



สถิติพื้นฐาน (Basic statistic)



วราภรณ์ เนตรพราว
นักวิชาการสถิติปฏิบัติการ

ประเภทของสถิติ

1. สถิติพรรณนา (Descriptive Statistics) เป็นสถิติที่ใช้อธิบายคุณลักษณะของสิ่งที่ต้องการศึกษากลุ่มใดกลุ่มหนึ่ง ไม่สามารถอ้างอิงไปยังกลุ่มอื่นๆ ได้ สถิติที่อยู่ในประเภทนี้ เช่น ค่าเฉลี่ย(Mean) ค่ามัธยฐาน(Median) ส่วนเบี่ยงเบนมาตรฐาน (SD) ค่าสูงสุด(Max) ค่าต่ำสุด(Min) พิสัยระหว่างควอไทล์ (IQR) ฯลฯ
2. สถิติอ้างอิง (Inferential statistics) เป็นสถิติที่ใช้อธิบายคุณลักษณะของสิ่งที่ต้องการศึกษากลุ่มใดกลุ่มหนึ่งหรือหลายกลุ่ม แล้วสามารถอ้างอิงไปยังกลุ่มประชากรได้

สถิติพรรณนา (Descriptive Statistics)

ตัวแปรเชิงคุณภาพ



จำนวน (n) , ร้อยละ (%)

ตัวแปรเชิงปริมาณ



ค่าเฉลี่ย (Mean) , ส่วนเบี่ยงเบนมาตรฐาน (SD)
ค่ามัธยฐาน (Median) , พิสัยระหว่างควอไทล์ (IQR)

หน้าต่างโปรแกรม SPSS

จำนวนจุดทศนิยม

ชื่อตัวแปร

Coding ตัวแปร

ชนิดตัวแปรที่ใช้บ่อยๆ

- Numeric ลักษณะข้อมูลเป็นตัวเลข
- String ลักษณะข้อมูลเป็นข้อความ

Data View Variable View

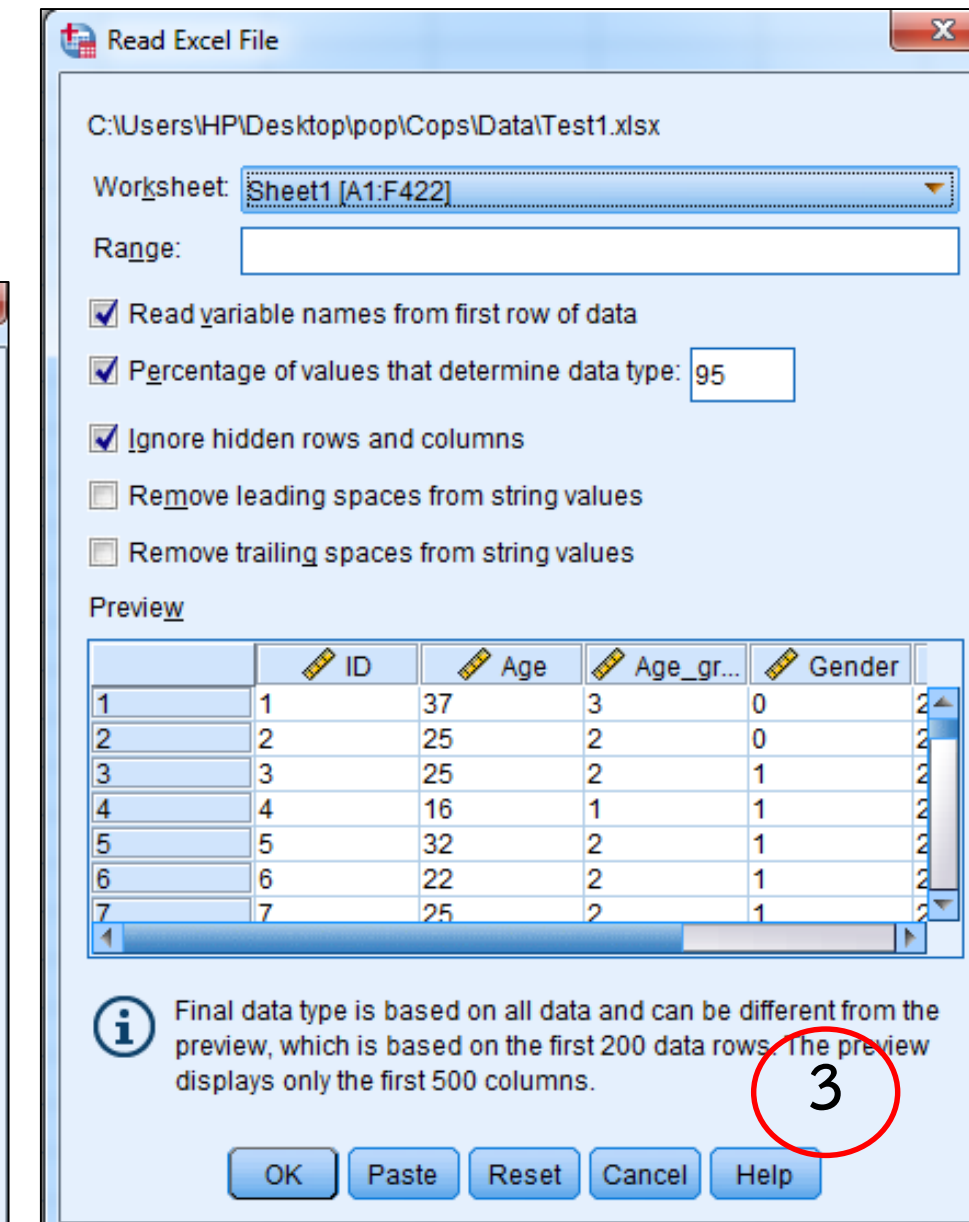
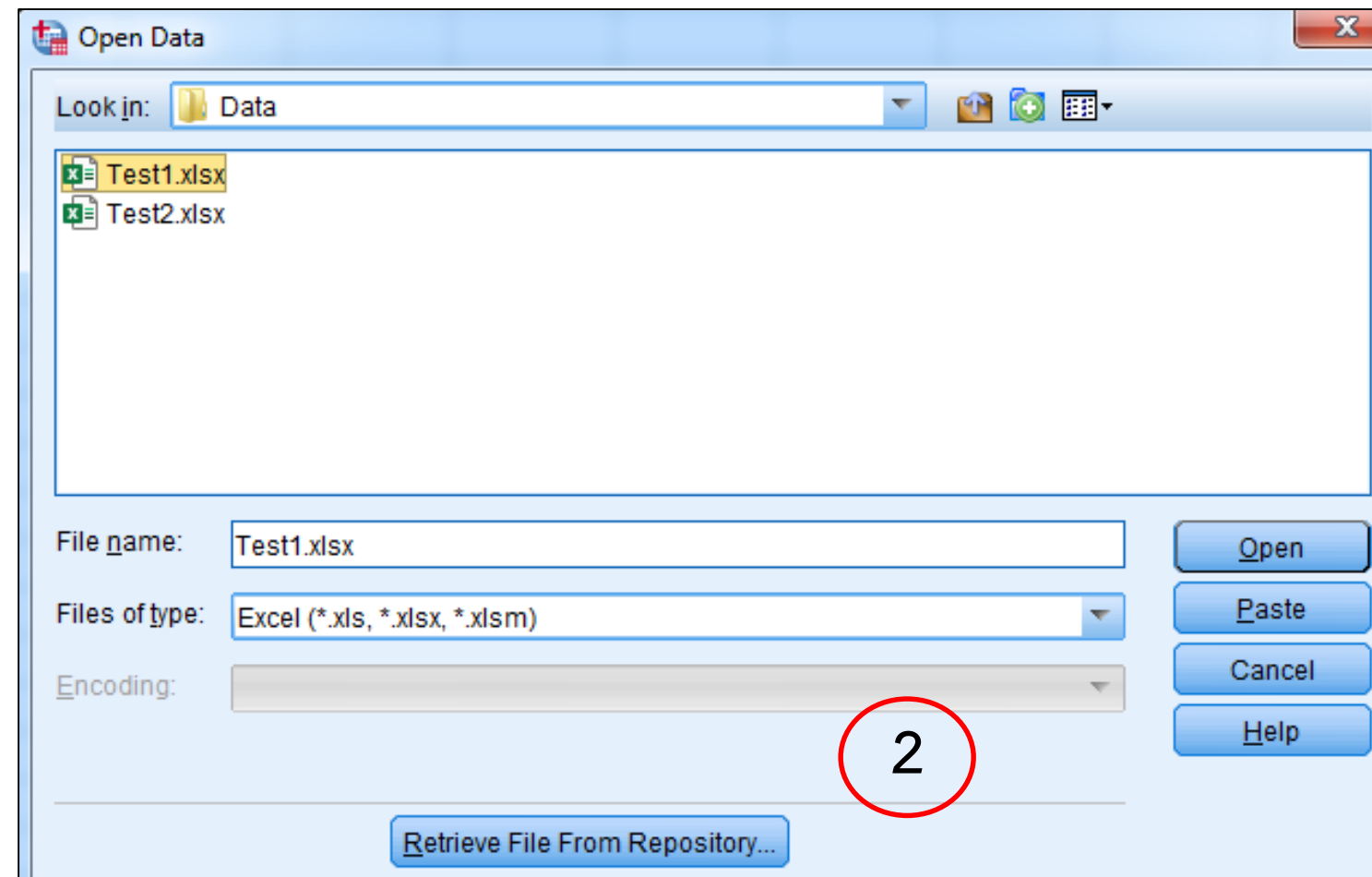
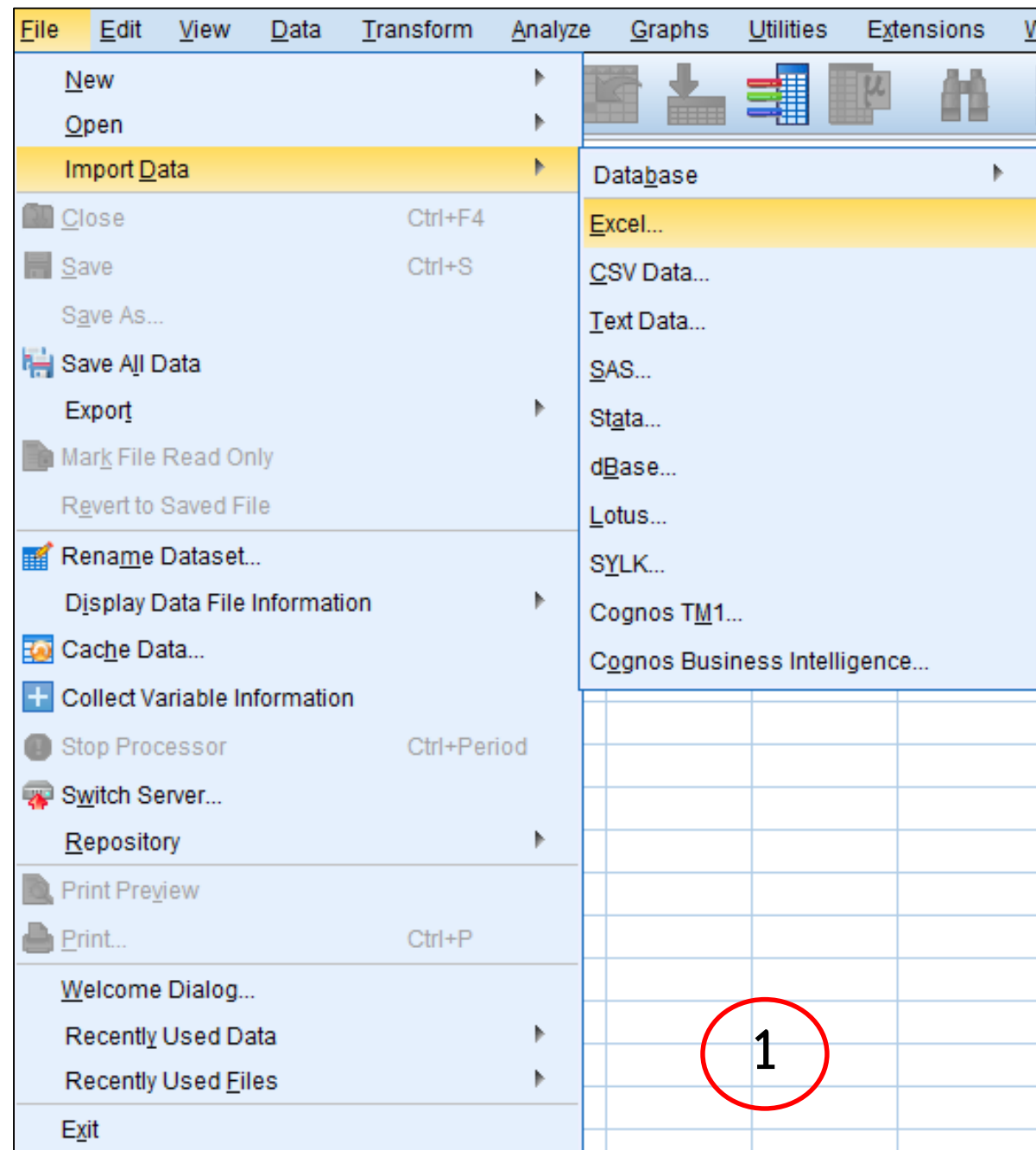
หน้าต่าง Variable View

	id	Age	Age_group	Gender	Method	Disease	var	var
1	1	37	3	0	2	1		
2	2	25	2	0	2	0		
3	3	25	2	1	2	1		
4	4	16	1	1	2	1		
5	5	32	2	1	2	1		
6	6	22	2	1	2	1		
7	7	25	2	1	2	1		
8	8	39	3	1	2	1		
9	9	34	2	1	2	1		
10	10	41	3	1	2	1		
11	11	26	2	1	2	1		
12	12	37	3	1	2	1		
13	13	20	2	1	2	1		
14	14	26	2	1	2	1		
15	15	26	2	1	2	0		
16	16	36	3	0	2	0		
17	17	26	2	0	2	0		
18	18	28	2	0	2	1		
19	19	24	2	1	2	1		
20	20	23	2	0	2	0		
21	21	24	2	1	2	0		
22	22	31	2	1	2	0		
23	23	30	2	1	2	1		
24	24	20	2	1	2	0		
25	25	23	2	0	2	0		
26	26	21	2	1	2	1		
27	27	21	2	1	2	1		
28	28	26	2	1	2	1		
29	29	38	3	1	2	1		

Data View Variable View

หน้าต่าง Data View

วิธีการนำเข้าข้อมูลจาก Excel เข้าโปรแกรม SPSS



ใช้คำสั่ง File > Import Data > Excel

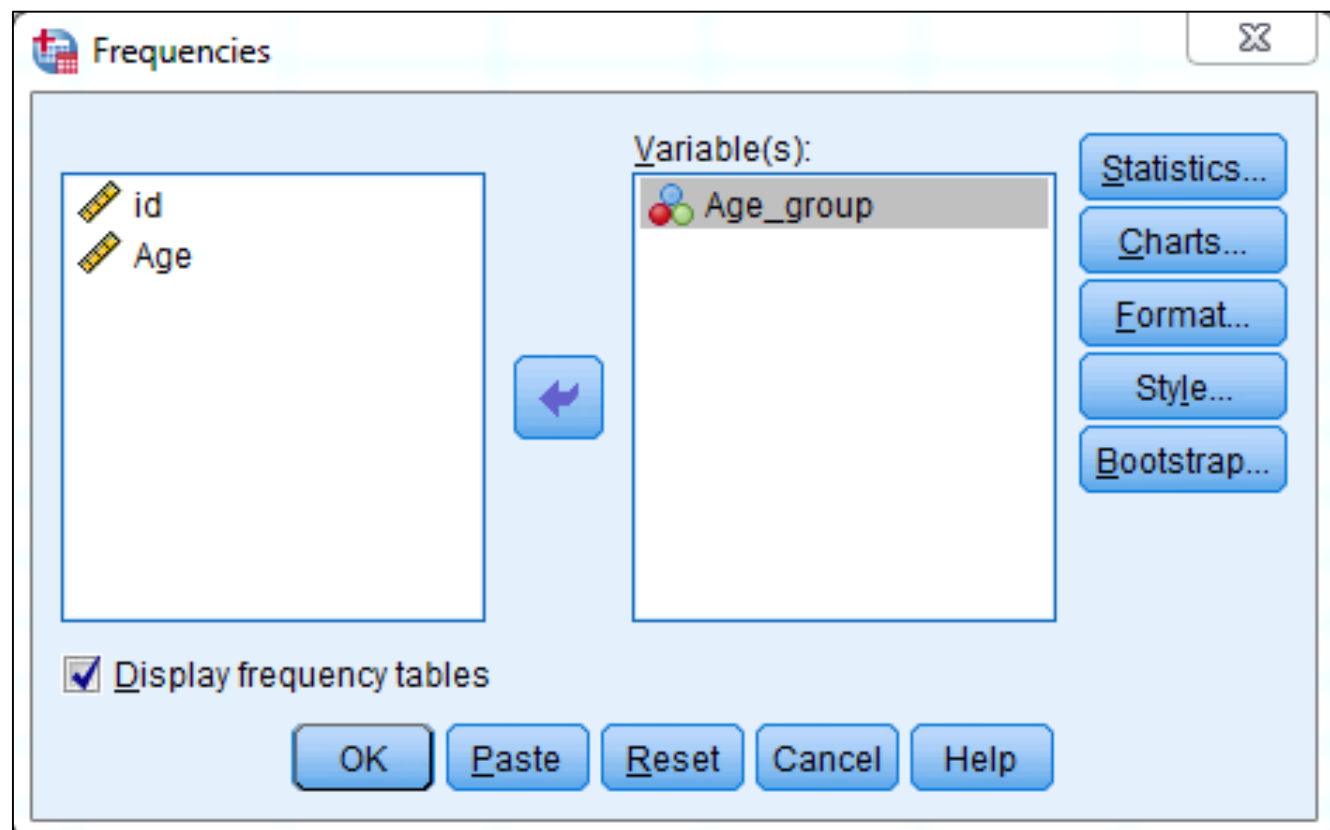
เลือกที่อยู่ไฟล์ Excel ที่ต้องการนำเข้า

เลือก ตามตัวอย่าง กด OK

ตัวอย่างการหาค่าจำนวน (n) ,ร้อยละ (%) จากโปรแกรม

คำสั่งในโปรแกรม

1. คลิกที่ Analyze > Descriptive Statistics > Frequencies
2. ปรากฏหน้าต่าง Frequencies > ช่อง Variable(s) : เลือกตัวแปรที่ต้องการตรวจสอบ (สามารถเลือกได้หลายตัวแปร)
ในตัวอย่างนี้คือตัวแปร Age_group



หน้าต่างคำสั่ง

กรณี ไม่มีค่า Missing

Age_group				
	Frequency	Percent	Valid Percent	Cumulative Percent
Valid <20	23	5.5	5.5	5.5
Valid 20-34	304	72.2	72.2	77.7
Valid >35	94	22.3	22.3	100.0
Total	421	100.0	100.0	

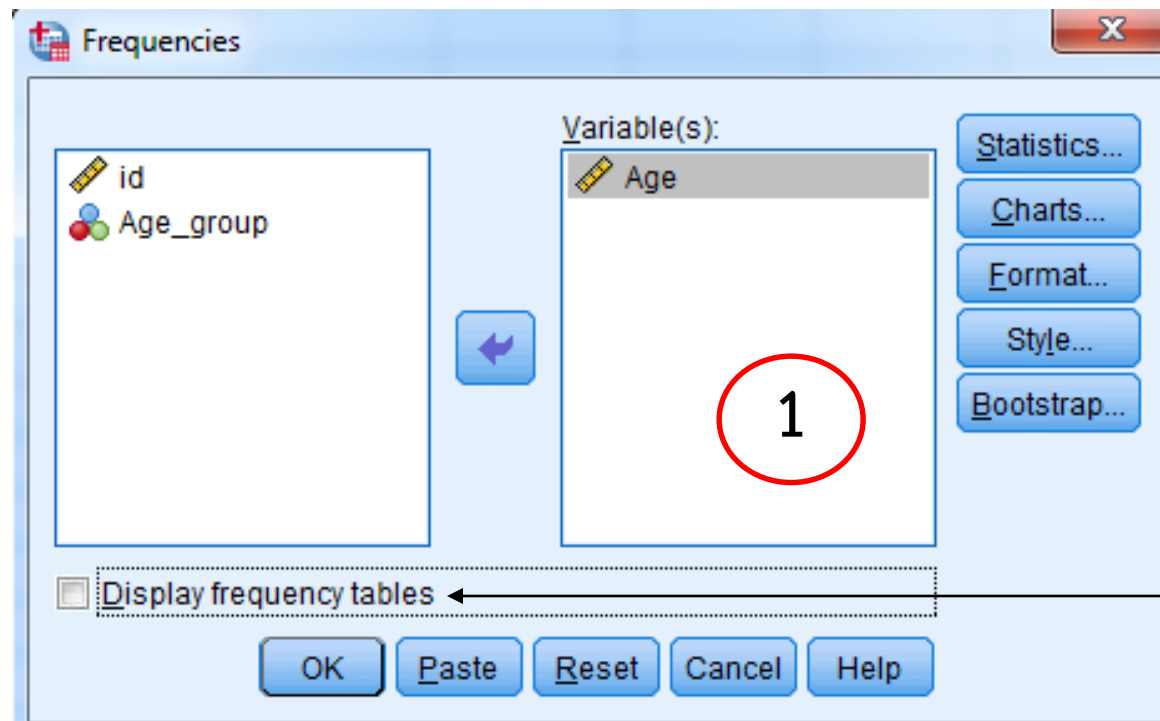
มีค่า Missing

Age_group				
	Frequency	Percent	Valid Percent	Cumulative Percent
Valid <20	22	5.2	5.3	5.3
Valid 20-34	301	71.5	72.4	77.6
Valid >35	93	22.1	22.4	100.0
Total	416	98.8	100.0	
Missing System	5	1.2		
Total	421	100.0		

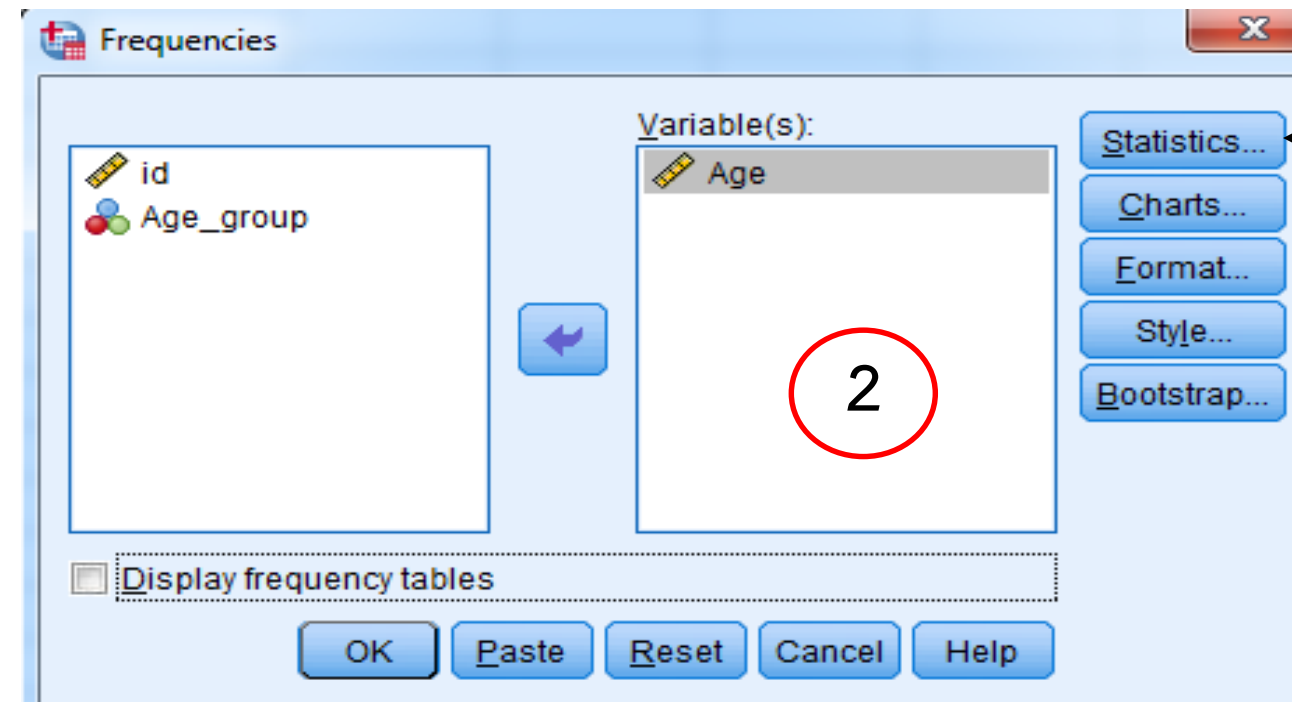
หน้าต่าง Output

ตัวอย่างการหาค่า Mean ,SD ,Median ,IQR จากโปรแกรม

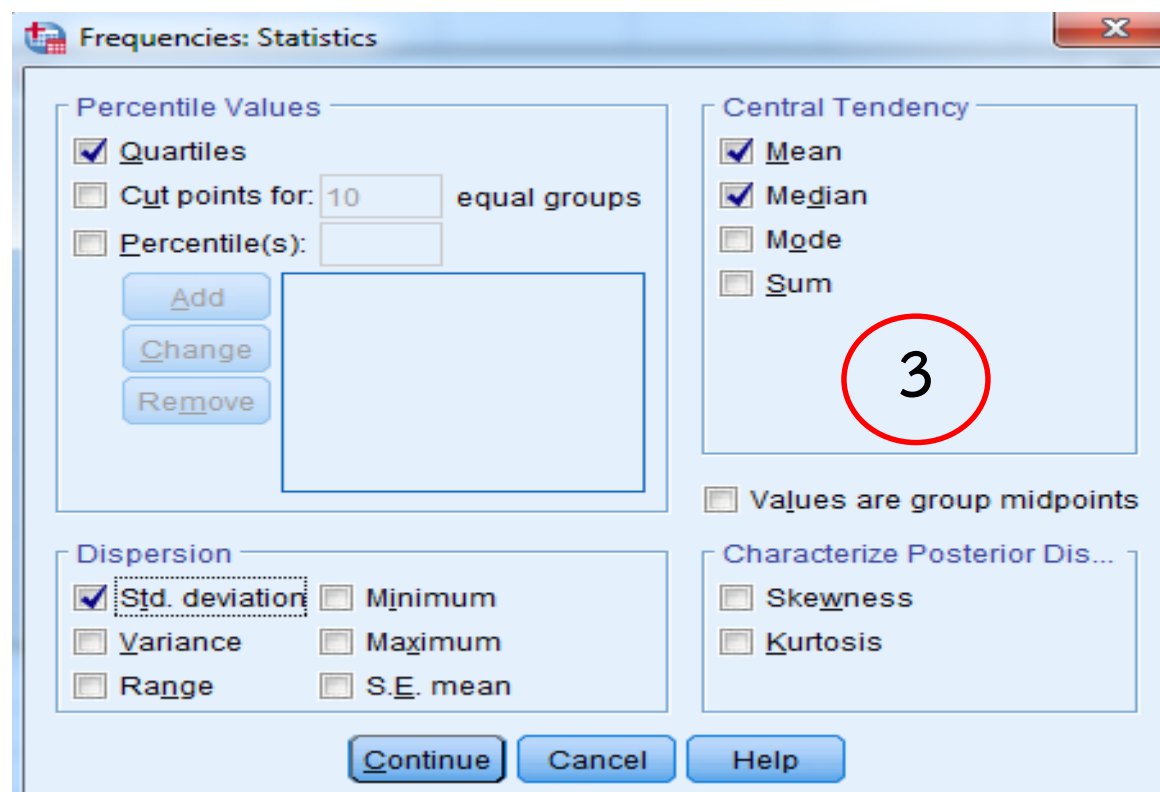
ใช้คำสั่ง Analyze > Descriptive Statistics > Frequencies



เอาเครื่องหมาย
 ออก



คลิกเมนู Statistics



→ **Frequencies**

Statistics		
Age	N	421
	Valid	421
	Missing	0
Mean		29.31
Median		29.00
Std. Deviation		6.330
Percentiles	25	25.00
	50	29.00
	75	34.00

Hint: IQR คำนวณได้จาก
Percentiles 75 ลบ Percentiles 25
จากตัวอย่าง IQR = 34-25 = 9

สถิติอ้างอิง (Inferential statistics)

มีข้อตกลงเบื้องต้น

Parametric

t-test (Paired)

t-test (Independent)

ANOVA



ไม่มีข้อตกลงเบื้องต้น

Non - Parametric

Wilcoxon Signed Ranks Test

Mann-Whitney U test

Kruskal-Wallis H test

ข้อตกลงเบื้องต้นข้อที่ 1 ข้อมูลมีการแจกแจงเป็นแบบโค้งปกติ

H_0 : ข้อมูลมีการแจกแจงเป็นแบบโค้งปกติ

H_1 : ข้อมูลไม่ได้มีการแจกแจงเป็นแบบโค้งปกติ

ปฏิเสธ H_0 ที่ p-value < 0.05

ข้อตกลงเบื้องต้นข้อที่ 2 ความแปรปรวนในแต่ละกลุ่มไม่แตกต่างกัน

H_0 : ความแปรปรวนในแต่ละกลุ่มไม่แตกต่างกัน

H_1 : ความแปรปรวนในแต่ละกลุ่มแตกต่างกัน

ปฏิเสธ H_0 ที่ p-value < 0.05

ข้อตกลงเบื้องต้น

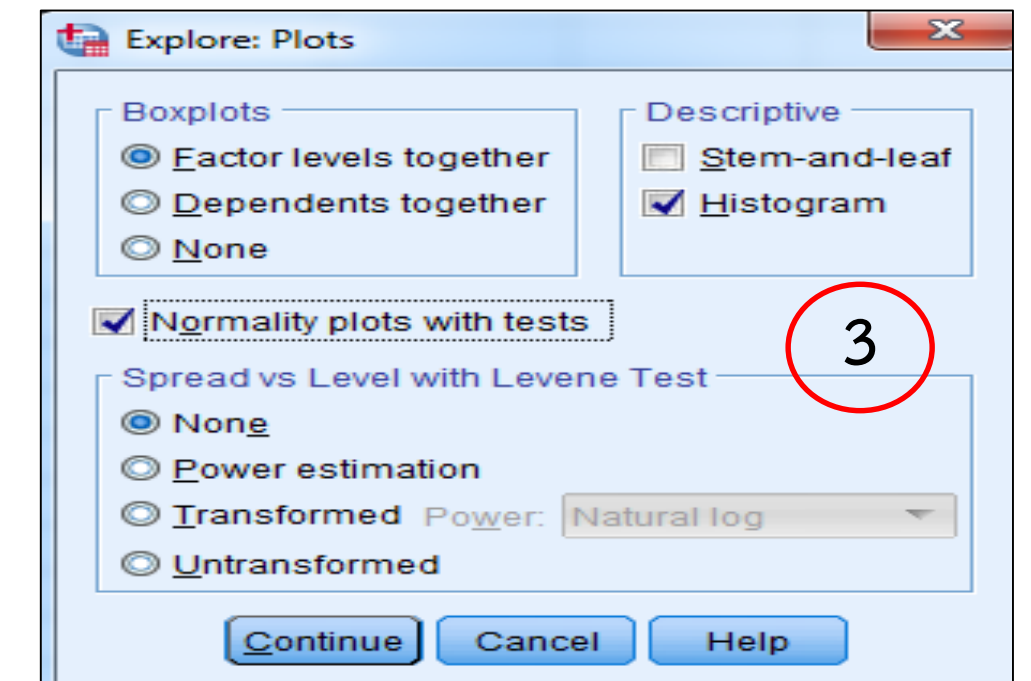
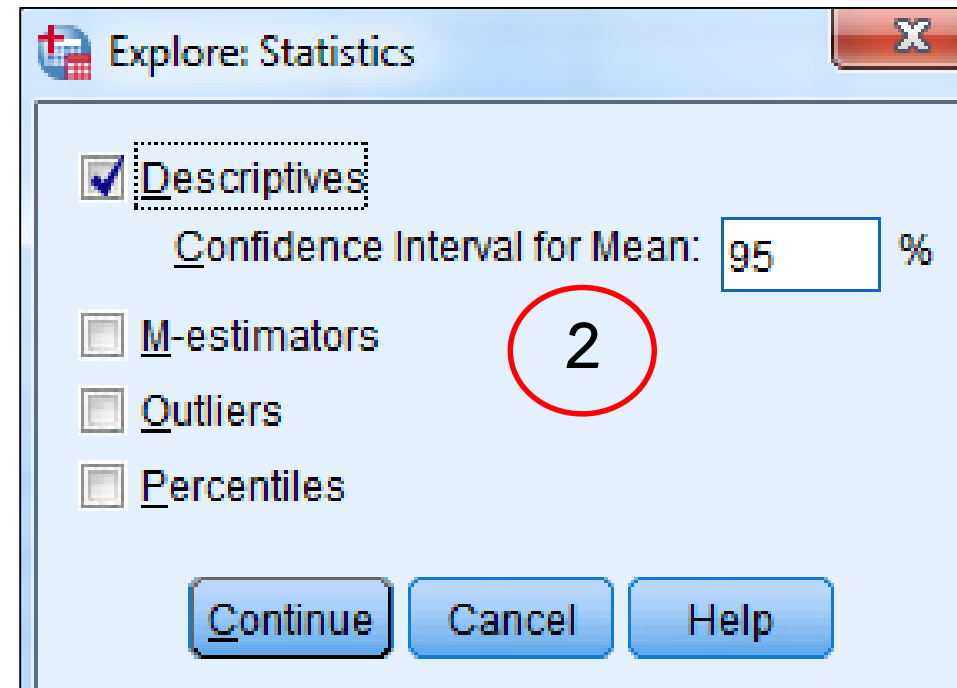
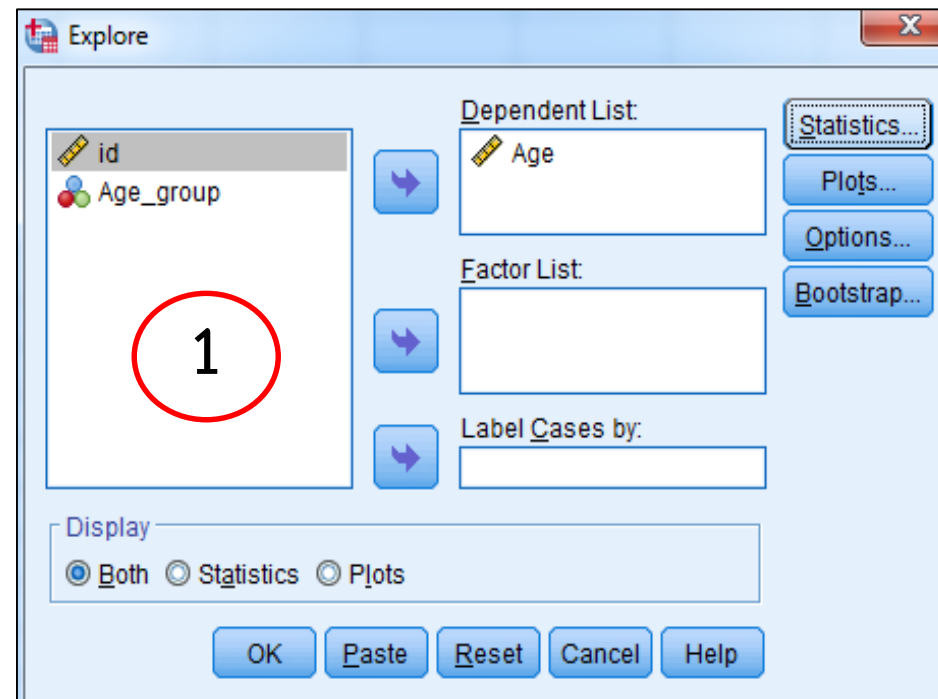
ข้อมูลมีการแจกแจงเป็นแบบโค้งปกติ

H_0 : ข้อมูลมีการแจกแจงเป็นแบบโค้งปกติ

H_1 : ข้อมูลไม่ได้มีการแจกแจงเป็นแบบโค้งปกติ

ใช้คำสั่ง Analyze > Descriptive Statistics > Explore

หน้าต่างคำสั่ง



หน้าต่าง output

1

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Age	.055	421	.004	.991	421	.009

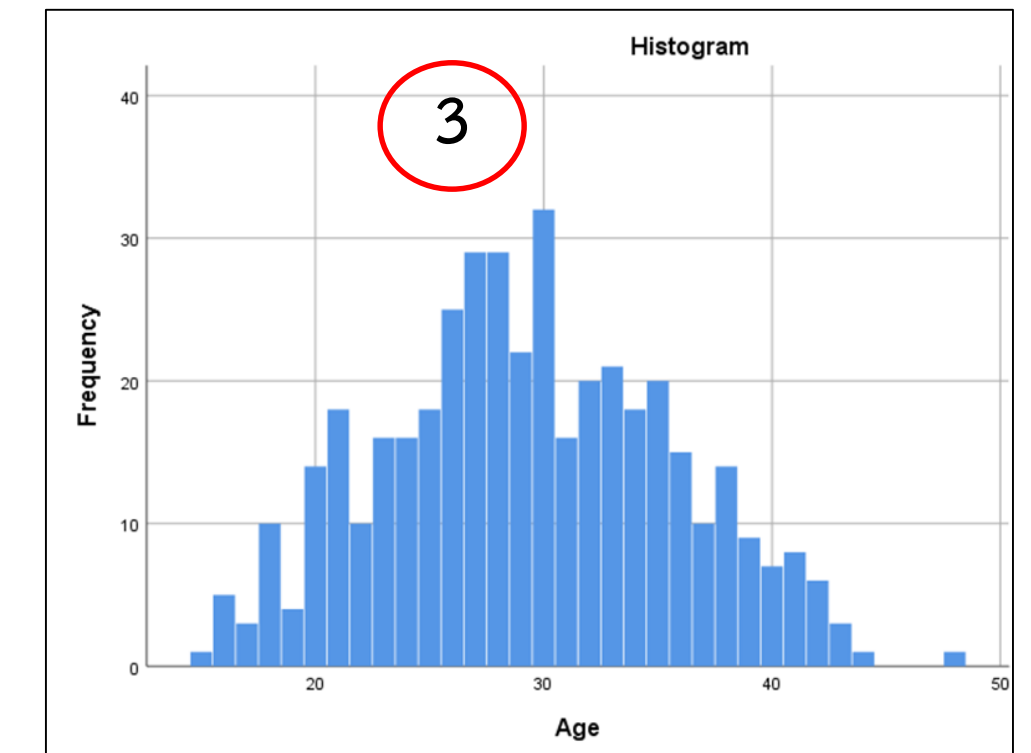
a. Lilliefors Significance Correction

ปฏิเสธ H_0 ที่ $p\text{-value} < 0.05$

2

Descriptives

		Statistic	Std. Error
Age	Mean	29.31	.309
	95% Confidence Interval for Mean		
	Lower Bound	28.70	
	Upper Bound	29.92	
	5% Trimmed Mean	29.27	
	Median	29.00	
	Variance	40.072	
	Std. Deviation	6.330	
	Minimum	15	
	Maximum	48	
	Range	33	
	Interquartile Range	9	
	Skewness	.099	.119
	Kurtosis	-.504	.237



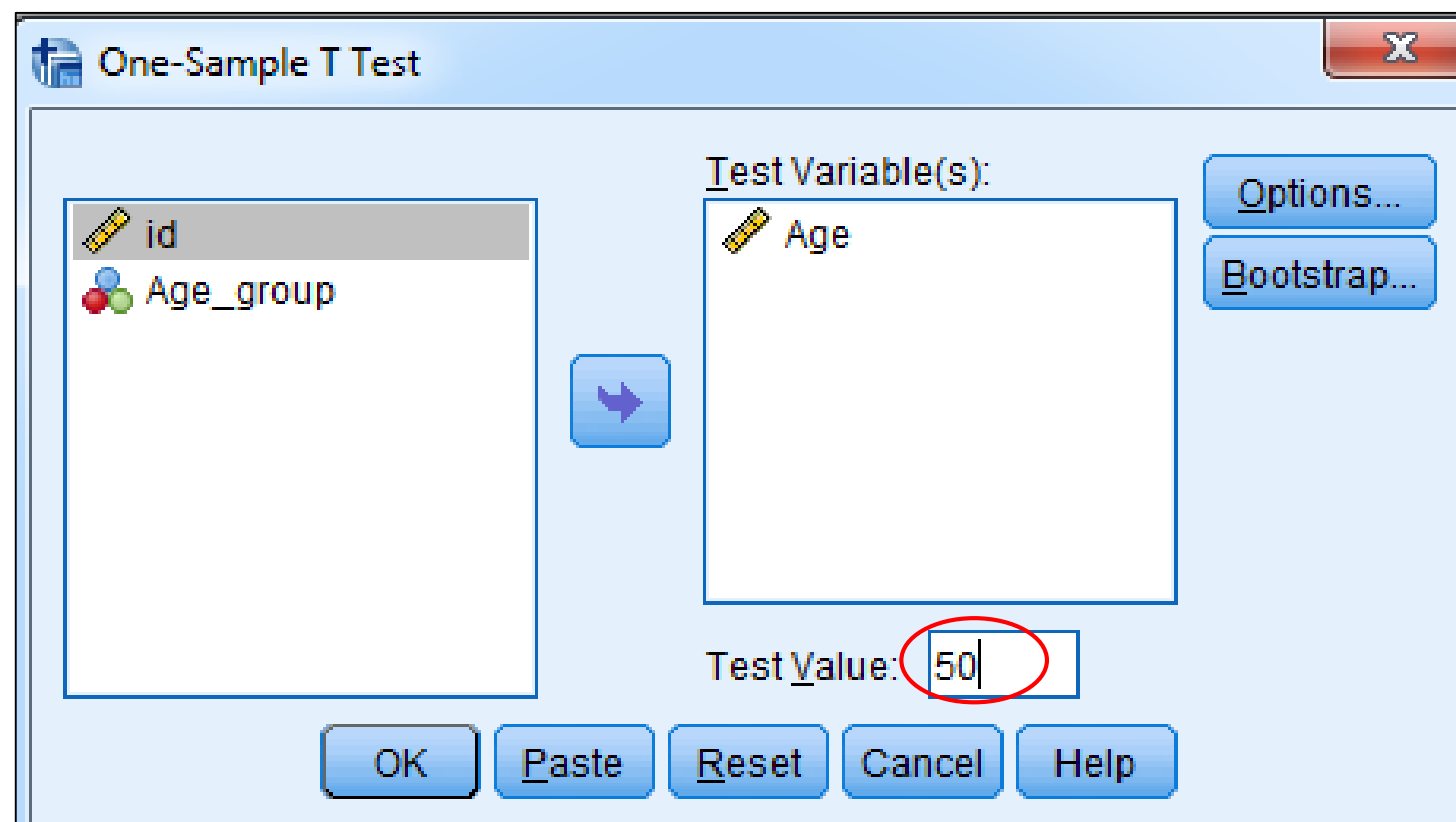
การวิเคราะห์ค่าเฉลี่ยกลุ่มตัวอย่างกลุ่มเดียว (One- Sample T-Test)

การวิเคราะห์ One- Sample T-Test เป็นการทดสอบกลุ่มตัวอย่างกลุ่มเดียวว่า ผลการทดสอบนั้น เป็นไปตามเกณฑ์ที่กำหนดหรือไม่

การทดสอบกำหนดสมมติฐานที่ทดสอบคือ

H_0 : อายุเฉลี่ยของผู้ป่วย เท่ากับ 50 ปี

H_1 : อายุเฉลี่ยของผู้ป่วย ไม่เท่ากับ 50 ปี



ใช้คำสั่ง Analyze > Compare Means > One-Sample T Test

วิธีการแปลผล One-Sample T-Test จากโปรแกรม

หน้าต่าง Output จากโปรแกรม

T-Test

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Age	421	29.31	6.330	.309

One-Sample Test

Test Value = 50

	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Age	-67.059	420	.000	-20.689	-21.30	-20.08

ปฏิเสธ H_0 ที่ $p\text{-value} < 0.05$

แปลผลจาก Output จากโปรแกรม

การทดสอบสมมติฐาน

H_0 : อายุเฉลี่ยของผู้ป่วย เท่ากับ 50 ปี

H_1 : อายุเฉลี่ยของผู้ป่วย ไม่เท่ากับ 50 ปี

ผลการทดสอบ : ค่า $p\text{-value} = <0.001$ ซึ่งมีค่าน้อยกว่า

ระดับนัยสำคัญที่กำหนดคือ 0.05 จึงปฏิเสธ H_0 ดังนั้น

อายุเฉลี่ยของผู้ป่วยไม่เท่ากับ 50 ปี อย่างมีนัยสำคัญทางสถิติ

การวิเคราะห์ค่าเฉลี่ยกลุ่มตัวอย่าง 2 กลุ่มอิสระจากกัน (Independent Sample T-Test)

การวิเคราะห์มีวัตถุประสงค์ของการทดสอบเพื่อทดสอบว่า 2 กลุ่มอิสระจากกันนั้นมีค่าเฉลี่ยแตกต่างกันหรือไม่ โดยมีข้อตกลงเบื้องต้น คือ ข้อมูลทั้งสองกลุ่มต้องมีการแจกแจงเป็นแบบโค้งปกติ

ตัวอย่างวัตถุประสงค์ที่ใช้สถิติ Independent Sample T-Test เช่น

- ต้องการเปรียบเทียบอายุเฉลี่ยในกลุ่มเพศหญิงและเพศชายแตกต่างกันหรือไม่
- ต้องการเปรียบเทียบค่า BMI เฉลี่ยในกลุ่มควบคุมและกลุ่มทดลองแตกต่างกันหรือไม่

ตัวอย่างการใช้ Independent Sample T-Test จากโปรแกรม

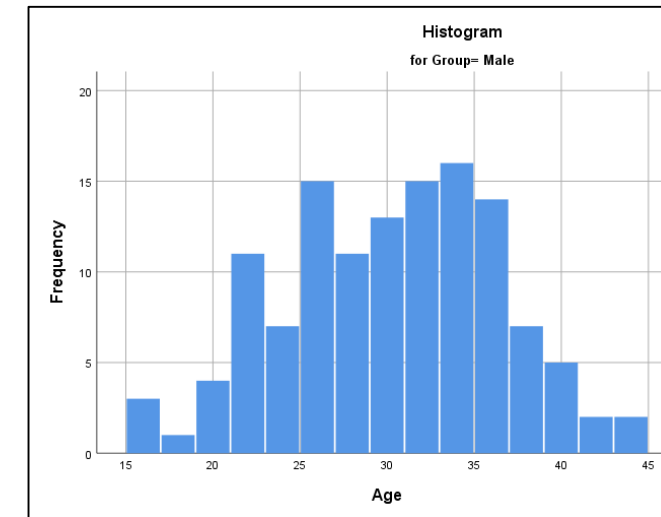
การทดสอบกำหนดสมมติฐานที่ทดสอบคือ

H_0 : อายุเฉลี่ยในกลุ่มเพศหญิงและเพศชายไม่แตกต่างกัน

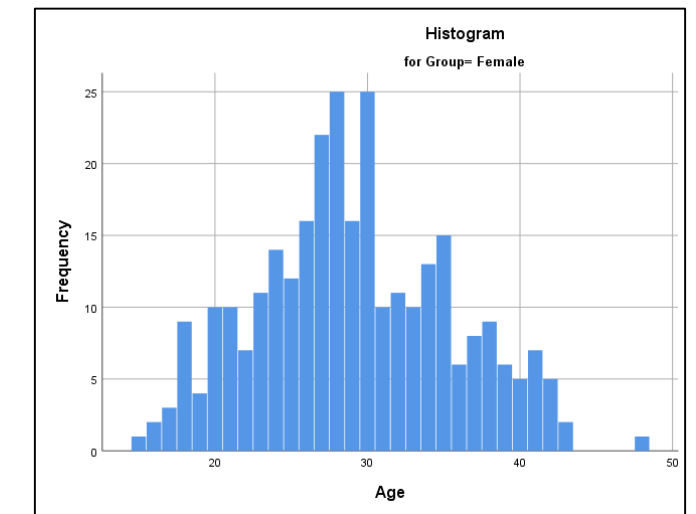
H_1 : อายุเฉลี่ยในกลุ่มเพศหญิงและเพศชายแตกต่างกัน

ใช้คำสั่ง Analyze > Compare Means > Independent-Sample T Test

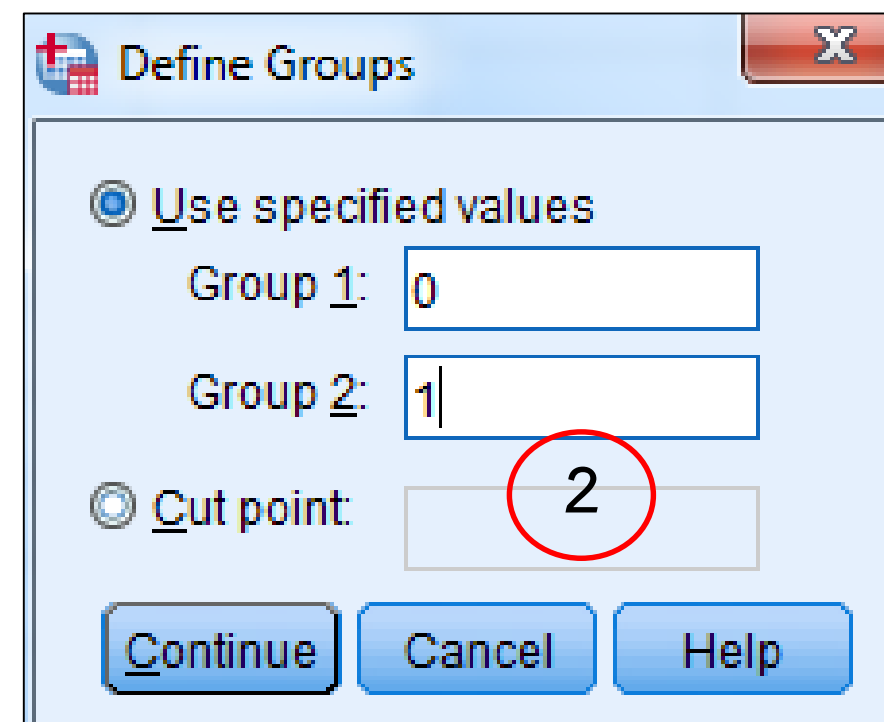
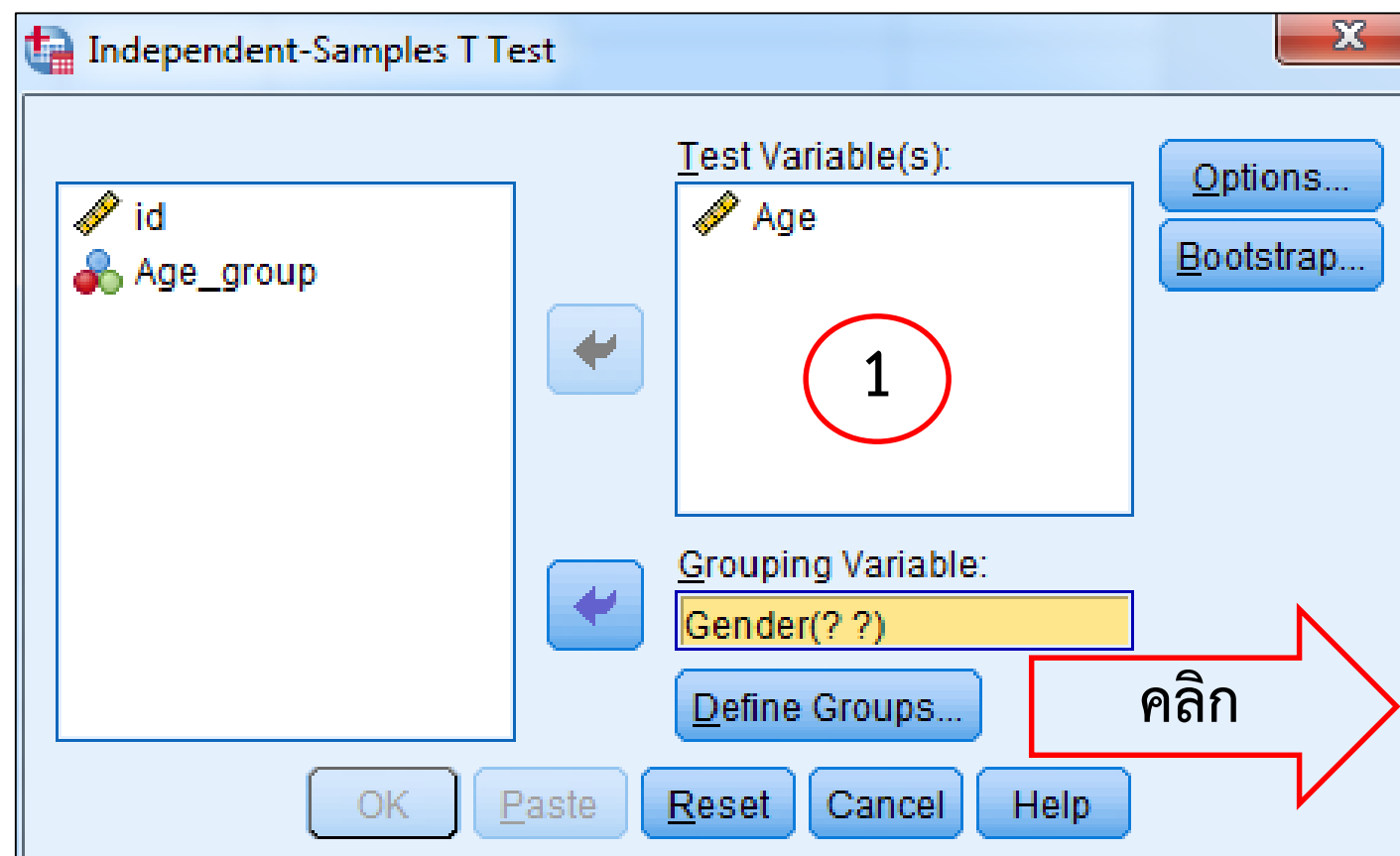
ตัวแปรอายุของทั้งเพศหญิงและเพศชายมีการแจกแจงเป็นแบบโค้งปกติ



Male



Female



ระบุเลข Coding จากข้อมูลดิบ

วิธีการแปลผล Independent Sample T-Test จากโปรแกรม

หน้าต่าง Output จากโปรแกรม

T-Test

Group Statistics					
	Gender	N	Mean	Std. Deviation	Std. Error Mean
Age	Male	126	29.75	6.211	.553
	Female	295	29.12	6.382	.372

การทดสอบกำหนดสมมติฐานที่ทดสอบคือ

H_0 : ความแปรปรวนในแต่ละกลุ่มไม่แตกต่างกัน

H_1 : ความแปรปรวนในแต่ละกลุ่มแตกต่างกัน

ปฏิเสธ H_0 ที่ $p\text{-value} < 0.05$ ในที่นี้ $p > 0.05$ จึงยอมรับ H_0 แสดงว่าความแปรปรวนในแต่ละกลุ่มไม่แตกต่างกัน

Levene's Test for Equality of Variances				t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Age	Equal variances assumed	.003	.957	.938	419	.349	.632	.674	-.693	1.956
	Equal variances not assumed			.948	242.224	.344	.632	.666	-.681	1.945

การทดสอบกำหนดสมมติฐานที่ทดสอบคือ

H_0 : อายุเฉลี่ยในกลุ่มเพศหญิงและเพศชายไม่แตกต่างกัน

H_1 : อายุเฉลี่ยในกลุ่มเพศหญิงและเพศชายแตกต่างกัน

ปฏิเสธ H_0 ที่ $p\text{-value} < 0.05$

แปลผลจาก Output จากโปรแกรม

ผลการทดสอบ : ค่า $p\text{-value} = 0.349$ ซึ่งมีค่ามากกว่าระดับนัยสำคัญที่กำหนดคือ 0.05 จึงยอมรับ H_0 ดังนั้นอายุเฉลี่ยในกลุ่มเพศหญิงและเพศชายไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

การวิเคราะห์ค่าเฉลี่ยกลุ่มตัวอย่าง 2 กลุ่มไม่อิสระจากกัน (Dependent Sample T-Test or Paired T-Test)

การวิเคราะห์ค่าเฉลี่ยของกลุ่มตัวอย่าง 2 กลุ่ม เป็นการทดสอบซ้ำกับกลุ่มตัวอย่างเดิม วัดจากกลุ่มตัวอย่าง 2 ครั้ง (pretest และ posttest) โดยมีข้อตกลงเบื้องต้น คือ ข้อมูลต้องมีการแจกแจงเป็นแบบโค้งปกติ

ตัวอย่างวัตถุประสงค์ที่ใช้สถิติ Paired T-Test เช่น

- ต้องการเปรียบเทียบปริมาณน้ำหนักเฉลี่ยก่อนและหลังเข้ารับโปรแกรมลดกระตุ่นน้ำหนักว่าแตกต่างกันหรือไม่
- ต้องการทดสอบค่า BMI เฉลี่ยก่อนและหลังเข้ารับโปรแกรมออกกำลังกายว่าแตกต่างกันหรือไม่

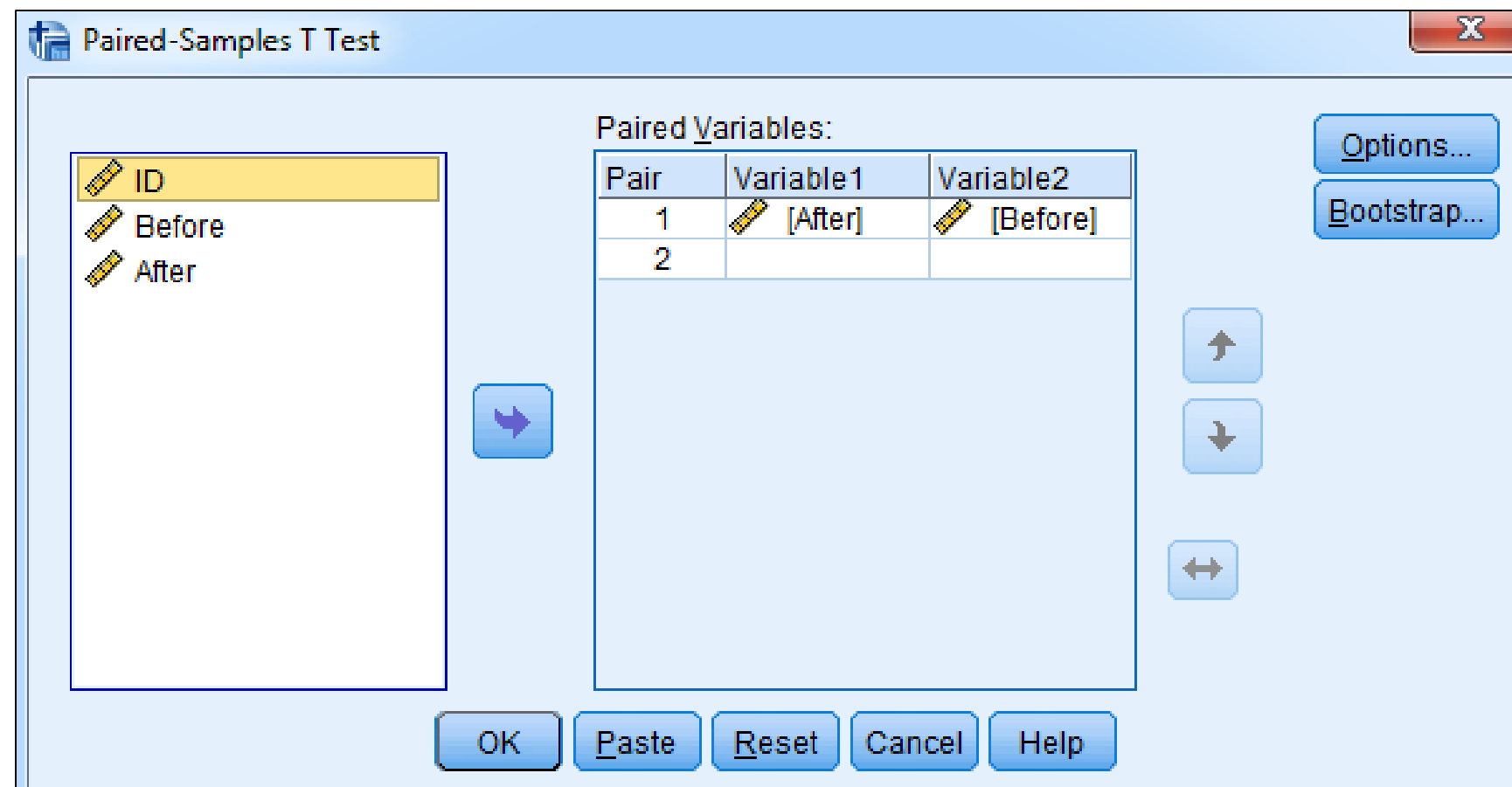
ตัวอย่างการใช้ Paired T-Test จากโปรแกรม

การทดสอบกำหนดสมมติฐานที่ทดสอบคือ

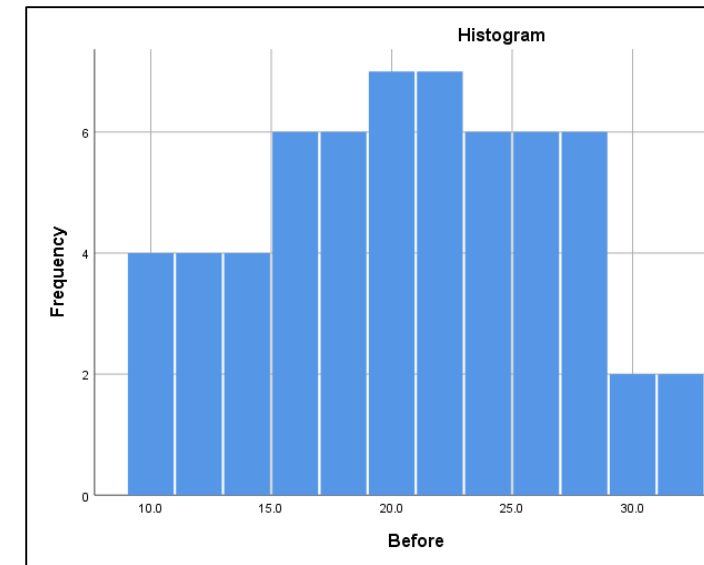
H_0 : ปริมาณน้ำนมเฉลี่ยก่อนและหลังได้รับโปรแกรมไม่แตกต่างกัน

H_1 : ปริมาณน้ำนมเฉลี่ยก่อนและหลังได้รับโปรแกรมแตกต่างกัน

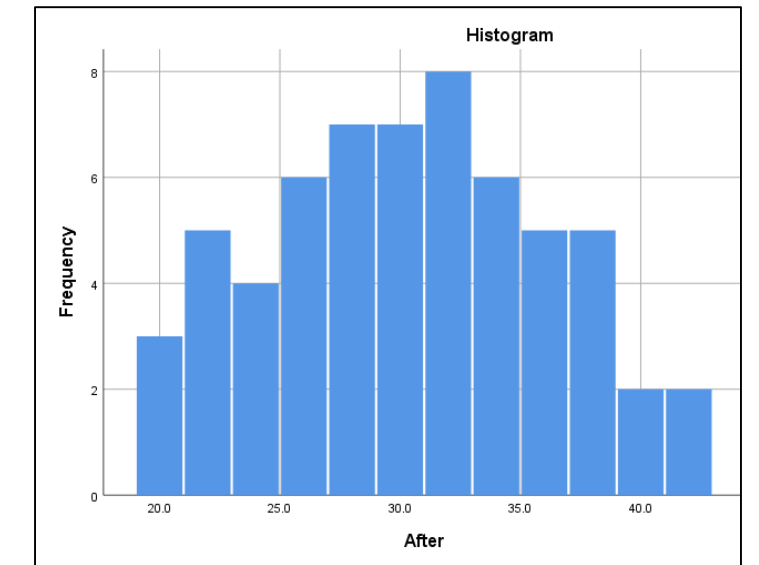
ใช้คำสั่ง Analyze > Compare Means > Paired-Sample T Test



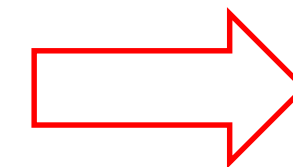
ตัวแปรปริมาณน้ำนมทั้งก่อนและหลังมีการแจกแจงเป็นแบบโค้งปกติ



Before



After



ใส่ตัวแปร After ในช่อง Variable1
ใส่ตัวแปร Before ในช่อง Variable2
กด OK

วิธีการแปลผล Paired T-Test จากโปรแกรม

หน้าต่าง Output จากโปรแกรม

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	After	29.967	60	5.8105	.7501
	Before	20.167	60	5.9523	.7684

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	After & Before	60	.974	.000

Paired Samples Test									
		Mean	Std. Deviation	Std. Error Mean	Paired Differences		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	After - Before	9.8000	1.3505	.1743	9.4511	10.1489	56.211	59	.000

การทดสอบกำหนดสมมติฐานที่ทดสอบคือ

H_0 : ปริมาณน้ำนมเฉลี่ยก่อนและหลังได้รับโปรแกรมไม่แตกต่างกัน

H_1 : ปริมาณน้ำนมเฉลี่ยก่อนและหลังได้รับโปรแกรมแตกต่างกัน

ปฏิเสธ H_0 ที่ $p\text{-value} < 0.05$

แปลผลจาก Output จากโปรแกรม

ผลการทดสอบ : ค่า $p\text{-value} = < 0.001$ ซึ่งมีค่าน้อยกว่าระดับนัยสำคัญที่กำหนดคือ 0.05 จึงปฏิเสธ H_0 ดังนั้นปริมาณน้ำนมเฉลี่ยก่อนและหลังได้รับโปรแกรมแตกต่างกันอย่างมีนัยสำคัญทางสถิติ

การวิเคราะห์ค่าเฉลี่ยของกลุ่มตัวอย่าง มากกว่า 2 กลุ่ม (One-way ANOVA)

การวิเคราะห์ One-way ANOVA เป็นการวิเคราะห์ความแตกต่างของค่าเฉลี่ยของกลุ่มตัวอย่างมากกว่า 2 กลุ่ม ที่มีตัวแปรอิสระ 1 ตัวแปร และเป็นตัวแปรเชิงกลุ่ม ส่วนตัวแปรตามนั้นเป็นตัวแปรปริมาณ โดยมีข้อตกลงเบื้องต้น 2 ข้อ คือ ข้อมูลทุกกลุ่มต้องมีการแจกแจงเป็นแบบโค้งปกติ และความแปรปรวนในแต่ละกลุ่มไม่แตกต่างกัน

ตัวอย่างวัตถุประสงค์ที่ใช้สถิติ One-way ANOVA เช่น

- ต้องการเปรียบเทียบปริมาณน้ำนมเฉลี่ยที่ได้รับวิธีนวดกระตุ้นน้ำนมทั้ง 3 วิธี ว่าแตกต่างกันหรือไม่
- ต้องการเปรียบเทียบคะแนนเฉลี่ยที่ได้รับการสอนทั้ง 3 วิธี ว่าแตกต่างกันหรือไม่

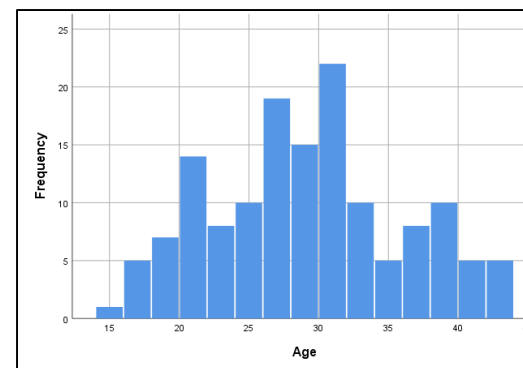
ตัวอย่างการใช้ One-way ANOVA จากโปรแกรม

ตัวแปรอายุของทั้ง 3 วิธีมีการแจกแจงเป็นแบบโค้งปกติ

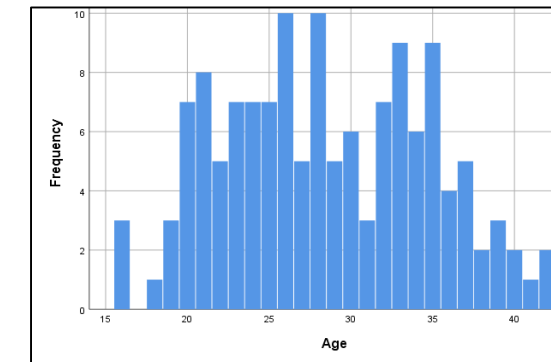
การทดสอบกำหนดสมมติฐานที่ทดสอบคือ

H_0 : อายุเฉลี่ยของทั้ง 3 วิธี ไม่แตกต่างกัน

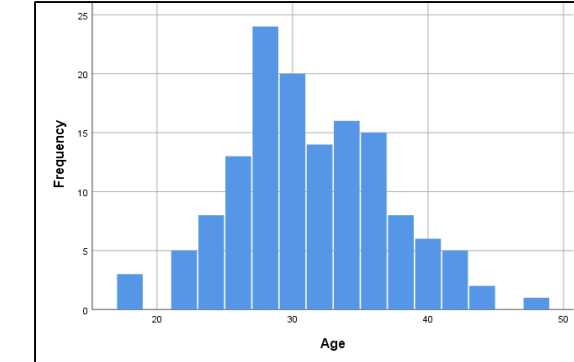
H_1 : อายุเฉลี่ยแตกต่างกันอย่างน้อย 1 คู่



วิธีที่1

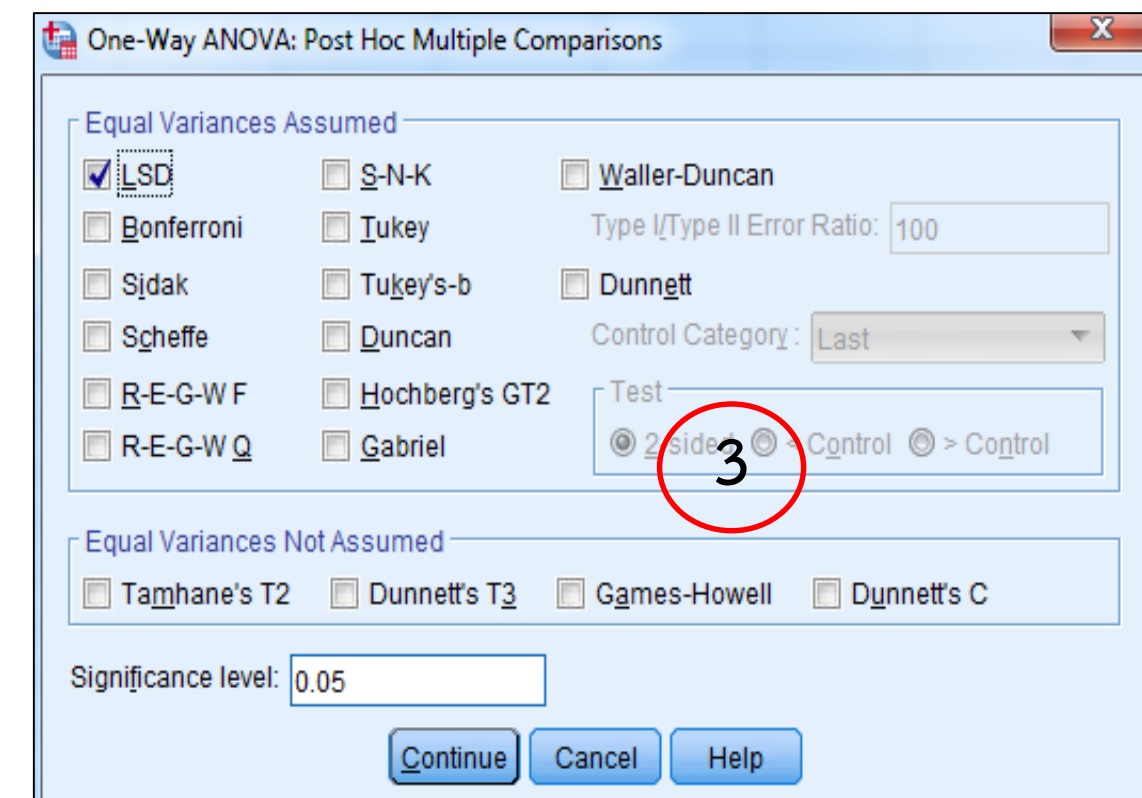
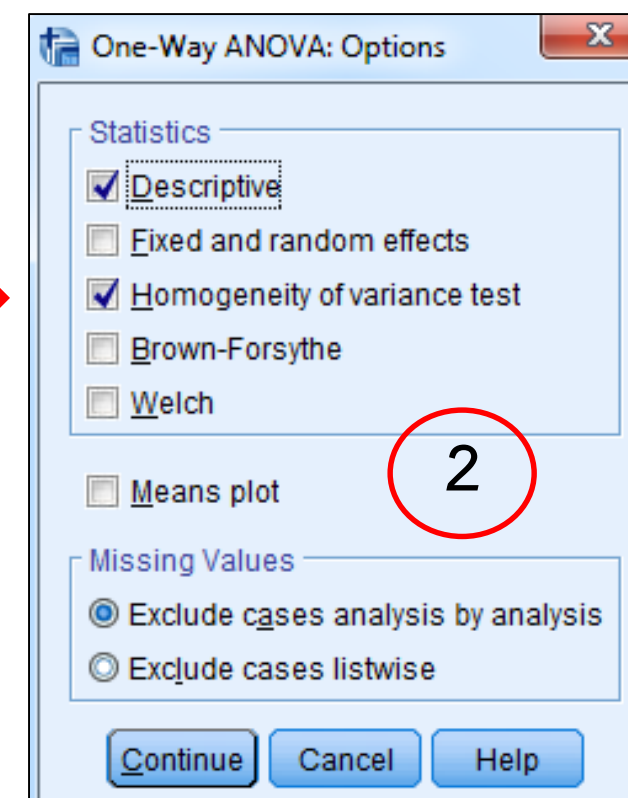
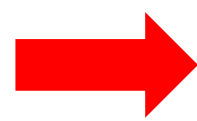
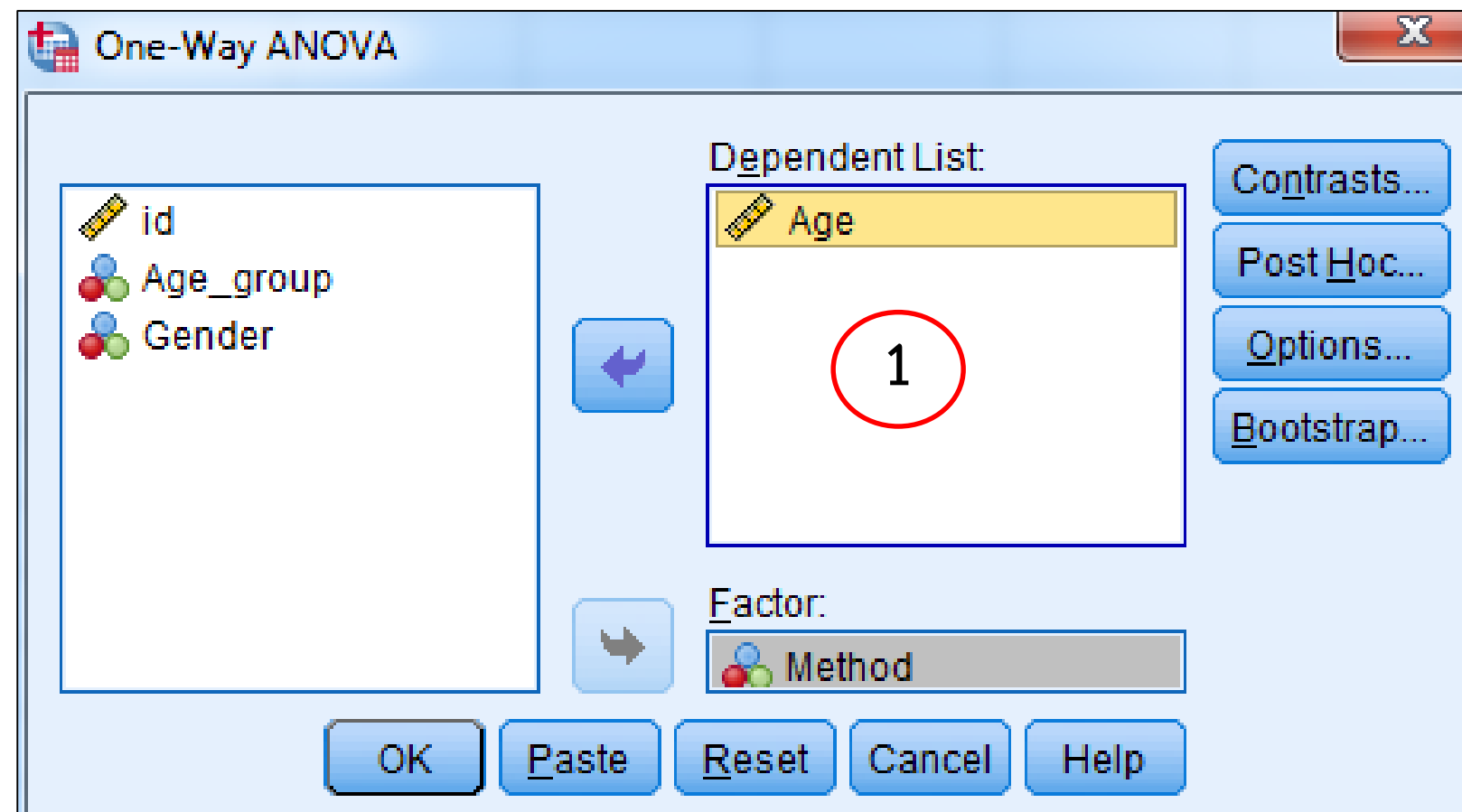


วิธีที่2



วิธีที่3

ใช้คำสั่ง Analyze > Compare Means > One- Way ANOVA



วิธีการแปลผล One-way ANOVA จากโปรแกรม

Descriptives								
Age	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
วิธีที่1	144	28.60	6.792	.566	27.49	29.72	15	43
วิธีที่2	137	28.50	6.230	.532	27.45	29.56	16	42
วิธีที่3	140	30.83	5.672	.479	29.88	31.78	18	48
Total	421	29.31	6.330	.309	28.70	29.92	15	48

Test of Homogeneity of Variances					
Age	Based on Mean	Levene Statistic	df1	df2	Sig.
		2.339	2	418	.098

ANOVA					
Age	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	483.666	2	241.833	6.184	.002
Within Groups	16346.571	418	39.107		
Total	16830.238	420			

การทดสอบกำหนดสมมติฐานที่ทดสอบคือ

H_0 : อายุเฉลี่ยของทั้ง 3 วิธี ไม่แตกต่างกัน

H_1 : อายุเฉลี่ยแตกต่างกันอย่างน้อย 1 คู่

แปลผลจาก Output จากโปรแกรม

ผลการทดสอบ : ค่า p-value = 0.002 ซึ่งมีค่าน้อยกว่า ระดับนัยสำคัญที่กำหนดคือ 0.05 จึงปฏิเสธ H_0 ดังนั้นอายุเฉลี่ยแตกต่างกันอย่างน้อย 1 คู่ อย่างมีนัยสำคัญทางสถิติ จึงไปทำการเปรียบเทียบเชิงพหุ (Multiple comparisons) ต่อไป

ปฏิเสธ H_0 ที่ p-value < 0.05

วิธีการแปลผล Multiple comparisons จากโปรแกรม

Multiple Comparisons

Dependent Variable: Age
LSD

(I) Method	(J) Method	Mean Difference (I-J)	Std. Error	Sig.	Lower Bound	Upper Bound
วิธีที่1	วิธีที่2	.101	.746	.893	-1.37	1.57
	วิธีที่3	-2.224 [*]	.742	.003	-3.68	-.77
วิธีที่2	วิธีที่1	-.101	.746	.893	-1.57	1.37
	วิธีที่3	-2.325 [*]	.752	.002	-3.80	-.85
วิธีที่3	วิธีที่1	2.224 [*]	.742	.003	.77	3.68
	วิธีที่2	2.325 [*]	.752	.002	.85	3.80

*. The mean difference is significant at the 0.05 level.

การทดสอบกำหนดสมมติฐานที่ทดสอบคือ

H_0 : อายุเฉลี่ยกลุ่ม i และ j ไม่แตกต่างกัน

H_1 : อายุเฉลี่ยกลุ่ม i และ j แตกต่างกัน

โดยที่ $i = 1,2,3$ และ $j = 1,2,3$

แปลผลจาก Output จากโปรแกรม

- อายุเฉลี่ยกลุ่ม 1 และ 3 แตกต่างกัน

- อายุเฉลี่ยกลุ่ม 2 และ 3 แตกต่างกัน

สรุปได้ว่าอายุเฉลี่ยกลุ่ม 3 แตกต่างจากกลุ่มอื่น

อย่างมีนัยสำคัญทางสถิติ

ปฏิเสธ H_0 ที่ $p\text{-value} < 0.05$

การวิเคราะห์ความสัมพันธ์ของข้อมูลเชิงคุณภาพ (Pearson Chi-square)

วิธีการวิเคราะห์ Pearson Chi-square เป็นการเปรียบเทียบความสัมพันธ์ระหว่างตัวแปรเชิงคุณภาพ 2 ตัวแปร

- ตัวแปรตาม และ ตัวแปรอิสระ เป็นระดับ Nominal เช่น เพศ หรือ ระดับ Ordinal เช่น ระดับการศึกษา

ตัวอย่างวัตถุประสงค์ที่ใช้สถิติ Pearson Chi-square เช่น

- เพศมีความสัมพันธ์กับการเป็นโรคหรือไม่
- กลุ่มผู้ป่วยเบาหวานมีความสัมพันธ์กับภาวะความดันโลหิตสูงหรือไม่

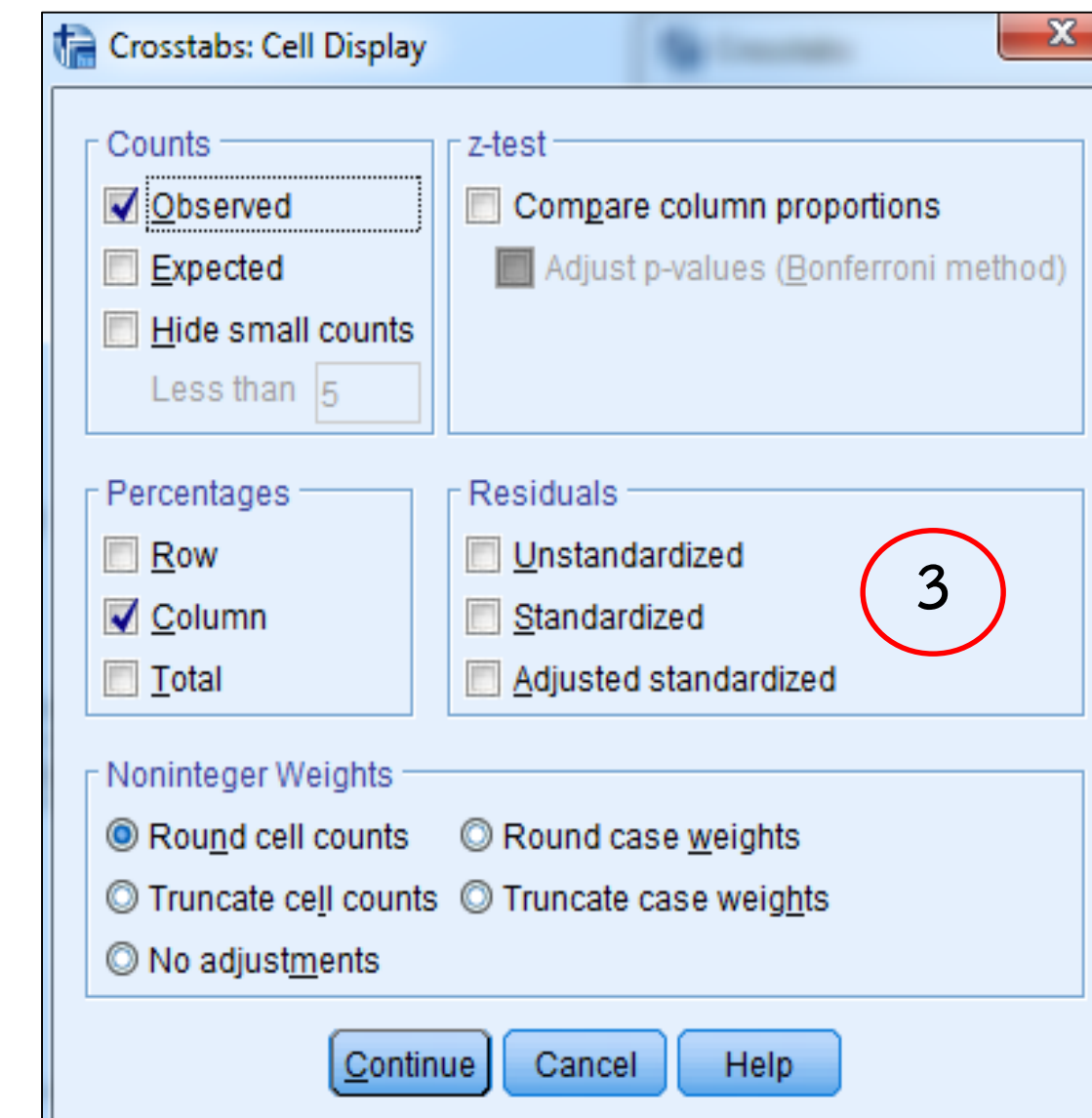
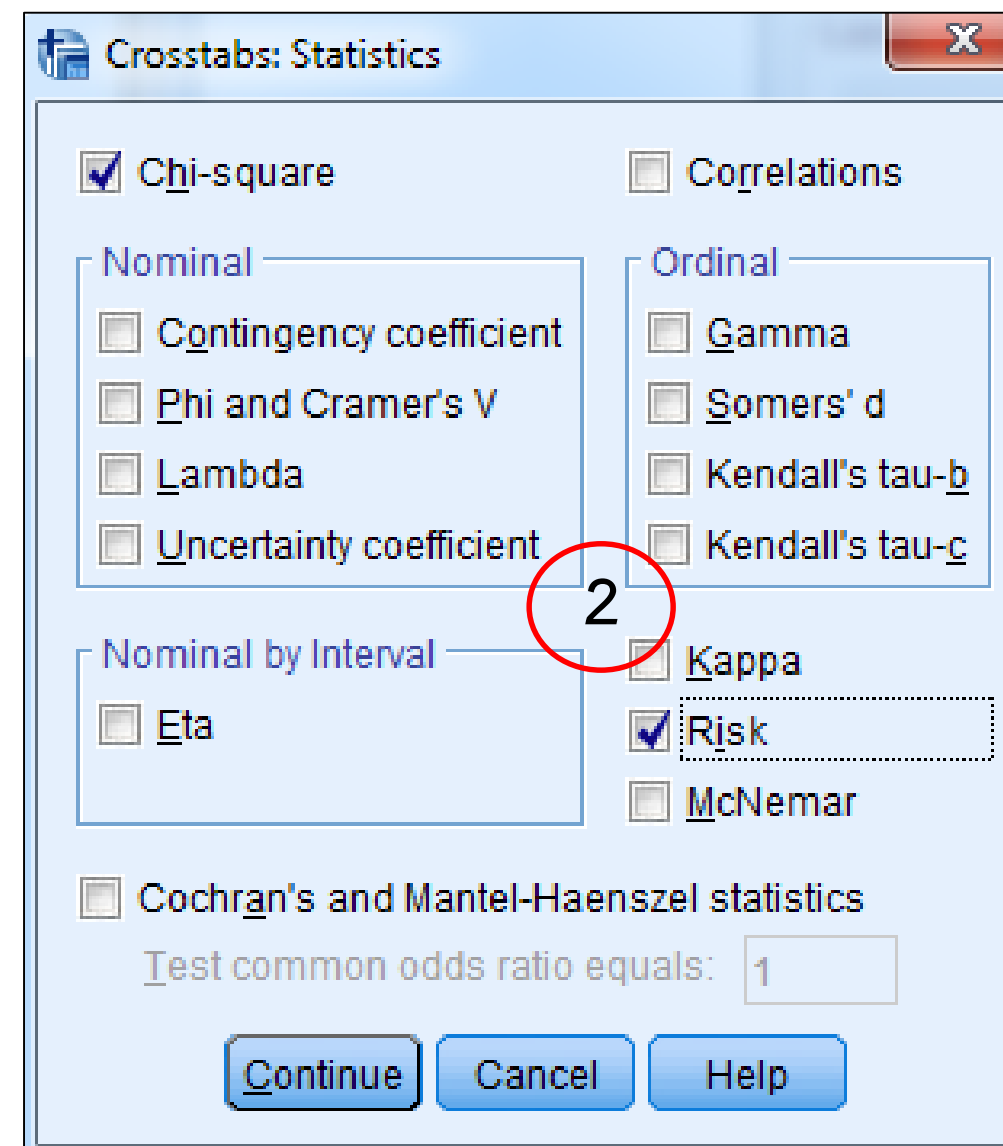
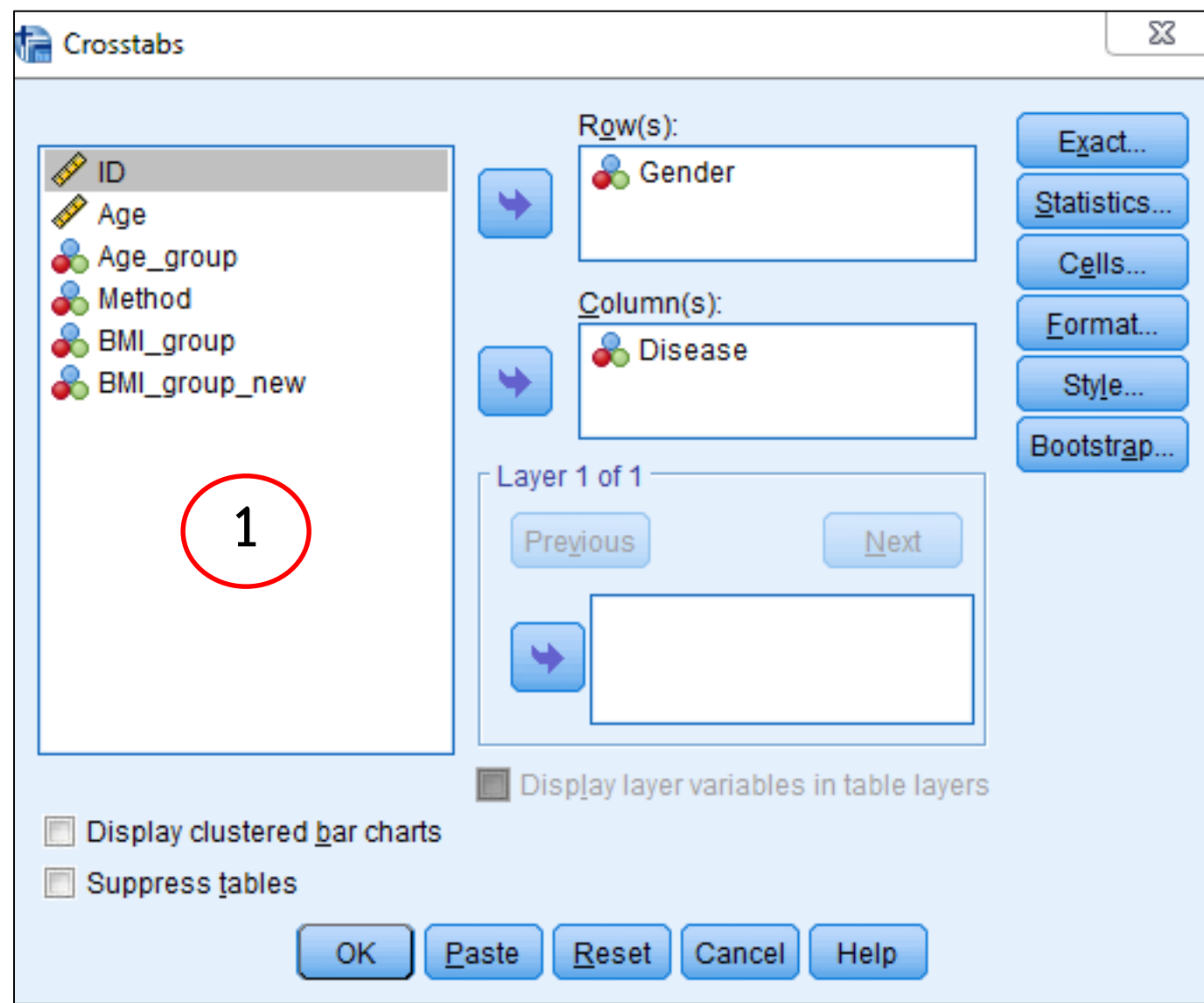
ตัวอย่างการใช้ Pearson Chi-square จากโปรแกรม

ใช้คำสั่ง Analyze > Descriptive Statistics > Crosstabs

การทดสอบกำหนดสมมติฐานที่ทดสอบคือ

H_0 : เพศไม่มีความสัมพันธ์กับการเป็นโรค

H_1 : เพศมีความสัมพันธ์กับการเป็นโรค



วิธีการแปลผล Pearson Chi-square กรณี ตาราง 2*2 จากโปรแกรม

		Disease		Total	
		No	Yes		
Gender	Male	Count	99	27	126
		% within Disease	50.3%	12.1%	29.9%
	Female	Count	98	197	295
		% within Disease	49.7%	87.9%	70.1%
Total		Count	197	224	421
		% within Disease	100.0%	100.0%	100.0%

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	72.935 ^a	1	.000		

	Value	95% Confidence Interval	
		Lower	Upper
Odds Ratio for Gender (Male / Female)	7.371	4.517	12.027

การทดสอบกำหนดสมมติฐานที่ทดสอบคือ

H_0 : เพศไม่มีความสัมพันธ์กับการเป็นโรค

H_1 : เพศมีความสัมพันธ์กับการเป็นโรค

ปฏิเสธ H_0 ที่ p-value < 0.05

แปลผลจาก Output จากโปรแกรม

ผลการทดสอบ : ค่า p-value = <0.001 ซึ่งมีค่าน้อยกว่าระดับนัยสำคัญที่กำหนดคือ 0.05 จึงปฏิเสธ H_0 ดังนั้น

เพศมีความสัมพันธ์กับการเป็นโรค โดยที่เพศหญิงมีโอกาสเสี่ยงเป็นโรคมากกว่าเพศชาย 7.37 เท่าอย่างมีนัยสำคัญทางสถิติ

วิธีการยุบ cell กรณี มี cell ใด cell หนึ่งมีค่าเท่ากับ 0 ในตาราง ROW*COL

		Disease		Total	
		No	Yes		
BMI_group	Underweight	Count	64	80	144
		% within Disease	32.5%	35.7%	34.2%
	Normal	Count	72	75	147
		% within Disease	36.5%	33.5%	34.9%
	Obese1	Count	57	69	126
		% within Disease	28.9%	30.8%	29.9%
	Obese2	Count	4	0	4
		% within Disease	2.0%	0.0%	1.0%
Total		Count	197	224	421
		% within Disease	100.0%	100.0%	100.0%

ก่อนยุบ cell

		Disease		Total	
		No	Yes		
BMI_new	Underweight	Count	64	80	144
		% within Disease	32.5%	35.7%	34.2%
	Normal	Count	72	75	147
		% within Disease	36.5%	33.5%	34.9%
	Obese	Count	61	69	130
		% within Disease	31.0%	30.8%	30.9%
Total		Count	197	224	421
		% within Disease	100.0%	100.0%	100.0%

หลังยุบ cell

***กรณียุบ cell จะสามารถยุบได้ในระดับที่ใกล้เคียงกันเท่านั้น
 เช่น ระดับ Obese1 และ Obese2 สามารถปรับเป็นระดับ Obese 1 ขึ้นไปได้
 แต่ ถ้ายกตัวอย่างกรณี ระดับ Underweight มี cell เท่ากับ 0 จะไม่สามารถ ยุบ cell รวมกับระดับอื่นได้ เนื่องจากลักษณะ
 ของผู้ป่วยอยู่ในระดับแตกต่างจากผู้ป่วยกลุ่มอื่นๆ

วิธีการแปลผล Pearson Chi-square กรณี ตาราง ROW*COL จากโปรแกรม

		Disease		Total	
		No	Yes		
BMI_new	Underweight	Count	64	80	144
		% within Disease	32.5%	35.7%	34.2%
	Normal	Count	72	75	147
		% within Disease	36.5%	33.5%	34.9%
	Obese	Count	61	69	130
		% within Disease	31.0%	30.8%	30.9%
Total		Count	197	224	421
		% within Disease	100.0%	100.0%	100.0%

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	.602 ^a	2	.740
Likelihood Ratio	.603	2	.740
Linear-by-Linear Association	.185	1	.667
N of Valid Cases	421		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 60.83.

การทดสอบกำหนดสมมติฐานที่ทดสอบคือ

H_0 : กลุ่ม BMI ไม่มีความสัมพันธ์กับการเป็นโรค

H_1 : กลุ่ม BMI มีความสัมพันธ์กับการเป็นโรค

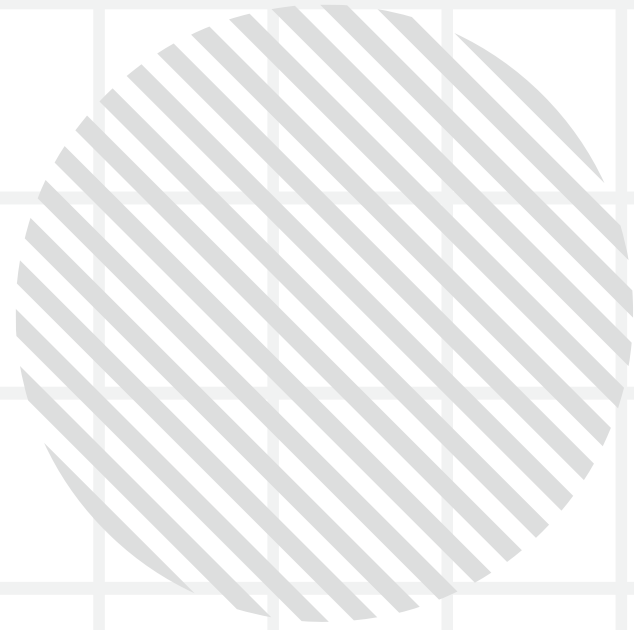
ปฏิเสธ H_0 ที่ $p\text{-value} < 0.05$

แปลผลจาก Output จากโปรแกรม

ผลการทดสอบ : ค่า $p\text{-value} = 0.740$ ซึ่งมีค่ามากกว่าระดับนัยสำคัญที่กำหนดคือ 0.05 จึงไม่สามารถปฏิเสธ H_0 ดังนั้น กลุ่ม BMI ไม่มีความสัมพันธ์กับการเป็นโรคอย่างมีนัยสำคัญทางสถิติ

ตารางสรุปการใช้สถิติเบื้องต้น

ชนิดของตัวแปร	ลักษณะตัวอย่าง	Parametric	Non Parametric
ตัวแปรตามเป็นตัวแปรเชิงปริมาณ	1 กลุ่มตัวอย่าง	One-sample T Test	Sign test
ตัวแปรตามเป็นตัวแปรเชิงปริมาณ ตัวแปรต้นเป็นตัวแปรเชิงคุณภาพ	2 กลุ่มตัวอย่างอิสระกัน	Independent T test	Wilcoxon rank sum or Mann- Whitney U test
ตัวแปรตามเป็นตัวแปรเชิงปริมาณ ตัวแปรต้นเป็นตัวแปรเชิงคุณภาพ	2 กลุ่มตัวอย่างไม่อิสระกัน	Dependent T test or Paired T test	Wilcoxon sign rank test
ตัวแปรตามเป็นตัวแปรเชิงปริมาณ ตัวแปรต้นเป็นตัวแปรเชิงคุณภาพ	>2 กลุ่มตัวอย่างอิสระกัน	One – way ANOVA	Kruskal-Wallis test
ตัวแปรตามเป็นตัวแปรเชิงคุณภาพ ตัวแปรต้นเป็นตัวแปรเชิงคุณภาพ	2 กลุ่มตัวอย่างอิสระกัน		Chi-square test or Fisher’s exact test



Thank you

