世代研究的描述性統計(2)

醫學研究部 醫學研究部 生統小組

徐倩儀

2024/11/19

內容大綱

- 單組及兩組以上連續型變數的描述及比較
  (Parametric test)

- 單組及兩組以上連續型變數的描述及比較
  (Non-parametric test)

醫學研究常用統計

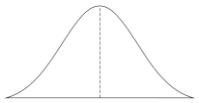
Table 2. Percentage of articles in corpus (N = 113,450) citing the following 16 Delphi panel–derived statistical methods groups

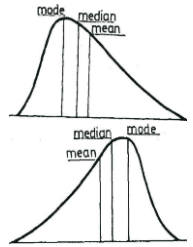
Method groups	Percentage of articles citing specific method group overall and by strata				Relative percent change in method group over time[a]			Examples of commonly occurring terms in each statistical method group
	Overall[b]	95-05[c]	06-10[d]	11-15[e]	06-10 vs. 95-05[f]	11-15 vs. 06-10[g]	11-15 vs. 95-05[h]	
Numeric summary measures	96.4	92.8	96.4	96.8	3.9	0.4	4.3	Average, mode, percentage, standard deviation
Epidemiological measures of risk/effect	53.5	38.0	52.3	55.7	37.5	6.5	46.5	Prevalence, incidence, odds ratio, hazard ratio
Statistical inference concepts	52.9	36.4	52.1	55.0	43.2	5.7	51.4	*P*-value, confidence interval, multiple comparisons
Epidemiological concepts of classification	40.1	43.7	39.8	39.9	−9.0	0.2	−8.8	Sensitivity, specificity, ROC curve
Specific hypothesis test	28.8	23.8	29.0	29.2	21.8	0.7	22.7	*t*-test, Fisher exact test, chi-square test
ANOVA	23.2	14.9	22.2	24.4	49.0	9.8	63.6	ANOVA, ANCOVA, RMANOVA
Regression	22.6	11.9	21.6	24.1	82.1	11.5	103.0	Linear, logistic, poisson regression
Graphics	8.8	8.5	8.7	8.8	2.7	1.3	4.0	Histogram, scatter plot, box plot
Survival analysis	6.8	3.0	6.6	7.3	123.2	11.0	147.6	Cox regression, Kaplan–Meier
Missing data	6.8	2.4	5.9	7.6	148.4	28.0	217.9	Missing data, multiple imputation, LOCF
Computationally intensive algorithms	6.3	3.8	6.2	6.5	63.5	4.5	70.9	Simulation, bootstrap, Monte Carlo, MCMC
Multivariate methods	5.9	3.1	5.5	6.4	73.6	16.6	102.5	Cronbach α, factor analysis, PCA, cluster analysis
Correlated data analysis	4.1	1.6	3.7	4.6	137.3	23.1	192.2	GEE, LMM, GLMM, multilevel model
Machine learning	3.1	1.8	3.0	3.3	66.0	9.9	82.5	Lasso, wavelet, neural network
Time series	1.4	0.9	1.3	1.5	48.8	12.0	66.7	ARIMA, forecasting, spectral analysis
Causal inference observational studies	1.0	0.2	0.8	1.2	268.3	57.8	481.2	Propensity score, instrumental variable

Meaney C, Moineddin R, Voruganti T, O'Brien MA, Krueger P, Sullivan F. (2016). Text mining describes the use of statistical and epidemiological methods in published medical research. *Journal of Clinical Epidemiology*, *74*, 124-132.

Statistical method (Continuous data)

Continuous data	Two groups		Three groups	Correlation
	獨立樣本	相依樣本	獨立樣本	
Parametric	Independent T-test	Paired T-test	ANOVA	Pearson correlation
Nonparametric	Mann-Whitney U test	Wilcoxon signed-rank test	Kruskal-Wallis test	Spearman rank correlation

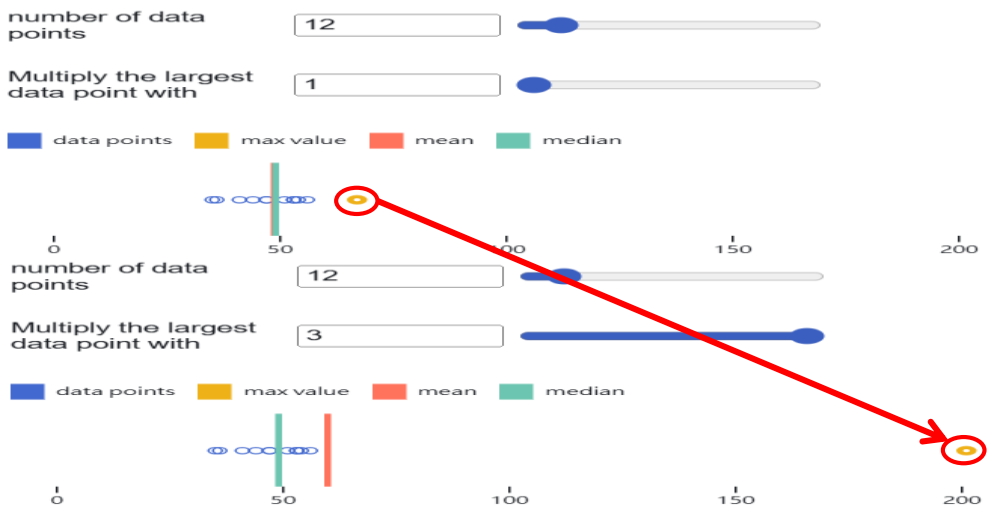




Mean and Median

Mean

- Sensitive to extreme values



Nonparametric

- Rank

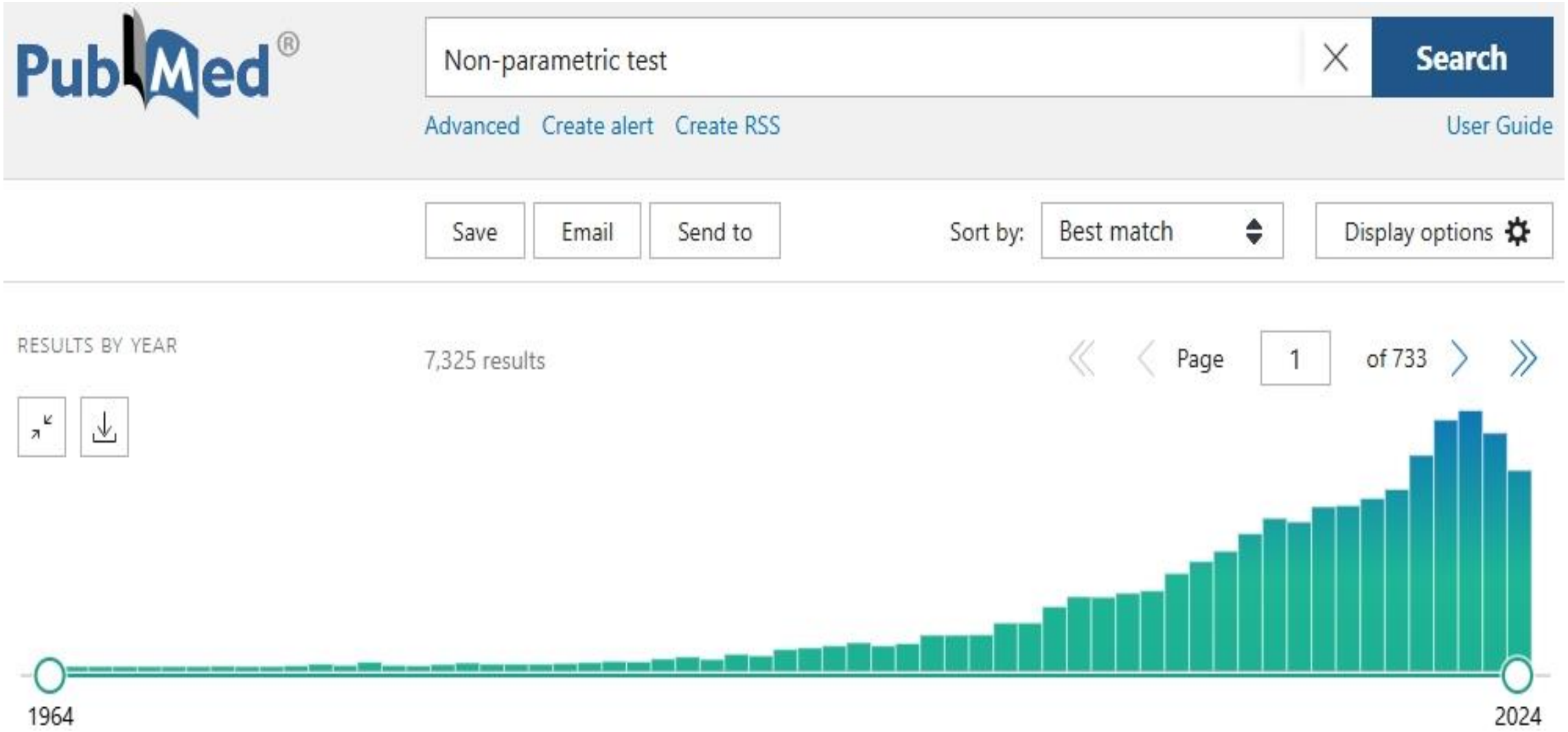
NO	Height	Rank	Group
1	212	9	1
2	151	1	1
3	155	4	1
4	159	8	1
5	152	2	2
6	153	3	2
7	156	5	2
8	157	6	2
9	158	7	2

https://observablehq.com/@mattiasvillani/how-the-mean-and-median-are-affected-by-extreme-observation

Parametric vs Non-parametric test

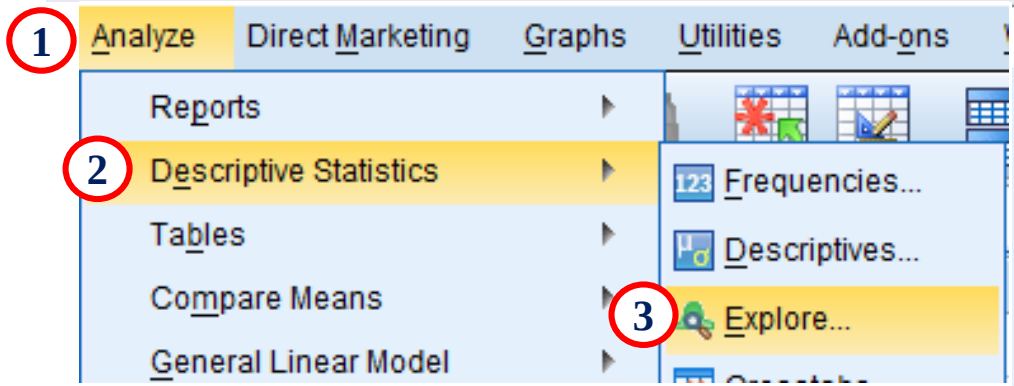
	Parametric test	Non-parametric test
優點	• Power較高 • 以真實觀察值統計	• 不需要為常態分佈 • 樣本數小 • 對於測量誤差較不敏感
缺點	• 需要常態分佈 • 樣本數大 • 容易有測量誤差	• Power較低，且需要較多樣本數才能達顯著差異 • 計算依據Rank而非真實數值

PubMed : non-parametric test



Test of Normality

點選步驟：分析→描述性統計→探索



H_0：常態分佈

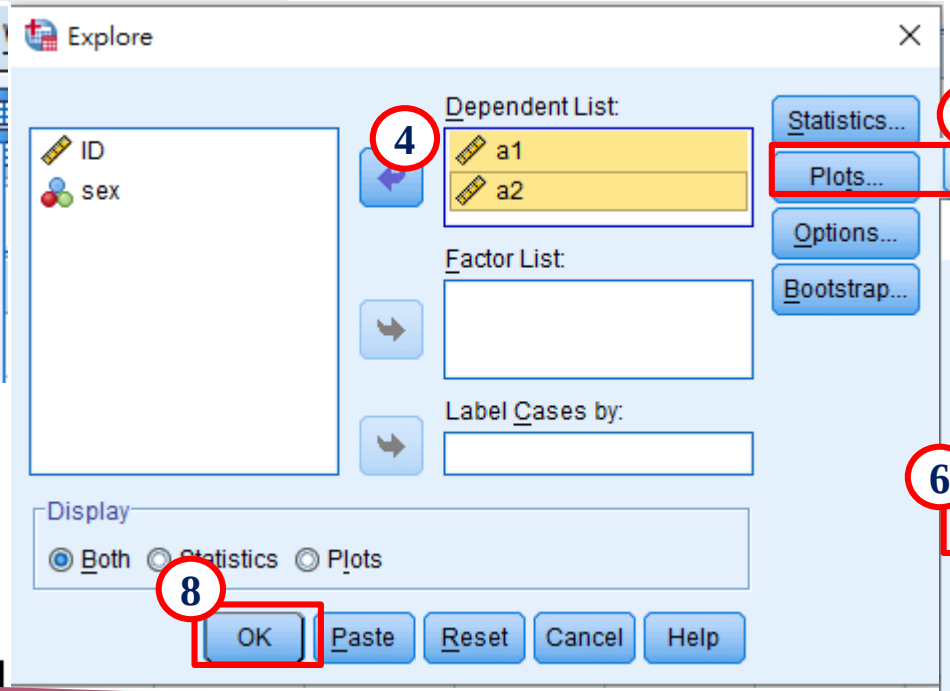
H_1：非常態分佈

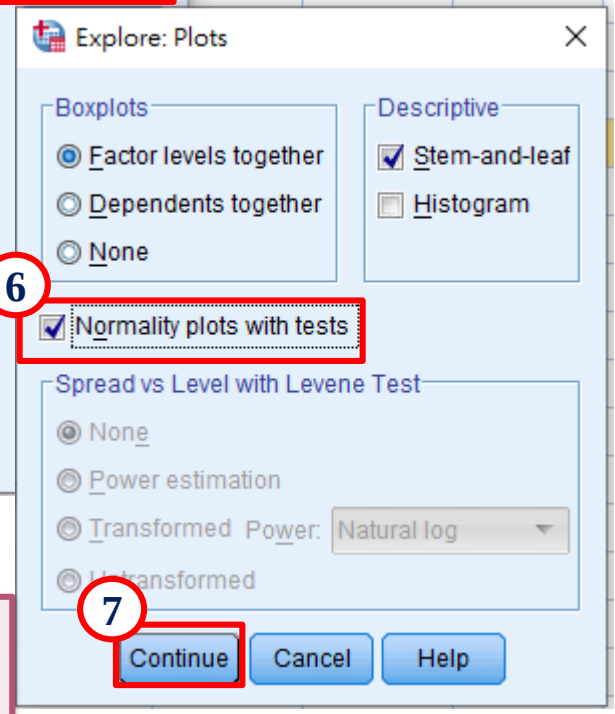
Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
a1	.059	49	.200*	.992	49	.980
a2	.073	49	.200*	.969	49	.215

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction





Kolmogorov-Smirnov test, n ≥ 50

Shapiro-Wilk test, n < 50

Two groups-獨立樣本

	Independent T-test	Mann-Whitney U test
Test	• Mean	• Median
Data	• Normal distribution	• Non-normal distribution

Independent T-test

Example

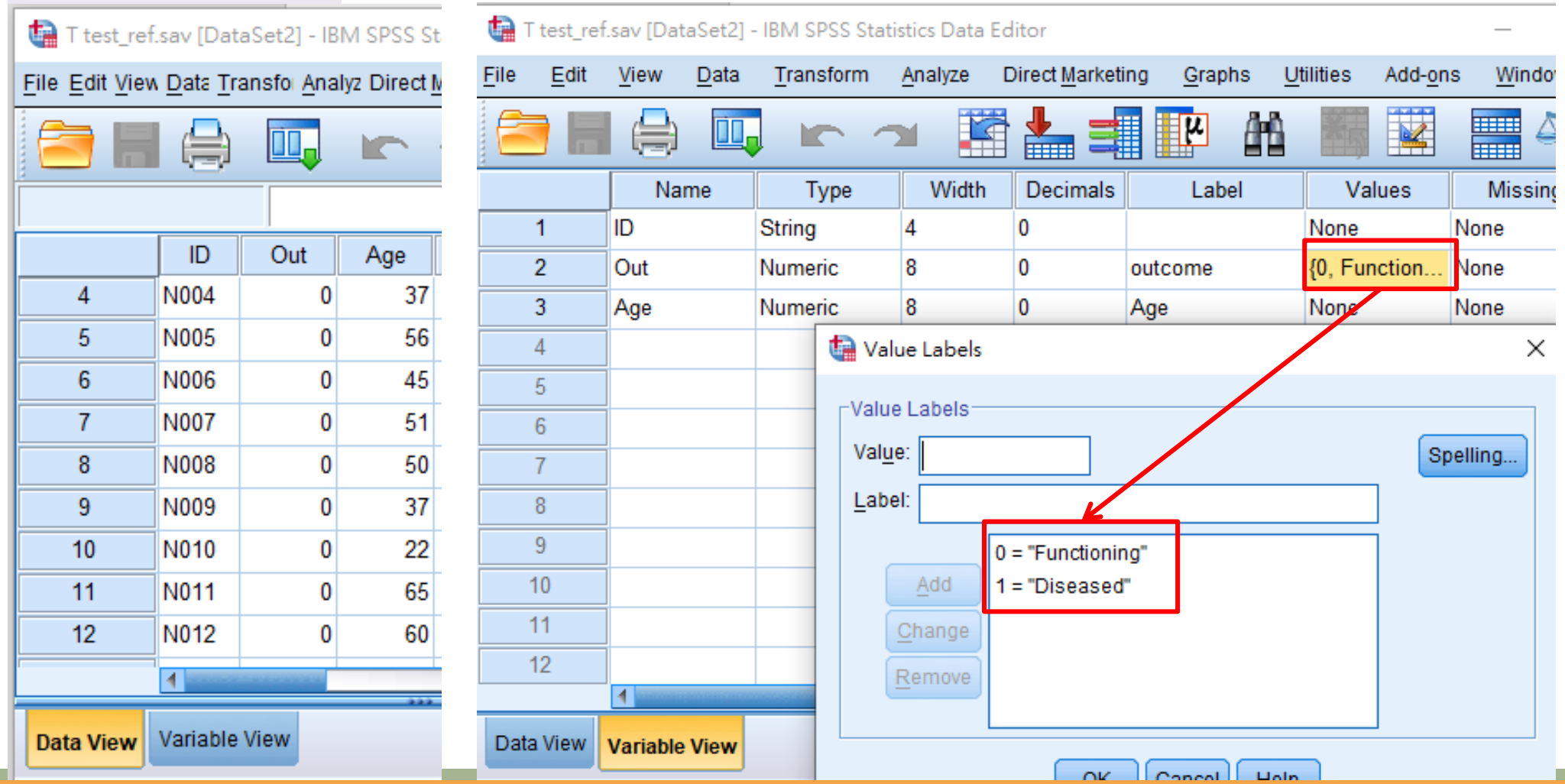
Table 2 Characteristics of the study sample and bivariate associations between investigated variables and treatment outcome.

	Teeth (n = 215), n (%)		Outcome, n (%)				p value
			Diseased (n = 50)		Functioning (n = 165)		
Age (Mean ± SD)t	44.3	±12.4	44.9	±13.3	44.2	±12.2	0.724

- Independent variable
  - Diseased / Functioning (Two groups)
- Dependent variable
  - Age (Continuous)

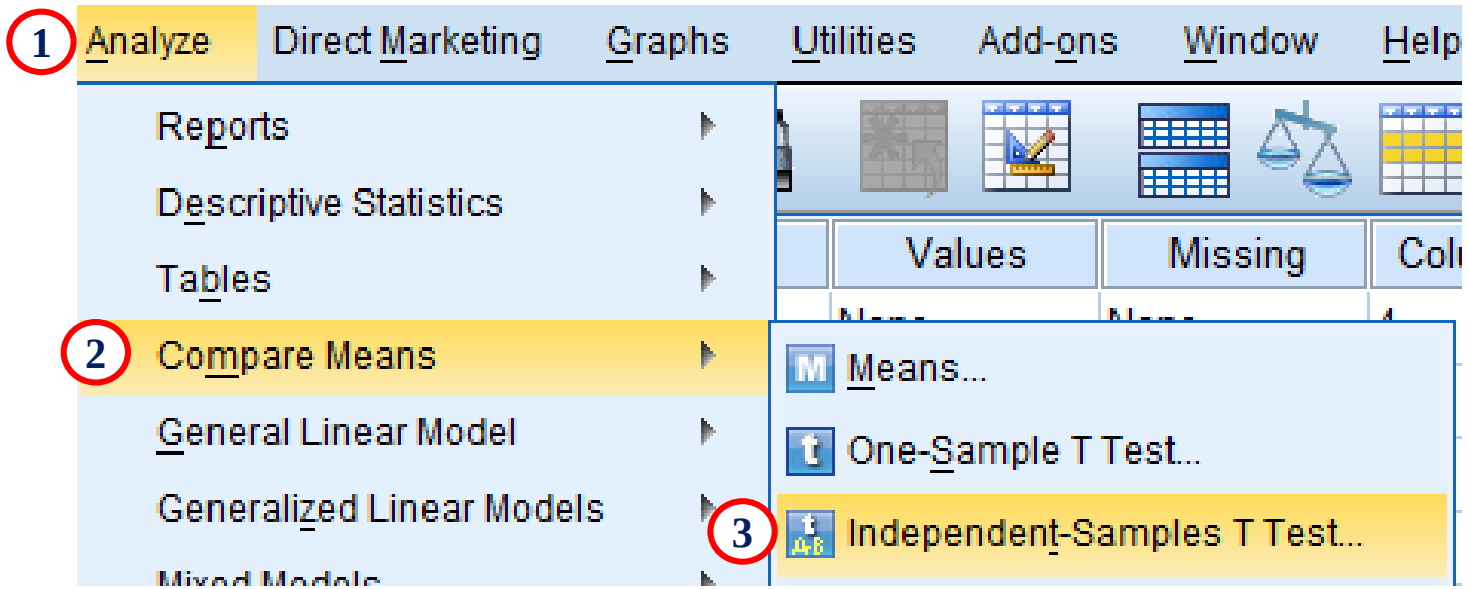
Independent T-test

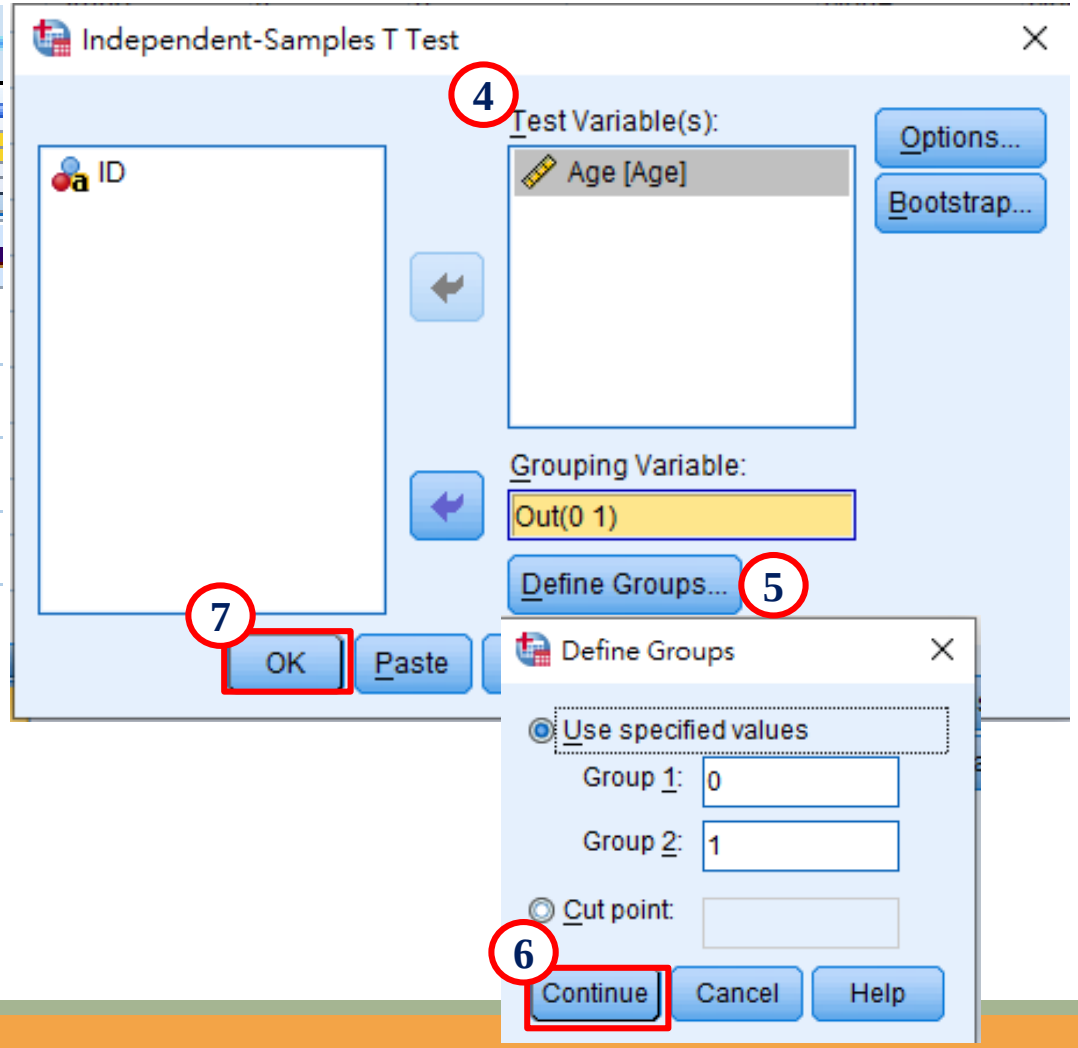
SPSS dataset



Independent T-test (SPSS analysis)

點選步驟：分析→比較平均數法→獨立樣本T檢定





Independent T-test (SPSS output)

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Age	.057	215	.088	.988	215	.060

a. Lilliefors Significance Correction

Group Statistics

	outcome	N	Mean	Std. Deviation	Std. Error Mean
Age	Functioning	165	44.17	12.178	.948
	Diseased	50	44.88	13.304	1.882

Levene's test

H_0：變異數相等

H_1：變異數不相等

T test result：

t = -0.354, p = 0.724

Independent Samples Test

		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	.109	.742	-.354	213	.724	-.710	2.009	-4.671	3.250
	Equal variances not assumed			-.337	75.586	.737	-.710	2.107	-4.907	3.486

Mann-Whitney U test

Example

Table 1. Demographic and clinical characteristics of participants.

	Experimental Group (n = 17)		Control Group (n = 21)		p-Value
	Median	IQR	Median	IQR	
Age (yr)	68.0	(56.5–74.5)	70.0	(63.5–74.5)	0.527
Sex (n,%)					0.444
Male	11	(64.71%)	11	(52.38%)	
Female	6	(35.29%)	10	(47.62%)	
Height (cm)	158.00	(155.50–168.25)	160.00	(158.10–165.50)	0.499
Weight (kg)	69.40	(57.65–70.95)	72.00	(60.60–77.30)	0.122
BMI	24.60	(21.90–27.35)	26.70	(23.56–29.48)	0.082
NIHSS score	6	(3–10)	7	(4.5–11.5)	0.594
HbA1C (%)	6.30	(5.60–7.15)	6.60	(6.20–7.55)	0.181
LDL (mg/dL)	105.00	(78.00–135.00)	121.00	(100.50–179.00)	0.072
Homocysteine (μmol/L)	11.05	(7.99–12.52)	11.00	(9.50–16.00)	0.310

- Independent variable
  - Experimental Group / Control Group (Two groups)
- Dependent variable
  - Age, Height (cm), Weight (kg), BMI, NIHSS score, HbA1C (%)… (Continuous)

Lee YH, Ko LW, Hsu CY, Cheng YY. (2023). Therapeutic Effects of Robotic-Exoskeleton-Assisted Gait Rehabilitation and Predictive Factors of Significant Improvements in Stroke Patients: A Randomized Controlled Trial. *Bioengineering*, *10*(5), 585. https://doi.org/10.3390/bioengineering10050585

Mann-Whitney U test

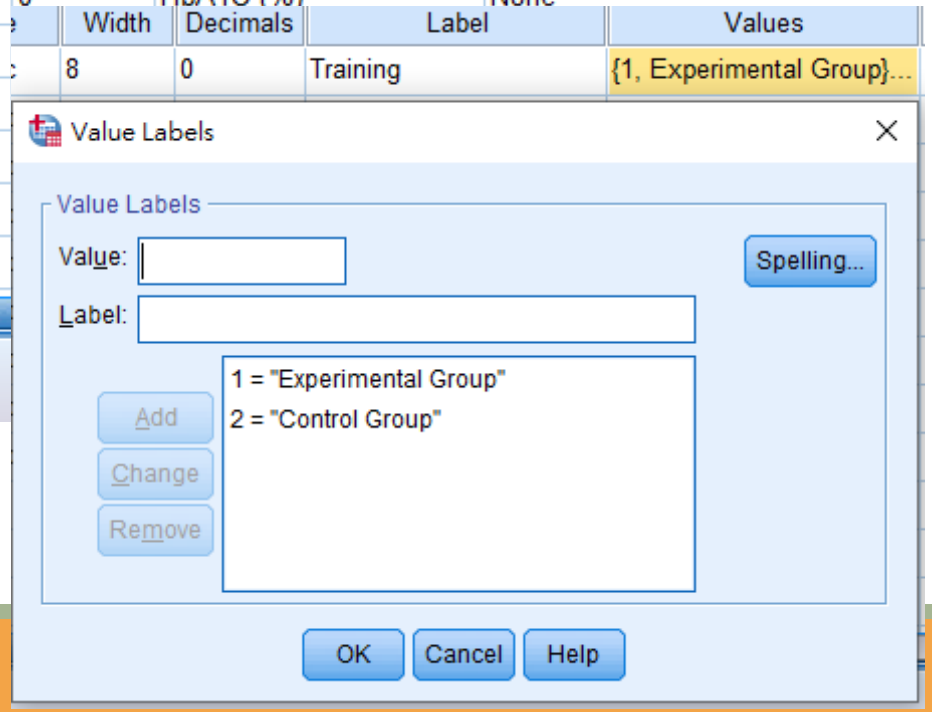
SPSS dataset

	gp	age	bh	bw	BMI
1	1	72	152	60	26
2	1	52	175	70	23
3	1	46	156	71	29
4	1	68	170	71	25
5	1	68	161	63	24
6	1	88	156	55	23
7	1	43	163	54	20
8	1	64	163	69	26
9	1	61	167	74	27
10	1	67	158	63	25
11	1	72	158	70	28

Data View | Variable View

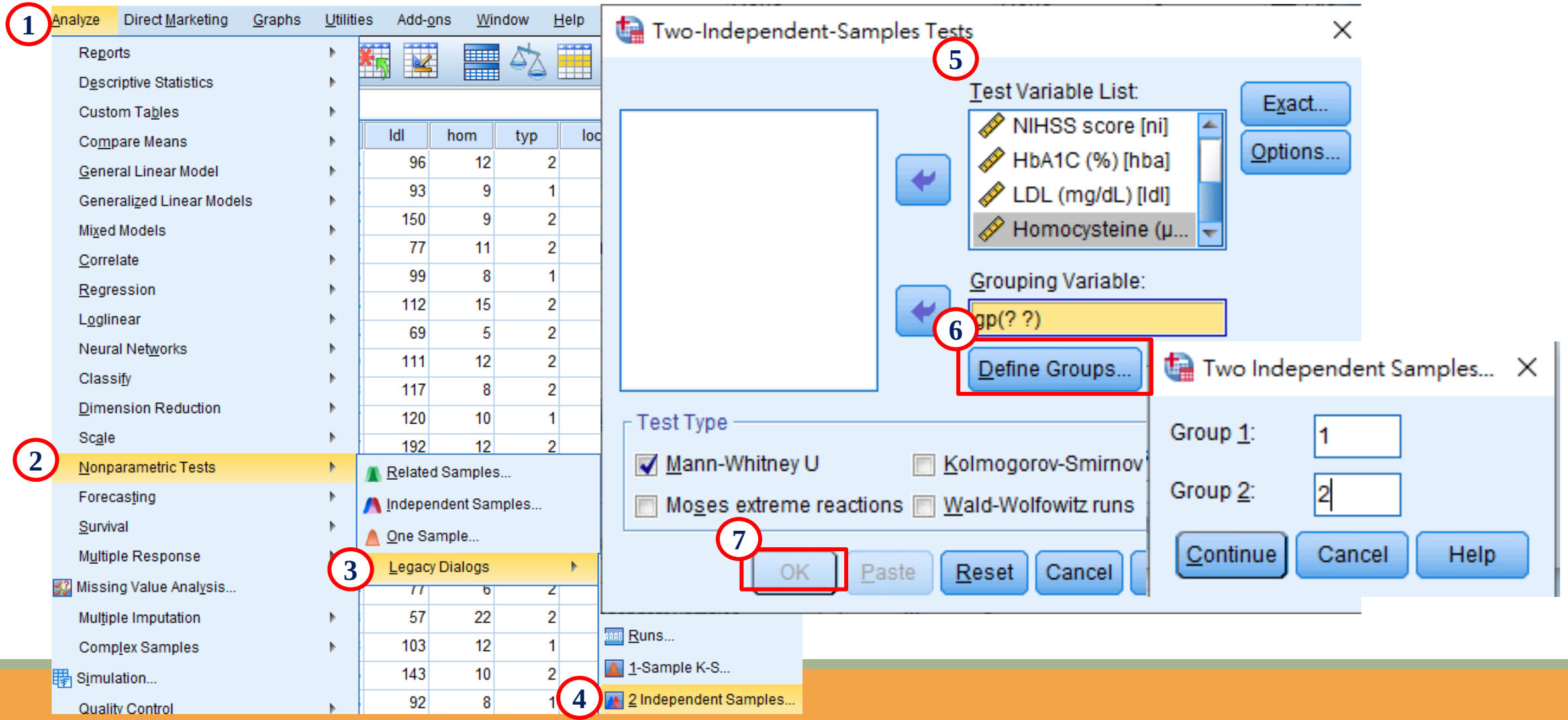
	Name	Type	Width	Decimals	Label	Values
1	gp	Numeric	8	0	Training	{1, Experimental Group}...
2	age	Numeric	8	0	Age (yr)	None
3	bh	Numeric	8	0	Height (cm)	None
4	bw	Numeric	8	0	Weight (kg)	None
5	BMI	Numeric	8	0	BMI	None
6	ni	Numeric	8	0	NIHSS score	None
7	hba	Numeric	8	0	HbA1C (%)	None
8	ldl	Numeric	8			
9	hom	Numeric	8			
10						
11						
12						

Data View | Variable View



Mann-Whitney U test (SPSS analysis)

點選步驟：分析→無母數檢定→歷史對話記錄→2個獨立樣本



Mann-Whitney U test (SPSS output)

Ranks

	Training	N	Mean Rank	Sum of Ranks
Age (yr)	Experimental Group	17	18.21	309.50
	Control Group	21	20.55	431.50
	Total	38		
Height (cm)	Experimental Group	17	18.12	308.00
	Control Group	21	20.62	433.00
	Total	38		
Weight (kg)	Experimental Group	17	16.38	278.50
	Control Group	21	22.02	462.50
	Total	38		
BMI	Experimental Group	17	16.00	272.00
	Control Group	21	22.33	469.00
	Total	38		
NIHSS score	Experimental Group	17	18.41	313.00
	Control Group	21	20.38	428.00
	Total	38		
HbA1C (%)	Experimental Group	17	16.79	285.50
	Control Group	21	21.69	455.50
	Total	38		
LDL (mg/dL)	Experimental Group	17	15.88	270.00
	Control Group	21	22.43	471.00
	Total	38		
Homocysteine (µmol/L)	Experimental Group	17	17.44	296.50
	Control Group	21	21.17	444.50
	Total	38		

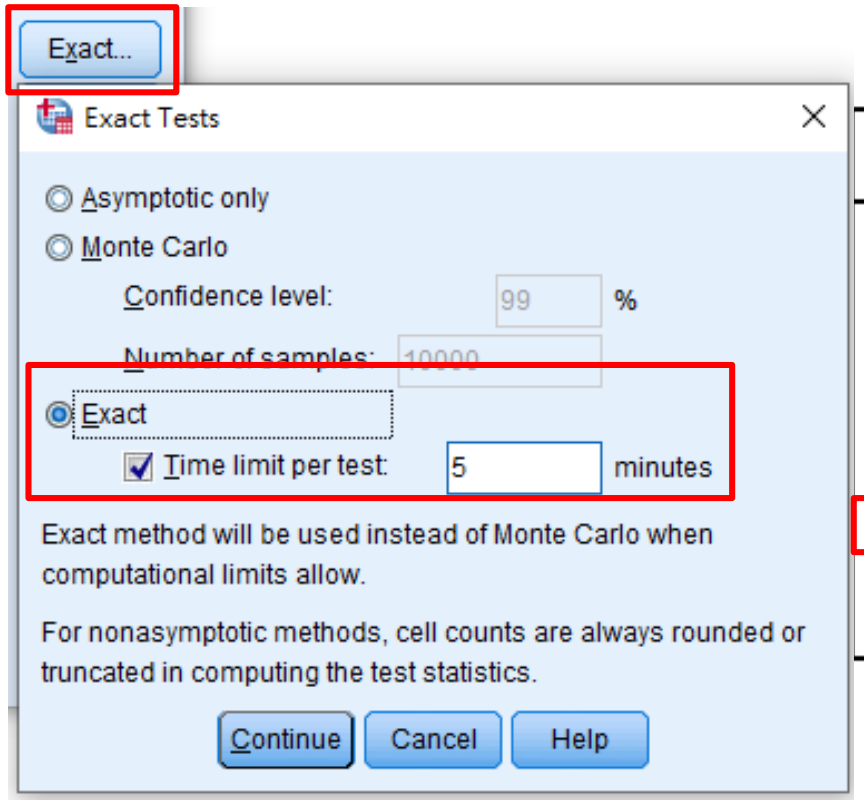
Test Statistics[a]

	Age (yr)	Height (cm)	Weight (kg)	BMI	NIHSS score	HbA1C (%)	LDL (mg/dL)	Homocysteine (µmol/L)
Mann-Whitney U	156.500	155.000	125.500	119.000	160.000	132.500	117.000	143.500
Wilcoxon W	309.500	308.000	278.500	272.000	313.000	285.500	270.000	296.500
Z	-.647	-.691	-1.556	-1.747	-.546	-1.351	-1.806	-1.030
Asymp. Sig. (2-tailed)	.518	.490	.120	.081	.585	.177	.071	.303
Exact Sig. [2*(1-tailed Sig.)]	.523[b]	.504[b]	.121[b]	.083[b]	.601[b]	.179[b]	.073[b]	.308[b]

a. Grouping Variable: Training

b. Not corrected for ties.

Mann-Whitney U test (SPSS output)



Test Statistics[a]

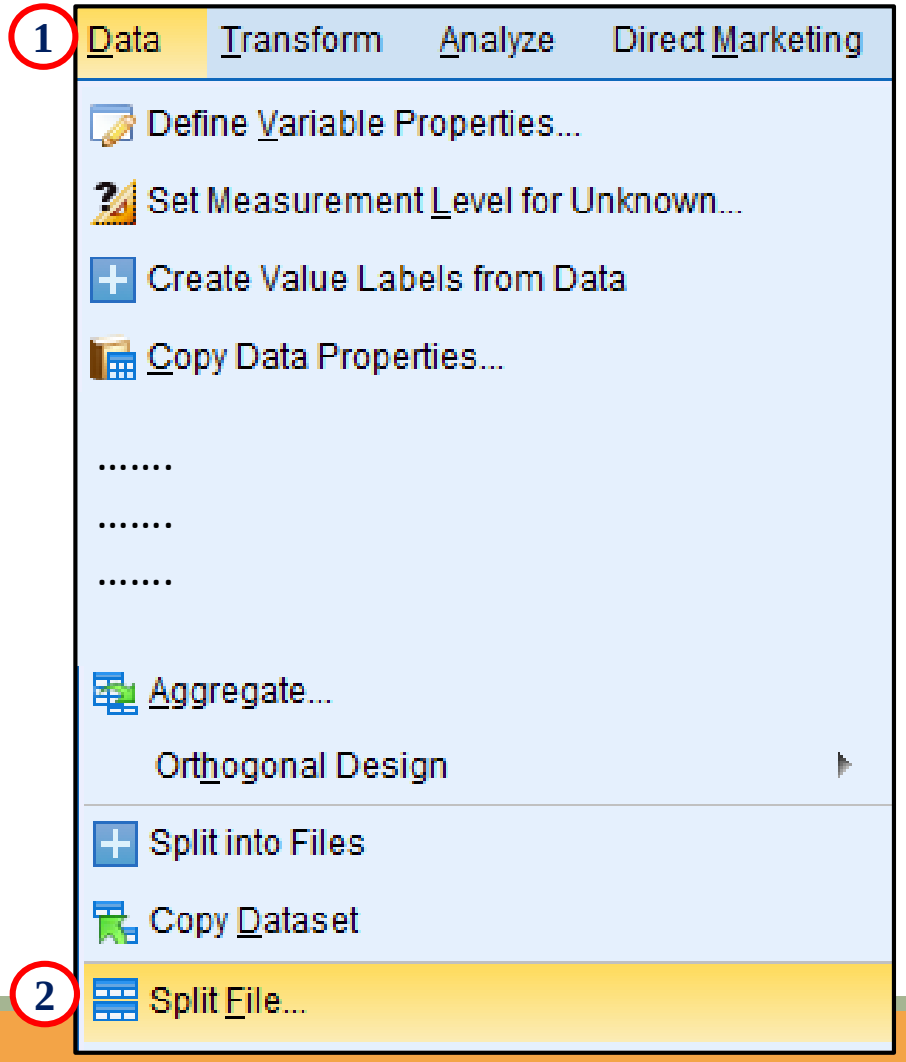
	Age (yr)	Height (cm)	Weight (kg)	BMI	NIHSS score	HbA1C (%)	LDL (mg/dL)	Homocysteine (μmol/L)
Mann-Whitney U	156.500	155.000	125.500	119.000	160.000	132.500	117.000	143.500
Wilcoxon W	309.500	308.000	278.500	272.000	313.000	285.500	270.000	296.500
Z	-.647	-.691	-1.556	-1.747	-.546	-1.351	-1.806	-1.030
Asymp. Sig. (2-tailed)	.518	.490	.120	.081	.585	.177	.071	.303
Exact Sig. [2*(1-tailed Sig.)]	.523[b]	.504[b]	.121[b]	.083[b]	.601[b]	.179[b]	.073[b]	.308[b]
Exact Sig. (2-tailed)	.527	.499	.122	.082	.594	.181	.072	.310
Exact Sig. (1-tailed)	.263	.249	.061	.041	.297	.091	.036	.155
Point Probability	.005	.005	.002	.001	.005	.002	.001	.003

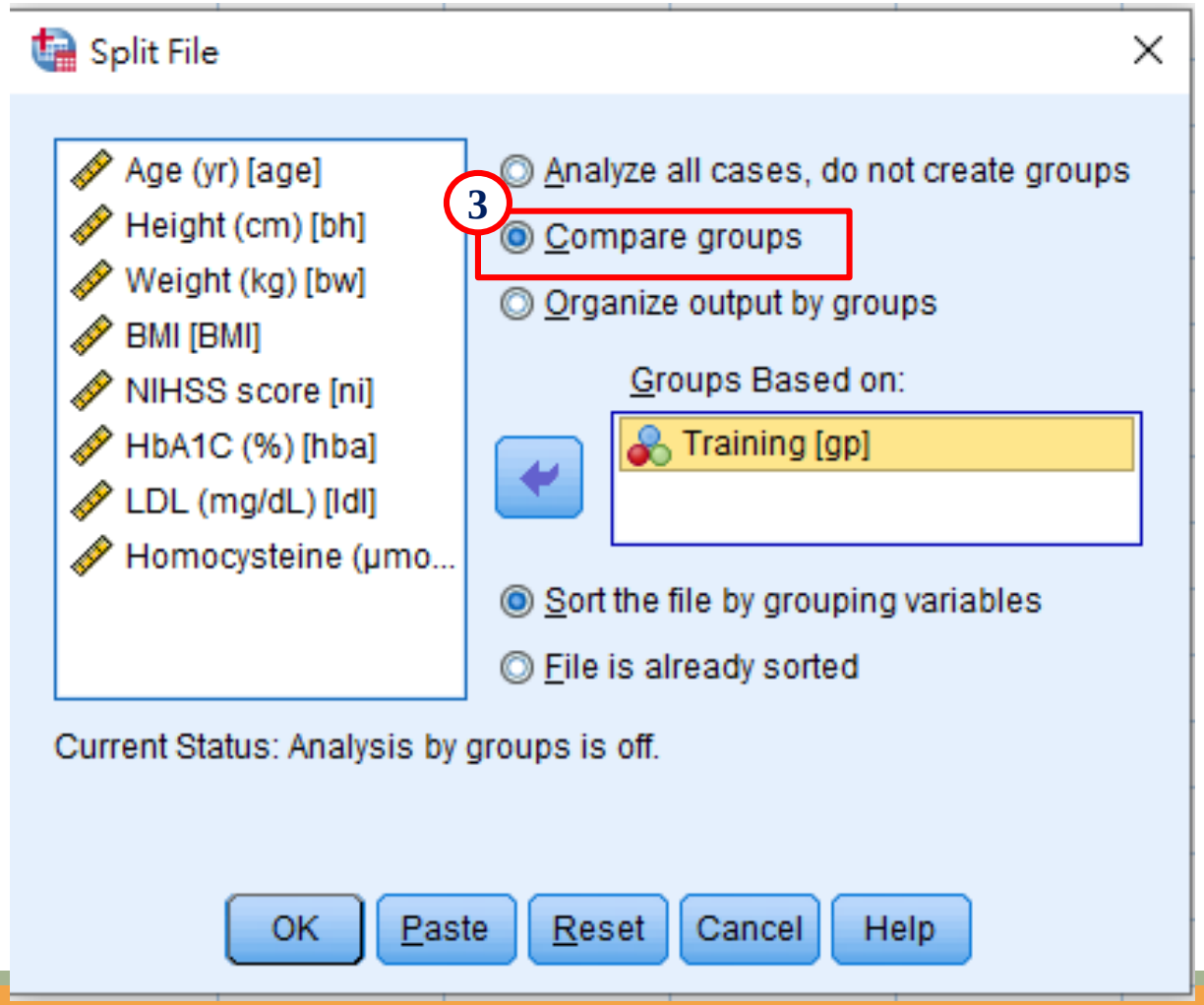
a. Grouping Variable: Training

b. Not corrected for ties.

描述性資料 (SPSS analysis)

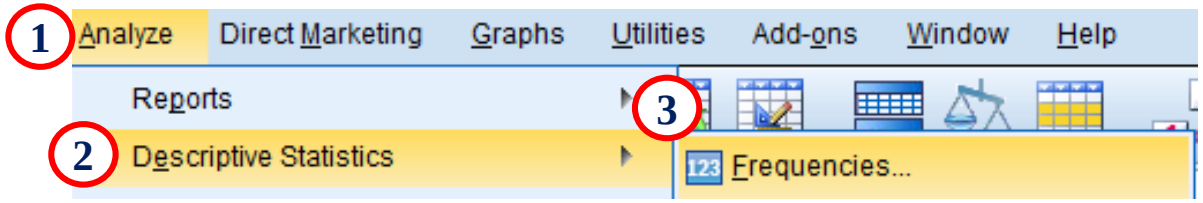
點選步驟1：資料→分割檔案

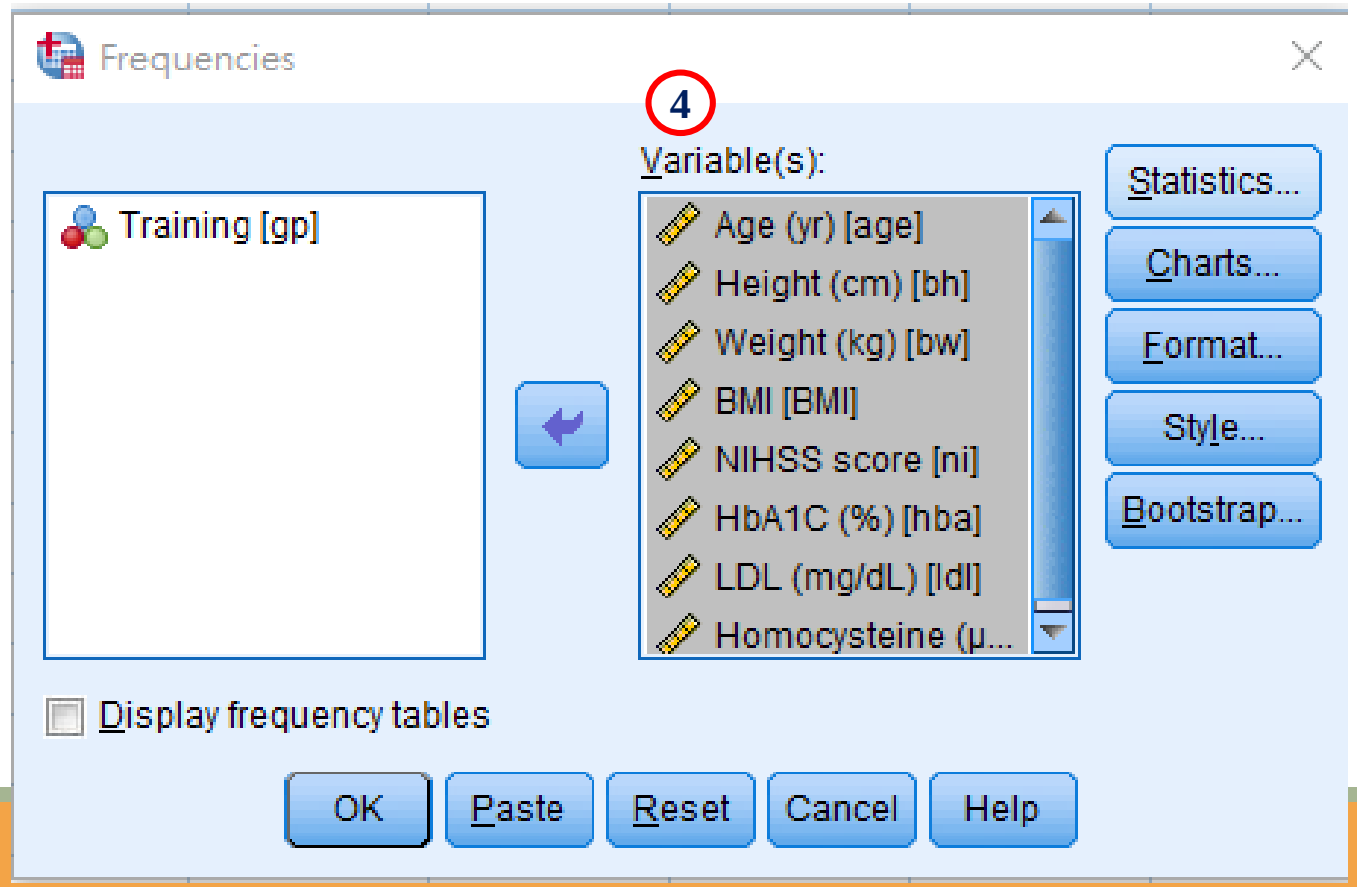


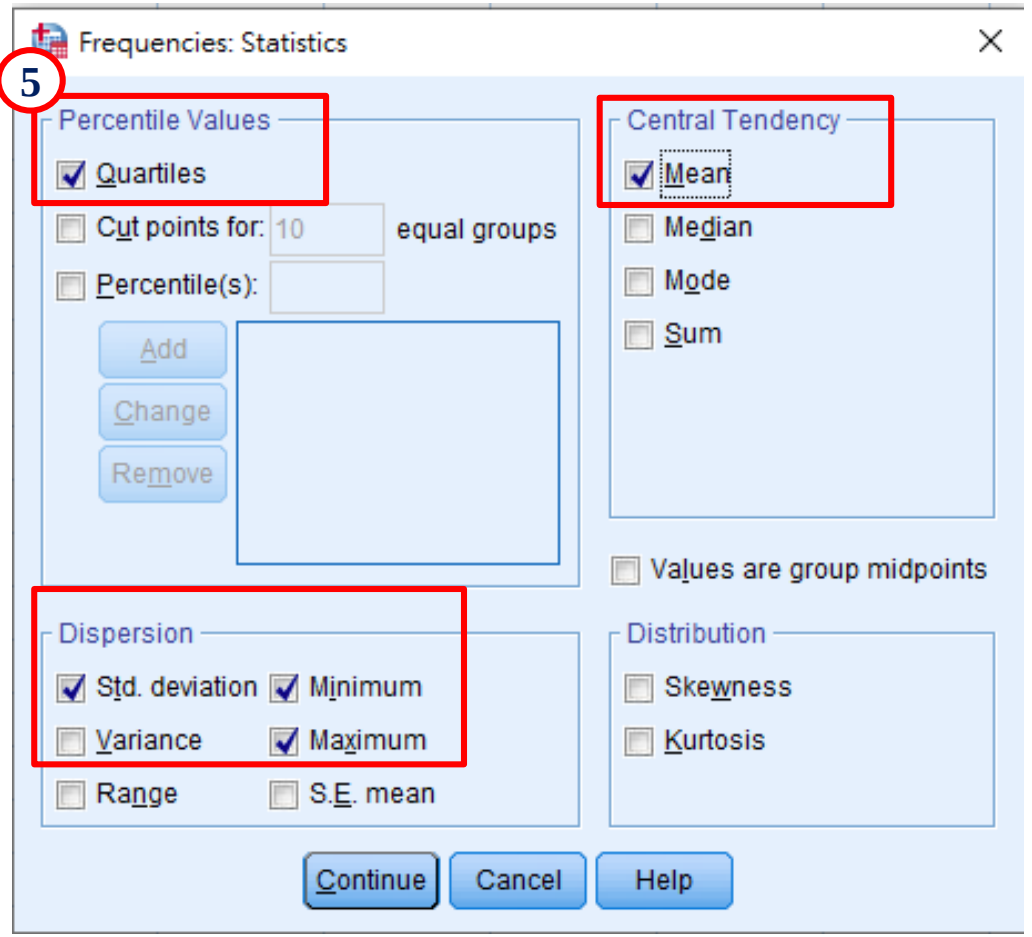


描述性資料 (SPSS analysis)

點選步驟2：分析→描述性資料→次數分配表







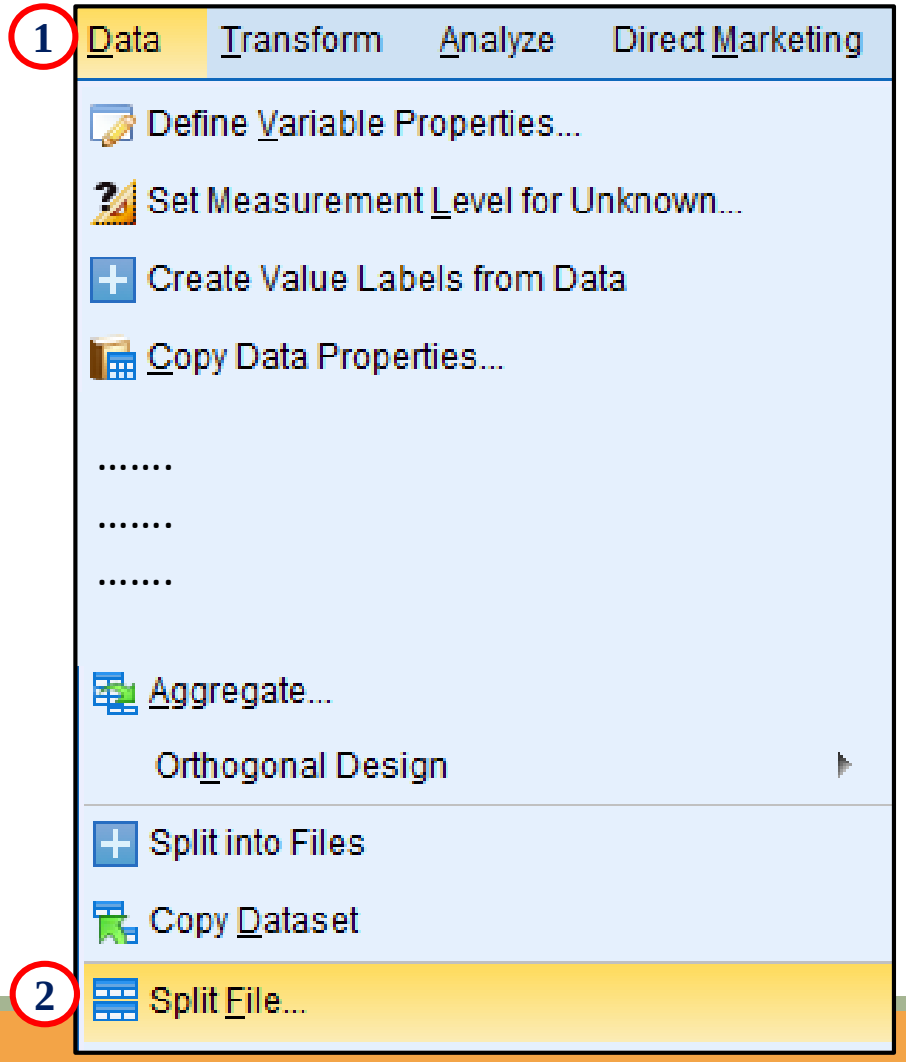
Mann-Whitney U test (SPSS output)

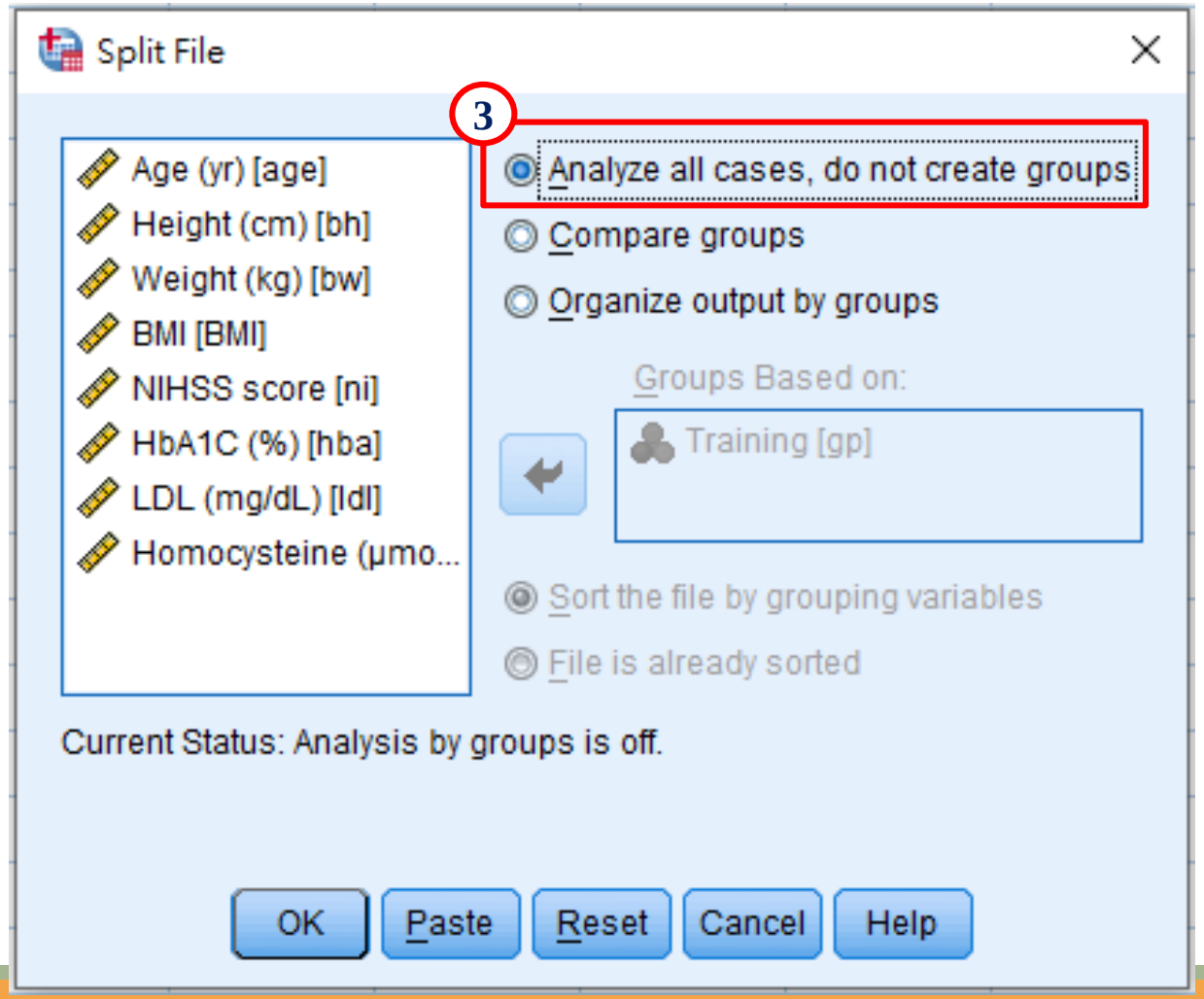
Statistics

Training			Age (yr)	Height (cm)	Weight (kg)	BMI	NIHSS score	HbA1C (%)	LDL (mg/dL)	Homocysteine (μmol/L)
Experimental Group	N	Valid	17	17	17	17	17	17	17	17
		Missing	0	0	0	0	0	0	0	0
	Mean		66.00	161.29	64.56	24.78	8.00	6.48	112.47	10.94
	Std. Deviation		13.620	7.341	9.421	3.585	5.723	1.323	40.693	4.033
	Minimum		43	152	46	19	2	5	57	5
	Maximum		88	175	79	33	19	10	192	22
	Percentiles	25	56.50	155.50	57.65	21.90	3.00	5.60	78.00	7.98
		50	68.00	158.00	69.40	24.60	6.00	6.30	105.00	11.05
		75	74.50	168.25	70.95	27.35	10.00	7.15	135.00	12.52
Control Group	N	Valid	21	21	21	21	21	21	21	21
		Missing	0	0	0	0	0	0	0	0
	Mean		68.57	161.69	69.58	26.58	8.24	6.83	138.62	12.81
	Std. Deviation		8.681	5.954	9.731	3.518	4.206	.839	52.103	5.046
	Minimum		46	148	52	20	3	6	86	7
	Maximum		82	173	84	32	17	9	250	25
	Percentiles	25	63.50	158.10	60.60	23.56	4.50	6.20	100.50	9.50
		50	70.00	160.00	72.00	26.70	7.00	6.60	121.00	11.00
		75	74.50	165.50	77.30	29.48	11.50	7.55	179.00	16.00

描述性資料 (SPSS analysis)

點選步驟3(取消分割)：資料→分割檔案





Two groups-相依樣本

	Paired T-test	Wilcoxon signed-rank test
Test	• Mean	• Median
Data	• 配對樣本 • Normal distribution	• 配對樣本 • Non-normal distribution

Paired T-test

Example

	Before	After	*p* value
	mean±SD	mean±SD	
BMI	22.11±1.66	22.11±1.66	0.993

Paired t test. *$p<0.05$.

$H_0 : \mu_d=0$
$H_1 : \mu_d\neq0$

- Independent variable
  - Before / After (Paired)
- Dependent variable
  - BMI (Continuous)

Paired T-test

SPSS dataset

	ID	BMI0	BMI1
1	001	19	21
2	002	19	22
3	003	20	22
4	004	20	22
5	005	20	22
6	006	20	22
7	007	20	22
8	008	20	22
9	009	21	21
10	010	21	19

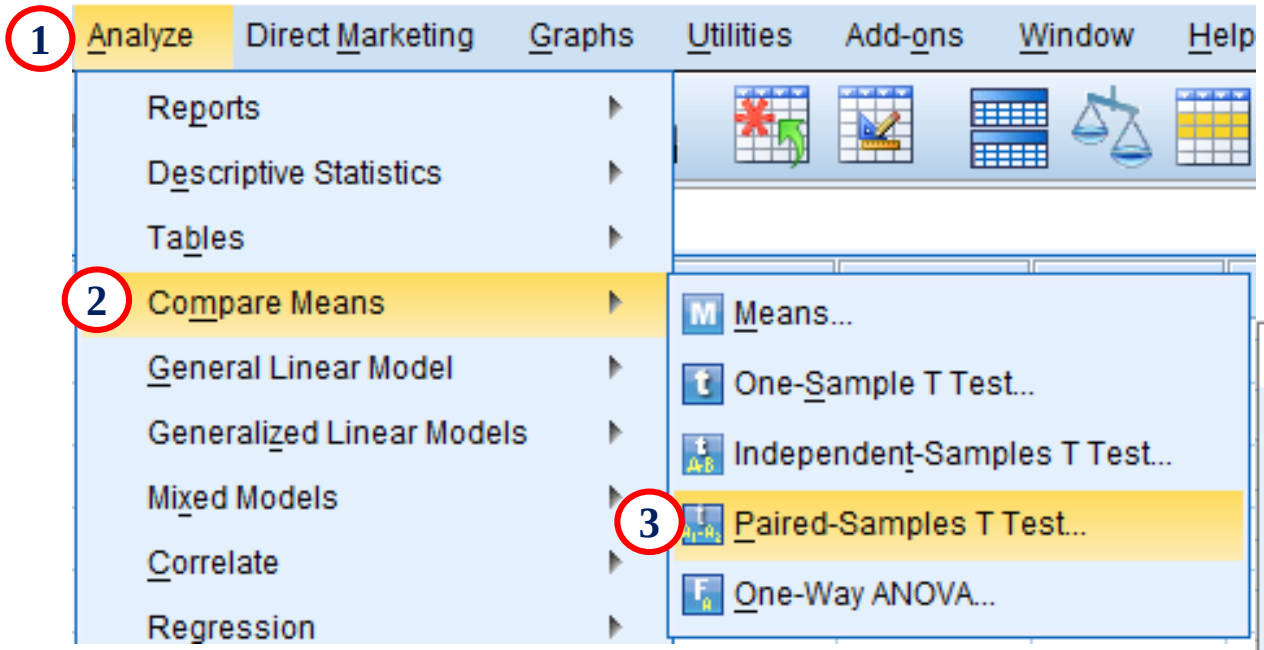
Data View | Variable View

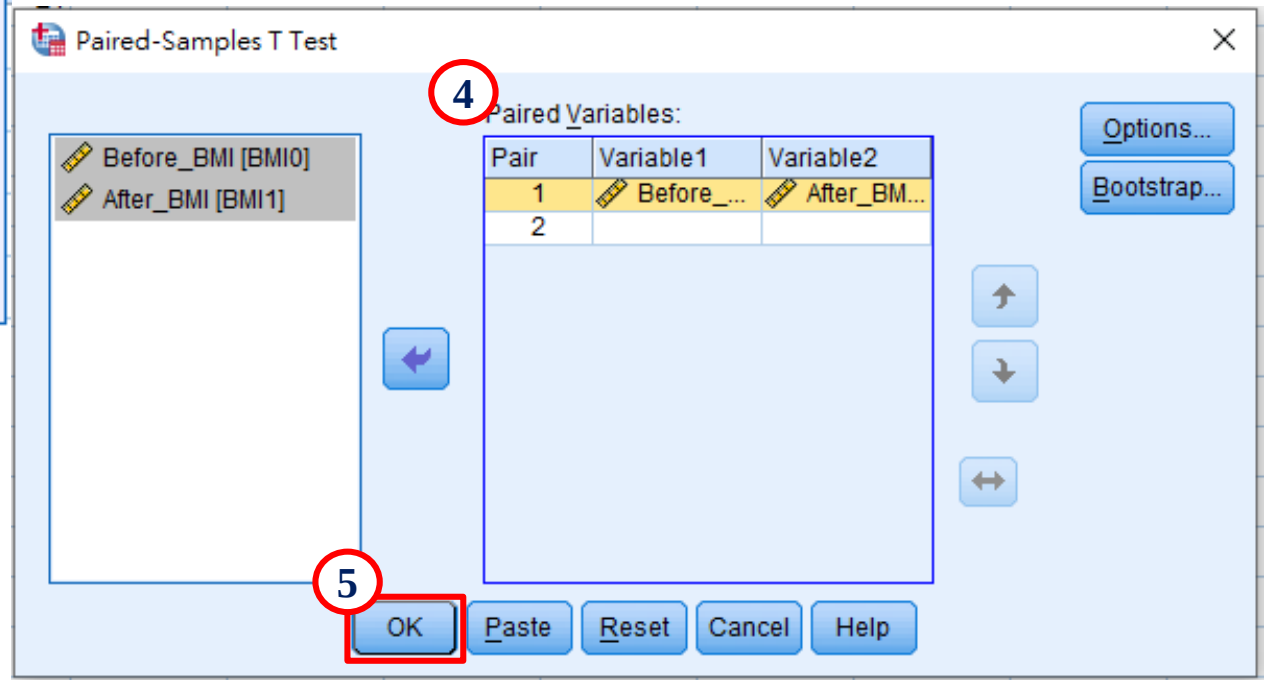
	Name	Type	Width	Decimals	Label
1	ID	String	3	0	
2	BMI0	Numeric	8	0	Before_BMI
3	BMI1	Numeric	8	0	After_BMI
4					
5					
6					
7					
8					
9					

Data View | Variable View

Paired T-test (SPSS analysis)

點選步驟：分析→比較平均數法→成對樣本T檢定





Paired T-test (SPSS output)

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Before_BMI	22.11	42	1.660	.256
	After_BMI	22.11	42	1.660	.256

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Before_BMI - After_BMI	-.002	1.529	.236	-.478	.474	-.008	41	.993

Wilcoxon signed-rank test

Example

- Independent variable
  - Admission / Discharge (Paired)
- Dependent variable
  - MRS, ADLs, IADLs, FOIS, MNA, BBS, FMA-modified sensation, FMA-motor, MMSE… (Continuous)

TABLE 2 | Effect of PAC on functional performance and quality of life in patients with stroke.

	Admission	Discharge	*p*-Value
MRS	4.0 (3.0–4.0)	3.0 (2.0–4.0)	<0.001
ADLs	40.0 (20.0–65.0)	70.0 (50.0–90.0)	<0.001
IADLs	1.0 (0.0–2.0)	2.0 (1.0–4.0)	<0.001
FOIS	6.0 (5.0–7.0)	7.0 (6.0–7.0)	<0.001
MNA	18.3 (14.0–22.5)	21.0 (16.5–24.0)	<0.001
BBS	20.0 (4.0–38.0)	40.5 (19.0–51.0)	<0.001
FMA-modified sensation	39.0 (24.0–44.0)	44.0 (36.0–44.0)	<0.001
FMA-motor	45.0 (15.5–59.0)	57.0 (33.5–62.0)	<0.001
MMSE	20.0 (12.0–26.0)	25.0 (19.0–29.0)	<0.001
CCAT	10.5 (7.3–11.5)	11.0 (8.9–11.9)	<0.001
EQ-5D-3L			
Mobility	2.0 (2.0–2.0)	2.0 (1.0–2.0)	<0.001
Self-care	2.0 (2.0–3.0)	2.0 (1.0–2.0)	<0.001
Usual activities	2.0 (2.0–3.0)	2.0 (2.0–2.0)	<0.001
Pain/discomfort	2.0 (1.0–2.0)	1.0 (1.0–2.0)	<0.001
Anxiety/depression	2.0 (1.0–2.0)	1.0 (1.0–2.0)	<0.001

Continuous data were expressed as median (IQR, interquartile range) and analyzed by the Wilcoxon signed ranks test.

Weng SC, Hsu CY, Shen CC, Huang JA, Chen PL, Lin SY. (2022). Combined Functional Assessment for Predicting Clinical Outcomes in Stroke Patients After Post-acute Care: A Retrospective Multi-Center Cohort in Central Taiwan. *Front Aging Neurosci*, *14*, 834273.

Wilcoxon signed-rank test

SPSS dataset

	NO	P01_0	P01_1	P02_0	P02_1	P03_0	P03_1	P04_0	P04_1	P05_0	P05_1
1	1	4	4	20	65	0	2	6	6	12	12
2	2	4	4	50	85	3	6	7	7	26	26
3	3	4	2	50	85	0	1	7	7	23	23
4	4	3	3	50	80	1	4	7	7	19	19
5	5	4	3	60	90	2	2	7	7	23	23
6	6	4	2	70	95	5	8	.	.	12	12
7	7	4	4	10	30	7	7	7	6	15	15
8	8	4	3	45	65	.	.	5	5	24	27
9	9	3	3	50	70	0	0	7	7	22	22
10	10	3	3	70	90	.	.	7	7	19	19
11	11	2	2	80	100	5	5	7	0	14	13
12	12	4	4	15	30	0	0	2	3	12	12
13	13	4	3	60	75	0	0	6	6	15	15
14	14	4	4	0	10	.	.	2	6	17	18
15	15	4	4	10	20	.	.	2	3	20	20
16	16	4	4	10	20	0	0	5	5	17	17

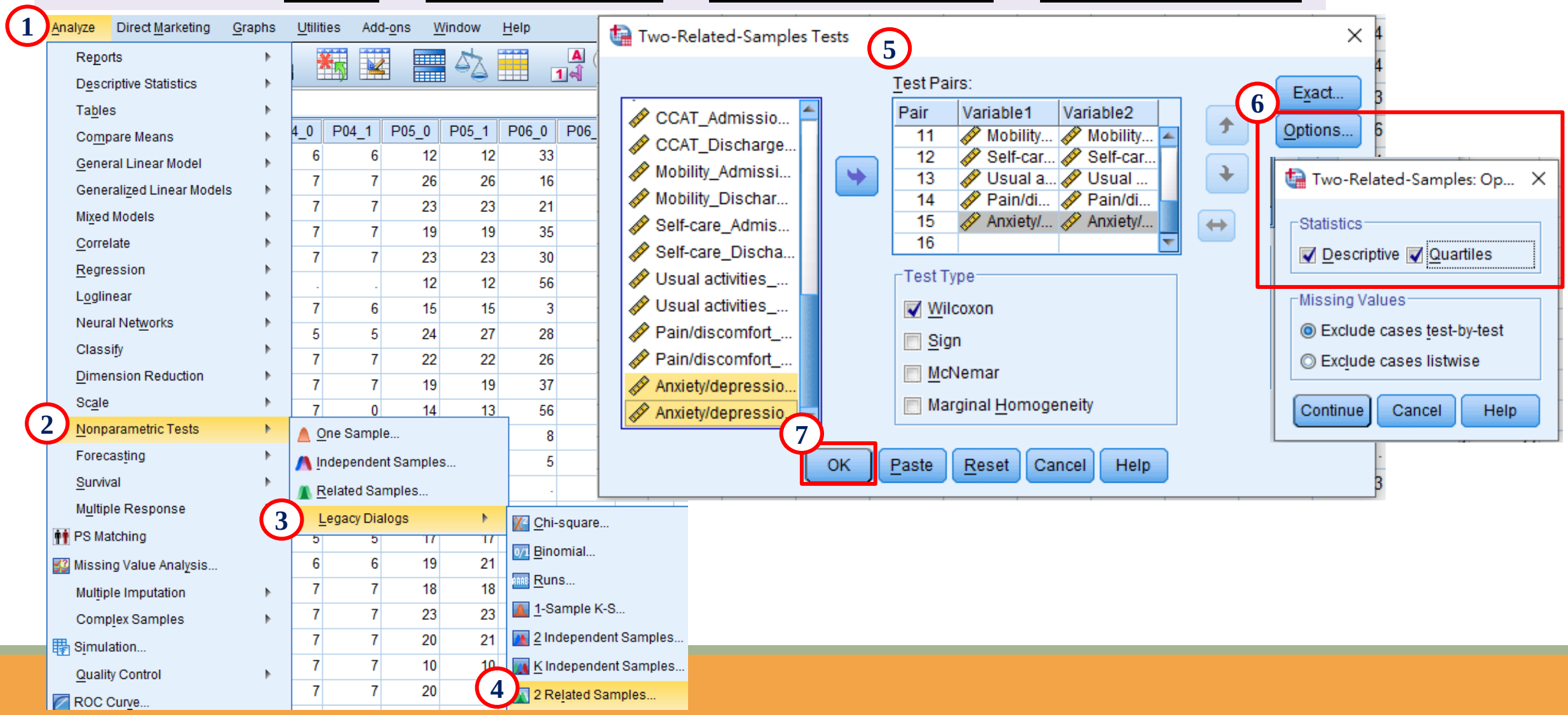
Data View | Variable View

	Name	Type	Width	Decimals	Label
1	NO	Numeric	8	0	
2	P01_0	Numeric	8	0	MRS_Admission
3	P01_1	Numeric	8	0	MRS_Discharge
4	P02_0	Numeric	8	0	ADLs_Admission
5	P02_1	Numeric	8	0	ADLs_Discharge
6	P03_0	Numeric	8	0	IADLs_Admission
7	P03_1	Numeric	8	0	IADLs_Discharge
8	P04_0	Numeric	8	0	FOIS_Admission
9	P04_1	Numeric	8	0	FOIS_Discharge
10	P05_0	Numeric	8	0	MNA_Admission
11	P05_1	Numeric	8	0	MNA_Discharge

Data View | Variable View

Wilcoxon signed-rank test (SPSS analysis)

點選步驟：分析→無母數檢定→歷史對話記錄→2個相關樣本



Wilcoxon signed-rank test (SPSS output)

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles		
						25th	50th (Median)	75th
MRS_Admission	267	3.65	.615	1	5	3.00	4.00	4.00
ADLs_Admission	267	43.18	26.332	0	100	20.00	40.00	65.00
IADLs_Admission	247	1.49	1.585	0	8	.00	1.00	2.00
FOIS_Admission	262	5.40	2.012	0	7	5.00	6.00	7.00
MNA_Admission	266	18.02	5.090	7	27	14.00	18.25	22.50
MRS_Discharge	267	2.94	1.011	0	4	2.00	3.00	4.00
ADLs_Discharge	267	66.20	27.480	0	100	50.00	70.00	90.00
IADLs_Discharge	247	2.46	2.087	0	11	1.00	2.00	4.00
FOIS_Discharge	262	6.11	1.744	0	7	6.00	7.00	7.00
MNA_Discharge	266	19.78	5.263	0	29	16.50	21.00	24.00

Test Statistics[a]

	MRS_Discharge - MRS_Admission	ADLs_Discharge - ADLs_Admission	IADLs_Discharge - IADLs_Admission	FOIS_Discharge - FOIS_Admission	MNA_Discharge - MNA_Admission
Z	-10.908[b]	-12.910[c]	-9.100[c]	-6.915[c]	-9.978[c]
Asymp. Sig. (2-tailed)	.000	.000	.000	.000	.000

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

c. Based on negative ranks.

Three groups-獨立樣本

	One-way ANOVA test	Kruskal-Wallis test
Test	• Mean	• Median
Data	• Normal distribution	• Non-normal distribution

One-way ANOVA test

Example

	Non-tx	Tx-A	Tx-B	*p* value
	mean±SD	mean±SD	mean±SD	
Age	27.00±3.95	26.00±3.95	29.00±3.95	<0.001

ANOVA test. *$p<0.05$.

H_0：$\mu_0=\mu_1=\mu_2(=\mu)$
H_1：μ_i不全等

- Independent variable
  - Non-tx / Tx-A / Tx-B (Three groups)
- Dependent variable
  - Age (Continuous)

One-way ANOVA test

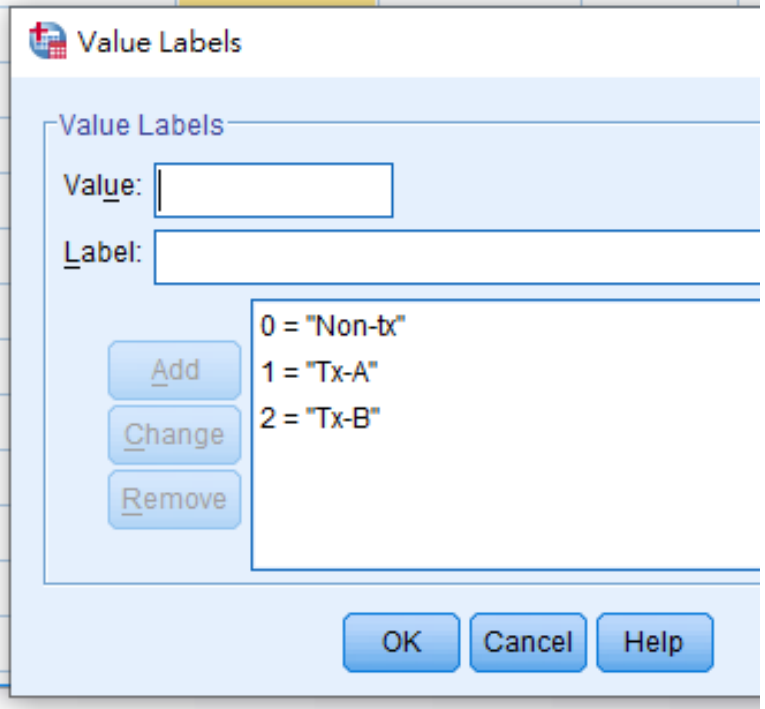
SPSS dataset

	ID	tx	age
1	001	0	19
2	002	0	19
3	003	0	19
4	004	0	20
5	005	0	20
6	006	0	20
7	007	0	20
8	008	0	21
9	009	0	21
10	010	0	21
11	011	0	21
12	012	0	21
13	013	0	22

Data View | Variable View

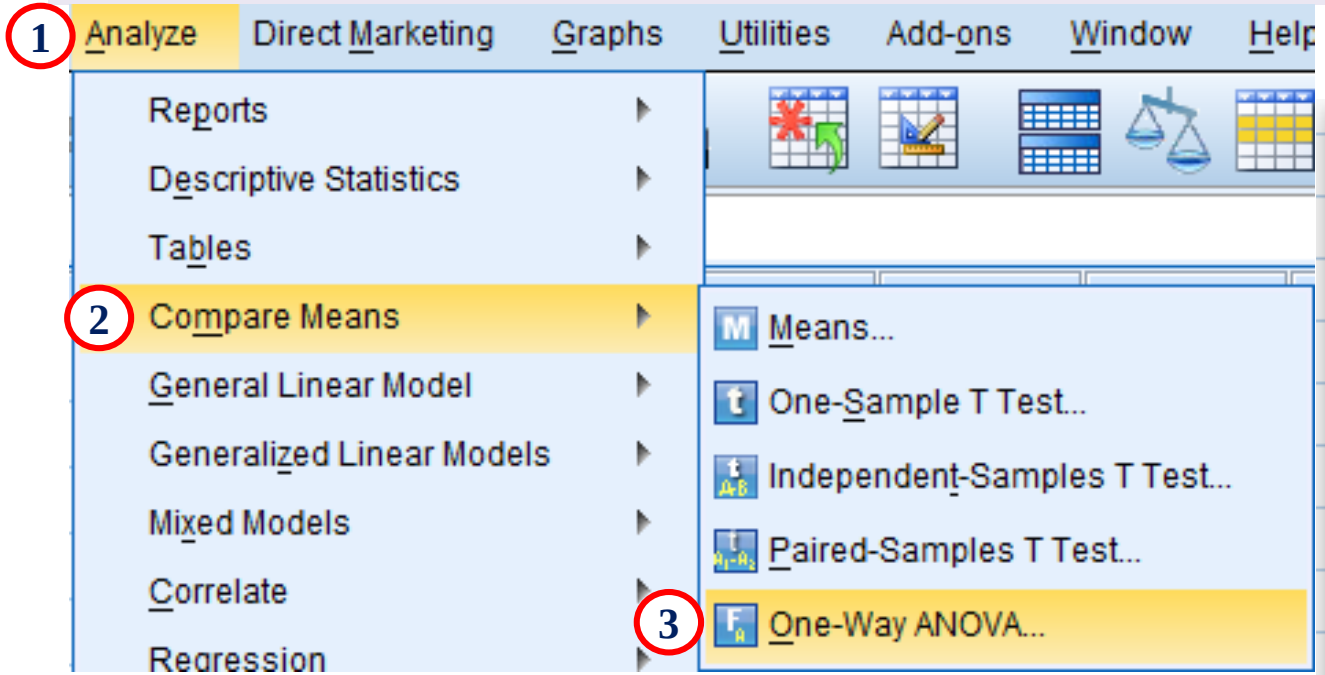
	Name	Type	Width	Decimals	Label	Values	Missing	Columns
1	ID	String	3	0		None	None	5
2	tx	Numeric	11	0	Treatment	{0, Non-tx}...	None	5
3	age	Numeric	11	0	Age			
4								
5								
6								
7								
8								
9								
10								
11								
12								
13								
14								

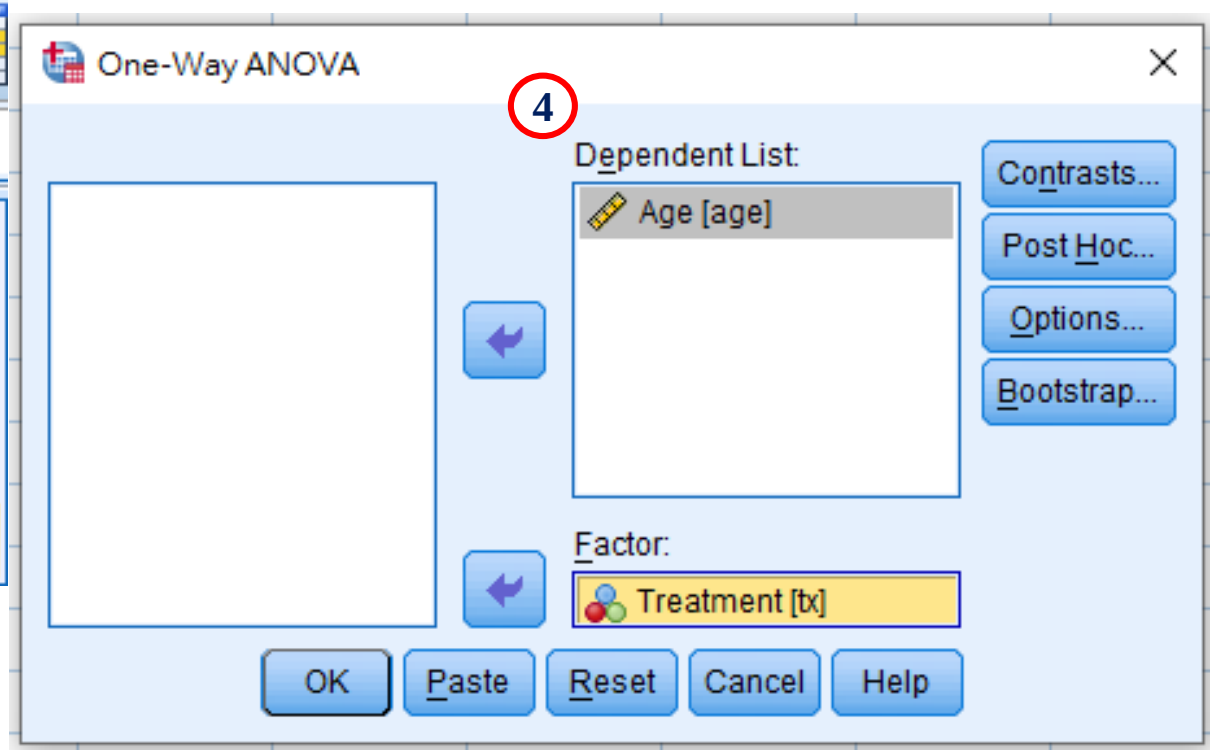
Data View | Variable View



One-way ANOVA test (SPSS analysis)

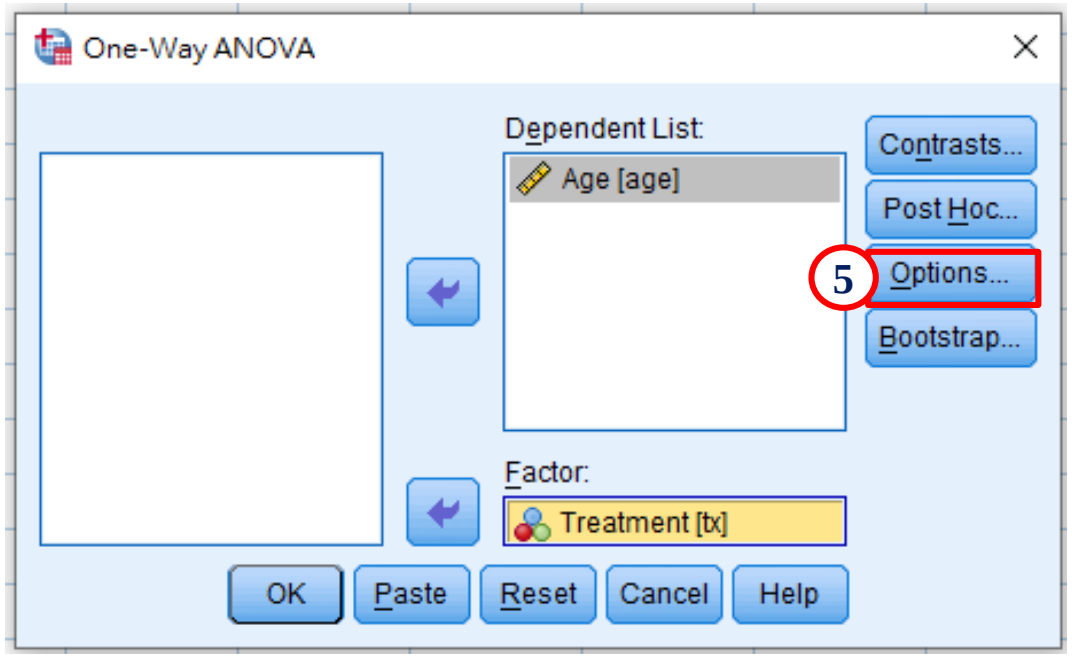
點選步驟：分析→比較平均數法→單因子變異數分析

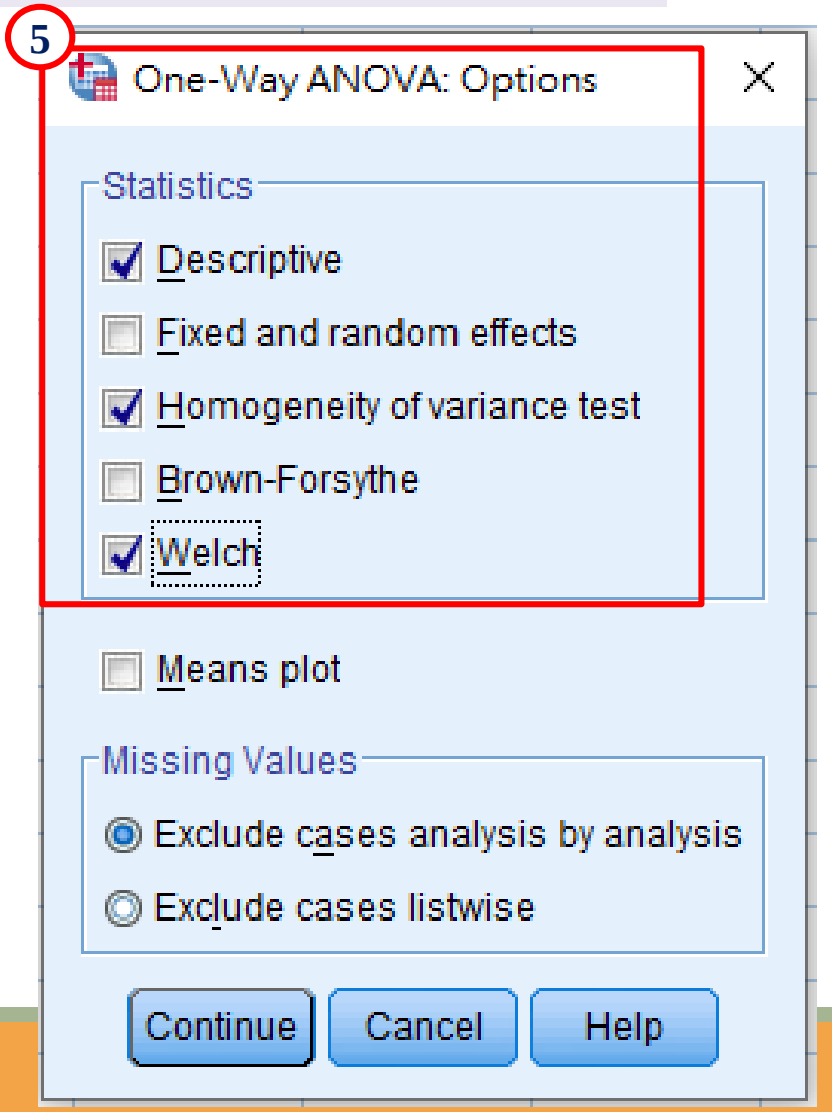




One-way ANOVA test (SPSS analysis)

點選步驟：分析→比較平均數法→單因子變異數分析





One-way ANOVA test (SPSS output)

Age

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Non-tx	129	27.00	3.949	.348	26.31	27.69	19	35
Tx-A	129	26.00	3.949	.348	25.31	26.69	18	34
Tx-B	129	29.00	3.949	.348	28.31	29.69	21	37
Total	387	27.33	4.132	.210	26.92	27.75	18	37

Test of Homogeneity of Variances

Age

Levene Statistic	df1	df2	Sig.
.000	2	384	1.000

$p > 0.05$→ 三組變異數無差異(同質性)

Robust Tests of Equality of Means

Age

	Statistic[a]	df1	df2	Sig.
Welch	19.252	2	256.000	.000

a. Asymptotically F distributed.

ANOVA

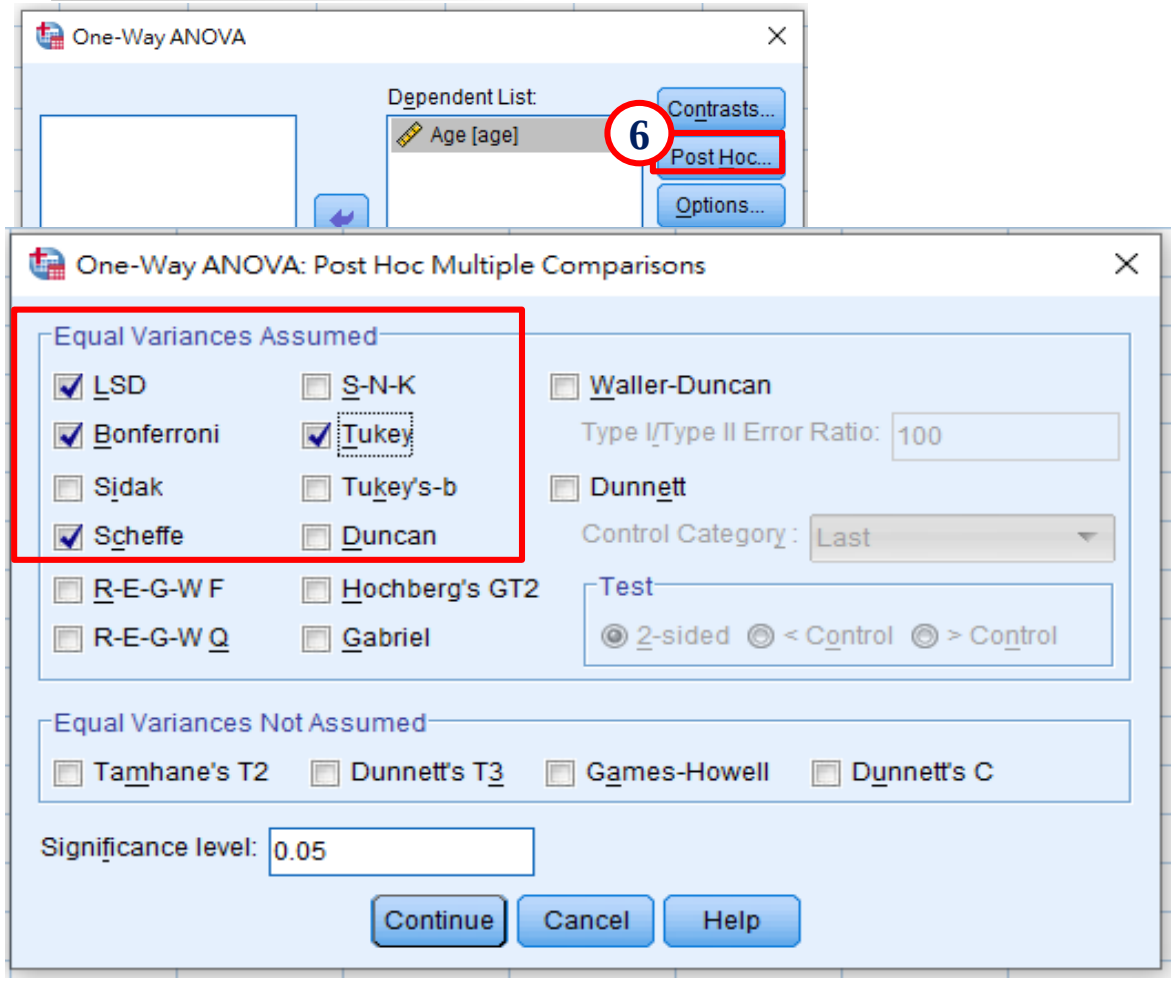
Age

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	602.000	2	301.000	19.303	.000
Within Groups	5988.000	384	15.594		
Total	6590.000	386			

Non-tx、Tx-A、Tx-B的年齡不全等，可得知哪兩組不相等嗎?

如果違反同質性假設，則用Welch結果，取代ANOVA結果

One-way ANOVA test (SPSS output)



Multiple Comparisons

Dependent Variable: Age

	(I) Treatment	(J) Treatment	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	Non-tx	Tx-A	1.000	.492	.106	-.16	2.16
		Tx-B	-2.000*	.492	.000	-3.16	-.84
	Tx-A	Non-tx	-1.000	.492	.106	-2.16	.16
		Tx-B	-3.000*	.492	.000	-4.16	-1.84
	Tx-B	Non-tx	2.000*	.492	.000	.84	3.16
		Tx-A	3.000*	.492	.000	1.84	4.16
Scheffe	Non-tx	Tx-A	1.000	.492	.128	-.21	2.21
		Tx-B	-2.000*	.492	.000	-3.21	-.79
	Tx-A	Non-tx	-1.000	.492	.128	-2.21	.21
		Tx-B	-3.000*	.492	.000	-4.21	-1.79
	Tx-B	Non-tx	2.000*	.492	.000	.79	3.21
		Tx-A	3.000*	.492	.000	1.79	4.21
LSD	Non-tx	Tx-A	1.000*	.492	.043	.03	1.97
		Tx-B	-2.000*	.492	.000	-2.97	-1.03
	Tx-A	Non-tx	-1.000*	.492	.043	-1.97	-.03
		Tx-B	-3.000*	.492	.000	-3.97	-2.03
	Tx-B	Non-tx	2.000*	.492	.000	1.03	2.97
		Tx-A	3.000*	.492	.000	2.03	3.97
Bonferroni	Non-tx	Tx-A	1.000	.492	.128	-.18	2.18
		Tx-B	-2.000*	.492	.000	-3.18	-.82
	Tx-A	Non-tx	-1.000	.492	.128	-2.18	.18
		Tx-B	-3.000*	.492	.000	-4.18	-1.82
	Tx-B	Non-tx	2.000*	.492	.000	.82	3.18
		Tx-A	3.000*	.492	.000	1.82	4.18

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

組數超過兩組 (Post Hoc Tests)

ANOVA

Age

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	602.000	2	301.000	19.303	.000
Within Groups	5988.000	384	15.594		
Total	6590.000	386			

H_0：$\mu_0=\mu_1=\mu_2$ $(=\mu)$

H_1：μ_i不全等

→至少有一組平均值不相等
非所有平均值均不相等

- Pairwise comparison (兩兩相比)
  - 3組(A,B,C)檢定比較
    - 重複檢定3次 (A/B、A/C、B/C)
  - 4組(A,B,C,D)檢定比較
    - 重複檢定6次 (A/B、A/C 、A/D、B/C、B/D、C/D)

誤用T test

組數超過兩組 (Post Hoc Tests)

- 應用ANOVA test，誤用T test
  - T test正確機率：**1-α**
  - 當檢定重複檢定**k**次
    - Accurate：$(1-\alpha)^k$
    - Type I error：$1-(1-\alpha)^k$

		真實情形	
		H_0 is true	H_1 is true
研究結果	H_0 is true	**Accurate (1-α)**	Type II error (β)
	H_1 is true	**Type I error (α)**	Accurate (1-β)

# of popl.	n	Overall Type I error
3	3	0.1426
4	6	0.2649
5	10	0.4013
6	15	0.5367

組數超過兩組 (Post Hoc Tests)

- Post Hoc Tests (常見)
  - Tukey HSD
    - 兩兩單純比較
  - Scheffe method
    - 複雜多重比較 (較保守)
  - Bonferroni method
    - 重覆檢定次數不多 (較嚴格)
  - Least squares difference method (LSD)
    - 不會調整多重比較的誤差率 (α error上升)

	(I) Treatment	(J) Treatment	Sig.
Tukey HSD	Non-tx	Tx-A	.106
		Tx-B	.000
	Tx-A		
		Tx-B	.000
Scheffe	Non-tx	Tx-A	.128
		Tx-B	.000
	Tx-A		
		Tx-B	.000
LSD	Non-tx	Tx-A	.043
		Tx-B	.000
	Tx-A		
		Tx-B	.000
Bonferroni	Non-tx	Tx-A	.128
		Tx-B	.000
	Tx-A		
		Tx-B	.000

Kruskal-Wallis test

Example

- Independent variable
  - Normal / Osteopenia / Osteoporosis (Three groups)
- Dependent variable
  - Charlson Comorbidity Index, Functional capacity, Handgrip strength, 6-meter walking speed, Timed Up and Go test (sec)…(Continuous)

Table 3 Comparison between normal, osteopenia, and osteoporosis

Comprehensive geriatric assessment	Normal		Osteopenia		Osteoporosis		*p* value
	(*n* = 23)		(*n* = 124)		(*n* = 190)		
Charlson Comorbidity Index, Median (IQR)	3.0	(2.0–4.0)	2.0	(1.0–4.0)	2.0	(1.0–3.0)	0.064
Functional capacity							
Barthel index of activities of daily living, Median (IQR)	100.0	(90.0-100.0)	95.0	(90.0-100.0)	95.0	(85.0-100.0)	0.068
Lawton instrumental activities of daily living, Median (IQR)	7.0	(5.0–8.0)	7.0	(4.0–8.0)	6.0	(3.0–8.0)	0.164
Physical performance							
Handgrip strength, Median (IQR)	20.3	(15.0-32.1)	21.0	(16.8–27.5)	16.9	(13.5–22.1)	< 0.001*
Low muscle strength, n (%)	11	(47.8%)	58	(46.8%)	129	(67.9%)	0.001*
6-meter walking speed, Median (IQR)	1.0	(0.9–1.2)	1.0	(0.7–1.2)	0.9	(0.7–1.1)	0.182
Low physical performance, n (%)	8	(40.0%)	57	(55.3%)	105	(70.0%)	0.006*
Timed Up and Go test (sec), Median (IQR)	9.9	(8.6–13.2)	11.8	(9.2–15.0)	12.7	(10.5–16.2)	0.009*
Mini-mental state examination, Median (IQR)	28.0	(24.0–29.0)	27.0	(23.0–29.0)	25.0	(19.0–29.0)	0.005*
Cognitive impairment, n (%)	5	(21.7%)	33	(26.6%)	58	(30.5%)	0.572
Five-item geriatric depression scale, Median (IQR)	0.0	(0.0–1.0)	1.0	(0.0–1.0)	1.0	(0.0–2.0)	0.552
Depressive symptoms, n (%)	4	(17.4%)	27	(22.0%)	49	(25.8%)	0.558
Mini-nutritional assessment-short form, Median (IQR)	14.0	(12.0–14.0)	13.0	(12.0–14.0)	13.0	(11.0–14.0)	0.012*
Mini-nutritional assessment-short form, n (%)							0.399
Normal	19	(82.6%)	91	(75.8%)	124	(68.9%)	
At risk	4	(17.4%)	25	(20.8%)	45	(25.0%)	
Malnutrition	0	(0.0%)	4	(3.4%)	11	(6.1%)	
One-year mortality, n (%)	0	(0.0%)	1	(0.8%)	4	(2.1%)	0.756

*$p < 0.05$

Chou YY, Lin CF, Lee YS, Weng SC, Kuo FH, Hsu CY, Lin SY. (2023). The associations of osteoporosis and possible sarcopenia with disability, nutrition, and cognition in community-dwelling older adults. *BMC Geriatr*, *23* (1), 730. https://doi.org10.1186/s12877-023-04431-x.

Kruskal-Wallis test

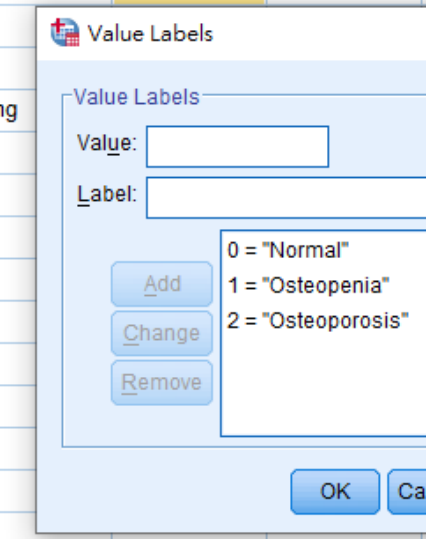
SPSS dataset

	PERSON_ID	ost	cci	adl	iadl	hgs	ws
1	2504890	2	1	100	8	18	1
2	538559	2	0	100	7	18	1
3	1029080	2	0	100	6	19	1
4	1432098	2	0	100	8	19	2
5	2716208	2	1	100	8	19	1
6	1446402	1	1	100	7	19	1
7	970130	1	2	100	7	19	1
8	1895979	1	1	100	8	19	1
9	2709037	1	7	100	8	19	1
10	949198	2	0	100	8	20	1
11	1413812	1	1	100	8	20	1
12	278690	0	2	100	8	20	1

Data View | Variable View

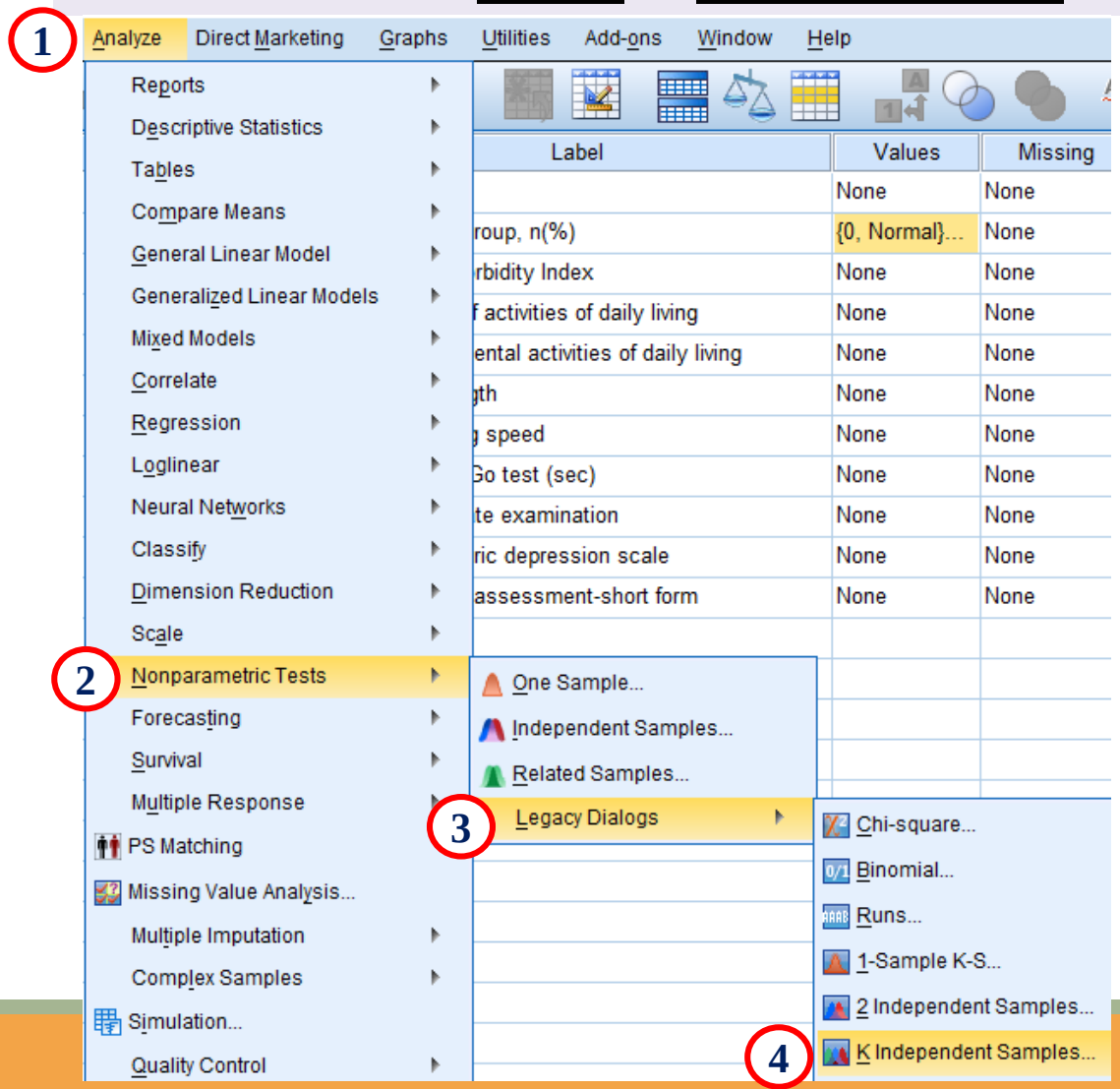
	Name	Type	Width	Decimals	Label	Values	Missing
1	PERSON_ID	Numeric	8	0		None	None
2	ost	Numeric	8	0	Osteoporosis group, n(%)	{0, Normal}...	None
3	cci	Numeric	8	0	Charlson Comorbidity Index		
4	adl	Numeric	8	0	Barthel index of activities of daily living		
5	iadl	Numeric	8	0	Lawton instrumental activities of daily living		
6	hgs	Numeric	8	0	Handgrip strength		
7	ws	Numeric	8	0	6-meter walking speed		
8	tug	Numeric	8	0	Timed Up and Go test (sec)		
9	mms	Numeric	8	0	Mini-mental state examination		
10	gds	Numeric	8	0	Five-item geriatric depression scale		
11	mna	Numeric	8	0	Mini-nutritional assessment-short form		
12							
13							
14							
15							
16							
17							
18							

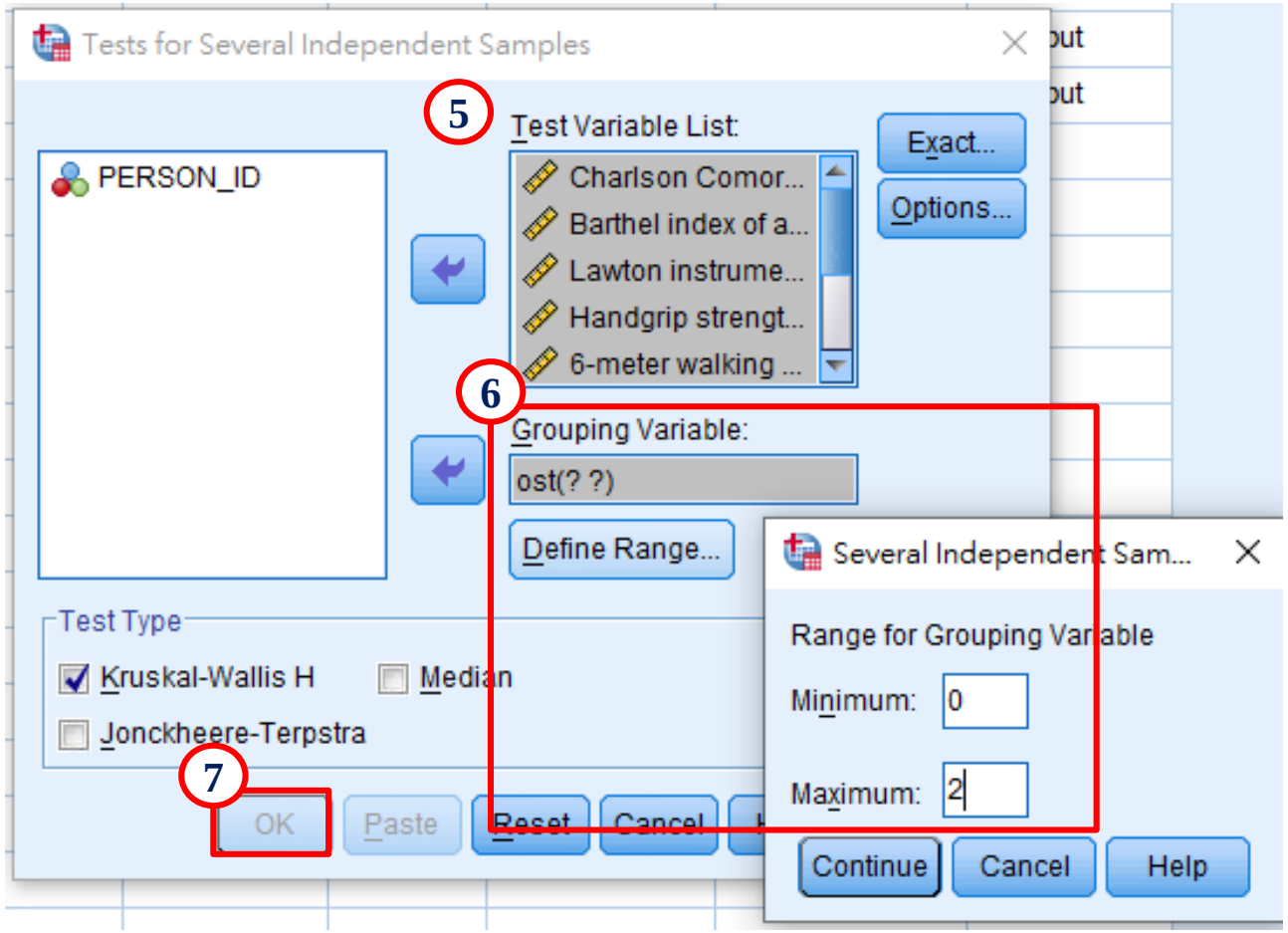
Data View | Variable View



Kruskal-Wallis test (SPSS analysis)

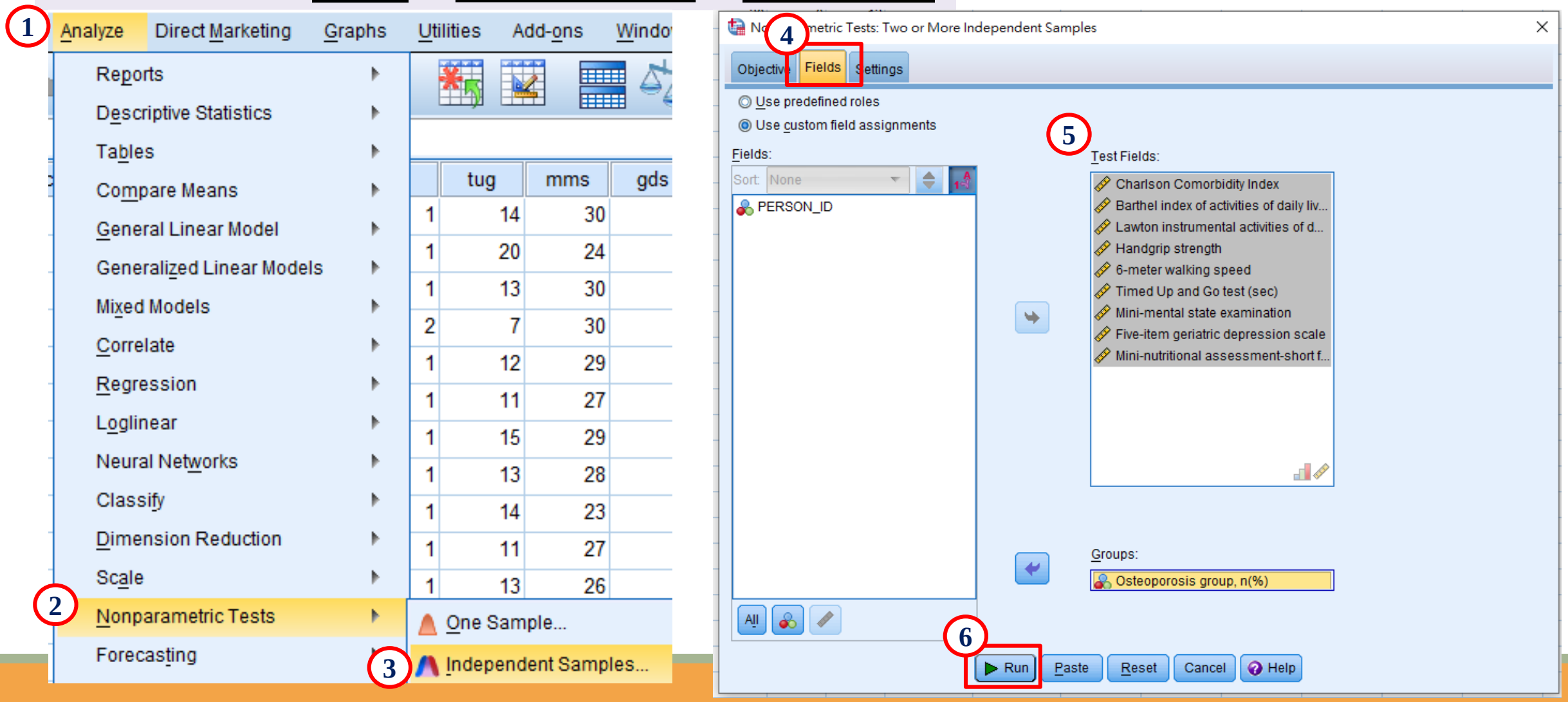
點選步驟1：分析→無母數檢定→歷史對話記錄→K個獨立樣本





Kruskal-Wallis test (SPSS analysis)

點選步驟2：分析→無母數檢定→獨立樣本



Kruskal-Wallis test (SPSS output)

點選步驟1	Charlson Comorbidity Index	Barthel index of activities of daily living	Lawton instrumental activities of daily living	Handgrip strength	6-meter walking speed	Timed Up and Go test (sec)	Mini-mental state examination	Five-item geriatric depression scale	Mini-nutritional assessment-short form
Chi-Square	5.504	5.388	3.613	24.772	3.402	9.375	10.552	1.188	8.889
df	2	2	2	2	2	2	2	2	2
Asymp. Sig.	.064	.068	.164	.000	.182	.009	.005	.552	.012

a. Kruskal Wallis Test

點選步驟2

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Charlson Comorbidity Index is the same across categories of Osteoporosis group, n(%).	Independent-Samples Kruskal-Wallis Test	.064	Retain the null hypothesis.
2	The distribution of Barthel index of activities of daily living is the same across categories of Osteoporosis group, n(%).	Independent-Samples Kruskal-Wallis Test	.068	Retain the null hypothesis.
3	The distribution of Lawton instrumental activities of daily living is the same across categories of Osteoporosis group, n(%).	Independent-Samples Kruskal-Wallis Test	.164	Retain the null hypothesis.
4	The distribution of Handgrip strength is the same across categories of Osteoporosis group, n(%).	Independent-Samples Kruskal-Wallis Test	.000	Reject the null hypothesis.
5	The distribution of 6-meter walking speed is the same across categories of Osteoporosis group, n (%).	Independent-Samples Kruskal-Wallis Test	.182	Retain the null hypothesis.
6	The distribution of Timed Up and Go test (sec) is the same across categories of Osteoporosis group, n (%).	Independent-Samples Kruskal-Wallis Test	.009	Reject the null hypothesis.
7	The distribution of Mini-mental state examination is the same across categories of Osteoporosis group, n (%).	Independent-Samples Kruskal-Wallis Test	.005	Reject the null hypothesis.
8	The distribution of Five-item geriatric depression scale is the same across categories of Osteoporosis group, n(%).	Independent-Samples Kruskal-Wallis Test	.552	Retain the null hypothesis.
9	The distribution of Mini-nutritional assessment-short form is the same across categories of Osteoporosis group, n(%).	Independent-Samples Kruskal-Wallis Test	.012	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Table 3 Comparison between normal, osteopenia, and osteoporosis

Comprehensive geriatric assessment	Normal (*n* = 23)		Osteopenia (*n* = 124)		Osteoporosis (*n* = 190)		*p* value
Charlson Comorbidity Index, Median (IQR)	3.0	(2.0–4.0)	2.0	(1.0–4.0)	2.0	(1.0–3.0)	0.064
Functional capacity							
Barthel index of activities of daily living, Median (IQR)	100.0	(90.0-100.0)	95.0	(90.0-100.0)	95.0	(85.0-100.0)	0.068
Lawton instrumental activities of daily living, Median (IQR)	7.0	(5.0–8.0)	7.0	(4.0–8.0)	6.0	(3.0–8.0)	0.164
Physical performance							
Handgrip strength, Median (IQR)	20.3	(15.0-32.1)	21.0	(16.8–27.5)	16.9	(13.5–22.1)	< 0.001*
6-meter walking speed, Median (IQR)	1.0	(0.9–1.2)	1.0	(0.7–1.2)	0.9	(0.7–1.1)	0.182
Timed Up and Go test (sec), Median (IQR)	9.9	(8.6–13.2)	11.8	(9.2–15.0)	12.7	(10.5–16.2)	0.009*
Mini-mental state examination, Median (IQR)	28.0	(24.0–29.0)	27.0	(23.0–29.0)	25.0	(19.0–29.0)	0.005*
Five-item geriatric depression scale, Median (IQR)	0.0	(0.0–1.0)	1.0	(0.0–1.0)	1.0	(0.0–2.0)	0.552
Mini-nutritional assessment-short form, Median (IQR)	14.0	(12.0–14.0)	13.0	(12.0–14.0)	13.0	(11.0–14.0)	0.012*

Kruskal-Wallis test (SPSS output)

點選
步驟2

Hypothesis Test Summary

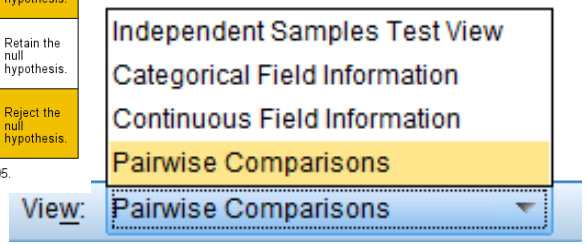
	Hypothesis	Test	Sig.	Decision
1	n of Charlson Comorbidity Index is the same across categories of Osteoporosis group, n(%).	Independent-Samples Kruskal-Wallis Test	.064	Retain the null hypothesis.
2	The distribution of Barthel index of activities of daily living is the same across categories of Osteoporosis group, n(%).	Independent-Samples Kruskal-Wallis Test	.068	Retain the null hypothesis.
3	The distribution of Lawton instrumental activities of daily living is the same across categories of Osteoporosis group, n(%).	Independent-Samples Kruskal-Wallis Test	.164	Retain the null hypothesis.
4	The distribution of Handgrip strength is the same across categories of Osteoporosis group, n(%).	Independent-Samples Kruskal-Wallis Test	.000	Reject the null hypothesis.
5	The distribution of 6-meter walking speed is the same across categories of Osteoporosis group, n (%).	Independent-Samples Kruskal-Wallis Test	.182	Retain the null hypothesis.
6	The distribution of Timed Up and Go test (sec) is the same across categories of Osteoporosis group, n (%).	Independent-Samples Kruskal-Wallis Test	.009	Reject the null hypothesis.
7	The distribution of Mini-mental state examination is the same across categories of Osteoporosis group, n (%).	Independent-Samples Kruskal-Wallis Test	.005	Reject the null hypothesis.
8	The distribution of Five-item geriatric depression scale is the same across categories of Osteoporosis group, n(%).	Independent-Samples Kruskal-Wallis Test	.552	Retain the null hypothesis.
9	The distribution of Mini-nutritional assessment-short form is the same across categories of Osteoporosis group, n(%).	Independent-Samples Kruskal-Wallis Test	.012	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Each node shows the sample average rank of Osteoporosis group, n(%).

Sample1-Sample2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj.Sig.
Osteoporosis-Osteopenia	52.538	11.248	4.671	.000	.000
Osteoporosis-Normal	56.911	21.509	2.646	.008	.024
Osteopenia-Normal	4.373	22.119	.198	.843	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.
Asymptotic significances (2-sided tests) are dis|



The post hoc analysis revealed that these differences persisted between the possible osteosarcopenia group and normal group. Moreover, daily activities, walking speed, cognitive score and nutrition score were found to be significantly lower in the possible osteosarcopenia group in comparison with the osteoporosis group. However, there were no significant differences between the possible osteosarcopenia group and the possible sarcopenia only (data not shown).

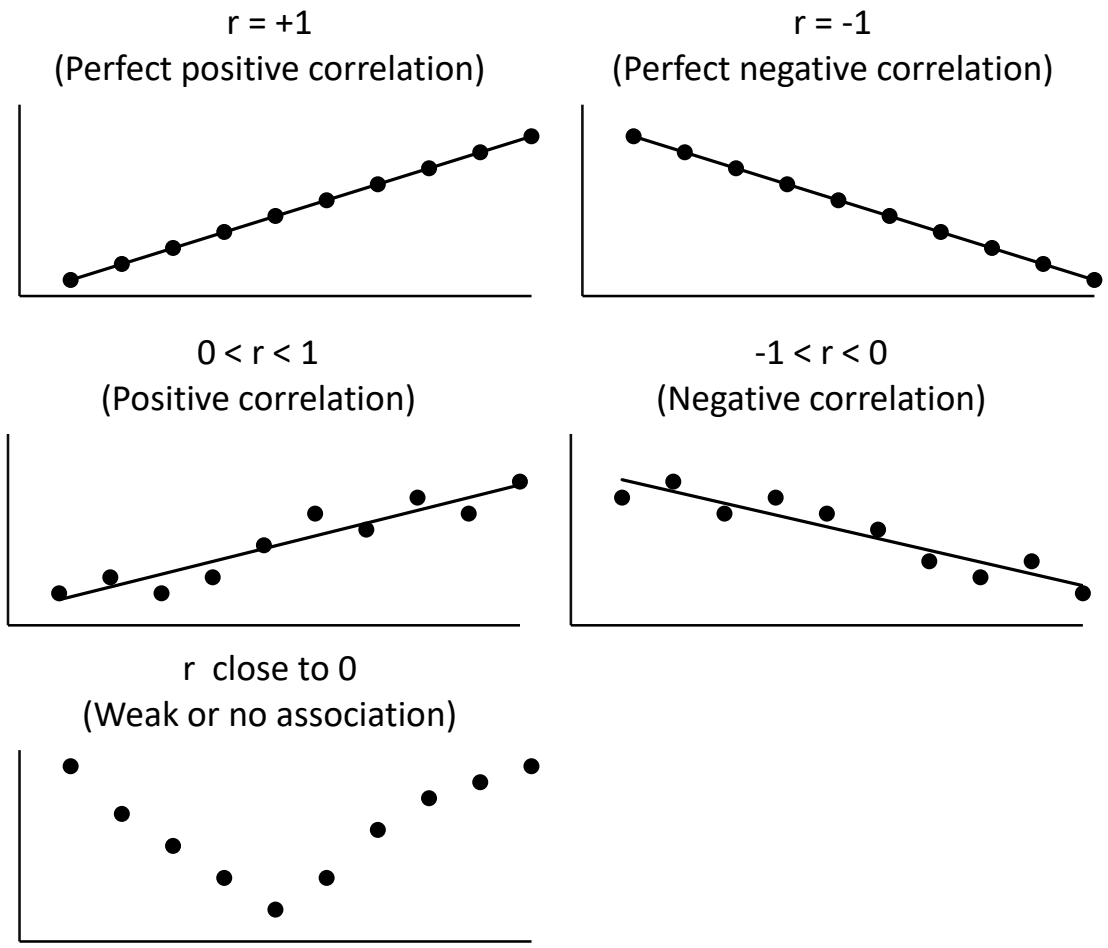
Correlation

	Pearson correlation	Spearman rank correlation
Distribution	• Normal distribution	• Skewed distribution
Types of Data	• Continuous	• Ordinal • Continuous

相關

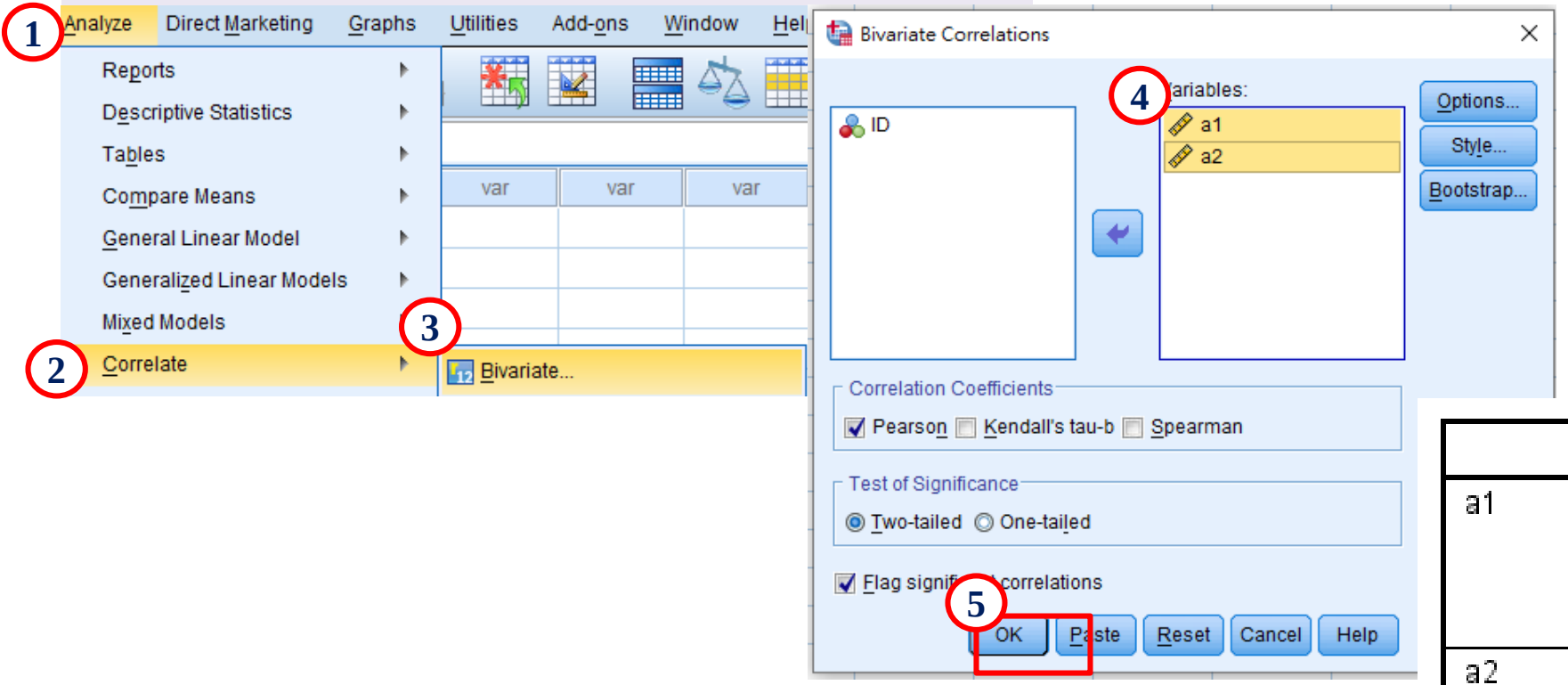
- 兩個連續變項之間的關係
- 相關強度
  - Range (-1 to 1)

相關係數	相關	
1	完全	
0.70 - 0.99	高度	
0.50 - 0.69	中度	相關
0.25 - 0.49	低度	
0.00 - 0.24	無	



Pearson correlation (SPSS analysis)

點選步驟：分析→相關→雙變數

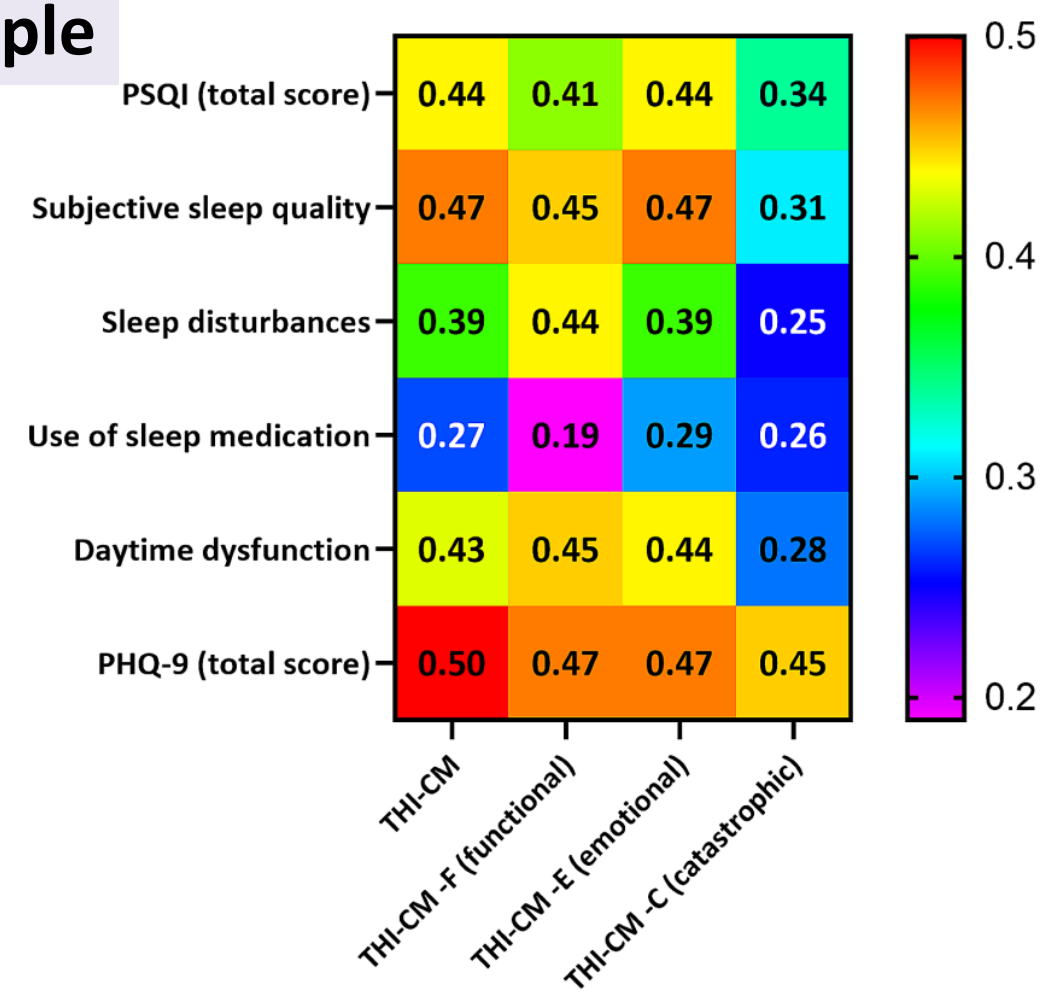


		a1	a2
a1	Pearson Correlation	1	.502**
	Sig. (2-tailed)		.000
	N	49	49
a2	Pearson Correlation	.502**	1
	Sig. (2-tailed)	.000	
	N	49	49

**. Correlation is significant at the 0.01 level (2-tailed).

Spearman rank correlation

Example



Correlations between the THI-CM scores and different variables (n=65)

	THI-CM		THI-CM -F (functional)		THI-CM -E (emotional)		THI-CM -C (catastrophic)	
	r_s	*p* value	r_s	*p* value	r_s	*p* value	r_s	*p* value
PSQI (total score)	0.44	<0.001**	0.41	0.001**	0.44	<0.001**	0.34	0.005**
Subjective sleep quality	0.47	<0.001**	0.45	<0.001**	0.47	<0.001**	0.31	0.011*
Sleep disturbances	0.39	0.001**	0.44	<0.001**	0.39	0.001**	0.25	0.049*
Use of sleep medication	0.27	0.032*	0.19	0.130	0.29	0.017*	0.26	0.039*
Daytime dysfunction	0.43	<0.001**	0.45	<0.001**	0.44	<0.001**	0.28	0.023*
PHQ-9 (total score)	0.50	<0.001**	0.47	<0.001**	0.47	<0.001**	0.45	<0.001**

Spearman's rank correlation was used to obtain the data. *p<0.05, **p<0.01.

Fig. 1 the correlations between the Chinese-Mandarin version of the Tinnitus Handicap Inventory score (THI-CM), Pittsburgh Sleep Quality Index score (PSQI) and Patient Health Questionnaire-9 score (PHQ-9)

Chang TG, Yao YT, Hsu CY, Yen TT. (2024). Exploring the interplay of depression, sleep quality, and hearing in tinnitus-related handicap: insights from polysomnography and pure-tone audiometry. *BMC Psychiatry*, *24* (1), 459. https://doi.org10.1186/s12888-024-05912-y.

Spearman rank correlation

SPSS dataset

	NO	tin	t01	t02	t03	psq
1	65	14	4	0	10	11
2	47	14	6	4	4	3
3	43	12	2	2	8	4
4	34	12	4	2	6	4
5	12	12	6	0	6	9
6	4	10	4	2	4	6
7	30	10	6	0	4	4
8	31	10	4	0	6	8
9	13	8	4	0	4	5
10	42	6	2	0	4	1
11	24	4	0	0	4	13

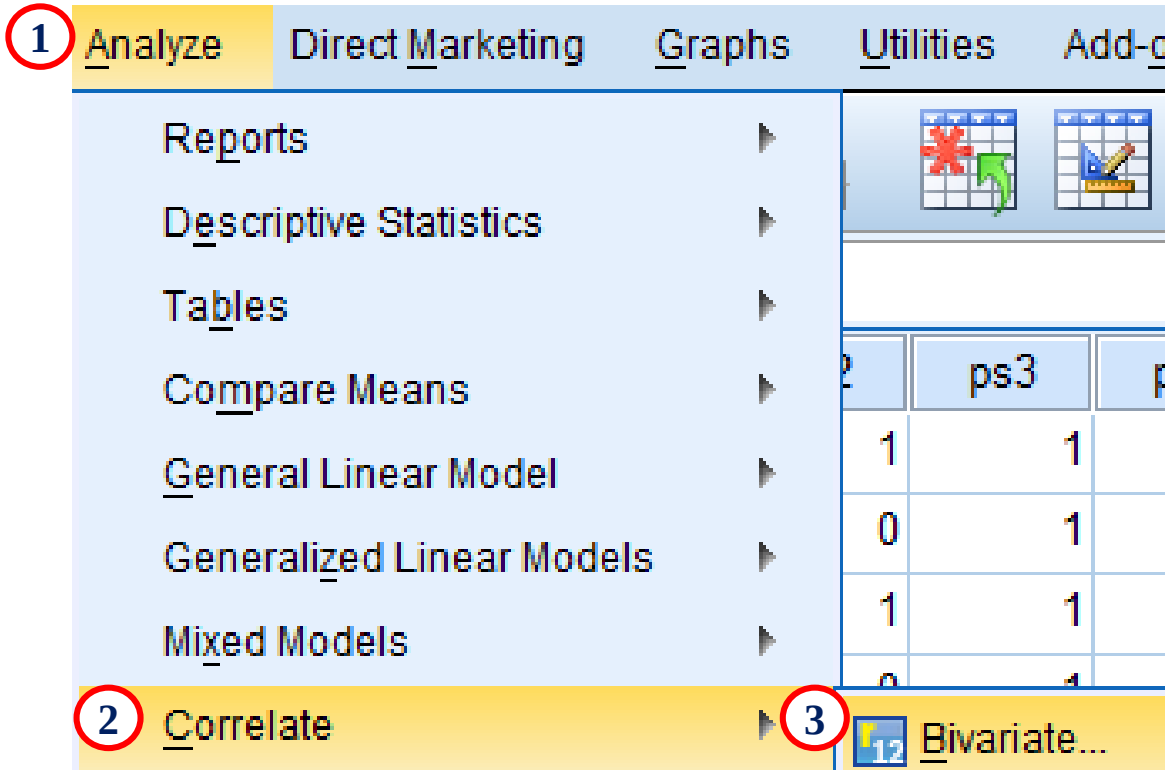
Data View | Variable View

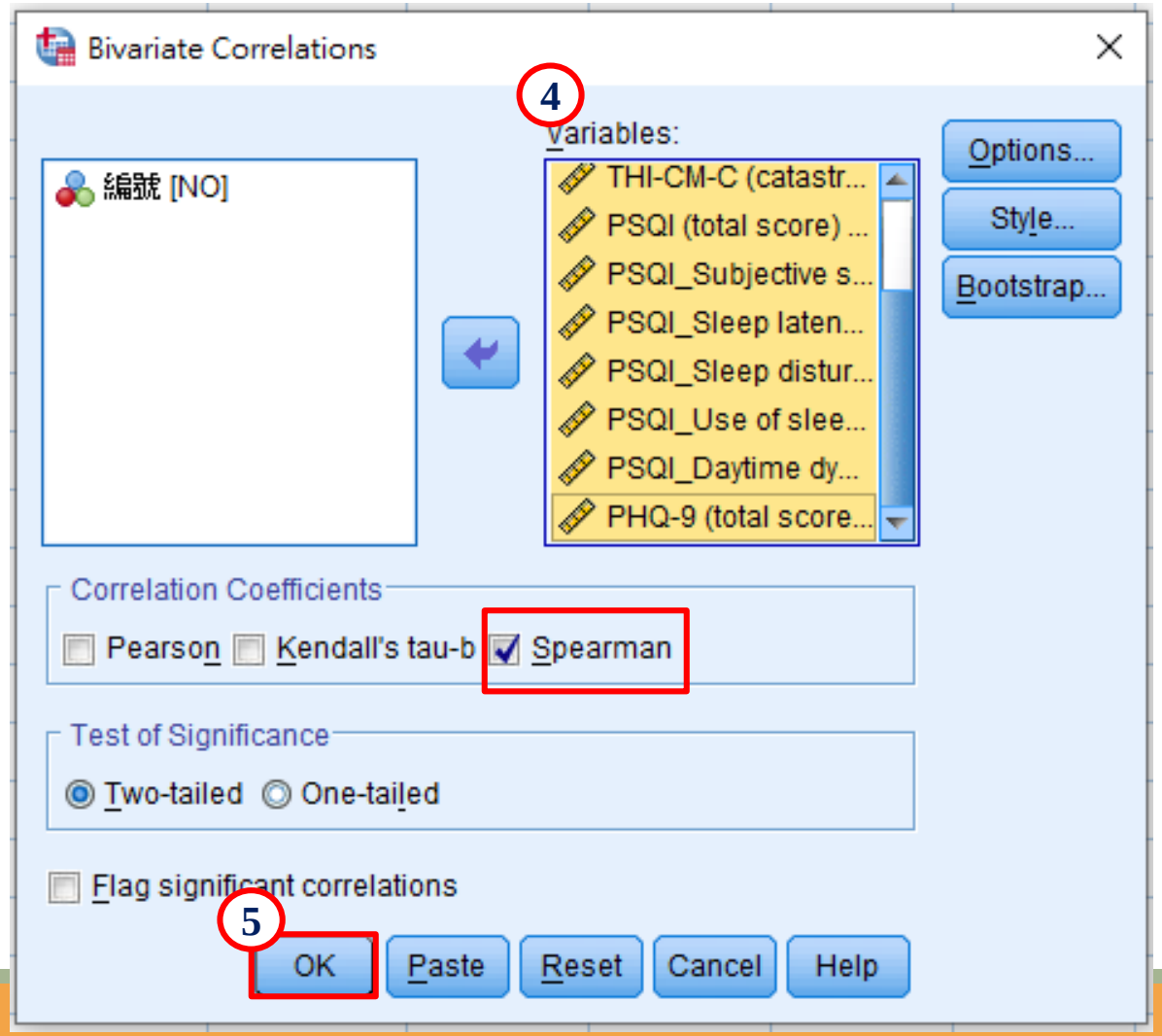
	Name	Type	Width	Decimals	Label
1	NO	Numeric	11	0	編號
2	tin	Numeric	8	0	THI-CM
3	t01	Numeric	8	0	THI-CM-F (functional)
4	t02	Numeric	8	0	THI-CM E (emotional)
5	t03	Numeric	8	0	THI-CM-C (catastrophic)
6	psq	Numeric	8	0	PSQI (total score)
7	ps1	Numeric	8	0	PSQI_Subjective sleep quality
8	ps2	Numeric	8	0	PSQI_Sleep latency
9	ps3	Numeric	8	0	PSQI_Sleep disturbances
10	ps4	Numeric	8	0	PSQI_Use of sleep medication
11	ps5	Numeric	8	0	PSQI_Daytime dysfunction
12	ph	Numeric	8	0	PHQ-9 (total score)
13					

Data View | Variable View

Spearman rank correlation (SPSS analysis)

點選步驟：分析→相關→雙變數





Spearman rank correlation (SPSS output)

			THI-CM	THI-CM-F (functional)	THI-CM E (emotional)	THI-CM-C (catastrophic)	PSQI (total score)	PSQI_Subjective sleep quality	PSQI_Sleep latency	PSQI_Sleep disturbances	PSQI_Use of sleep medication	PSQI_Daytime dysfunction	PHQ-9 (total score)
Spearman's rho	THI-CM	Correlation Coefficient	1.000	.952	.951	.847	.443	.467	.204	.391	.266	.431	.501
		Sig. (2-tailed)	.	.000	.000	.000	.000	.000	.104	.001	.032	.000	.000
		N	65	65	65	65	65	65	65	65	65	65	65
	THI-CM-F (functional)	Correlation Coefficient	.952	1.000	.867	.728	.414	.446	.194	.438	.190	.454	.465
		Sig. (2-tailed)	.000	.	.000	.000	.001	.000	.121	.000	.130	.000	.000
		N	65	65	65	65	65	65	65	65	65	65	65
	THI-CM E (emotional)	Correlation Coefficient	.951	.867	1.000	.726	.438	.465	.162	.388	.294	.438	.474
		Sig. (2-tailed)	.000	.000	.	.000	.000	.000	.197	.001	.017	.000	.000
		N	65	65	65	65	65	65	65	65	65	65	65
	THI-CM-C (catastrophic)	Correlation Coefficient	.847	.728	.726	1.000	.343	.315	.176	.246	.256	.281	.450
		Sig. (2-tailed)	.000	.000	.000	.	.005	.011	.161	.049	.039	.023	.000
		N	65	65	65	65	65	65	65	65	65	65	65
	PSQI (total score)	Correlation Coefficient	.443	.414	.438	.343	1.000	.745	.702	.556	.614	.502	.566
		Sig. (2-tailed)	.000	.001	.000	.005	.	.000	.000	.000	.000	.000	.000
		N	65	65	65	65	65	65	65	65	65	65	65
	PSQI_Subjective sleep quality	Correlation Coefficient	.467	.446	.465	.315	.745	1.000	.489	.455	.414	.372	.572
		Sig. (2-tailed)	.000	.000	.000	.011	.000	.	.000	.000	.001	.002	.000
		N	65	65	65	65	65	65	65	65	65	65	65
	PSQI_Sleep latency	Correlation Coefficient	.204	.194	.162	.176	.702	.489	1.000	.364	.228	.168	.366
		Sig. (2-tailed)	.104	.121	.197	.161	.000	.000	.	.003	.067	.182	.003
		N	65	65	65	65	65	65	65	65	65	65	65
	PSQI_Sleep disturbances	Correlation Coefficient	.391	.438	.388	.246	.556	.455	.364	1.000	.170	.473	.395
		Sig. (2-tailed)	.001	.000	.001	.049	.000	.000	.003	.	.176	.000	.001
		N	65	65	65	65	65	65	65	65	65	65	65
	PSQI_Use of sleep medication	Correlation Coefficient	.266	.190	.294	.256	.614	.414	.228	.170	1.000	.159	.258
		Sig. (2-tailed)	.032	.130	.017	.039	.000	.001	.067	.176	.	.206	.038
		N	65	65	65	65	65	65	65	65	65	65	65
	PSQI_Daytime dysfunction	Correlation Coefficient	.431	.454	.438	.281	.502	.372	.168	.473	.159	1.000	.473
		Sig. (2-tailed)	.000	.000	.000	.023	.000	.002	.182	.000	.206	.	.000
		N	65	65	65	65	65	65	65	65	65	65	65
	PHQ-9 (total score)	Correlation Coefficient	.501	.465	.474	.450	.566	.572	.366	.395	.258	.473	1.000
		Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.003	.001	.038	.000	.
		N	65	65	65	65	65	65	65	65	65	65	65

Summary

	Parametric test	Non-parametric test
Two groups (Independent)	• Independent T-test	• Mann-Whitney U test
Two groups (Dependent)	• Paired T-test	• Wilcoxon signed-rank test
Three groups (Independent)	• ANOVA test	• Kruskal-Wallis test
Correlation	• Pearson correlation	• Spearman rank correlation

生統小組：統計方法教育訓練

Thank you

課程問卷

11/19問卷

For your attention!!