล

การทดสอบการแจกแจงข้อมูล (Distribution Tests) คือ การทดสอบลักษณะข้อมูลของกลุ่มตัวอย่างที่ได้มาจากประชากรมีการแจกแจงปกติ (normal distributions) หรือไม่ และมีการแจกเป็นไปอย่างสุ่ม (randomness) หรือไม่

1.1 การทดสอบการแจกแจงของข้อมูลแบบรันส์

การทดสอบการแจกแจงข้อมูลแบบรันส์ (Runs test/ Wald–Wolfowitz runs test) คือ การทดสอบข้อมูลภายในกลุ่ม (one group) และระหว่างกลุ่ม (two group) ของกลุ่มตัวอย่างที่มีลักษณะการจัดเรียง (sequence) มีการแจกแจงเป็นไปอย่างสุ่มหรือไม่ (randomness) โดยใช้ค่ามัธยฐาน (median) ค่าฐานนิยม (mode) ค่าเฉลี่ย (mean) และค่าที่ผู้ทดสอบกำหนดเอง (custom)

ตัวอย่างคำถามการวิจัย
จากการสำรวจข้อมูลของพนักงานในหน่วยงานต้องการทราบว่า ระดับการศึกษาของพนักงานมีการแจกแจกแบบสุ่มหรือไม่

1.1.1 ข้อตกลงเบื้องต้น

ขั้นตอน	**ข้อตกลง**
การออกแบบการวิจัย	1. ข้อมูลได้มาจากการสุ่มตัวอย่างแบบสุ่ม (random sampling) และเป็นอิสระจากกัน (independent sample) 2. ข้อมูลเป็นตัวแปรแบบสองประเภท (dichotomous variables)
การวิเคราะห์ข้อมูล	ข้อมูลประเภทตัวอักษร (string data) ต้องแปลงค่า (recode) ให้เป็นข้อมูลประเภทตัวเลข (numeric data)

1.1.2 ตัวแปรและวิธีการทดสอบ

ตัวแปร		การทดสอบ
ค่าต่อเนื่อง	->	การแจกแจงข้อมูล
การศึกษา: ปี (educ)		แจกแจงแบบสุ่ม

1.1.3 การกำหนดสมมติฐานและเกณฑ์การทดสอบ

จากข้อมูลที่ทำการสำรวจจากพนักงานในหน่วยงาน ผู้วิจัยต้องการทราบว่า ข้อมูลการศึกษาของพนักงานมีการแจกแจงแบบสุ่มหรือไม่ โดยมีสมมติฐานดังนี้

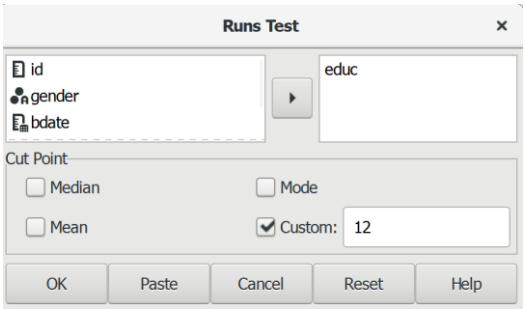
H_0: ข้อมูลการศึกษาของพนักงานมีการแจกแจงแบบสุ่ม

H_1: ข้อมูลการศึกษาของพนักงานมีการแจกแจงแบบไม่สุ่ม

เกณฑ์ในการทดสอบสมมุติฐาน คือ หากการทดสอบการแจกแจงข้อมูลแบบรันส์ (Runs Test) ได้ค่าความน่าจะเป็น (p-value) มากกว่าค่านัยสำคัญทางสถิติ (α = 0.05) จะยอมรับสมมติฐาน H_0 แต่หากได้ค่าความน่าจะเป็น (p-value) น้อยกว่าหรือเท่ากับค่านัยสำคัญทางสถิติ (α = 0.05) จะปฏิเสธสมมติฐาน H_0 และยอมรับสมมติฐาน H_1

1.1.4 การใช้โปรแกรมคอมพิวเตอร์วิเคราะห์ข้อมูล

- เปิดเมนู **File -> Open** -> Employee data.sav
- เลือกเมนู **Analyze -> Non-Parametric Statistics -> Runs ...**
  - **Test Variable List:** เลือกตัวแปรทดสอบที่เป็นตัวแปรแบบค่าจัดประเภทหรือตัวแปรแบบค่าต่อเนื่อง 1 ตัวแปร จากรายการตัวแปรด้านซ้าย
  - **Cut Point:** กำหนดกลุ่มวิธีการแบ่งกลุ่มในกลุ่มคำสั่ง
  - กดปุ่ม **OK**



1.1.5 ผลการวิเคราะห์

Runs Test

	Educational Level (years)
Test Value	12.00
Cases < Test Value	53
Cases ≥ Test Value	421
Total Cases	474
Number of Runs	97
Z	.43
Asymp. Sig. (2-tailed)	.667

1.1.6 การสร้างตารางแสดงผลการทดสอบ

ตารางผลการทดสอบรันส์แบบฐานนิยม (Mode)

	ระดับการศึกษา (ปี)
ค่าทดสอบ (Test Value)	12
จำนวนตัวอย่างที่มีค่าน้อยกว่าค่าทดสอบ (Cases < Test Value)	53
จำนวนตัวอย่างที่มีค่ามากกว่าหรือเท่ากับค่าทดสอบ (Cases >= Test Value)	421
จำนวนตัวอย่างทั้งหมด (Total Cases)	474
จำนวนของรันส์ (Number of Runs)	97
Z	.430
Asymp. Sig. (2-tailed)	.667

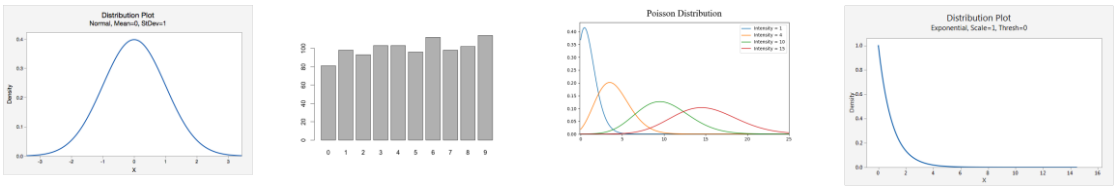
1.1.7 การแปลผลและสรุป

จากตารางที่ (#) ผลการทดสอบแสดงให้เห็นว่า ค่าความน่าจะเป็น (Asymp. Sig.) = .667 มีค่ามากกว่าค่านัยสำคัญทางสถิติที่กำหนดไว้ (α = 0.05) จึงยอมรับสมมติฐาน H_0

สรุป ข้อมูลการศึกษาของพนักงานมีการแจกแจงเป็นไปอย่างสุ่ม การทดสอบการแจกแจงข้อมูลความเป็นสุ่มของข้อมูลตัวแปรการศึกษาของพนักงานมีค่าทดสอบ (ระดับการศึกษา) ที่ 12 ปี แสดงว่าจำนวนตัวอย่างที่มีค่าน้อยกว่าค่าทดสอบมีจำนวน 53 คน จำนวนตัวอย่างที่มีค่ามากกว่าหรือเท่ากับค่าทดสอบมีจำนวน 421 คน ระดับการศึกษาที่อยู่ตรงกลางของผู้ตอบแบบสอบถามหรือมีค่าฐานนิยม (mode) 12 ปี และจากค่าการทดสอบแสดงให้เห็นว่า ผู้ตอบแบบสอบถามมีการศึกษาต่ำกว่า 12 ปี จำนวน 53 คน และมีการศึกษาเท่ากับและมากกว่า 12 ปี มีจำนวน 421 คน

1.2 การทดสอบการแจกแจงของข้อมูลแบบคอลมอโกรอฟ-สเมอร์นอฟ

การทดสอบการแจกแจงข้อมูลแบบคอลมอโกรอฟ-สเมอร์นอฟ (Kolmogorov-Smirnov: K-S Test) คือ การทดสอบการแจกแจงข้อมูลแบบสะสม (cumulative distributions) ระหว่างข้อมูลที่ได้จากการสังเกต (observed) กับการแจกแจงเชิงทฤษฎี (theoretical) ซึ่งอาจเป็นการแจกแจงแบบปกติ (normal/gaussian distribution)[19] การแจกแจงแบบเอกรูปหรือสม่ำเสมอหรือสี่เหลี่ยมผืนผ้า (uniform/rectangular distribution)[20] การแจกแจงแบบปัวซอง (poisson distribution)[21] การแจกแจงแบบเอ็กโปเนนเชียลหรือแบบเลขชี้กำลัง (exponential distribution)[22]



การแจกแจงแบบปกติ	การแจกแจงแบบเอกรูป	การแจกแจงแบบปัวซอง	การแจกแจงแบบเลขยกกำลัง

รูปที่ 13.1 การแจกแจงความถี่

ตัวอย่างคำถามการวิจัย

จากการสำรวจข้อมูลของพนักงานในหน่วยงานต้องการทราบว่า เงินเดือนเริ่มต้นของพนักงานในหน่วยงานมีการแจกแจกแบบปกติหรือไม่

[19] การแจกแจงของข้อมูลจากเหตุการณ์ที่เกิดซ้ำๆกัน

[20] การแจกแจงของข้อมูลจากเหตุการณ์ที่เกิดขึ้นเหมือนๆกัน หรือแต่ละเหตุการณ์มีช่วงเท่ากัน เช่น จำนวนสมาชิกเฉลี่ยของแต่ละครอบครัว หรือจำนวนน้ำหนักเฉลี่ยที่เพิ่มขึ้นของแต่ละคนในฤดูหนาว (ประมาณ 0-30 ปอนด์)

[21] การแจกแจงของข้อมูลจากหลายเหตุการณ์และไม่ต่อเนื่องกันหรือเป็นอิสระจากกัน เช่น จำนวนรถยนต์ที่ผ่านด่านแต่ละนาที

[22] การแจกแจงของข้อมูลจากเหตุการณ์ที่เกิดขึ้นแบบสะสมตั้งแต่เริ่มต้นจนถึงเกิดเหตุการณ์ เช่น จำนวนเงินที่เก็บสะสม ระยะเวลาการเกิดแผ่นดินไหม

1.2.1 ข้อตกลงเบื้องต้น

ขั้นตอน	ข้อตกลง
การออกแบบการวิจัย	1. ข้อมูลได้มาจากการสุ่มตัวอย่างแบบสุ่ม (random sampling) และเป็นอิสระจากกัน (independent sample) 2. ข้อมูลเป็นตัวแปรแบบค่าต่อเนื่อง (continuous variable) หรือตัวแปรแบบมาตราจัดอันดับ (ordinal variable) บางตำราบอกว่า ตัวแปรมาตราแบบจัดกลุ่ม (nominal variable)
การวิเคราะห์ข้อมูล	-

1.2.2 ตัวแปรและวิธีการทดสอบ

ตัวแปร		การทดสอบ
ค่าต่อเนื่อง	->	การแจกแจงของข้อมูล
เงินเดือนเริ่มต้น: ดอลลาร์ (salbegin)		แจกแจงแบบปกติ

1.2.3 การกำหนดสมมติฐานและเกณฑ์การทดสอบ

จากข้อมูลที่ทำการสำรวจจากพนักงานในหน่วยงาน ผู้วิจัยต้องการทราบว่า ข้อมูลเงินเดือนเริ่มต้นของพนักงานมีการแจกแจกปกติหรือไม่ โดยมีสมมติฐานดังนี้

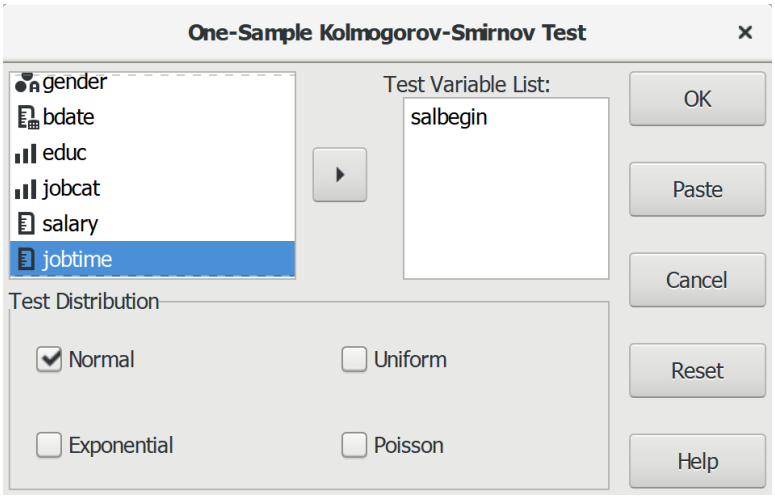
H_0: ข้อมูลเงินเดือนเริ่มต้นของพนักงานมีการแจกแจงปกติ

H_1: ข้อมูลเงินเดือนเริ่มต้นของพนักงานมีการแจกแจงไม่ปกติ

เกณฑ์ในการทดสอบสมมุติฐาน คือ หากการทดสอบการแจกแจงของข้อมูลแบบคอลมอโกรอฟ-สเมอร์นอฟ (Kolmogorov-Smirnov Test: K-S Test) ได้ค่าความน่าจะเป็น (p-value) มากกว่าค่านัยสำคัญทางสถิติ (α = 0.05) จะยอมรับสมมติฐาน H_0 แต่หากได้ค่าความน่าจะเป็น (p-value) น้อยกว่าหรือเท่ากับค่านัยสำคัญทางสถิติ (α = 0.05) จะปฏิเสธสมมติฐาน H_0 และยอมรับสมมติฐาน H_1

1.2.4 การใช้โปรแกรมคอมพิวเตอร์วิเคราะห์ข้อมูล

- เปิดแฟ้มข้อมูล **File -> Open** -> Employee data.sav
- เลือกเมนู **Analyze -> Non-Parametric Statistics -> 1-Sample K-S...**
  - **Test Variable List:** เลือกตัวแปรทดสอบที่เป็นตัวแปรแบบค่าต่อเนื่อง 1 ตัวแปร หรือหลายตัวแปร จากรายการตัวแปรด้านซ้าย
  - ในกลุ่มคำสั่ง **Test Distribution** เลือกสถิติ **Normal**
  - กดปุ่ม **OK**



1.2.5 ผลการวิเคราะห์

One-Sample Kolmogorov-Smirnov Test

		Beginning Salary
N		474
Normal Parameters	Mean	17016.09
	Std. Deviation	7870.64
Most Extreme Differences	Absolute	.25
	Positive	.25
	Negative	-.17
Kolmogorov-Smirnov Z		5.48
Asymp. Sig. (2-tailed)		.000

1.2.6 การสร้างตารางแสดงผลการทดสอบ

ตารางที่ (#) แสดงผลการทดสอบข้อมูล 1 กลุ่ม แบบคอลมอโกรอฟ-สเมอร์นอฟ

		เงินเดือนเริ่มต้น (ดอลลาร์)
จำนวน (N)		474
ค่าพารามิเตอร์การทดสอบการแจกแจงปกติ (Normal Parameters)	ค่าเฉลี่ย (Mean)	17,016.09
	ค่าเบี่ยงเบนมาตรฐาน (Std. Deviation)	7,870.64
ความแตกต่างของค่าโต่งมากที่สุด (Most Extreme Differences)	ค่าสัมบูรณ์ (Absolute)	.25
	ค่าบวก (Positive)	.25
	ค่าลบ (Negative)	-.170
Kolmogorov-Smirnov Z		5.48
Asymp. Sig. (2-tailed)		.000

1.2.7 การแปลผลและสรุป

จากตารางที่ (#) ผลการทดสอบแสดงให้เห็นว่า ค่าความน่าจะเป็น (Asymp. Sig.) = 0.000 มีค่าน้อยกว่าค่านัยสำคัญทางสถิติที่กำหนดไว้ (α = 0.05) จึงปฏิเสธสมมติฐาน H_0 และ*ยอมรับสมมติฐาน* H_a

สรุป ข้อมูลเงินเดือนเริ่มต้นของพนักงานมีการแจกแจงไม่ปกติ

2. การทดสอบสัดส่วนของกลุ่มตัวอย่าง

การทดสอบสัดส่วนของกลุ่มตัวอย่าง (Sample Proportion Test) คือ การทดสอบสัดส่วนของข้อมูลจากตัวอย่างกับสัดส่วนที่แท้จริงของประชากรทั้งหมดหรือค่าอ้างอิง สมมติฐานในการทดสอบอาจระบุเป็นสัดส่วน (proportions) แตกต่าง (difference) อัตราส่วน (ratio) หรืออัตราส่วนที่จะเกิด (odds ratio)

2.1 การทดสอบความแนบสนิทแบบไคสแควร์

การทดสอบความแนบสนิทแบบไคสแควร์ (X^2 goodness of fit) คือ การทดสอบสัดส่วนของกลุ่มตัวอย่าง 2 กลุ่ม หรือมากกว่า 2 กลุ่ม ที่ได้จากการสังเกต (observed value) กับค่าที่คาดหวัง (expected value) หรือค่าตามทฤษฎี

ตัวอย่างของข้อมูล เช่น ตำแหน่งงาน มีคำตอบคือ เจ้าหน้าที่ปฏิบัติงาน หัวหน้างาน และผู้จัดการ ตัวแปรการศึกษา มีคำตอบคือ ประถมศึกษา มัธยมศึกษา และอุดมศึกษา ตัวแปรศาสนา มีคำตอบคือ พุทธ คริสต์ และอิสลาม เป็นต้น

ตัวอย่างคำถามการวิจัย
จากข้อมูลการสำรวจพนักงานในหน่วยงานต้องการทราบว่า พนักงานของหน่วยงานมีสัดส่วนของตำแหน่งงานเท่ากันหรือไม่

2.1.1 ข้อตกลงเบื้องต้น

ขั้นตอน	**ข้อตกลง**
การออกแบบการวิจัย	1. ข้อมูลได้มาจากการสุ่มตัวอย่างแบบสุ่ม (random sampling) และเป็นอิสระจากกัน (independent sample) 2. มีค่าสังเกตอย่างน้อย 5 ตัวอย่าง ในแต่ละระดับของตัวแปร (level of the variable) 3. ตัวแปรเป็นตัวแปรแบบค่าจัดประเภท (categorical variables) 1 ตัวแปร
การวิเคราะห์ข้อมูล	-

2.1.2 ตัวแปรและวิธีการทดสอบ

ตัวแปร		การทดสอบ
ค่าจัดประเภท		สัดส่วน
ตำแหน่งงาน (Jobcat) - เจ้าหน้าที่ปฏิบัติงาน - หัวหน้างาน - ผู้จัดการ	->	เท่ากัน (Equal) หรือ ไม่เท่ากัน (Unequal)

2.1.3 การกำหนดสมมติฐานและเกณฑ์การทดสอบ

จากตารางที่ (#) ผลการศึกษาพบว่า กลุ่มตัวอย่างมีตำแหน่งงานเป็นเจ้าหน้าที่ปฏิบัติงานมากที่สุด คือ ร้อยละ 76.6 รองลงมามีตำแหน่งงานเป็นผู้จัดการ คือ ร้อยละ 17.7 และมีตำแหน่งงานเป็นหัวหน้างานน้อยที่สุด คือ ร้อยละ 5.7 ดังตารางที่ (#)

ตารางที่ (#) แสดงจำนวนและค่าร้อยละของกลุ่มตัวอย่าง จำแนกตามตำแหน่งงาน

ตำแหน่งงาน	จำนวน	ร้อยละ
เจ้าหน้าที่ปฏิบัติงาน	363	76.6
หัวหน้างาน	27	5.7
ผู้จัดการ	84	17.7
รวม	474	100.0

จากตารางที่ (#) แสดงให้เห็นว่า กลุ่มตัวอย่างมีตำแหน่งเป็นเจ้าหน้าที่ปฏิบัติงานมากที่สุด รองลงมาเป็นผู้จัดการ และเป็นหัวหน้างานน้อยที่สุด ดังนั้นจึงทำการทดสอบว่า พนักงานแต่ละตำแหน่งมีสัดส่วนเท่ากันหรือไม่ โดยมีสมมติฐานดังนี้

H_0: สัดส่วนของตำแหน่งงานเท่ากัน หรือ $\Sigma(O - E)^2 = 0$

H_1: สัดส่วนของตำแหน่งงานไม่เท่ากัน หรือ $\Sigma(O - E)^2 \neq 0$

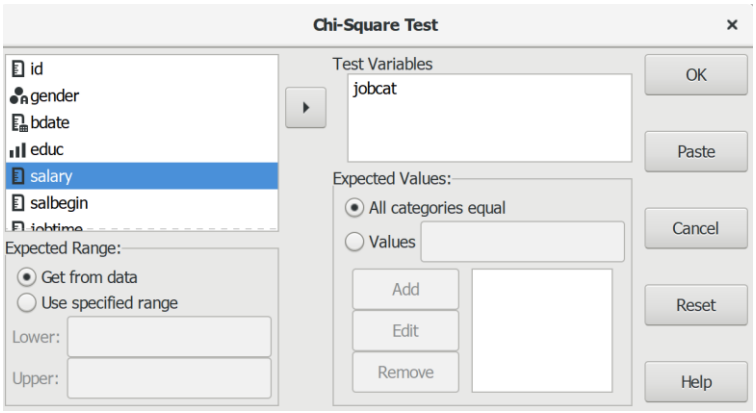
เกณฑ์ในการทดสอบสมมุติฐาน คือ หากการทดสอบสัดส่วนของข้อมูลด้วยวิธีไคสแควร์ (X^2 Test) ได้ค่าความน่าจะเป็น (p-value) มากกว่าค่านัยสำคัญทางสถิติ ($\alpha = 0.05$) จะยอมรับสมมติฐาน H_0 แต่

หากได้ค่าความน่าจะเป็น (p-value) น้อยกว่าหรือเท่ากับค่านัยสำคัญทางสถิติ (α = 0.05) จะปฏิเสธสมมติฐาน H_0 และยอมรับสมมติฐาน H_1

2.1.4 การใช้โปรแกรมวิเคราะห์ข้อมูล

- เปิดแฟ้มข้อมูล **File -> Open** -> Employee Data.sav
- เลือกเมนู **Analyze -> Non-Parametric Statistics -> Chi-square...**
  - **Test Variable:** เลือกตัวแปรทดสอบที่เป็นตัวแปรแบบจัดประเภท 1 ตัวแปร หรือหลายตัวแปร จากรายการตัวแปรด้านซ้าย
  - ในกลุ่มคำสั่ง **Expected Range** กำหนดค่าตัวแปรหรือกลุ่มข้อมูล
    - Get from data (ทั้งหมดจากข้อมูล)
    - Use specified range (กำหนดช่วงข้อมูลกรณีเป็นตัวแปรแบบค่าต่อเนื่อง)
  - ในกลุ่มคำสั่ง **Expected Values** กำหนดสัดส่วน
    - All categories equal (ทุกกลุ่มเท่ากัน)
    - Values: (กำหนดสัดส่วนในช่องว่างและเพิ่มจำนวนให้สัมพันธ์กับกลุ่มของตัวแปร เช่น มี 2 กลุ่ม กำหนดเป็น 2 ค่า มี 3 กลุ่ม กำหนดเป็น 3 ค่า)
  - กดปุ่ม **OK**

<u>หมายเหตุ</u> ค่าคาดหวัง (Expected frequencies) มาจากข้อมูลในอดีต ความรู้ที่มีอยู่ ทฤษฎี หรืองานวิจัย



2.1.5 ผลการวิเคราะห์

Employment Category

Value	Observed N	Expected N	Residual
Clerical	363	158.00	205.00
Custodial	27	158.00	-131.00
Manager	84	158.00	-74.00
Total	474		

Test Statistics

	Chi-square	df	Asymp. Sig.
Employment Category	409.25	2	.000

2.1.6 การสร้างตารางแสดงผลการทดสอบ

ตารางที่ (#) แสดงผลการทดสอบเปรียบเทียบสัดส่วนระหว่างค่าสังเกตและค่าคาดหวังของตำแหน่งงาน

ตำแหน่งงาน	ค่าสังเกต (Observed N)	ค่าคาดหวัง (Expected N)	ส่วนเหลือ (Residual)	สถิติทดสอบ (Test Statistics)
เจ้าหน้าที่ปฏิบัติงาน	363	158.0	205.0	X^2 = 409.25
หัวหน้างาน	27	158.0	-131.0	df = 2
ผู้จัดการ	84	158.0	-74.0	Asymp. Sig. = .000
รวม	474			

2.1.7 การแปลผลและสรุป

จากตารางที่ (#) ผลการทดสอบแสดงให้เห็นว่า ค่าความน่าจะเป็น (Asymp. Sig.) = 0.00 น้อยกว่าค่านัยสำคัญทางสถิติที่กำหนดไว้ (α = 0.05) จึงปฏิเสธสมมติฐาน H_0 และยอมรับสมมติฐาน H_1

สรุป ตำแหน่งงานของพนักงานมีสัดส่วนไม่เท่ากัน

2.2 การทดสอบสัดส่วนของกลุ่มตัวอย่างแบบทวินาม

การทดสอบสัดส่วนของกลุ่มตัวอย่าง (Sample Proportion Test) แบบทวินาม (Binomial Test) คือ การทดสอบสัดส่วนของข้อมูลที่ได้จากการสังเกต (observed value) ของตัวแปรแบบค่าจัดประเภท 2 ประเภท (dichotomous variables) หรือตัวแปรแบบค่าทวิลักษณะ (binary variables) กับค่าพารามิเตอร์ที่เป็นค่าสัดส่วนที่คาดหวัง (expected value) หรือค่าสัดส่วนที่ใช้ทดสอบ (test proportion)

ตัวอย่างคำถามการวิจัย
จากการสำรวจข้อมูลของพนักงานในหน่วยงานต้องการทราบว่า พนักงานมีสัดส่วนระหว่างเพศชายกับเพศหญิงคล้องกับสัดส่วนประชากรของประเทศหรือไม่ (ชาย:หญิง = 49:51)

2.2.1 ข้อตกลงเบื้องต้น

ขั้นตอน	**ข้อตกลง**
การออกแบบการวิจัย	1. ข้อมูลได้มาจากการสุ่มตัวอย่างแบบสุ่ม (random sampling) และเป็นอิสระจากกัน (independent sample) 2. ตัวแปรเป็นตัวแปรแบบค่าสองประเภท (dichotomous variables)
การวิเคราะห์ข้อมูล	-

2.2.2 ตัวแปรและวิธีการทดสอบ

ตัวแปร		**การทดสอบ**
ค่าจัดประเภท		สัดส่วน
เพศ (gender) - ชาย - หญิง	->	ชาย:หญิง = 49:51

2.2.3 การกำหนดสมมติฐานและเกณฑ์การทดสอบ

จากผลศึกษาพบว่า กลุ่มตัวอย่างส่วนใหญ่เป็นพนักงานเพศชาย มีจำนวนร้อยละ 54.43 มากกว่าพนักงานเพศหญิง ที่มีจำนวนร้อยละ 45.57 ดังตารางที่ (#)

ตารางที่ (#) แสดงจำนวนและร้อยละของพนักงาน จำแนกตามเพศ

เพศ	จำนวน	ร้อยละ
หญิง	216	45.57
ชาย	258	54.43
รวม	474	100.0

จากตารางที่ (#) แสดงให้เห็นว่า กลุ่มตัวอย่างเป็นพนักงานเพศชายมากกว่าเพศหญิง ดังนั้นจึงทำการทดสอบว่า พนักงานเพศชายมีสัดส่วนเท่ากับพนักงานเพศหญิงหรือไม่ โดยมีสมมติฐานดังนี้

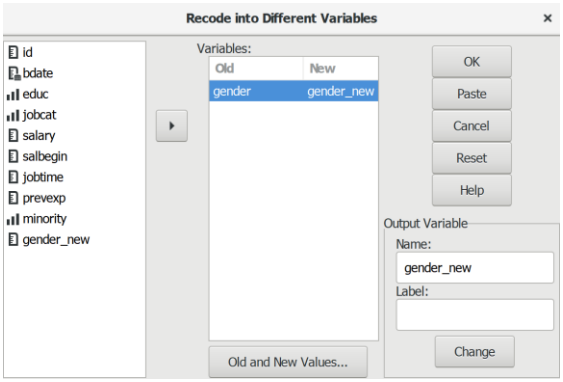
H_0: สัดส่วนของพนักงานเพศชายกับเพศหญิงเท่ากับ 49:51 (P1=0.49)

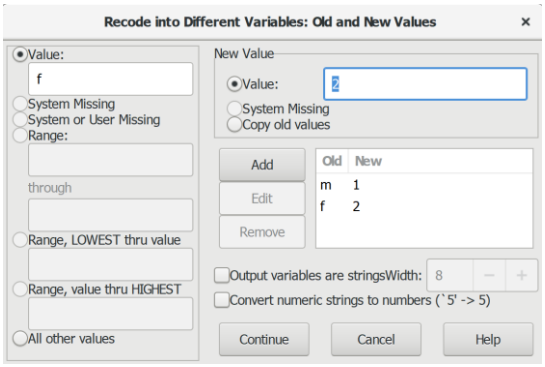
H_1: สัดส่วนของพนักงานเพศชายกับเพศหญิงไม่เท่ากับ 49:51 (P≠0.49)

เกณฑ์ในการทดสอบสมมุติฐาน คือ หากทดสอบสัดส่วนของข้อมูลแบบทวินาม (Binomial Test)ได้ค่าความน่าจะเป็น (p-value) มากกว่าค่านัยสำคัญทางสถิติ (α = 0.05) จะยอมรับสมมติฐาน H_0 แต่หากได้ค่าความน่าจะเป็น (p-value) น้อยกว่าหรือเท่ากับค่านัยสำคัญทางสถิติ (α = 0.05) จะปฏิเสธสมมติฐาน H_0 และยอมรับสมมติฐาน H_1

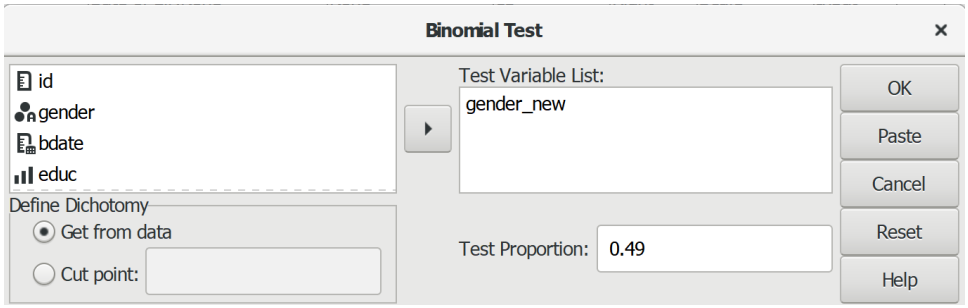
2.2.4 การใช้โปรแกรมคอมพิวเตอร์วิเคราะห์ข้อมูล

- เปิดแฟ้มข้อมูล **File -> Open** -> Employee data.sav
  - สร้างตัวแปร gender_new โดยใช้คำสั่ง **Recode into Different Variables** จากตัวแปร gender กำหนดให้ค่า m = 1 และ f = 2
  - กำหนดค่าตัวแปร 1 = Male และ 2 = Female





- เลือกเมนู **Analyze -> Non-Parametric Statistics -> Binomial ...**
  - **Test Variable List:** เลือกตัวแปรทดสอบที่เป็นตัวแปรแบบค่าจัดประเภทหรือตัวแปรแบบค่าต่อเนื่อง 1 ตัวแปร จากรายการตัวแปรด้านซ้าย
  - ในกลุ่มคำสั่ง **Define Dichotomy** กำหนดกลุ่มข้อมูล
    - Get from data (เฉพาะตัวแปรที่มีค่า 2 กลุ่ม)
    - Cut point (ใส่ค่าแบ่งข้อมูลของตัวแปรให้เป็น 2 กลุ่ม)
  - **Test Proportion:** กำหนดสัดส่วนในการทดสอบ
  - กดปุ่ม **OK**



2.2.5 ผลการวิเคราะห์

Binomial Test

		Category	N	Observed Prop.	Test Prop.	Exact Sig. (1-tailed)
gender_new	Group 1	Male	258	.54	.49	.010
	Group 2	Female	216	.46		
	Total		474	1.00		

2.2.6 การสร้างตารางแสดงผลการทดสอบ

ตาราง (#) แสดงการทดสอบสัดส่วนเพศของพนักงาน

เพศ	จำนวน (N)	สัดส่วนค่าสังเกต (Observed Prop.)	สัดส่วนการทดสอบ (Test Prop.)	Exact Sig.
ชาย	258	.54		
หญิง	216	.46	.49	.01
รวม	474	1.00		

2.2.7 การแปลผลและสรุป

จากตารางที่ (#) ผลการทดสอบแสดงให้เห็นว่า ค่าความน่าจะเป็น (Exact Sig.) = 0.01 น้อยกว่าค่านัยสำคัญทางสถิติที่กำหนดไว้ (α = 0.05) จึงปฏิเสธรับสมมติฐาน H_0 และยอมรับสมมติฐาน H_1

สรุป สัดส่วนของพนักงานเพศชายเท่ากับพนักงานเพศหญิงไม่เท่ากับ 49:51

สรุป

การทดสอบข้อมูล 1 ตัวแปร คือ การใช้ค่าของตัวแปรหรือข้อมูลภายในตัวแปรเดียวที่เป็นข้อมูลจริงที่ได้จากการสำรวจหรือเป็นข้อมูลที่มีอยู่ (observed value) แล้วมาทำการทดสอบค่าที่คาดหวัง (expected value) ซึ่งอาจเป็นข้อมูลในอดีต ความรู้ที่มีอยู่ ทฤษฎี หรืองานวิจัย

แบบฝึกหัดท้ายบท

ให้ผู้เรียนเปิดแฟ้มข้อมูล 1991 US. General Social Survey.sav แล้วประมวลผลข้อมูล เพื่อทดสอบและแปลผล ดังนี้

1. เปรียบเทียบกลุ่มตัวอย่างแต่ละภูมิภาค (region) มีสัดส่วนที่เท่ากันหรือไม่

2. ใช้ตัวแปรกลุ่มชน (race) สร้างเป็นตัวแปรใหม่ชื่อ กลุ่มชนใหม่ (newrace) มี 2 กลุ่ม คือ ผิวขาว และไม่ใช่ผิวขาว และเปรียบเทียบว่ากลุ่มชนผิวขาวมีสัดส่วนเท่ากับ 0.72 ของกลุ่มตัวอย่างทั้งหมดหรือไม่

3. ทดสอบข้อมูลการศึกษาของผู้ตอบ (educ) มีการแจกแจงเป็นไปอย่างสุ่มหรือไม่

4. ทดสอบข้อมูลการศึกษาของผู้ตอบ (educ) มีการแจกแจงปกติหรือไม่

บทที่ 14

การทดสอบข้อมูลแบบไม่อิงค่าพารามิเตอร์ (2 ตัวแปร)

ความมุ่งหมายของบทเรียน

ผู้เรียนสามารถใช้โปรแกรมประมวลผลทดสอบข้อมูลของตัวแปร 2 ตัวแปร ด้วยสถิติเชิงอ้างอิงแบบไม่อิงค่าพารามิเตอร์ และเขียนบรรยายผลการทดสอบได้

เนื้อหา

1. การทดสอบความเป็นอิสระของกลุ่มตัวอย่าง
2. การทดสอบค่ากลางของกลุ่มตัวอย่างที่เป็นอิสระจากกัน 2 กลุ่ม
3. การทดสอบค่ากลางของกลุ่มตัวอย่างที่เป็นอิสระจากกันมากกว่า 2 กลุ่ม
4. การทดสอบค่ากลางของกลุ่มตัวอย่างที่ไม่เป็นอิสระจากกัน 2 กลุ่ม
5. การทดสอบค่ากลางของกลุ่มตัวอย่างที่ไม่เป็นอิสระจากกันมากกว่า 2 กลุ่ม
6. การทดสอบความสัมพันธ์รายคู่แบบสเปียร์แมน

กิจกรรมและวิธีสอน

บรรยาย อธิบาย แสดงตัวอย่าง และทำแบบฝึกหัดท้ายบท

อุปกรณ์การสอน

เอกสารคำสอน ภาพเลื่อน คอมพิวเตอร์ โปรแกรม PSPP โปรแกรมสอนออนไลน์ (สอนแบบออนไลน์) และเครื่องฉาย (สอนในชั้นเรียน)

งานที่มอบหมาย

ทำแบบฝึกหัดท้ายบท

การวัดผลและประเมินผล

การเข้าชั้นเรียน สังเกตความสนใจจากการเรียน สอบถามความเข้าใจจากการบรรยาย และประเมินจากการทำแบบฝึกหัดท้ายบท

เนื้อหา

การวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรอิสระกับตัวแปรตามหากมีตัวแปรใดตัวแปรหนึ่งเป็นตัวแปรแบบค่าแบ่งกลุ่มต้องเลือกใช้การทดสอบแบบไม่อิงค่าพารามิเตอร์ตามลักษณะของตัวแปรดังต่อไปนี้

1. การทดสอบความเป็นอิสระของกลุ่มตัวอย่าง

การทดสอบความเป็นอิสระของกลุ่มตัวอย่าง (Test of Independence) คือ การทดสอบความแนบสนิท (goodness-of-fit test) โดยใช้ตารางไขว้ (contingency table) ทำการทดสอบค่าที่ได้จากการสังเกต (observed values) กับค่าที่คาดหวัง (expected values) หรือค่าที่ควรจะเป็นตามทฤษฎี เพื่อตัดสินใจว่า ค่าสังเกตแต่ละค่ามีความเป็นอิสระจากกันหรือไม่เป็นอิสระจากกัน

การทดสอบความเป็นอิสระของกลุ่มตัวอย่างเป็นการทดสอบความสัมพันธ์ (related) ระหว่างตัวแปรที่มีระดับการวัดแบบจัดประเภท 2 ตัวแปร เพื่อดูว่าตัวแปร 2 ตัว มีความสัมพันธ์กันหรือมีความเป็นอิสระจากกันหรือไม่ มีหลายวิธีดังนี้

1. ไคสแควร์ (Chi-square หรือ X^2) ใช้วิเคราะห์ตารางข้อมูลที่ตัวแปรอิสระมีค่าเป็นข้อมูล 2 กลุ่ม และตัวแปรตามมีค่าเป็นข้อมูล 2 กลุ่ม หรือที่เรียกว่า ตาราง 2x2 มีสถิติทดสอบดังนี้

1.1 ไคสแควร์แบบเพียร์สัน (Pearson chi-square)

1.2 ไคสแควร์แบบความเป็นไปได้หรือความน่าจะเป็นของอัตราส่วน (The likelihood-ratio chi-square)

1.3 ความเที่ยง/แม่นยำของฟิชเชอร์ (Fisher's exact test)

1.4 ไคสแควร์แบบแก้ไขของเยตส์ (Yates' corrected chi-square)

2. สหสัมพันธ์ (Correlations) ใช้วิเคราะห์ตารางข้อมูลที่ตัวแปรอิสระมีค่าเป็นข้อมูลจัดอันดับ (อันดับมาตรา) และตัวแปรอิสระมีค่าเป็นข้อมูลแบบจัดอันดับ (มาตราแบบจัดอันดับ) มีสถิติทดสอบดังนี้

2.1 สัมประสิทธิ์สหสัมพันธ์ของสเปียร์แมน (Spearman's correlation coefficient)

2.2 โรท์ (rho) เฉพาะข้อมูลที่เป็นตัวเลข (numeric data only)

3. ข้อมูลแบบจัดกลุ่ม (Nominal) ใช้วิเคราะห์ข้อมูลจัดกลุ่ม (nominal data) มีสถิติทดสอบดังนี้

3.1 สัมประสิทธิ์ของสถานการณ์ที่สอดคล้องกัน (Contingency coefficient)

3.2 ฟีและแครมเมอร์ วี (Phi and Cramer’s V)

3.3 แลมดา (Lambda)

3.4 สัมประสิทธิ์ความไม่แน่นอน (Uncertainty coefficient)

4. ข้อมูลแบบจัดอันดับ (Ordinal) ใช้วิเคราะห์ข้อมูลที่ตัวแปรอิสระและตัวแปรตามเป็นมีค่าจัดอันดับ (ordinal value) มีสถิติทดสอบดังนี้

4.1 แกรมมา (Gamma) ลบคู่ที่ไม่สอดคล้องกันออกและนับเฉพาะคู่ที่สอดคล้องกันในตารางไขว้

4.2 เดลต้าของโสเมอร์ส (Somers' d)

4.3 สัมประสิทธิ์สหสัมพันธ์ของเคนดอล์ล (Kendall's coefficient) คำนวณเหมือนแกรมมา แต่มีการแบ่งส่วน (denominator) ที่ซับซ้อนกว่า และให้ค่าสถิติที่ต่ำกว่าค่าแกรมมา

- แบบทาว-บี (Kendall's tau-b) ตารางแบบ 2x2 (square table)
- แบบทาว-ซี (Kendall's tau-c) ตารางแบบ NxN (rectangular tables)

5. ข้อมูลแบบจัดกลุ่มกับข้อมูลแบบช่วง (Nominal by Interval) ใช้วิเคราะห์ข้อมูลเมื่อตัวแปรหนึ่งเป็นตัวแปรแบบค่าแบบจัดประเภท (categorical) ส่วนอีกตัวแปรเป็นตัวแปรเชิงปริมาณ (quantitative) มีสถิติทดสอบดังนี้

5.1 อีตา (Eta) วัดความสัมพันธ์ระหว่าง 0-1 โดย 0 แสดงว่าไม่มีความสัมพันธ์ระหว่างตัวแปรแถว (row) และสดมภ์ (column) 1 แสดงว่ามีความสัมพันธ์มาก เหมาะกับการวัดตัวแปรตามที่มีการวัดเป็นมาตราแบบช่วง และตัวแปรอิสระที่มีการวัดเป็นมาตราแบบจัดกลุ่ม

6. แคปปา (Kappa) หรือสัมประสิทธิ์แคปปาของโคเฮน (Cohen's kappa coefficient) ใช้วิเคราะห์ความคิดเห็น (agreement) จากผู้ให้คะแนน 2 ฝ่าย (two raters) ในเรื่องเดียวกัน โดยกำหนดให้ค่า 1 หมายถึง เห็นด้วยทั้งหมด (perfect agreement) และให้ค่า 0 หมายถึง ไม่เห็นด้วยทั้งหมด ตัวแปรที่นำมาวิเคราะห์ทั้งสองตัวแปรต้องเป็นตัวแปรแบบค่าจัดประเภทเหมือนกันและมีการแบ่งจำนวนกลุ่มของค่าตัวแปรเท่ากันด้วย

7. ความเสี่ยง (Risk) ใช้วิเคราะห์ความเข้มของความสัมพันธ์ในรูปของตารางแบบ 2x2 ในการนำเสนอปัจจัยหนึ่งกับเหตุการณ์ที่เกิด หากค่าสถิติที่ได้มีช่วงของความเชื่อมั่นเป็น 1 แสดงว่า ปัจจัยนั้นไม่มีความสัมพันธ์กับเหตุการณ์ที่เกิด อัตราส่วนของโอกาสที่จะเกิด (odds ratio) สามารถใช้ในการประมาณการ (estimate) ความเสี่ยงที่เหตุการณ์นั้นยากที่จะเกิดขึ้นมาจากปัจจัยนั้นๆ

8. แมคนีมาร์ (McNemar) ใช้วิเคราะห์ความสัมพันธ์ของตัวแปรสองค่า (dichotomous) เพื่อทดสอบการตอบสนองต่อการเปลี่ยนแปลงจากปัจจัยแทรกในการทดลอง (experimental intervention) ที่ออกแบบเป็นก่อนการทดลองและหลังการทดลอง

9. คอครานและแมนเทล-แฮนส์เซล (Cochran's and Mantel-Haenszel) ใช้ในการวิเคราะห์ความเป็นอิสระระหว่างตัวแปรปัจจัยแบบสองค่า (dichotomous factor variable) และตัวแปร

ตอบสนองแบบสองค่า (dichotomous response variable) บนเงื่อนไขที่กำหนดให้เป็นรูปแบบของการใช้ตัวแปรร่วม (covariate patterns) โดยมีตัวแปรควบคุม 1 ตัวแปรหรือมากกว่า 1 ตัวแปร

วิธีการทดสอบความเป็นอิสระของกลุ่มตัวอย่างระหว่างตัวแปรแบบจัดประเภท 2 ตัวแปรที่นิยมกันอย่างแพร่หลาย คือ การทดสอบความเป็นอิสระของกลุ่มตัวอย่างแบบไคสแควร์ (X^2 Test for Independence)

ตัวอย่างคำถามการวิจัย
จากข้อมูลการสำรวจพนักงานในหน่วยงานต้องการทราบว่า พนักงานที่มีเพศแตกต่างกันมีตำแหน่งงานแตกต่างกันหรือไม่ หรือเพศมีความสัมพันธ์กับตำแหน่งงานหรือไม่

1.1 ข้อตกลงเบื้องต้น

ขั้นตอน	ข้อตกลง
การออกแบบการวิจัย	1. ตัวอย่างได้มาจากการสุ่มแบบสุ่ม (random sampling) และเป็นอิสระจากกัน (independent sample) 2. ค่าสังเกตแต่ละค่ามีความเป็นอิสระจากกัน (independence of observations) หรือค่าของตัวแปรแบ่งเป็นกลุ่มและไม่มีความสัมพันธ์ระหว่างกัน 3. มีกลุ่มตัวอย่างไม่น้อยกว่า 10 ตัวอย่าง หรือมีขนาดใหญ่พอให้แต่ละเซลล์มีอย่างน้อย 1 ตัวอย่าง และเซลล์ส่วนใหญ่มีอย่างน้อย 5 ตัวอย่าง 4. ตัวแปรเป็นแบบค่าจัดประเภท (categorical)
การวิเคราะห์ข้อมูล	1. ข้อมูลมีการแจกแจงไม่ปกติ (asymmetrical distribution) 2. ตัวแปรแบบค่าจัดประเภท 2 กลุ่ม แต่ละเซลล์ต้องมีความถี่อย่างน้อย 5 ตัวอย่าง และตัวแปรแบบค่าจัดประเภทที่มีจำนวนมากกว่า 2 กลุ่ม แต่ละเซลล์ต้องมีความถี่ไม่น้อยกว่า 1 ตัวอย่าง และความถี่ที่น้อยกว่า 5 ตัวอย่าง ต้องไม่เกินร้อยละ 20

1.2 รูปแบบการทดสอบตัวแปร

ตัวแปรอิสระ		ตัวแปรตาม
ค่าจัดประเภท		ค่าจัดประเภท
เพศ (gender) - ชาย - หญิง	->	ตำแหน่งงาน (jobcat) - เจ้าหน้าที่ปฏิบัติงาน - หัวหน้างาน - ผู้จัดการ

1.3 การกำหนดสมมติฐานและเกณฑ์การทดสอบ

ตารางที่ (#) แสดงจำนวนและค่าร้อยละของพนักงาน จำแนกตามเพศและตำแหน่งงาน

เพศ	ตำแหน่งงาน			รวม
	เจ้าหน้าที่ปฏิบัติงาน	หัวหน้างาน	ผู้จัดการ	
หญิง	205 (95.4)	0 (0.0)	10 (4.6)	216 (100.0)
ชาย	157 (60.9)	27 (10.5)	74 (28.7)	258 (100.00)
รวม	363 (76.6)	27 (5.7)	84 (17.7)	474 (100.0)

จากตารางที่ (#) แสดให้เห็นว่า เพศหญิงส่วนใหญ่เป็นเจ้าหน้าที่ปฏิบัติงาน รองลงมา คือ ผู้จัดการ และไม่มีผู้ดำรงตำแหน่งหัวหน้างานเลย ส่วนเพศชายส่วนใหญ่เป็นเจ้าหน้าที่ปฏิบัติงาน รองลงมา คือ ผู้จัดการ และหัวหน้างาน ตามลำดับ

เมื่อเปรียบเทียบตำแหน่งระหว่างเพศชายกับเพศหญิงแสดงให้เห็นว่า เพศหญิงชายมีตำแหน่งงานเป็นเจ้าหน้าที่ปฏิบัติงานมากกว่าเพศชาย ส่วนเพศชายมีตำแหน่งงานเป็นหัวหน้างานและผู้จัดการมากกว่าเพศหญิง หรืออาจกล่าวได้ว่า เพศชายมีตำแหน่งงานแตกต่างกับเพศหญิง (เพศชายมีตำแหน่งงานและสูงกว่าเพศหญิง)

ดังนั้น เพื่อเป็นการยืนยันข้อสรุปดังกล่าว จึงนำข้อสรุปมาทำการทดสอบความเป็นอิสระของข้อมูลด้วยวิธีการทางสถิติแบบไคสแควร์ (Chi-square หรือ X^2) โดยกำหนดสมมติฐาน ดังนี้

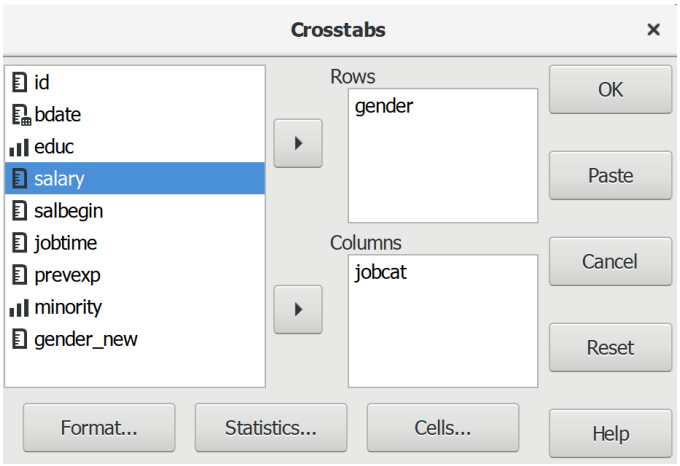
H_0: เพศชายกับเพศหญิงมีตำแหน่งงานไม่แตกต่างกัน

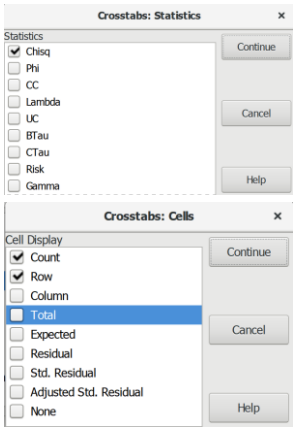
H_1: เพศชายกับเพศหญิงมีตำแหน่งงานแตกต่างกัน

เกณฑ์ในการทดสอบสมมุติฐาน คือ หากการทดสอบความสัมพันธ์ (เปรียบเทียบความแตกต่าง) ด้วยวิธีไคสแควร์ (Chi-square Test) ได้ค่าความน่าจะเป็น (p-value) มากกว่าค่านัยสำคัญทางสถิติ (α = 0.05) จะยอมรับสมมติฐาน H_0 แต่หากได้ค่าความน่าจะเป็น (p-value) น้อยกว่าหรือเท่ากับค่านัยสำคัญทางสถิติ (α = 0.05) จะปฏิเสธสมมติฐาน H_0 และยอมรับสมมติฐาน H_1

1.4 การใช้โปรแกรมคอมพิวเตอร์วิเคราะห์ข้อมูล

- เปิดแฟ้มข้อมูล **File -> Open** -> Employee data.sav
- เลือกเมนู **Analyze -> Descriptive Statistics -> Crosstabs...**
  - **Rows:** เลือกตัวแปรแบบค่าจัดประเภท 1 ตัวแปร หรือหลายตัวแปร (ที่ต้องการให้เป็นตัวแปรอิสระ) จากรายการตัวแปรด้านซ้าย
  - **Column:** เลือกตัวแปรแบบค่าจัดประเภท 1 ตัวแปร หรือหลายตัวแปร (ที่ต้องการให้เป็นตัวแปรตาม) จากรายการตัวแปรด้านซ้าย
  - กดปุ่ม **Statistics** เลือกสถิติ Chisq -> กดปุ่ม **Continue**
  - กดปุ่ม **Cells...** ในกลุ่มคำสั่ง **Percentages** (เลือก **Row**) -> กดปุ่ม **Continue**
  - กดปุ่ม **OK**





1.5 ผลการวิเคราะห์

Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Gender × Employment Category	474	100.0%	0	.0%	474	100.0%

Gender × Employment Category

			Employment Category			
			Clerical	Custodial	Manager	Total
Gender	Female	Count	206	0	10	216
		Row %	95.4%	.0%	4.6%	100.0%
	Male	Count	157	27	74	258
		Row %	60.9%	10.5%	28.7%	100.0%
Total		Count	363	27	84	474
		Row %	76.6%	5.7%	17.7%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Sig. (2-tailed)
Pearson Chi-Square	79.28	2	.000
Likelihood Ratio	95.46	2	.000
N of Valid Cases	474		

1.6 การสร้างตารางแสดงผลการทดสอบ

ตารางที่ (#) แสดงผลการทดสอบเปรียบเทียบตำแหน่งงานจำแนกตามเพศ

เพศ	ตำแหน่งงาน			รวม
	เจ้าหน้าที่ปฏิบัติงาน	หัวหน้างาน	ผู้จัดการ	
หญิง	206 (95.4)	0 (0.0)	10 (4.6)	216 (100.0)
ชาย	157 (60.9)	27 (10.5)	74 (28.7)	258 (100.00)
รวม	363 (76.6)	27 (5.7)	84 (17.7)	474 (100.0)

X^2 =79.28 df = 2 Sig = 0.00

1.7 การแปลผลและสรุป

จากตารางที่ (#) ผลการทดสอบแสดงให้เห็นว่า ค่าความน่าจะเป็น (Sig.) = 0.00 มีค่าน้อยกว่าค่านัยสำคัญทางสถิติที่กำหนดไว้ (α = 0.05) จึงปฏิเสธ H_0 ยอมรับสมมติฐาน H_1

สรุป เพศชายกับเพศหญิงมีตำแหน่งงานแตกต่างกัน

ตัวอย่างเพิ่มเติม: การทดสอบ การแปลผลและสรุป

#1

ตารางที่ (#) แสดงผลการทดสอบเปรียบเทียบตำแหน่งงานจำแนกตามระดับการศึกษา

ระดับการศึกษา	ตำแหน่งงาน			รวม
	เจ้าหน้าที่ปฏิบัติงาน	หัวหน้างาน	ผู้จัดการ	
ต่ำกว่าปริญญา	88.9 (216)	10.7 (26)	0.4 (1)	100.0 (243)
ปริญญา	63.6 (147)	0.4 (1)	35.9 (83)	100.0 (231)
รวม	76.6 (363)	5.7 (27)	17.7 (84)	100.0 (474)

X^2 =116.08 df = 2 Sig = 0.00

จากตารางที่ (#) แสดงให้เห็นว่า กลุ่มตัวอย่างพนักงานที่มีการศึกษาต่ำกว่าปริญญา ส่วนใหญ่ดำรงตำแหน่งงานเป็นเจ้าหน้าที่ปฏิบัติงาน รองลงมา เป็นหัวหน้างาน และผู้จัดการ ส่วนกลุ่มตัวอย่างพนักงานที่มีการศึกษาระดับปริญญา ส่วนใหญ่ดำรงตำแหน่งงานเป็นเจ้าหน้าที่ปฏิบัติงาน รองลงมา เป็นผู้จัดการ และหัวหน้างาน และเมื่อเปรียบเทียบตำแหน่งงานระหว่างพนักงานที่มีการศึกษาแตกต่างกัน แสดงให้เห็นว่า พนักงานที่มีการศึกษาระดับปริญญา มีตำแหน่งงานแตกต่างกับพนักงานที่มีการศึกษาต่ำกว่าปริญญา ดังนั้นเพื่อเป็นการยืนยันข้อสรุปที่ว่า พนักงานที่มีการศึกษาระดับปริญญา มีตำแหน่งงานแตกต่างกับพนักงานที่มีการศึกษาต่ำกว่าปริญญา จึงทำการทดสอบด้วยวิธีการทางสถิติ โดยกำหนดสมมติฐาน ดังนี้

H_0: พนักงานที่มีการศึกษาระดับปริญญากับพนักงานที่มีการศึกษาต่ำกว่าปริญญามีตำแหน่งงานไม่แตกต่างกัน

H_1: พนักงานที่มีการศึกษาระดับปริญญากับพนักงานที่มีการศึกษาต่ำกว่าปริญญามีตำแหน่งงานแตกต่างกัน

เกณฑ์ในการทดสอบสมมุติฐาน คือ หากการทดสอบความสัมพันธ์ (เปรียบเทียบความแตกต่าง) ด้วยวิธีไคสแควร์ (Chi-square Test) ได้ค่าความน่าจะเป็น (p-value) มากกว่าค่านัยสำคัญทางสถิติ (α = 0.05) จะยอมรับสมมติฐาน H_0 แต่หากได้ค่าความน่าจะเป็น (p-value) น้อยกว่าหรือเท่ากับค่านัยสำคัญทางสถิติ (α = 0.05) จะปฏิเสธสมมติฐาน H_0 และยอมรับสมมติฐาน H_1

จากตารางที่ (#) ผลการทดสอบแสดงให้เห็นว่า ค่าความน่าจะเป็น (Sig.) = 0.00 มีค่าน้อยกว่าค่านัยสำคัญทางสถิติที่กำหนดไว้ (α = 0.05) จึงปฏิเสธ H_0 ยอมรับสมมติฐาน H_1 กล่าวคือ พนักงานที่มีการศึกษาระดับปริญญากับพนักงานที่มีการศึกษาต่ำกว่าระดับปริญญามีตำแหน่งงานแตกต่างกัน

#2

ตารางที่ (#) แสดงผลการทดสอบเปรียบเทียบตำแหน่งงานจำแนกตามระดับการศึกษา

ระดับการศึกษา	ตำแหน่งงาน			รวม
	เจ้าหน้าที่ปฏิบัติงาน	หัวหน้างาน	ผู้จัดการ	
ต่ำกว่าปริญญา	88.9 (216)	10.7 (26)	0.4 (1)	100.0 (243)
ปริญญา	63.6 (147)	0.4 (1)	35.9 (83)	100.0 (231)
รวม	76.6 (363)	5.7 (27)	17.7 (84)	100.0 (474)

X^2 =116.08 df = 2 Sig = 0.00

จากตารางที่ (#) แสดงให้เห็นว่า กลุ่มตัวอย่างพนักงานที่มีการศึกษาต่ำกว่าปริญญา ส่วนใหญ่ดำรงตำแหน่งงานเป็นพนักงาน รองลงมา เป็นหัวหน้างาน และผู้จัดการ ส่วนกลุ่มตัวอย่างพนักงานที่มีการศึกษาระดับปริญญา ส่วนใหญ่ดำรงตำแหน่งงานเป็นเจ้าหน้าที่ปฏิบัติงาน รองลงมา เป็นผู้จัดการ และหัวหน้างาน และเมื่อเปรียบเทียบตำแหน่งงานระหว่างพนักงานที่มีการศึกษาแตกต่างกัน แสดงให้เห็นว่าพนักงานที่มีการศึกษาระดับปริญญา มีตำแหน่งงานสูงกว่าพนักงานที่มีการศึกษาต่ำกว่าปริญญา อย่างมีนัยสำคัญทางสถิติ

#3

ตารางที่ (#) แสดงผลการเปรียบเทียบตำแหน่งงานจำแนกตามระดับการศึกษา

ระดับการศึกษา	ตำแหน่งงาน			รวม
	เจ้าหน้าที่ปฏิบัติงาน	หัวหน้างาน	ผู้จัดการ	
ต่ำกว่าปริญญา	88.9 (216)	10.7 (26)	0.4 (1)	100.0 (243)
ปริญญา	63.6 (147)	0.4 (1)	35.9 (83)	100.0 (231)
รวม	76.6 (363)	5.7 (27)	17.7 (84)	100.0 (474)

X^2 =116.08 df = 2 Sig = 0.00

จากตารางที่ (#) แสดงให้เห็นว่า กลุ่มตัวอย่างพนักงานที่มีการศึกษาต่ำกว่าปริญญา ส่วนใหญ่ดำรงตำแหน่งงานเป็นเจ้าหน้าที่ปฏิบัติงาน รองลงมา เป็นหัวหน้างาน และผู้จัดการ ส่วนกลุ่มตัวอย่าง

พนักงานที่มีการศึกษาระดับปริญญา ส่วนใหญ่ดำรงตำแหน่งงานเป็นเจ้าหน้าที่ปฏิบัติงาน รองลงมา เป็นผู้จัดการ และหัวหน้างาน

ผลการศึกษาดังกล่าว เมื่อพิจารณาจากสมมติฐานการศึกษาตามทฤษฎีที่กำหนดไว้ว่า พนักงานที่มีการศึกษาแตกต่างกันมีตำแหน่งงานที่แตกต่างกัน แสดงให้เห็นว่า พนักงานที่มีการศึกษาระดับปริญญามีตำแหน่งงานแตกต่างกับพนักงานที่มีการศึกษาต่ำกว่าปริญญา อย่างมีนัยสำคัญทางสถิติ ดังนั้นจึงยอมรับสมมติฐานดังกล่าว

2. การทดสอบค่ากลางของกลุ่มตัวอย่างที่เป็นอิสระจากกัน 2 กลุ่ม

การทดสอบแบบไม่อิงค่าพารามิเตอร์มีข้อสมมติพื้นฐาน (assumption) คือ ข้อมูลมีการแจกแจงไม่ปกติ (non-normal distribution) ทำให้ไม่สามารถใช้วิธีการเปรียบเทียบความแตกต่างระหว่างกลุ่มตัวอย่างด้วยค่าเฉลี่ย (mean) เหมือนกับวิธีการทดสอบแบบอิงค่าพารามิเตอร์ ดังนั้นในการทดสอบข้อมูลระหว่างกลุ่มตัวอย่างจึงใช้วิธีการเปรียบเทียบจากค่ากลาง (median)

การทดสอบค่ากลางของกลุ่มตัวอย่างที่เป็นอิสระจากกัน 2 กลุ่ม (2 Independent Samples Test) คือ การเปรียบเทียบตำแหน่งข้อมูล (location) โดยใช้ค่ากลาง (median) ของกลุ่มตัวอย่างที่ไม่มีความสัมพันธ์กันหรือเป็นอิสระจากกัน 2 กลุ่ม เพื่อตัดสินใจว่าข้อมูลระหว่าง 2 กลุ่มตัวอย่างมีความเหมือนกันหรือแตกต่างกัน

วิธีการทดสอบแบบนี้ใช้แทนกรณีที่ข้อมูลไม่เป็นไปตามข้อตกลงเบื้องต้นของการทดสอบค่าเฉลี่ยแบบ Independent t-Test โดยมีวิธีดังนี้

1. การทดสอบของแมน-วิทนี ยู (Mann-Whitney U Test) อาจเรียกอีกอย่างว่า การทดสอบของแมน-วิทนี-วิลคอกซัน (Mann-Whitney-Wilcoxon Test) หรือการทดสอบแบบผลรวมของอันดับที่ของวิลคอกซัน (Wilcoxon rank-sum Test) เป็นการทดสอบความเท่ากันของการกระจายของข้อมูลจากตำแหน่งหรือแนวโน้มเข้าสู่ส่วนกลางของข้อมูลจากกลุ่มตัวอย่าง 2 กลุ่มที่เป็นอิสระจากกัน

2. การทดสอบของคอลมอโกรอฟ-สเมอร์นอฟ (Kolmogorov-Smirnov Z) เป็นการทดสอบความแตกต่างกันของการกระจายทั้งจากตำแหน่งและรูปร่างของข้อมูลจากกลุ่มตัวอย่าง 2 กลุ่มที่เป็นอิสระจากกัน จากค่าสูงสุดสัมบูรณ์ (absolute maximum) ของข้อมูลการกระจายสะสม (cumulative distribution)

3. การทดสอบค่าต่างสุดของโมเซส (Moses extreme reactions) เป็นการทดสอบความแตกต่างกันของการกระจาย (distribution) ของตัวแปรแบบสองทิศในทางตรงข้าม (opposite direction) ที่ได้จากการทดลอง 2 กลุ่ม โดยนำเอาข้อมูลที่ได้จากการสังเกตมารวมกัน (combined) และทำการจัดตำแหน่ง (ranked) เพื่อวัดช่วง (span) ของตำแหน่งสูงสุดและต่ำสุด แล้วทำการเปรียบเทียบค่าต่างสุดกับกลุ่มทดลอง

4. การทดสอบของวอร์ค–วูล์ฟอวิทซ์ (Wald-Wolfowitz runs) เป็นการทดสอบความแตกต่างกันของการกระจายทั้งจากตำแหน่งและรูปร่างของข้อมูลจากกลุ่มตัวอย่าง 2 กลุ่มที่เป็นอิสระจากกัน จากการนำเอาข้อมูลจากกลุ่มตัวอย่าง 2 กลุ่มมารวมกัน (combines) และจัดตำแหน่ง (ranks) แล้วนับจำนวนความต่อเนื่อง (runs)

วิธีการทดสอบค่ากลางของกลุ่มตัวอย่างที่เป็นอิสระจากกัน 2 กลุ่ม ที่นิยมกันอย่างแพร่หลาย คือ การทดสอบแบบแมน-วิทนี ยู (Mann-Whitney U Test)

ตัวอย่างคำถามการวิจัย
จากข้อมูลการสำรวจพนักงานในหน่วยงานต้องการทราบว่าพนักงานที่มีเป็น กลุ่มชนแตกต่างกันมีประสบการณ์ก่อนการทำงานแตกต่างกันหรือไม่

2.1 ข้อตกลงเบื้องต้น

ขั้นตอน	ข้อตกลง
การออกแบบการวิจัย	1. ตัวอย่างได้มาจากการสุ่มแบบสุ่ม (random sampling) และเป็นอิสระจากกัน (independent sample) 2. ค่าสังเกตแต่ละค่ามีความเป็นอิสระจากกัน (independence of observations) หรือค่าของตัวแปรแบ่งเป็น 2 กลุ่มและไม่มีความสัมพันธ์ระหว่างกัน 3. ตัวแปรอิสระมีการวัดแบบมาตราจัดอันดับ (ordinal level) หรือเป็นค่าต่อเนื่อง (continuous level)
การวิเคราะห์ข้อมูล	1. ข้อมูลของแต่ละกลุ่มมีการแจกแจงไม่ปกติ (asymmetrical distribution) 2. ข้อมูลของแต่ละกลุ่มมีรูปแบบการแจกแจงเหมือนกัน (same distribution)

2.2 รูปแบบการทดสอบตัวแปร

ตัวแปรอิสระ		ตัวแปรตาม
ค่าจัดประเภท (มาตราแบบจัดกลุ่ม 2 กลุ่ม)		ค่าต่อเนื่อง
กลุ่มชน (minority) - ชนกลุ่มหลัก - ชนกลุ่มน้อย	->	ประสบการณ์ก่อนทำงาน (prevexp)

2.3 การกำหนดสมมติฐานและเกณฑ์การทดสอบ

H_0: พนักงานชนกลุ่มหลักกับพนักงานชนกลุ่มน้อยมีประสบการณ์ก่อนทำงานไม่แตกต่างกัน

H_1: พนักงานชนกลุ่มหลักกับพนักงานชนกลุ่มน้อยมีประสบการณ์ก่อนทำงานแตกต่างกัน

เกณฑ์ในการทดสอบสมมุติฐาน คือ หากการทดสอบความแตกต่างของการกระจายของข้อมูลโดยใช้วิธีของแมน-วิทนี ยู (Mann-Whitney U Test) ได้ค่าความน่าจะเป็น (p-value) มากกว่าค่านัยสำคัญทางสถิติ (α = 0.05) จะยอมรับสมมติฐาน H_0 แต่หากได้ค่าความน่าจะเป็น (p-value) น้อยกว่าหรือเท่ากับค่านัยสำคัญทางสถิติ (α = 0.05) จะปฏิเสธสมมติฐาน H_0 และยอมรับสมมติฐาน H_1

2.4 การใช้โปรแกรมคอมพิวเตอร์วิเคราะห์ข้อมูล

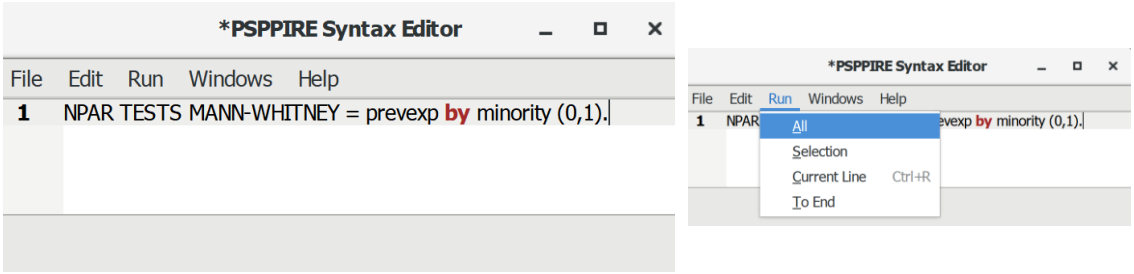
- วิเคราะห์ผ่าน **Syntax Editor** ด้วยคำสั่งดังนี้

  NPAR TESTS

  [/MANN-WHITNEY = var_list BY var (group1, group2)]

- เปิดแฟ้มข้อมูล **File -> Open** -> Employee data.sav
- เลือกเมนู **File -> New -> Syntax**
  - **NPAR TESTS**

    /MANN-WHITNEY = var list (พิมพ์ชื่อตัวแปรทดสอบหรือตัวแปรตามที่เป็นตัวแปรแบบค่าต่อเนื่องแทนคำว่า var list) **BY var** (พิมพ์ชื่อตัวแปรอิสระที่เป็นตัวแปรแบบค่าจัดประเภทแทนคำว่า var) และ (**group1** เช่น 1 หรือ 4, **group2** เช่น 2 หรือ 5).
  - เมนู **Run -> All**



2.5 ผลการวิเคราะห์

Ranks

		N	Mean Rank	Sum of Ranks
Previous Experience (months)	No	370	222.70	82399.50
	Yes	104	290.15	30175.50
	Total	474		

Test Statistics

	Mann-Whitney U	Wilcoxon W	Z	Asymp. Sig. (2-tailed)
Previous Experience (months)	13764.50	82399.50	-4.44	.000

2.6 การสร้างตารางแสดงผลการทดสอบ

ตารางที่ (#) แสดงผลการทดสอบเปรียบเทียบลำดับที่ค่าเฉลี่ยของประสบการณ์ก่อนทำงานของพนักงานระหว่างกลุ่มชน

กลุ่มชน	ลำดับที่ค่าเฉลี่ย (Mean Rank)	ผลรวมของลำดับที่ (Sum of Ranks)	Mann-Whitney U Test	Z	*Sig.*
กลุ่มหลัก	222.70	82399.50	13764.50	-4.44	.000
กลุ่มน้อย	290.15	30175.50			

2.7 การแปลผลและสรุป

จากตารางที่ (#) ผลการทดสอบการเปรียบเทียบลำดับที่ค่าเฉลี่ยของประสบการณ์ก่อนทำงานของพนักงานระหว่างกลุ่มชน แสดงให้เห็นว่า พนักงานที่เป็นชนกลุ่มหลักมีลำดับที่ค่าเฉลี่ยของประสบการณ์ก่อนทำงานน้อยกว่าพนักงานที่เป็นชนกลุ่มน้อย และจากผลการทดสอบความแตกต่างของการกระจายของข้อมูลโดยวิธีของแมน-วิทนี ยู (Mann-Whitney U Test) ได้ค่าความน่าจะเป็น (Sig.) = 0.00 ซึ่งมีค่าน้อยกว่าค่านัยสำคัญทางสถิติที่กำหนดไว้ (α = 0.05) จึงปฏิเสธสมมติฐาน H_0 และยอมรับสมมติฐาน H_1

สรุป พนักงานชนกลุ่มหลักกับพนักงานชนกลุ่มน้อยมีประสบการณ์ก่อนทำงานแตกต่างกัน

3. การทดสอบค่ากลางของกลุ่มตัวอย่างที่เป็นอิสระจากกันมากกว่า 2 กลุ่ม

การทดสอบค่ากลางของกลุ่มตัวอย่างที่เป็นอิสระจากกันมากกว่า 2 กลุ่ม (K Independent Samples Test) คือ การเปรียบเทียบตำแหน่งของข้อมูล (location) โดยใช้ค่ากลาง (median) ของกลุ่มตัวอย่างที่ไม่มีความสัมพันธ์กันหรือเป็นอิสระจากกันมากกว่า 2 กลุ่ม เพื่อตัดสินใจว่าข้อมูลระหว่างกลุ่มตัวอย่างมีความเหมือนกันหรือแตกต่างกัน

วิธีการทดสอบแบบนี้ใช้แทนกรณีที่ข้อมูลไม่เป็นไปตามข้อตกลงเบื้องต้นของการทดสอบค่าเฉลี่ยด้วยวิธีการวิเคราะห์ความแปรปรวนแบบจำแนกทางเดียว (One-way Analysis of Variance: One-way ANOVA) โดยมีวิธีดังนี้

1. การทดสอบของครัสคาลและวัลลิส (Kruskal Wallis Test) เป็นการทดสอบความเหมือนกันของรูปร่างของข้อมูลที่ไม่มีการจัดเรียง (no a priori ordering) จากกลุ่มตัวอย่างมากกว่า 2 กลุ่มที่เป็นอิสระจากกัน (multiple/several independent)

2. การทดสอบของจองฮีรี (Jonckheere Test) หรืออาจเรียกว่า การทดสอบแนวโน้มของจองฮีรี (Jonckheere's Trend Test) หรือการทดสอบของจองฮีรี-เทิร์พสตรา (Jonckheere-Terpstra Test) เป็นการทดสอบความเหมือนกันของรูปร่างของข้อมูลที่มีการจัดเรียง (priori ordered) จากกลุ่มตัวอย่างมากกว่า 2 กลุ่มที่เป็นอิสระจากกัน ข้อมูลที่มีการจัดเรียงจากกลุ่มตัวอย่างอาจเป็นการจัดเรียงแบบจากน้อยไปมาก (ascending) หรือจากมากไปน้อย (descending) ก็ได้ เช่น การเพิ่มขึ้นของอุณหภูมิ การลดของอุณหภูมิ เป็นต้น

3. การทดสอบค่ามัธยฐาน (Median Test) เป็นการทดสอบความเหมือนกันของค่ามัธยฐาน (median) ของข้อมูลจากกลุ่มตัวอย่างมากกว่า 2 กลุ่มที่เป็นอิสระจากกัน

วิธีการทดสอบข้อมูลเชิงคุณภาพของกลุ่มตัวอย่างที่เป็นอิสระจากกันมากกว่า 2 กลุ่ม ที่นิยมกันอย่างแพร่หลาย คือ การทดสอบแบบครัสคาลและวัลลิส (Kruskal Wallis Test) โดยใช้วิธีการเปรียบเทียบค่ามัธยฐาน (median) ของแต่ละกลุ่มตัวอย่าง

ตัวอย่างคำถามการวิจัย
จากข้อมูลการสำรวจพนักงานในหน่วยงานต้องการทราบว่า พนักงานที่มีตำแหน่งงานแตกต่างกันมีประสบการณ์ก่อนการทำงานแตกต่างกันหรือไม่

3.1 ข้อตกลงเบื้องต้น

ขั้นตอน	ข้อตกลง
การออกแบบการวิจัย	1. ตัวอย่างได้มาจากการสุ่มแบบสุ่ม (random sampling) และเป็นอิสระจากกัน (independent sample) 2. ค่าสังเกตแต่ละค่ามีความเป็นอิสระจากกัน (independence of observations) หรือค่าของตัวแปรแบ่งเป็น 3 กลุ่ม หรือมากกว่า และไม่มีความสัมพันธ์ระหว่างกัน 3. ตัวแปรตามมีการวัดแบบมาตราจัดอันดับ (ordinal level) หรือเป็นค่าต่อเนื่อง (continuous level)
การวิเคราะห์ข้อมูล	1. ข้อมูลของแต่ละกลุ่มมีการแจกแจงไม่ปกติ (asymmetrical distribution) 2. ข้อมูลของแต่ละกลุ่มมีการแจกแจงเหมือนกัน (same distribution)

3.2 รูปแบบการทดสอบตัวแปร

ตัวแปรอิสระ		**ตัวแปรตาม**
ค่าจัดประเภท (มาตราแบบจัดอันดับ 3 กลุ่มหรือมากกว่า)		ค่าต่อเนื่อง
ตำแหน่งงาน (jobcat) - เจ้าหน้าที่ปฏิบัติงาน - หัวหน้างาน - ผู้จัดการ	->	ค่าเฉลี่ยของประสบการณ์ก่อนทำงาน (prevexp)

3.3 การกำหนดสมมติฐานและเกณฑ์การทดสอบ

H_0: พนักงานที่มีตำแหน่งงานแตกต่างกันมีประสบการณ์ก่อนทำงานไม่แตกต่างกัน

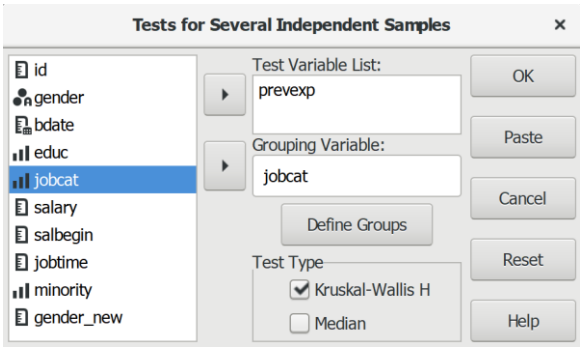
H_1: พนักงานที่มีตำแหน่งงานแตกต่างกันมีประสบการณ์ก่อนทำงานแตกต่างกัน

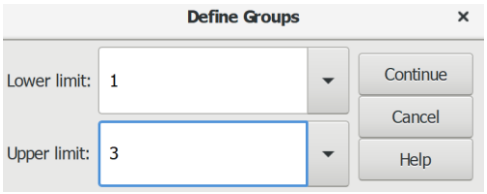
เกณฑ์ในการทดสอบสมมุติฐาน คือ หากการเปรียบเทียบรูปร่างของข้อมูลโดยใช้วิธีของครัสคาลและวัลลิส (Kruskal Wallis Test) ได้ค่าความน่าจะเป็น (p-value) มากกว่าค่านัยสำคัญทางสถิติ (α =

0.05) จะยอมรับสมมติฐาน H_0 แต่หากได้ค่าความน่าจะเป็น (p-value) น้อยกว่าหรือเท่ากับค่านัยสำคัญทางสถิติ (α = 0.05) จะปฏิเสธสมมติฐาน H_0 และยอมรับสมมติฐาน H_1

3.4 การใช้โปรแกรมคอมพิวเตอร์วิเคราะห์ข้อมูล

- เปิดแฟ้มข้อมูล **File -> Open** -> Employee data.sav
- เลือกเมนู **Analyze -> Non-Parametric Statistics -> K Independent Samples...**
  - **Test Variable List:** เลือกตัวแปรทดสอบ (ตัวแปรตาม) ที่เป็นตัวแปรแบบค่าจัดอันดับหรือค่าต่อเนื่อง 1 ตัวแปร หรือมากกว่า 1 ตัวแปร จากรายการตัวแปรด้านซ้าย
  - **Grouping Variable:** เลือกตัวแปรอิสระที่เป็นตัวแปรแบบค่าจัดประเภท 1 ตัวแปร จากรายการตัวแปรด้านซ้าย
  - กดปุ่ม **Define Range...**
  - กำหนดค่าของตัวแปรแบบค่าจัดประเภทในช่อง **Lower limit:** (ค่าต่ำสุดของกลุ่ม) และ **Upper limit:** (ค่าสูงสุดของกลุ่ม) -> กดปุ่ม **Continue**
  - ในกลุ่มคำสั่ง **Test Type** เลือกสถิติ **Kruskal-Wallis H**
  - กดปุ่ม **OK**





3.5 ผลการวิเคราะห์

Ranks

		N	Mean Rank
Previous Experience (months)	Clerical	363	224.53
	Custodial	27	431.26
	Manager	84	231.27
	Total	474	

Test Statistics

	Previous Experience (months)
Chi-Square	57.47
df	2
Asymp. Sig.	.000

3.6 การสร้างตารางแสดงผลการทดสอบ

ตารางที่ (#) แสดงผลการทดสอบเปรียบเทียบลำดับที่ค่าเฉลี่ยของประสบการณ์ก่อนทำงานของพนักงานระหว่างตำแหน่งงาน

ตำแหน่งงาน	จำนวน	ค่าเฉลี่ย	ค่าเบี่ยงเบนมาตรฐาน	ลำดับที่ค่าเฉลี่ย (Mean Rank)	Kruskal Wallis Test
เจ้าหน้าที่ปฏิบัติงาน	363	85.04	95.275	224.53	X^2 = 57.466
หัวหน้างาน	27	298.11	101.426	431.26	df = 2
ผู้จัดการ	84	77.62	73.260	231.27	Asymp. Sig = 0.00
รวม	474	95.86	104.586		

3.7 การแปลผลและสรุป

จากการเปรียบเทียบลำดับที่ค่าเฉลี่ยของประสบการณ์ก่อนทำงานของพนักงานระหว่างตำแหน่งงาน แสดงให้เห็นว่า พนักงานที่เป็นชนกลุ่มหลักมีลำดับที่ค่าเฉลี่ยของประสบการณ์ก่อนทำงานหัวหน้างานมีลำดับที่ค่าเฉลี่ยของประสบการณ์ก่อนทำงานมากที่สุด รองลงมา คือ เจ้าหน้าที่ปฏิบัติงาน และน้อยที่สุด คือ ผู้จัดการ และจากการทดสอบการเปรียบเทียบรูปร่างของข้อมูลโดยใช้วิธีของครัสคาลและวัลลิส (Kruskal Wallis Test) ได้ค่าความน่าจะเป็น (Sig.) = 0.00 ซึ่งมีค่าน้อยกว่าค่านัยสำคัญทางสถิติที่กำหนดไว้ (α = 0.05) จึงปฏิเสธสมมติฐาน H_0 และยอมรับสมมติฐาน H_1

สรุป พนักงานที่มีตำแหน่งงานแตกต่างกันมีประสบการณ์ก่อนทำงานแตกต่างกัน

4. การทดสอบค่ากลางของกลุ่มตัวอย่างที่มีความสัมพันธ์กัน 2 กลุ่ม

การทดสอบค่ากลางของกลุ่มตัวอย่างที่มีความสัมพันธ์กัน 2 กลุ่ม (2 Related Samples Test) คือ การทดสอบเปรียบเทียบตำแหน่งข้อมูล (location) โดยใช้ค่ากลาง (median) ของกลุ่มตัวอย่างที่มีความสัมพันธ์กันหรือไม่เป็นอิสระจากกันมากกว่า 2 กลุ่ม เพื่อตัดสินใจว่าข้อมูลระหว่างกลุ่มตัวอย่างมีความเหมือนกันหรือแตกต่างกัน

การทดสอบแบบนี้อาจมีชื่อเรียกในชื่ออื่น เช่น Paired Samples Test, Matched Samples Test และ Dependent Samples Test

วิธีการทดสอบแบบนี้ใช้แทนกรณีที่ข้อมูลไม่เป็นไปตามข้อตกลงเบื้องต้น (not meet basic assumptions) ของการทดสอบค่าเฉลี่ย 2 กลุ่มที่สัมพันธ์กัน (Pair Sample t-test) โดยมีวิธีการดังนี้

1. การทดสอบแบบเครื่องหมาย (Sign Test) เป็นการทดสอบความแตกต่างกันของค่ามัธยฐาน (median) ของข้อมูลต่อเนื่อง (continuous data) จากเครื่องหมาย + และเครื่องหมาย – ที่ได้จากข้อมูลที่นำมาจับคู่กัน ($X_i > Y_i$ ได้เครื่องหมาย + แต่ถ้า $X_i < Y_i$ จะได้เครื่องหมาย -)

2. การทดสอบแบบวิลคอกซัน (Wilcoxon Test หรือ Wilcoxon Singed Ranks Test) เป็นการทดสอบความเหมือนกันของค่ามัธยฐาน (median) ของข้อมูลต่อเนื่อง (continuous data) ที่มีความสัมพันธ์กัน (related) จากกลุ่มตัวอย่างเดียวกันหรือคู่เดียวกัน (same/matched sample) ที่พัฒนามาจากการทดสอบเครื่องหมาย โดยการนำค่าของความแตกต่างมาจัดอันดับในการทดสอบ

3. การทดสอบแบบแมคนีมาร์ (McNemar Test) เป็นการทดสอบความไม่เปลี่ยนแปลง (unchanged) ของข้อมูลแบบ 2 ค่า (binary response) หรือตัวแปรที่มีค่า 2 ค่า (data: binary) เป็นการทดสอบที่สามารถใช้ได้ทั้งตัวแปรแบบมาตราจัดกลุ่ม (nominal) และจัดอันดับ (ordinal)

4. การทดสอบความเป็นเอกพันธ์ของส่วนเพิ่ม (Marginal Homogeneity Test) เป็นการทดสอบความเหมือนกันของการกระจายของส่วนเพิ่มของตัวแปรพหุวิภาคหรือตัวแปรหลายค่า (polytomous variables) จากกลุ่มตัวอย่าง 2 กลุ่มที่มีความสัมพันธ์กัน

วิธีการทดสอบกลุ่มตัวอย่าง 2 กลุ่มที่มีความสัมพันธ์กันที่นิยมกันอย่างแพร่หลาย คือ การทดสอบแบบวิลคอกซัน (Wilcoxon Test หรือ Wilcoxon Singed Ranks Test)

ตัวอย่างคำถามการวิจัย
จากข้อมูลการสำรวจพนักงานในหน่วยงานต้องการทราบว่า พนักงานที่ทำงานอยู่ขณะนี้ได้รับเงินเดือนแตกต่างไปจากเงินเดือนเริ่มต้นหรือไม่

4.1 ข้อตกลงเบื้องต้น

ขั้นตอน	ข้อตกลง
การออกแบบการวิจัย	1. ตัวอย่างได้มาจากการสุ่มแบบสุ่ม (random sampling) และเป็นอิสระจากกัน (independent sample) 2. ข้อมูลได้มาจากกลุ่มตัวอย่างเดียวกันหรือคู่เดียวกัน (same/matched sample) และเป็นสถานการณ์ที่มีความเกี่ยวข้องกันหรือสัมพันธ์กัน 2 สถานการณ์ (2 related samples or dependent/paired sample) หรือเป็นการวัดซ้ำ (repeated measures data) 3. ตัวแปรต้องมีระดับการแบบค่าจัดอันดับหรือค่าต่อเนื่อง (ordinal or continuous level) 4. ตัวแปรทั้งคู่ต้องมีระดับการวัดอยู่ในระดับเดียวกัน (มาตราแบบจัดอันดับกับมาตราแบบจัดอันดับ หรือค่าต่อเนื่องกับค่าต่อเนื่อง)
การวิเคราะห์ข้อมูล	1. ข้อมูลของ 2 ตัวแปร มีการแจกแจงไม่ปกติ (asymmetrical distribution) 2. ข้อมูลของ 2 ตัวแปร มีการแจกแจงเหมือนกัน (same distribution)

4.2 รูปแบบการทดสอบตัวแปร

ตัวแปรทดสอบ (การวัดครั้งที่ 1)	วัดซ้ำ	ตัวแปรทดสอบ (การวัดครั้งที่ 2)
ค่าต่อเนื่อง	->	ค่าต่อเนื่อง
เงินเดือนเริ่มต้น (salbegin)		เงินเดือนปัจจุบัน (salary)

4.3 การกำหนดสมมติฐานและเกณฑ์การทดสอบ

H_0: พนักงานมีเงินเดือนปัจจุบันเท่ากับเงินเดือนเริ่มต้น

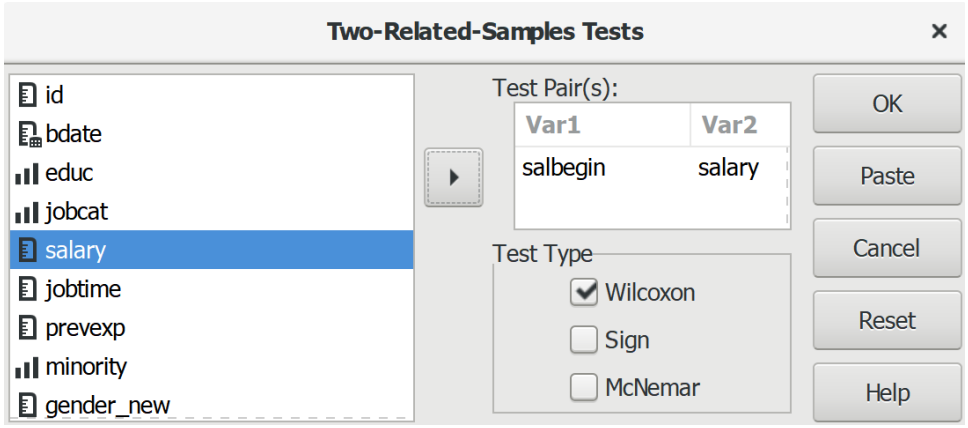
H_1: พนักงานมีเงินเดือนปัจจุบันไม่เท่ากับเงินเดือนเริ่มต้น

เกณฑ์ในการทดสอบสมมุติฐาน คือ หากการทดสอบเปรียบเทียบค่ามัธยฐานโดยใช้วิธีของวิลคอกซัน (Wilcoxon Test) ได้ค่าความน่าจะเป็น (p-value) มากกว่าค่านัยสำคัญทางสถิติ (α = 0.05) จะ

ยอมรับสมมติฐาน H_0 แต่หากได้ค่าความน่าจะเป็น (p-value) น้อยกว่าหรือเท่ากับค่านัยสำคัญทางสถิติ (α = 0.05) จะปฏิเสธสมมติฐาน H_0 และยอมรับสมมติฐาน H_1

4.4 การใช้โปรแกรมคอมพิวเตอร์วิเคราะห์ข้อมูล

- เปิดแฟ้มข้อมูล **File -> Open** -> Employee data.sav
- เลือกเมนู **Analyze -> Non-Parametric Statistics -> 2 Related Samples...**
  - **Test Pair(s):** เลือกตัวแปรทดสอบตัวที่ 1 (ตัวแปรแบบค่าต่อเนื่อง/เหตุการณ์ก่อน) จากช่องรายการตัวแปรใส่ในช่อง **Var1**
  - **Test Pair(s):** เลือกตัวแปรทดสอบตัวที่ 2 (ตัวแปรแบบค่าต่อเนื่อง/เหตุการณ์หลัง) จากช่องรายการตัวแปรใส่ในช่อง **Var2**
  - ในกลุ่มคำสั่ง **Test Type** เลือกสถิติ **Wilcoxon**
  - กดปุ่ม **OK**



4.5 ผลการวิเคราะห์

Ranks

		N	Mean Rank	Sum of Ranks
Beginning Salary - Current Salary	Negative Ranks	474	237.50	112575.0
	Positive Ranks	0	NaN	.00
	Ties	0		
	Total	474		

Test Statistics

	Beginning Salary - Current Salary
Z	-18.86
Asymp. Sig. (2-tailed)	.000

4.6 การสร้างตารางแสดงผลการทดสอบ

ตารางที่ (#) แสดงผลการทดสอบเปรียบเทียบลำดับที่ค่าเฉลี่ยระหว่างเงินเดือนเริ่มต้นกับเงินเดือนปัจจุบัน

เงินเดือน	จำนวน		ลำดับที่ค่าเฉลี่ย (Mean Rank)	ผลรวมของลำดับที่ (Sum of Ranks)	Wilcoxon Test
เงินเดือนปัจจุบัน-เงินเดือนเริ่มต้น	อันดับค่าที่มีค่าน้อยกว่า (Negative Ranks)	474	237.50	112575.00	Z = -18.86
	อันดับค่าที่มีค่ามากกว่า (Positive Ranks)	0	0	.00	Sig = 0.00
	อันดับค่าที่มีค่าเท่ากัน (Ties)	0			
รวม		474			

4.7 การแปลผลและสรุป

จากตารางที่ (#) ผลการทดสอบการเปรียบเทียบค่ามัธยฐานโดยใช้วิธีของวิลคอกซัน (Wilcoxon Test) ได้ค่าความน่าจะเป็น (Sig.) = 0.00 มีค่าน้อยกว่าค่านัยสำคัญทางสถิติที่กำหนดไว้ (α = 0.05) จึงปฏิเสธสมมติฐาน H_0 และยอมรับสมมติฐาน H_1

สรุป พนักงานมีเงินเดือนปัจจุบันมากกว่าเงินเดือนเริ่มต้น พนักงานมีเงินเดือนปัจจุบันมากกว่าเงินเดือนเริ่มต้นจำนวน 474 คน ไม่มีพนักงานที่มีเงินเดือนปัจจุบันน้อยกว่าเงินเดือนเริ่มต้น และไม่มีพนักงานที่มีเงินเดือนปัจจุบันเท่ากับเงินเดือนเริ่มต้น

5. การทดสอบค่ากลางของกลุ่มตัวอย่างที่มีความสัมพันธ์กันมากกว่า 2 กลุ่ม

การทดสอบค่ากลางของกลุ่มตัวอย่างที่มีความสัมพันธ์กันมากกว่า 2 กลุ่ม (K Related Samples Tests) คือ การเปรียบเทียบตำแหน่ง (location) โดยใช้ค่ากลาง (median) ของกลุ่มตัวอย่างที่มีความสัมพันธ์กันมากกว่า 2 กลุ่ม เพื่อตัดสินใจว่าข้อมูลระหว่างกลุ่มตัวอย่างมีความเหมือนกันหรือแตกต่างกัน

วิธีการทดสอบแบบนี้ใช้แทนกรณีที่ข้อมูลไม่เป็นไปตามข้อตกลงเบื้องต้น (not meet basic assumptions) ของการทดสอบค่าเฉลี่ยมากกว่า 2 กลุ่มที่สัมพันธ์กัน (One-Way repeated-measures ANOVA) โดยมีวิธีดังนี้

1. การทดสอบแบบฟรีดแมน (Friedman Test) เป็นการทดสอบความเหมือนของข้อมูลจากตัวแปรแบบมาตราจัดอันดับแบบพหุ (multiple ordinal responses) ที่ได้จากการวัดซ้ำแบบเดียวกัน (same measure) จากกลุ่มตัวอย่างเดียวกันแบบพหุ (multiple matched samples)

2. การทดสอบแบบเคนดอลล์ (Kendall's W Test) เป็นวิธีการทดสอบสัมประสิทธิ์ของความสอดคล้องในการจัดลำดับที่ (coefficient of concordance for ranks: W) จากคำตอบความคิดเห็น (agreement) ของผู้ตอบที่มีค่าแบบมาตราจัดอันดับช่วง 0 ถึง n-1 โดยค่า 0 = ไม่เห็นด้วย (no agreement) และ 1 = เห็นด้วยทั้งหมด (complete agreement) เช่น ค่า 1=0 ค่า 2=1 และค่า 9=8 เป็นต้น

3. การทดสอบแบบคอคราน (Cochran's Q Test) เป็นการทดสอบความเหมือนกันของสัดส่วนของข้อมูลจากตัวแปรแบบสองค่า (dichotomous) ที่มีความสัมพันธ์กันแบบพหุ (multiple related) ที่พัฒนามาจากการทดสอบแบบแมคนีมาร์ (McNemar) ที่ใช้ในการทดสอบข้อมูล 2 ค่า

วิธีการทดสอบความสัมพันธ์ของกลุ่มตัวอย่างมากกว่า 2 กลุ่มที่ใช้เป็นตัวอย่างการทดสอบ คือ การทดสอบแบบฟรีดแมน (Friedman Test)

ตัวอย่างคำถามการวิจัย
จากคะแนนการตัดสินของคณะกรรมการทั้งหมดการทราบว่า คณะกรรมการของประเทศอิตาลี ประเทศเกาหลีใต้ และประเทศโรมาเนีย ให้คะแนนแตกต่างกันหรือไม่

5.1 ข้อตกลงเบื้องต้น

ขั้นตอน	ข้อตกลง
การออกแบบการวิจัย	1. ตัวอย่างได้มาจากการสุ่มแบบสุ่ม (random sampling) และเป็นอิสระจากกัน (independent sample) 2. ข้อมูลของตัวแปรทดสอบมีความเกี่ยวข้องหรือสัมพันธ์กันมากกว่า 2 สถานการณ์ (K related sample) 3. ข้อมูลของตัวแปรทดสอบมีระดับการวัดด้วยมาตราแบบจัดอันดับหรือค่าต่อเนื่อง (ordinal or continuous level) 4. ข้อมูลของตัวแปรทดสอบทุกตัวต้องมีระดับการวัดอยู่ในระดับเดียวกัน (มาตราแบบจัดอันดับทุกตัวแปร หรือค่าต่อเนื่องทุกตัวแปร)
การวิเคราะห์ข้อมูล	ข้อมูลมีการแจกแจงไม่ปกติ (asymmetrical distribution)

5.2 รูปแบบการทดสอบตัวแปร

ตัวแปรทดสอบ (การวัดครั้งที่ 1)	ตัวแปรทดสอบ (การวัดครั้งที่ 2)	ตัวแปรทดสอบ (การวัดครั้งที่ n)
ค่าต่อเนื่อง	ค่าต่อเนื่อง	ค่าต่อเนื่อง
คะแนนของกรรมการคนที่ 1 อิตาลี่ (judge1)	คะแนนของกรรมการคนที่ 2 เกาหลีใต้ (judge3)	คะแนนของกรรมการคนที่ n โรมาเนีย (judge3)

5.3 การกำหนดสมมติฐานและเกณฑ์การทดสอบ

จากการแข่งขันยิมนาสติก ต้องการทดสอบความเหมือนกันของคะแนนการตัดสินนักยิมนาสติกของกรรมการจากประเทศอิตาลี เกาหลีใต้ และโรมาเนีย

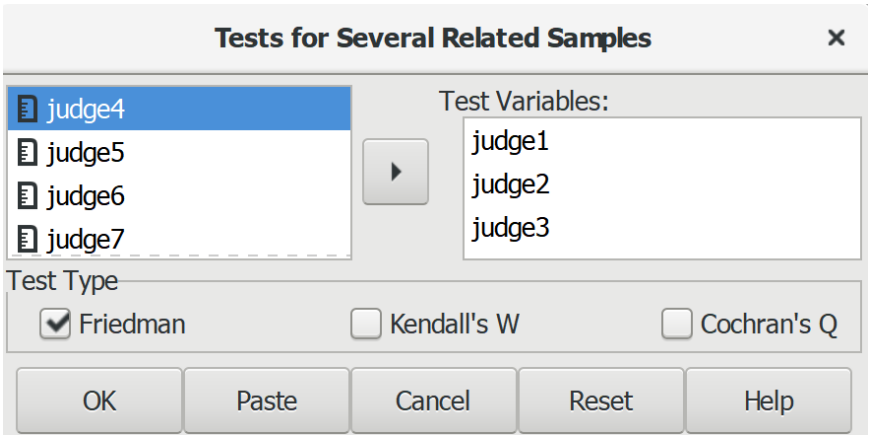
H_0: คะแนนการตัดสินของกรรมการจากประเทศอิตาลี เกาหลีใต้ และโรมาเนียเหมือนกัน

H_1: คะแนนการตัดสินของกรรมการจากประเทศอิตาลี เกาหลีใต้ และโรมาเนียไม่เหมือนกัน

เกณฑ์ในการทดสอบสมมุติฐาน คือ หากการทดสอบความเหมือนกันของลำดับที่คะแนนโดยใช้วิธีของฟรีดแมน (Friedman Test) ได้ค่าความน่าจะเป็น (p-value) มากกว่าค่านัยสำคัญทางสถิติ (α = 0.05) จะยอมรับสมมติฐาน H_0 แต่หากได้ค่าความน่าจะเป็น (p-value) น้อยกว่าหรือเท่ากับค่านัยสำคัญทางสถิติ (α = 0.05) จะปฏิเสธสมมติฐาน H_0 และยอมรับสมมติฐาน H_1

5.4 การใช้โปรแกรมคอมพิวเตอร์วิเคราะห์ข้อมูล

- เปิดแฟ้มข้อมูล **File -> Open** -> judges.sav
- เลือกเมนู **Analyze -> Non-Parametric Statistics -> K Related Samples...**
  - **Test Variables:** เลือกตัวแปรทดสอบ 3 ตัวแปรหรือมากกว่า (ตัวแปรแบบค่าต่อเนื่อง/เหตุการณ์เหมือนกันหรือเป็นการวัดซ้ำ 3 เหตุการณ์หรือมากกว่า เช่น การให้คะแนนของกรรมการ 3 คนหรือมากกว่า) จากรายการตัวแปรด้านซ้าย
  - ในกลุ่มคำสั่ง **Test Type** เลือกสถิติ **Friedman**
  - กดปุ่ม **OK**



5.5 ผลการวิเคราะห์

Ranks

	Mean Rank
Italy	1.99
South Korea	2.90
Romania	1.11

Test Statistics

N	300
Chi-Square	497.80
df	2
Asymp. Sig.	.000

5.7 การสร้างตารางแสดงผลการทดสอบ

ตารางที่ (#) แสดงผลการทดสอบเปรียบเทียบลำดับที่ค่าเฉลี่ยของคะแนนจากกรรมการ 3 ประเทศ

กรรมการ	ลำดับที่ค่าเฉลี่ย (Mean Rank)	Friedman Test
อิตาลี	1.99	N = 300 X^2 = 497.803 df = 2 Sig = 0.000
เกาหลีใต้	2.90	
โรมาเนีย	1.11	

5.8 การแปลผลและสรุป

จากตารางที่ (#) ผลการทดสอบแสดงให้เห็นว่า ค่าความน่าจะเป็น (Sig.) = 0.00 มีค่าน้อยค่านัยสำคัญทางสถิติที่กำหนดไว้ (α = 0.05) จึงปฏิเสธสมมติฐาน H_0 และยอมรับสมมติฐาน H_1

สรุป คะแนนการตัดสินของกรรมการจากประเทศอิตาลี เกาหลีใต้ และโรมาเนียแตกต่างกัน โดยกรรมการจากประเทศเกาหลีใต้ให้คะแนนมากที่สุดเป็นอันดับ 1 กรรมการจากประเทศอิตาลีให้คะแนนน้อยรองลงมาเป็นอันดับที่ 2 และกรรมการจากประเทศโรมาเนียให้คะแนนน้อยที่สุดเป็นอันดับที่ 3

ตัวอย่างเพิ่มเติม: การทดสอบ การแปลผลและสรุป

#1

H_0: คะแนนการตัดสินของกรรมการจากประเทศอิตาลี เกาหลีใต้ และโรมาเนียไม่แตกต่างกัน

H_1: คะแนนการตัดสินของกรรมการจากประเทศอิตาลี เกาหลีใต้ และโรมาเนียแตกต่างกัน

เกณฑ์ในการทดสอบสมมุติฐาน คือ หากการทดสอบความสอดคล้องด้วยโดยใช้วิธีของเคนดอลล์ (Kendall's W Test)ได้ค่าความน่าจะเป็น (p-value) มากกว่าค่านัยสำคัญทางสถิติ (α = 0.05) จะยอมรับสมมติฐาน H_0 แต่หากได้ค่าความน่าจะเป็น (p-value) น้อยกว่าหรือเท่ากับค่านัยสำคัญทางสถิติ (α = 0.05) จะปฏิเสธสมมติฐาน H_0 และยอมรับสมมติฐาน H_1

ตารางที่ (#) แสดงผลการทดสอบเปรียบเทียบลำดับที่ค่าเฉลี่ยของคะแนนจากกรรมการ 3 ประเทศ

กรรมการ	ลำดับที่ค่าเฉลี่ย	Kendall's W Test
อิตาลี	1.99	N = 300 Kendall's W = .830
เกาหลีใต้	2.90	X^2 = 497.803 df = 2
โรมาเนีย	1.11	Asymp. Sig = 0.000

จากตารางที่ (#) ผลการทดสอบแสดงให้เห็นว่า ค่าความน่าจะเป็น (Sig.) = 0.00 มีค่าน้อยกว่าค่านัยสำคัญทางสถิติที่กำหนดไว้ (α = 0.05) จึงปฏิเสธสมมติฐาน H_0 และยอมรับสมมติฐาน H_1

สรุป คะแนนการตัดสินของกรรมการจากประเทศอิตาลี เกาหลีใต้ และโรมาเนียแตกต่างกัน โดยมีค่า Kendall's W = .830 หรือมีความสอดคล้องของคะแนนอยู่ระดับมาก กรรมการจากประเทศเกาหลีใต้ให้คะแนนมีลำดับที่ค่าเฉลี่ยมากที่สุด รองลงมาคือ กรรมการจากประเทศอิตาลี และน้อยที่สุดคือ กรรมการจากประเทศโรมาเนีย

6. การทดสอบความสัมพันธ์รายคู่แบบสเปียร์แมน

การทดสอบความสัมพันธ์กรณีข้อมูลของตัวแปรไม่เป็นค่าต่อเนื่อง มีวิธีการทดสอบความสัมพันธ์ระหว่างตัวแปร 2 ตัว ที่เป็นตัวแปรแบบค่าไม่ต่อเนื่อง (discontinuous variable) กับตัวแปรแบบค่าไม่ต่อเนื่อง (discontinuous variable) ด้วยการจัดอันดับของข้อมูล (ranks of the data) โดยมีวิธีดังนี้

1. การทดสอบความสัมพันธ์ด้วยสัมประสิทธิ์สหสัมพันธ์อันดับที่ของสเปียร์แมนหรือสเปียร์แมนโรท์ (Spearman's Rank Correlation Coefficient or Spearman's rho (ρ) Test) หรือ เป็นการวัดระดับความสัมพันธ์ (association) โดยตัวแปรมีการเรียงลำดับไปในทิศทางเดียวกัน (monotonicity) เหมาะกับตัวแปรแบบค่าต่อเนื่อง (continuous variable) ตัวแปรแบบค่าไม่ต่อเนื่อง (discrete variable) และตัวแปรแบบมาตราจัดอันดับ (order variable)

2. การทดสอบความสัมพันธ์ด้วยสัมประสิทธิ์สหสัมพันธ์อันดับของเคนดอล์ล (Kendall rank correlation coefficient) หรือ สัมประสิทธิ์แบบเทาว์ของเคนดอล์ล ทาว (Kendall Tau (τ) coefficient) เป็นการวัดระดับความสอดคล้อง (correspondence) ที่เกิดจากความแตกต่างระหว่างจำนวนคู่ที่เป็นสัดส่วน (concordant) หรือคู่ของลำดับที่มีลักษณะตรงกันกับจำนวนคู่ที่ไม่เป็นสัดส่วน (discordant) ประกอบด้วย

2.1 ทาว-เอ (Tau-a) เป็นสถิติทดสอบความเข้มของความสัมพันธ์ (strength of association) ของตารางไขว้ (cross tabulations) ตัวแปรทั้งสองเป็นตัวแปรแบบมาตราจัดอันดับ โดยการทดสอบแบบทาว-บี จะไม่ทำการปรับค่ามีค่าเท่ากัน (ties)

2.2 ทาว-บี (Tau-b) เป็นสถิติทดสอบที่แตกต่างจากสถิติทดสอบแบบทาว-เอ เพราะทำการปรับค่าที่มีค่าเท่ากัน (ties) ที่ได้จากการทดสอบความสัมพันธ์ของข้อมูลที่อยู่ในรูปของตารางสี่เหลี่ยมจตุรัส (square tables) ค่าของทาว-บี อยู่ในช่วง -1 (มีความสัมพันธ์เชิงลบ 100% หรือตรงข้ามอย่างสมบูรณ์) ถึง +1 (มีความสัมพันธ์เชิงบวก 100% หรือเห็นด้วยอย่างสมบูรณ์) และค่า 0 เป็นดัชนีของการไม่มีความสัมพันธ์

2.3 ทาว-ซี (Tau-c) แตกต่างจากทาว-บี คือ เหมาะกับการทดสอบข้อมูลที่อยู่ในรูปของตารางแบบสี่เหลี่ยมผืนผ้า (rectangular table) มากกว่าตารางแบบสี่เหลี่ยมจตุรัส (square tables)

การวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรหากข้อมูลไม่เป็นไปตามข้อตกลงของการวิเคราะห์ความสัมพันธ์แบบเพียร์สันสามารถใช้วิธีการทดสอบความสัมพันธ์แบบสเปียร์แมนแทน ผลการวิเคราะห์ให้ค่าความเข้มของความสัมพันธ์มากหรือน้อย และบอกทิศทางของความสัมพันธ์เชิงบวกหรือเชิงลบดังนี้

ตารางที่ 14.1 การแปลค่าสัมประสิทธิ์สหสัมพันธ์แบบสเปียร์แมน

ค่าสัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficient)	
ค่าของความสัมพันธ์	**ความเข้มและทิศทางของความสัมพันธ์**
1.00 ถึง 0.80	มีความสัมพันธ์เชิงบวกระดับมากอย่างยิ่ง
0.79 ถึง 0.60	มีความสัมพันธ์เชิงบวกระดับมาก
0.59 ถึง 0.40	มีความสัมพันธ์เชิงบวกระดับปานกลาง
0.39 ถึง 0.20	มีความสัมพันธ์เชิงบวกระดับน้อย
0.19 ถึง 0.01	มีความสัมพันธ์เชิงบวกระดับน้อยมาก
0.00	ไม่มีความสัมพันธ์
-0.01 ถึง -0.19	มีความสัมพันธ์เชิงลบระดับน้อยมาก
-0.39 ถึง -0.20	มีความสัมพันธ์เชิงลบระดับน้อย
-0.59 ถึง -0.40	มีความสัมพันธ์เชิงลบระดับปานกลาง
-0.79 ถึง -0.60	มีความสัมพันธ์เชิงลบระดับมาก
-1.00 ถึง 0.-80	มีความสัมพันธ์เชิงลบระดับมากอย่างยิ่ง

ปรับปรุงมาจาก: Evans, J. D. (1996). Straightforward statistics for the behavioral sciences. Pacific Grove, CA: Brooks/Cole Publishing.

6.1 การทดสอบความสัมพันธ์อันดับที่แบบง่าย

การทดสอบความสัมพันธ์อันดับที่แบบง่าย (Simple correlation) คือ การวิเคราะห์หาความสัมพันธ์ตัวแปร 1 คู่ โดยตัวแปรอิสระหรือตัวแปรตามที่ใช้ในการวิเคราะห์ความสัมพันธ์มีระดับการวัดแบบมาตราจัดอันดับหรือเป็นตัวแปรที่มีระดับการวัดแบบค่าต่อเนื่องแต่ข้อมูลไม่เป็นไปตามข้อตกลงเบื้องต้นของการวิเคราะห์ความสัมพันธ์แบบเพียร์สัน

ตัวอย่างคำถามการวิจัย
จากข้อมูลการสำรวจพนักงานในหน่วยงานต้องการทราบว่า ระดับการศึกษาของพนักงานมีความสัมพันธ์กับเงินเดือนเริ่มต้นหรือไม่

6.1.1 ข้อตกลงเบื้องต้น

ขั้นตอน	ข้อตกลง
การออกแบบการวิจัย	1. ตัวอย่างได้มาจากการสุ่มแบบสุ่ม (random sampling) และเป็นอิสระจากกัน (independent sample) 2. ตัวแปรอิสระเป็นตัวแปรแบบมาตราจัดอันดับ หรือตัวแปรแบบค่าต่อเนื่อง รวมถึงตัวแปรแบบค่าไม่ต่อเนื่อง (discrete variables) 1 ตัว 3. ตัวแปรตามเป็นตัวแปรแบบมาตราจัดอันดับ หรือตัวแปรแบบค่าต่อเนื่อง รวมถึงตัวแปรแบบค่าไม่ต่อเนื่อง (discrete variables) 1 ตัว
การวิเคราะห์ข้อมูล	ข้อมูลมีการแจกแจงไม่ปกติ (asymmetrical distribution)

6.1.2 รูปแบบการทดสอบตัวแปร

ตัวแปรอิสระ		ตัวแปรตาม
ค่าจัดประเภท (จัดอันดับ)		ค่าต่อเนื่อง
การศึกษา (educ) - ประถมศึกษา - มัธยมศึกษา - อุดมศึกษา	->	เงินเดือนเริ่มต้น (salbegin)

6.1.3 การกำหนดสมมติฐานและเกณฑ์การทดสอบ

ตารางที่ (#) แสดงค่าเฉลี่ยของเงินเดือนเริ่มต้น จำแนกตามระดับการศึกษา

ระดับการศึกษา	จำนวน	ค่าเฉลี่ยเงินเดือนเริ่มต้น	S.D.
มัธยมศึกษาตอนต้น	53	13,064.15	2,408.147
มัธยมศึกษาตอนปลาย	190	13,241.87	2,653.896
อุดมศึกษา	231	21,027.14	9,421.873
รวม	474	17,061.09	7870.638

จากตารางที่ (#) แสดงให้เห็นว่า พนักงานมีเงินเดือนเฉลี่ย 17,061.09 ดอลลาร์สหรัฐ โดยพนักงานที่มีการศึกษาระดับอุดมศึกษามีเงินเดือนเฉลี่ยมากที่สุด รองลงมา คือ มัธยมศึกษาตอนปลายและมัธยมศึกษาตอนต้น ตามลำดับ

จากข้อมูลที่กล่าวมาน่าจะมีความสัมพันธ์ระหว่างระดับการศึกษากับเงินเดือนเริ่มต้น ดังนั้นจึงทำการทดสอบความสัมพันธ์ด้วยสัมประสิทธิ์สหสัมพันธ์อันดับที่ของสเปียร์แมน (Spearman's Rank Correlation Coefficient Test) โดยมีสมมติฐานดังนี้

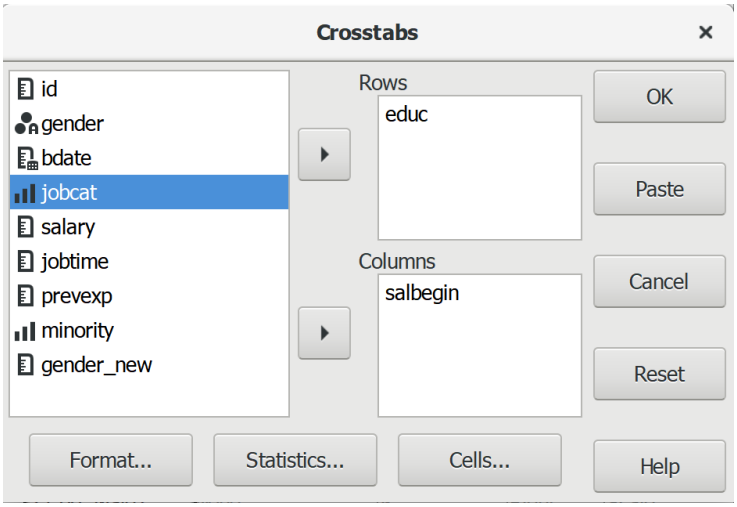
H_0: ระดับการศึกษากับเงินเดือนเริ่มต้นไม่มีความสัมพันธ์กัน

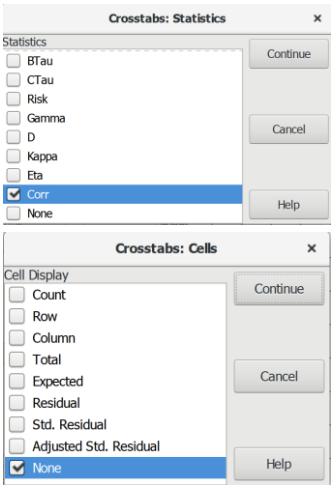
H_1: ระดับการศึกษากับเงินเดือนเริ่มต้นมีความสัมพันธ์กัน

เกณฑ์ในการทดสอบสมมุติฐาน คือ หากการทดสอบความสัมพันธ์ด้วยวิธีของสเปียร์แมน (Spearman's Test) ได้ค่าความน่าจะเป็น (p-value) มากกว่าค่านัยสำคัญทางสถิติ (α = 0.05) จะยอมรับสมมติฐาน H_0 แต่หากได้ค่าความน่าจะเป็น (p-value) น้อยกว่าหรือเท่ากับค่านัยสำคัญทางสถิติ (α = 0.05) จะปฏิเสธสมมติฐาน H_0 และยอมรับสมมติฐาน H_1

6.1.4 การใช้โปรแกรมคอมพิวเตอร์วิเคราะห์ข้อมูล

- เปิดแฟ้มข้อมูล **File -> Open ->** Employee data.sav
- เลือกเมนู **Analyze -> Descriptive Statistics -> Crosstabs...**
  - **Rows:** เลือกตัวแปรแบบค่าจัดประเภท (จัดอันดับ) หรือตัวแปรแบบค่าต่อเนื่อง/ตัวแปรแบบค่าไม่ต่อเนื่อง (discrete variables) 1 ตัวแปร หรือมากกว่า 1 ตัวแปร (ที่ต้องการให้เป็นตัวแปรตาม) ช่องรายการตัวแปรด้านซ้าย
  - **Column:** เลือกตัวแปรแบบค่าจัดประเภท (จัดอันดับ) หรือตัวแปรแบบค่าต่อเนื่อง 1 ตัวแปร หรือมากกว่า 1 ตัวแปร (ที่ต้องการให้เป็นตัวแปรตาม) จากรายการตัวแปรด้านซ้าย
  - กดปุ่ม **Statistics...** ไม่เลือกสถิติ Chisq และเลือกสถิติ Corr แทน -> กดปุ่ม **Continue**
  - กดปุ่ม **Cells...** ไม่เลือก Count, Row, Column, Total และเลือก None -> กดปุ่ม **Continue**
  - กดปุ่ม **OK**





6.1.5 ผลการวิเคราะห์

Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Educational Level (years) × Beginning Salary	474	100.0%	0	.0%	474	100.0%

Symmetric Measures

		Value	Asymp. Std. Error	Approx. T
Ordinal by Ordinal	Spearman Correlation	.69	.03	20.60
Interval by Interval	Pearson's R	.63	.03	17.77
N of Valid Cases		474		

6.1.6 การสร้างตารางแสดงผลการทดสอบ

ตาราง (#) แสดงการทดสอบความสัมพันธ์ระหว่างระดับการศึกษากับเงินเดือนเริ่มต้น

ตัวแปร	จำนวน	เงินเดือนเริ่มต้นของผู้ตอบ (ρ)	Sig. (2-tailed)
ระดับการศึกษา	474	.69	.03

6.1.7 การแปลผลและสรุป

จากตารางที่ (#) ผลการทดสอบแสดงให้เห็นว่า ค่าความน่าจะเป็น (Sig.) = 0.03 มีค่าน้อยกว่าค่านัยสำคัญทางสถิติที่กำหนดไว้ (α = 0.05) จึงปฏิเสธสมมติฐาน H_0 และยอมรับสมมติฐาน H_1 คือ ระดับการศึกษากับเงินเดือนเริ่มต้นมีความสัมพันธ์กัน

สรุป ระดับการศึกษาของพนักงานมีความสัมพันธ์เชิงบวกระดับปานกลางกับเงินเดือนเริ่มต้น

6.2 การทดสอบความสัมพันธ์อันดับที่แบบพหุ

การทดสอบความสัมพันธ์อันดับที่แบบพหุ (Multiple correlation) คือ การวิเคราะห์หาความสัมพันธ์ระหว่างตัวแปรมากกว่า 1 คู่ โดยตัวแปรอิสระหรือตัวแปรตามที่ใช้ในการวิเคราะห์ความสัมพันธ์มีระดับการวัดแบบมาตราจัดอันดับหรือเป็นตัวแปรที่มีระดับการวัดแบบค่าต่อเนื่องแต่ข้อมูลไม่เป็นไปตามข้อตกลงเบื้องต้นของการวิเคราะห์ความสัมพันธ์แบบเพียร์สัน

ตัวอย่างคำถามการวิจัย
จากข้อมูลการสำรวจพนักงานในหน่วยงานต้องการทราบว่า ระดับการศึกษาและประสบการณ์ก่อนการทำงานของพนักงานมีความสัมพันธ์กับเงินเดือนเริ่มต้นหรือไม่

6.2.1 ข้อตกลงเบื้องต้น

ขั้นตอน	**ข้อตกลง**
การออกแบบการวิจัย	1. ตัวแปรอิสระเป็นตัวแปรแบบมาตราจัดอันดับ หรือตัวแปรแบบค่าต่อเนื่อง รวมถึงตัวแปรแบบค่าไม่ต่อเนื่อง (discrete variables) 1 ตัวแปร หรือหลายตัวแปร 2. ตัวแปรตามเป็นตัวแปรแบบมาตราจัดอันดับ หรือตัวแปรแบบค่าต่อเนื่อง รวมถึงตัวแปรแบบค่าไม่ต่อเนื่อง (discrete variables) มากกว่า 1 ตัวแปร
การวิเคราะห์ข้อมูล	ข้อมูลมีการแจกแจงไม่ปกติ (asymmetrical distribution)

6.2.2 รูปแบบการทดสอบตัวแปร

ตัวแปรอิสระ		ตัวแปรตาม
ค่าจัดประเภท (จัดอันดับ)		ค่าต่อเนื่อง
การศึกษา (educ) - ประถมศึกษา - มัธยมศึกษา - อุดมศึกษา ประสบการณ์ก่อนทำงาน (prevexp_group) - ไม่มีประสบการณ์ - มีประสบการณ์น้อย - มีประสบการณ์ปานกลาง - มีประสบการณ์มาก	->	เงินเดือนเริ่มต้น (salbegin)

6.1.3 การกำหนดสมมติฐานและเกณฑ์การทดสอบ

ตารางที่ (#) แสดงค่าเฉลี่ยของเงินเดือนเริ่มต้น จำแนกตามระดับการศึกษาและตำแหน่งงาน

ตัวแปร	จำนวน	ค่าเฉลี่ยเงินเดือนเริ่มต้น	S.D.
ระดับการศึกษา			
- มัธยมศึกษาตอนต้น	53	13,064.15	2,408.147
- มัธยมศึกษาตอนปลาย	190	13,241.87	2,653.896
- อุดมศึกษา	231	21,027.14	9,421.873
รวม	474	17,061.09	7870.638
ประสบการณ์ก่อนทำงาน			
- ไม่มีประสบการณ์	24	11,331.25	1,308.402
- มีประสบการณ์น้อย (1-36 เดือน)	153	16,504.93	6,287.574
- มีประสบการณ์ปานกลาง (37-60 เดือน)	75	16,691.60	6,218.312
- มีประสบการณ์มาก (61 เดือนขึ้นไป)	222	18,092.57	9,347.561
รวม	474	17,016.09	7,870.638

จากตาราแสดงให้เห็นว่า พนักงานมีเงินเดือนเฉลี่ย 17,061.09 ดอลลาร์สหรัฐ โดยพนักงานที่มีการศึกษาระดับอุดมศึกษามีเงินเดือนเฉลี่ยมากที่สุด รองลงมาคือ มัธยมศึกษาตอนปลาย และมีเงินเดือนเฉลี่ยน้อยที่สุดคือ มัธยมศึกษาตอนต้น และเมื่อพิจารณาจากประสบการณ์ก่อนทำงานของพนักงานแสดงให้เห็นว่า พนักงานที่มีประสบการณ์มากมีเงินเดือนเฉลี่ยมากที่สุด รองลงมาคือ มีประสบการณ์ปานกลาง มีประสบการณ์น้อย และมีเงินเดือนเฉลี่ยน้อยที่สุดคือ ไม่มีประสบการณ์

จากข้อมูลที่กล่าวมาน่าจะมีความสัมพันธ์ระหว่างระดับการศึกษาและประสบการณ์การทำงานกับเงินเดือนเริ่มต้น ดังนั้นจึงทำการทดสอบความสัมพันธ์ด้วยสัมประสิทธิ์สหสัมพันธ์อันดับที่ของสเปียร์แมน (Spearman's Rank Correlation Coefficient Test) โดยมีสมมติฐานดังนี้

H_1: ระดับการศึกษากับเงินเดือนเริ่มต้นมีความสัมพันธ์กัน

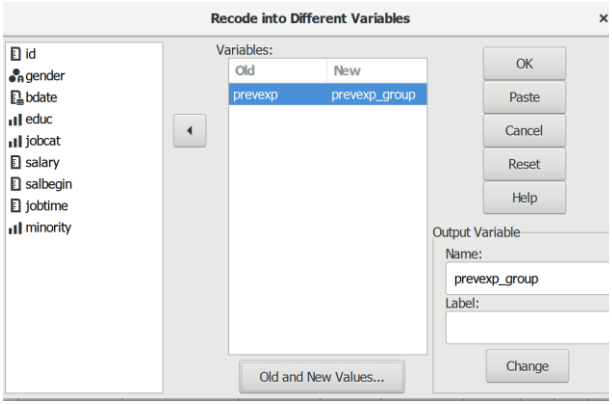
H_2: ประสบการณ์การทำงานกับเงินเดือนเริ่มต้นมีความสัมพันธ์กัน

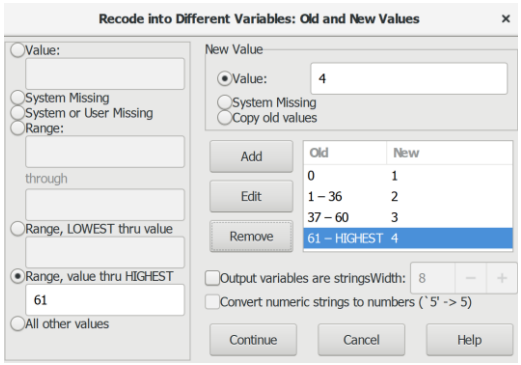
เกณฑ์ในการทดสอบสมมุติฐาน คือ หากการทดสอบความสัมพันธ์ด้วยวิธีของสเปียร์แมน (Spearman's Test) ได้ค่าความน่าจะเป็น (p-value) มากกว่าค่านัยสำคัญทางสถิติ (α = 0.05) จะปฏิเสธสมมติฐาน H_1 และ H_2 แต่หากได้ค่าความน่าจะเป็น (p-value) น้อยกว่าหรือเท่ากับค่านัยสำคัญทางสถิติ (α = 0.05) จะยอมรับสมมติฐาน H_1 และ H_2

6.2.4 การใช้โปรแกรมคอมพิวเตอร์วิเคราะห์ข้อมูล

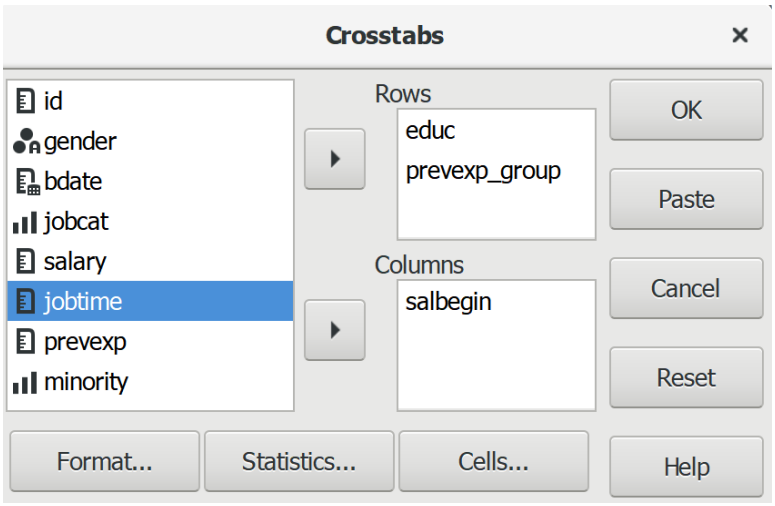
- เปิดแฟ้มข้อมูล **File -> Open** -> Employee data.sav
- สร้างตัวแปรกลุ่มประสบการณ์ก่อนทำงาน
  - เลือกเมนู **Transform -> Recode in Different Variables**
  - เลือกตัวแปร prevexp ใส่ในช่อง **Variables:**
  - ใต้คำว่า **Old** ในช่อง **Variables:** คลิกเมาส์ที่ชื่อตัวแปรเก่า
  - ในกลุ่มคำสั่ง **Output Variable** ช่อง **Name:** พิมพ์ prevexp_group -> กดปุ่ม **Change**
  - กดปุ่ม **Old and New Values...**

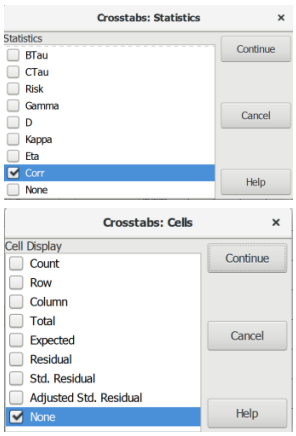
    เปลี่ยนค่า 0->1, 1-36 -> 2, 37-60 -> 3 และ 61-Highest -> 4
  - กดปุ่ม **Continue**
  - กดปุ่ม **OK**





- เลือกเมนู **Analyze -> Descriptive Statistics -> Crosstabs...**
  - **Rows:** เลือกตัวแปรแบบค่าจัดประเภท (จัดอันดับ) หรือตัวแปรแบบค่าต่อเนื่อง/ตัวแปรแบบค่าไม่ต่อเนื่อง (discrete variables) 1 ตัวแปร หรือมากกว่า 1 ตัวแปร (ที่ต้องการให้เป็นตัวแปรตาม) จากรายการตัวแปรด้านซ้าย
  - **Column:** เลือกตัวแปรแบบค่าจัดประเภท (จัดอันดับ) หรือตัวแปรแบบค่าต่อเนื่อง 1 ตัวแปร หรือมากกว่า 1 ตัวแปร (ที่ต้องการให้เป็นตัวแปรตาม) จากรายการตัวแปรด้านซ้าย
  - กดปุ่ม **Statistics...** ไม่เลือกสถิติ Chisq และเลือกสถิติ Corr แทน -> กดปุ่ม **Continue**
  - กดปุ่ม **Cells...** ไม่เลือก Count, Row, Column, Total และเลือก None -> กดปุ่ม **Continue**
  - กดปุ่ม **OK**





6.2.5 ผลการวิเคราะห์

Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Educational Level (years) × Beginning Salary	474	100.0%	0	.0%	474	100.0%
prevexp_group × Beginning Salary	474	100.0%	0	.0%	474	100.0%

Symmetric Measures

		Value	Asymp. Std. Error	Approx. T
Ordinal by Ordinal	Spearman Correlation	.69	.03	20.60
Interval by Interval	Pearson's R	.63	.03	17.77
N of Valid Cases		474		

Symmetric Measures

		Value	Asymp. Std. Error	Approx. T
Ordinal by Ordinal	Spearman Correlation	.20	.05	4.35
Interval by Interval	Pearson's R	.16	.04	3.50
N of Valid Cases		474		

6.2.6 การสร้างตารางแสดงผลการทดสอบ

ตารางที่ (#) แสดงผลการทดสอบความสัมพันธ์ระหว่างระดับการศึกษากับเงินเดือนเริ่มต้นและประสบการณ์ก่อนทำงาน

ตัวแปร	จำนวน	เงินเดือนเริ่มต้น (r_s)	Sig. (2-tailed)
ระดับการศึกษา	474	.69	.03
ประสบการณ์ก่อนทำงาน	474	.20	.05

6.2.7 การแปลผลและสรุป

จากตารางที่ (#) ผลการทดสอบแสดงให้เห็นว่า ความสัมพันธ์ระหว่างระดับการศึกษากับเงินเดือนเริ่มต้น และความสัมพันธ์ระหว่างประสบการณ์ก่อนทำงานกับเงินเดือนเริ่มต้น ได้ค่าความน่าจะเป็น (Sig.) = 0.00 มีค่าน้อยกว่าค่านัยสำคัญทางสถิติที่กำหนดไว้ (α = 0.05) จึงยอมรับสมมติฐาน H_1 และ H_2

สรุป ระดับการศึกษาของพนักงานมีความสัมพันธ์เชิงบวกระดับปานกลางกับเงินเดือนเริ่มต้น และประสบการณ์ก่อนทำงานมีความสัมพันธ์เชิงบวกระดับน้อยมากกับเงินเดือน

สรุป

การวิเคราะห์ข้อมูลเพื่อทดสอบสมมติฐานหากข้อมูลไม่เป็นไปตามข้อตกลงเบื้องต้น (not meet basic assumption) ของการทดสอบแบบอิงค่าพารามิเตอร์ (parametric tests) เช่น ข้อมูลมีการแจกแจงแบบไม่ปกติ ข้อมูลเป็นมาตราแบบจัดกลุ่ม (nominal scale) หรือจัดอันดับ (ordinal scale) หรือตัวอย่างที่สุ่มมีขนาดเล็กเกินไป (น้อยกว่า 30) ควรเลือกใช้การทดสอบแบบไม่อิงค่าพารามิเตอร์ (non-parametric statistics) เพราะเป็นวิธีการทดสอบที่มีข้อตกลงเพียงเล็กน้อยหรือไม่มีข้อตกลง (assumption-free test) เช่น ข้อมูลมีการวัดระดับใดก็ได้ ไม่จำเป็นต้องมีการแจกแจงเป็นโค้งปกติ ข้อมูลแต่ละกลุ่มไม่จำเป็นต้องมีการกระจายเท่ากัน และ/หรือไม่มีความเป็นเส้นตรงในการวิเคราะห์การถดถอยไม่ใช่เชิงเส้นตรง (nonlinear regression analysis)

แบบฝึกหัดท้ายบท

ให้ผู้เรียนเปิดแฟ้มข้อมูล 1991 US. General Social Survey.sav แล้วประมวลผลข้อมูล เพื่อทดสอบและแปลผล ดังนี้

1. ทดสอบความเป็นอิสระของข้อมูลระหว่างตัวแปรกลุ่มชนใหม่ (newrace) กับภูมิภาค (region) โดยใช้ตัวแปรกลุ่มชน (race) สร้างเป็นตัวแปรใหม่ชื่อ กลุ่มชนใหม่ (newrace) มี 2 กลุ่ม คือ ผิวขาว และไม่ใช่ผิวขาว เพื่อหาคำตอบว่า แต่ละภูมิภาคมีกลุ่มชนที่มีผิวขาวและไม่ใช่ผิวขาวแตกต่างกันหรือไม่

2. ทดสอบลำดับที่ค่าเฉลี่ยของตัวแปรระดับความสุข (happy) ด้วยตัวแปรเพศ (sex) โดยใช้วิธีการทดสอบของแมน-วิทนี ยู (Mann-Whitney U Test)

3. ทดสอบลำดับที่ค่าเฉลี่ยของตัวแปรระดับความสุข (happy) ด้วยตัวแปรภูมิภาค (region) โดยใช้วิธีการทดสอบของครัสคาลและวัลลิส (Kruskal Wallis Test)

4. ทดสอบความเหมือนกันของระดับการศึกษาสูงสุดของผู้ตอบแบบสำรวจ (edu) และบิดา (paeduc) ของผู้ตอบแบบสอบสำรวจ โดยใช้วิธีเครื่องหมาย (Sign Test)

5. ทดสอบความเหมือนกันของระดับการศึกษาสูงสุดของผู้ตอบแบบสำรวจ (edu) บิดา (paeduc) และมารดา (maeduc) ของผู้ตอบแบบสำรวจ โดยใช้วิธีฟรีดแมน (Friedman Test)

6. ทดสอบความสัมพันธ์ระหว่างจำนวนบุตร (childs) กับจำนวนพี่น้อง (sibs) กับระดับความสุข (happy)

7. ทดสอบความสัมพันธ์ระหว่างจำนวนบุตร (childs) กับจำนวนพี่น้อง (sibs) และระดับความสุข (happy)

บรรณานุกรม

ภาษาไทย

ราชบัณฑิตยสถาน. (2539). *พจนานุกรมฉบับราชบัณฑิตยสถาน พ.ศ. 2525*. กรุงเทพฯ: อักษรเจริญทัศน์.

เรวัต แสงสุริยงค์. (2541). *เอกสารประกอบการสอนวิชา 225101 สังคมวิทยาเบื้องต้น = Introduction to sociology.* ชลบุรี: ภาควิชาสังคมวิทยา คณะมนุษยศาสตร์และสังคมศาตร์ มหาวิทยาลัยบูรพา.

เรวัต แสงสุริยงค์. (2563*). ความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัยเชิงปริมาณด้านสังคมวิทยา : รายงานวิจัยฉบับสมบูรณ์ = Risk of error in the quantitative sociology research*. ชลบุรี : ภาควิชาสังคมวิทยา คณะมนุษยศาสตร์และสังคมศาสตร์.

อนุสรณ์ ลิ่มมณี. (2542). *การอธิบายกับการวิเคราะห์ทางการเมือง: ข้อพิจารณาเบื้องต้นในเชิงปรัชญาสังคมศาสตร์*. โครงการผลิตตำราและเอกสารการสอนคณะรัฐศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

ภาษาอังกฤษ

Aneshensel, C. S. (2013). *Theory-based data analysis for the social sciences*. 2nd ed. SAGE.

Babbie, E. (2008). *The basics of social research*. 4th ed. Thomson Wadsworth.

Babbie, E. (2007). *The practice of social research*. 11th ed. Thomson Wadsworth.

Babbie, E. (2013). *The practice of social research*. 13th ed. Wadsworth Cengage Learning.

Babones, S. J. (2014). *Methods for quantitative macro-comparative research*. SAGE.

Black, T. R. (1999). *Doing quantitative research in the social sciences: an integrated approach to research design, measurement and statistics*. SAGE.

Blaikie, N. (2010). *Designing social research*. Polity Press.

Bouma, G. D. & Ling, R. (2006). *The research process*. 5th ed. Oxford University.

Bryman, A. (2012). *Social research methods.* 4th ed. Oxford University Press.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. Academic Press.

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods Approaches*. SAGE.

Creswell, J. W. (2014). *Research design: Qualitative, Quantitative, and Mixed Methods Approaches (4 ed.)*. SAG.

Drost, E. A. (2011). Validity and Reliability in Social Science Research. *Education Research and Perspectives*, 38 (1), 105-123. Retrieved from: http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ942587&site=eds-live

Evans, A.N. (2014). *Using Basic Statistics in the Behavioral and Social Sciences.* SAGE.

Field, A. (2009). *Discovering statistics using SPSS (sex and drugs and rock 'n') roll*. SAGE.

Flynn, B. B., Sakakibara, S., Schroeder, R. G., Bates K. A., & Fynn, E.J. (1990). Empirical research methods in operations management. *Journal of Operations Management.* 9(2), 250-284. DOI: 10.1016/0272-6963(90)90098-X

Garner, M., Wagner, C., & Kawulich, B. (2009). *Teaching Research Methods in the Social Sciences*. Ashgate.

Garson, D. (2008). *Testing of Assumptions*. Retrieved January, 18, 2010 from: http://www2.chass.ncsu.edu/garson/pa765/assumpt.htm#normal

Glen, S. (2014). "Cronbach's Alpha: Simple Definition, Use and Interpretation" From StatisticsHowTo.com: Elementary Statistics for the rest of us! Retrieved January, 16, 2021 from: https://www.statisticshowto.com/cronbachs-alpha-spss/

GNU. (2019). *GNU PSPP — Frequently Asked Questions.* Retrieved January, 13, 2021 from: https://www.gnu.org/software/pspp/faq.html

GNU. (2016). *PSPP Users' Guide*. Retrieved January, 13, 2021 from: https://www.gnu.org/software/pspp/manual/pspp.pdf

GNU. (2020). *PSPP*. Retrieved January, 13, 2021 from: https://www.gnu.org/software/pspp/manual/pspp.html#Introduction

Greenfield, T. and Greener, S. (2016). *Research methods for postgraduates*. 3rd ed. John Wiley & Sons.

Hair, J. F. & Others. (2006). *Multivariate data analysis.* 6th ed. Pearson/Prentice Hall.

Halter, C. P. (2017). *The PSPP Guide: An Introduction to Statistical Analysis*. 2nd ed. CreativeMinds Press Group.

Hammond, M. & Wellington, J. (2013). *Research methods: the key concepts*. Routledge.

Hardy, M. & Bryman, A. (2004). *Handbook of data analysis*. SAGE.

Healey, J. F. (1999). *Statistics: A Tool for Social Research*. Wadsworth.

Hoaglin, D. C. (2003). John W. Tukey and Data Analysis. *Statistical Science*, 18(3), 311-318. Retrieved form: https://projecteuclid.org/euclid.ss/1076102418

Jary, D. & Jary, J. (1995). *Collins Dictionary of sociology*. 2nd ed. HaperCollins..

Kimberlin, C. L. & Winterstein, A. G. (2008). *Validity and reliability of measurement instruments used in research. Am J Health Syst Pharm*, 65(23), 2276-2284. DOI: 10.2146/ajhp070364.

Kumar, R. (2014). *Research Methodology: a step-by-step guide for beginners.* 4th ed. SAGE.

Kuper, A. & Jessica K. (2003). *The social science encyclopedia*. 2nd ed. Routledge.

Lash, T. L., Fox, M. P., & Fink, A. K. (2009). *Applying Quantitative Bias Analysis to Epidemiologic Data*. New York: Springer.

McHugh M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica*, 22(3), 276-282. Retrieved from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3900052/

Microsoft (n.d.) *About Microsoft R Open: The Enhanced R Distribution*. Retrieved November, 22, 2020 from: https://mran.microsoft.com/open

Neuman, W. L. (2006). *Social research methods: qualitative and quantitative approaches*. 6th ed. Pearson A and B.

Neuman, W. L. (2014). *Social research methods: qualitative and quantitative approaches*. 7th ed. Pearson.

Orloff, J. & Bloom, J. (2014). *Introduction to Probability and Statistics.* Retrieved October, 13, 2020 from: https://ocw.mit.edu/courses/mathematics/18-05-introduction-to-probability-and-statistics-spring-2014/readings/MIT18_05S14_Reading20.pdf

Osborne, J. W. (2008). *Best practices in quantitative methods*. Thousand Oaks, Calif.: SAGE.

Popenoe, D. (1993). *Sociology*. 9th ed. Prentice Hall.

Punch, K. F. (1998). *Introduction to social research: quantitative and qualitative approaches*. SAGE.

Rovinelli, R. J. & Hambleton, R. K. (1967). *On the Use of Content Specialists in the Assessment of Criterion-Referenced Test Item Validity*. Distributed by ERIC Clearinghouse. Retrieved January, 16, 2021 from: https://eric.ed.gov/?id=ED121845

Ritter, N. L. (2010). *Understanding a Widely Misunderstood Statistic: Cronbach's "Alpha"*. Retrieved January, 16, 2021 from: https://eric.ed.gov/?id=ED526237

Sawyer, R.K. (2005). *Social emergence: societies as complex systems*. Cambridge University Press.

Scanlon, M. (n.d). “Issues in research” Cited in Wilkinson, David (Editor). (2000). *The researcher’s toolkit: the complete guide to practitioner research*. London: Routledge Falmer.

SPSS. (2005). *SPSS Data validation™ 14.0*. SPSS Inc.

Stevens, S.S. (1946). On the theory of scales of measurement. *Science*, 103, 677-680

Struble, C. A. (2002). *Outlier Analysis*. Retrieved August, 24, 2008, from http://www.mscs.mu.edu/~cstruble/class/mscs282/fall2002/notes/outliers.ppt

Sullivan, L. (2015). *Nonparametric Tests.* Retrieved February, 19, 2021, from https://sphweb.bumc.bu.edu/otlt/MPHModules/BS/BS704_Nonparametric/BS704_Nonparametric_print.html.

Taherdoost, H. (2016). Validity and Reliability of the Research Instrument; How to Test the Validation of a Questionnaire/Survey in a Research. *International Journal of Academic Research in Management (IJARM)*, 5(3), 28-36. DOI: 10.2139/ssrn.3205040

Traub, Ross E.. (1994). *Reliability for the social sciences: theory and applications*. SAGE.

Validation of a Questionnaire/Survey in a Research. *International Journal of Academic Research in Management (IJARM)*, 5(3), 28-36. DOI: 10.2139/ssrn.3205040

Verma, J. P., G. & Abdel-Salam, A. G. (2019). *Testing Statistical Assumptions in Research*. Wiley.

Yang, K. (2010). *Making sense of statistical methods in social research.* SAGE.

Wikipedia. (2020). *List of statistical packages*. Retrieved November, 22, 2020 from: http://en.wikipedia.org/wiki/List_of_statistical_packages

Wikipedia. (2020). *PSPP*. Retrieved November, 22, 2020 from: https://en.wikipedia.org/wiki/PSPP

Williams, M. (2016). *Key concepts in the philosophy of social research*. SAGE.

ดัชนี

ก

กลุ่มตัวอย่าง

ขนาดเล็ก · 276

ความเป็นอิสระจากกัน · 276

แบบสุ่ม · 49

การจัดการข้อมูล

การกรอกข้อมูลเข้า · 112

การทำความสะอาดข้อมูล · 164

การทำให้ข้อมูลมีความสมบูรณ์ · 192

การปรับข้อมูล · 126

การเปลี่ยนรูปแบบข้อมูล · 126

การตรวจสอบการแจกแจงปกติ

กราฟกล่องและหนวด · 182

กราฟการกระจาย · 182

กราฟแจกแจงความถี่ · 182

กราฟลำต้นและใบ · 182

การทดสอบแบบคอลมอโกรอฟ-สเมอร์นอฟ · 183

การทดสอบแบบแชพิโล-วิลค์ · 183

การตรวจสอบข้อมูล

ความโด่ง · 182

ความเบ้ · 182

ค่าต่างสุด · 178

ค่าผิดปกติ · 178

ค่าสูญหาย · 173

การทดสอบความเป็นอิสระของกลุ่มตัวอย่าง

แกรมมา · 295

ความเที่ยง/แม่นยำของฟิชเชอร์ · 294

คอครานและแมนเทล-แฮนส์เซล · 295

แคปปา · 295

ไคสแควร์แบบแก้ไขของเยตส์ · 294

ไคสแควร์แบบความเป็นไปได้ · 294

ไคสแควร์แบบเพียร์สัน · 294

เดลต้าของโสเมอร์ส · 295

ฟีและแครมเมอร์ วี · 294

แมคนีมาร์ · 295

โรท์ · 294

แลมดา · 295

สัมประสิทธิ์ของสถานการณ์ที่สอดคล้องกัน · 294

สัมประสิทธิ์ความไม่แน่นอน · 295

สัมประสิทธิ์สหสัมพันธ์ของเคนดอล์ล · 295

สัมประสิทธิ์สหสัมพันธ์ของสเปียร์แมน · 294

อัตราส่วนของโอกาสที่จะเกิด · 295

อีตา · 295

การทดสอบค่ากลางของกลุ่มตัวอย่างที่เป็นอิสระจากกัน 2 กลุ่ม
การทดสอบของคอลมอโกรอฟ-สเมอร์นอฟ · 303
การทดสอบของแมน-วิทนี ยู · 303
การทดสอบของวอร์ค–วูล์ฟอวิทซ์ · 304
การทดสอบค่าต่างสุดของโมเซส · 303
การทดสอบค่ากลางของกลุ่มตัวอย่างที่เป็นอิสระจากกันมากกว่า 2 กลุ่ม
การทดสอบของครัสคาลและวัลลิส · 307
การทดสอบของจองฮีรี · 307
การทดสอบค่ามัธยฐาน · 307
การทดสอบค่ากลางของกลุ่มตัวอย่างที่มีความสัมพันธ์กันมากกว่า 2 กลุ่ม
การทดสอบแบบคอครัน · 315
การทดสอบแบบเคนดอลล์ · 315
การทดสอบแบบฟรีดแมน · 315
การทดสอบแบบไม่อิงค่าพารามิเตอร์
การทดสอบการแจกแจงข้อมูล · 277
การทดสอบความเป็นอิสระของกลุ่มตัวอย่าง · 294
การทดสอบความเป็นอิสระของกลุ่มตัวอย่างแบบไคสแควร์ · 296
การทดสอบความเป็นเอกพันธ์ของส่วนเพิ่ม · 311
การทดสอบความสัมพันธ์รายคู่แบบสเปียร์แมน · 320
การทดสอบค่ากลางของกลุ่มตัวอย่างที่เป็นอิสระจากกัน 2 กลุ่ม · 303
การทดสอบค่ากลางของกลุ่มตัวอย่างที่เป็นอิสระจากกันมากกว่า 2 กลุ่ม · 307
การทดสอบค่ากลางของกลุ่มตัวอย่างที่มีความสัมพันธ์กัน 2 กลุ่ม · 311
การทดสอบค่ากลางของกลุ่มตัวอย่างที่มีความสัมพันธ์กันมากกว่า 2 กลุ่ม · 315
การทดสอบแบบเครื่องหมาย · 311
การทดสอบแบบแมคนีมาร์ · 311
การทดสอบแบบวิลคอกซัน · 311
การทดสอบสัดส่วนของกลุ่มตัวอย่าง · 284, 288
การทดสอบแบบอิงค่าพารามิเตอร์
การทดสอบความสัมพันธ์รายคู่แบบเพียร์สัน · 264
การทดสอบค่าเฉลี่ยของกลุ่มตัวอย่าง 1 กลุ่ม · 238
การทดสอบค่าเฉลี่ยของกลุ่มตัวอย่างที่เป็นอิสระจากกัน 2 กลุ่ม · 247
การทดสอบค่าเฉลี่ยของกลุ่มตัวอย่างที่เป็นอิสระจากกันมากกว่า 2 กลุ่ม · 252
การทดสอบค่าเฉลี่ยของกลุ่มตัวอย่างที่ไม่เป็นอิสระจากกัน 2 กลุ่ม · 242
การทดสอบสัดส่วนของกลุ่มตัวอย่าง
แบบไคสแควร์ · 284
แบบทวินาม · 288

การเลือกตัวอย่าง

ความคลาดเคลื่อน · 58

แบบสุ่ม · 276

การวิเคราะห์ข้อมูล

1 ตัวแปร · 196

2 ตัวแปร · 212

การวิจัย

การวิจัยเชิงคุณภาพ · 11, 16

การวิจัยเชิงปริมาณ · 10, 16

การวิจัยเชิงผสมผสาน · 18

การวิจัยประยุกต์ · 10, 24

การวิจัยพื้นฐาน · 9, 24

ข

ข้อตกลงเบื้องต้นของการทดสอบแบบอิง

ค่าพารามิเตอร์

การแจกแจงแบบปกติ · 181

ความเท่ากันของความแปรปรวน · 186

ความเป็นเส้นตรง · 188

ค

ความคลาดเคลื่อน

เชิงระบบ · 58

เชิงสถิติ · 58

แบบที่ 1 · 59

แบบที่ 2 · 59

ความเที่ยงตรง

ดัชนีความสอดคล้อง · 149

ดัชนีแคปปาของโคเฮน · 154

ความน่าเชื่อถือ

การทดสอบ-ทดสอบซ้ำ · 158

การทดสอบแบบแบ่งครึ่ง · 158

การทดสอบหลากรูปแบบ · 158

ค่าสัมประสิทธิ์แอลฟาของครอนบัก · 158

ซ

ซอฟต์แวร์

JAMOVI · 82

JASP · 82

PSPP · 82

R · 82

SAS · 82

SPSS · 82

โปรแกรมมีลิขสิทธิ์ · 82

โปรแกรมรหัสเปิด · 82

ต

ตัวแปร (ความสัมพันธ์)

ความเข้ม/ขนาด · 264

ความสัมพันธ์เชิงผกผัน · 41

ความสัมพันธ์เชิงสาเหตุ · 40

ความสัมพันธ์เชิงหลายสาเหตุ · 41

เชิงเส้นโค้ง · 188
เชิงเส้นตรง · 188
ทิศทาง · 264

ตัวแปร (ประเภท)

ตัวแปรคุม · 33
ตัวแปรตาม · 32
ตัวแปรอิสระ · 32

ตัวแปร (ระดับการวัด)

ข้อมูลแบบเมตริก · 48
ข้อมูลเป็นแบบไม่ใช่เมตริก · 48
ตัวแปรแบบค่าจัดประเภท · 38
ตัวแปรแบบค่าต่อเนื่อง · 38
ตัวแปรแบบค่าทวิลักษณะ · 38
ตัวแปรแบบค่าไม่ต่อเนื่อง · 38
ตัวแปรแบบค่าวิมุต · 38
ตัวแปรแบบค่าสองประเภท · 38
ตัวแปรแบบค่าหลายประเภท · 39
ตัวแปรแบบตัวเลข · 39
ตัวแปรแบบตัวอักษร · 39
มาตราส่วนแบบจัดกลุ่ม · 37
มาตราส่วนแบบจัดอันดับ · 37
มาตราส่วนแบบช่วง · 37
มาตราส่วนแบบอัตราส่วน · 37

ท

ทฤษฎี · 24

ระดับกลาง · 26
ระดับจุลภาค · 26
ระดับมหภาค · 26

บ

แบบสอบถาม · 64, 148

คำถามแบบคำตอบเดียว · 67
คำถามแบบจัดอันดับ · 69
คำถามแบบถูกผิด · 71
คำถามแบบประเมินค่า · 68
คำถามแบบปลายเปิด · 68
คำถามแบบเปรียบเทียบ · 70
คำถามแบบเลือกตอบ · 67
คำถามแบบหลายคำตอบ · 67, 69

ร

ระบบติดต่อกับผู้ใช้งาน

แบบกราฟิก · 83
แบบข้อความ · 82

ศ

ศาสตร์ · 4

วิทยาศาสตร์ธรรมชาติ · 5
สังคมศาสตร์ · 5

ส

สถิติ · 44

เชิงการทำนาย · 47

เชิงความสัมพันธ์ · 47

เชิงพรรณนา · 45

เชิงอ้างอิง · 46

สถิติเชิงการทำนาย

การถดถอยเชิงโลจิสติก · 47

การถดถอยเชิงเส้นตรง · 47

สถิติเชิงความสัมพันธ์

สหสัมพันธ์แบบเพียร์สัน · 47

สหสัมพันธ์แบบสเปียร์แมน · 47

สถิติเชิงพรรณนา

การวัดการกระจายหรือความผันแปร · 45

การวัดการแจกแจงความถี่ · 45

การวัดตำแหน่ง · 45

การวัดแนวโน้มสู่ส่วนกลาง · 45

การวัดรูปร่าง · 45

สถิติเชิงอ้างอิง

การทดสอบแบบไม่อิงค่าพารามิเตอร์ · 46

การทดสอบแบบอิงค่าพารามิเตอร์ · 46

สำนักความคิดแบบเบย์ · 51

สำนักความคิดแบบฟรีเควนทิสต์ · 52

สมมติฐาน · 28

การทดสอบแบบ 2 หาง · 53

การทดสอบแบบหางเดียว · 53

เขตวิกฤต · 55

ค่านัยสำคัญทางสถิติ · 55, 57

ค่าวิกฤต · 55

สมมติฐานเชิงวิจัย · 29

สมมติฐานเชิงสถิติ · 29

สมมติฐานรอง · 30

สมมติฐานหลัก · 30

สังคมวิทยา · 6

สังคมศาสตร์ · 5

ห

หลักเหตุผล

หลักกลนัย · 8

หลักจารนัย · 8

หลักนิรนัย · 8

หลักอุปนัย · 8

อ

อำนาจของสถิติ

ขนาดของอิทธิพล · 59

ขนาดตัวอย่าง · 59

ระดับนัยสำคัญทางสถิติ · 59