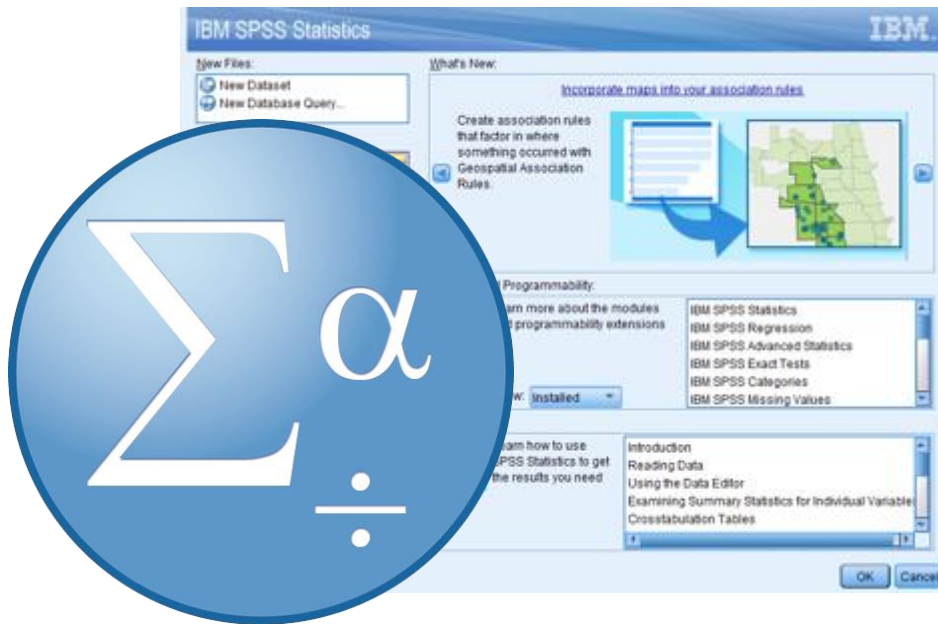


เอกสารประกอบการเรียนการสอน เรื่อง “การใช้โปรแกรม SPSS ขั้นพื้นฐาน”



ณ ห้องปฏิบัติการคอมพิวเตอร์ อปร. ชั้น 8
คณะแพทยศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
วันที่ 11 ธันวาคม 2558

อ.ดร.พญ.รุ่งฤดี ชัยธีรกิจ

Email: roon.chaiteerakij@chula.md, fb: roon chaiteerakij

พญ.ณิชา วงษ์จรรพษ์

Email: fai_716@hotmail.com

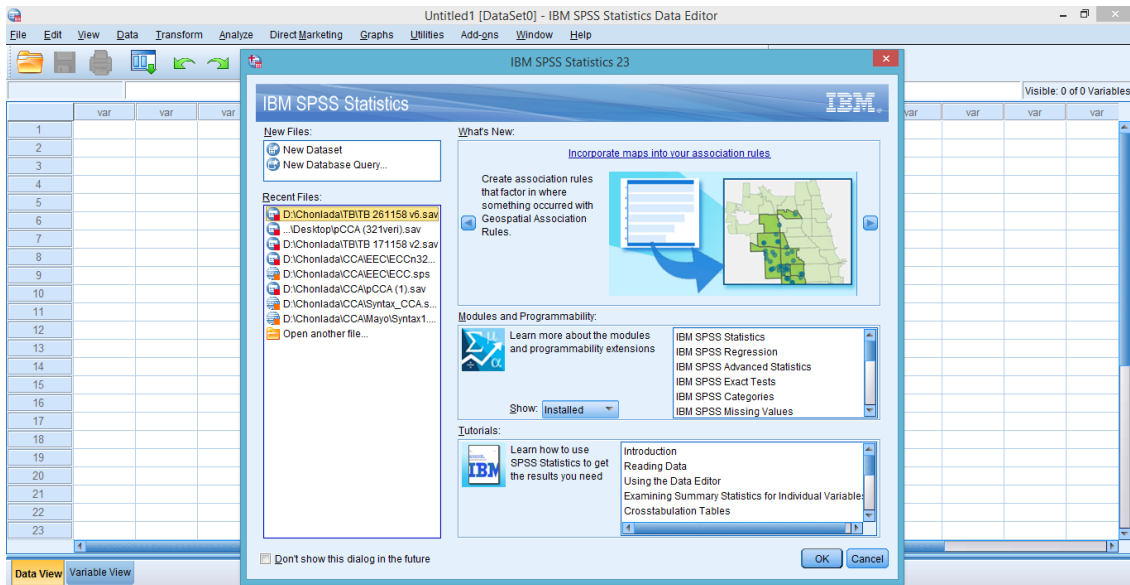
ชลดา ณะธง

Email: chonlada.pha@gmail.com

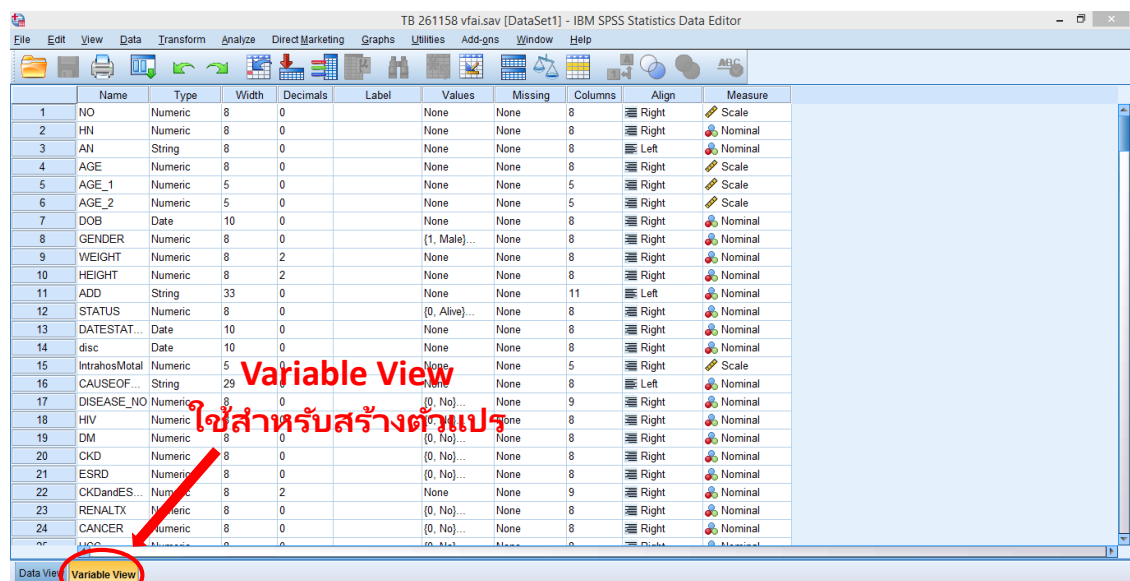
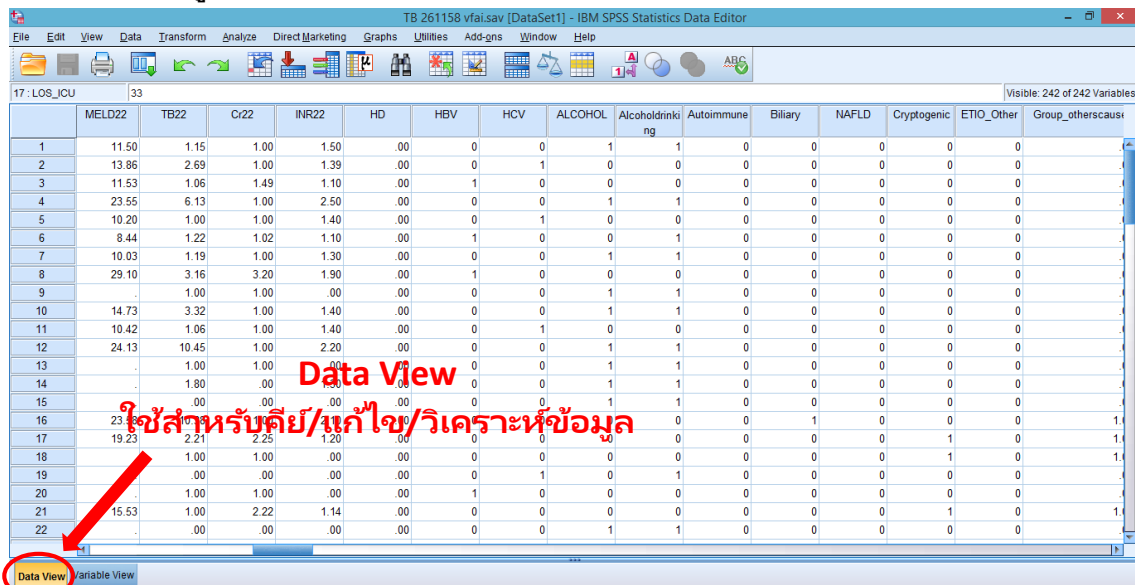
หน่วยโรคทางเดินอาหารและตับ (02-256 4265)

โปรแกรมสำเร็จรูป SPSS version 23

1. เปิดโปรแกรม



2. การสร้างแฟ้มข้อมูล



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	NO	Numeric	8	0		None	None	8	Right	Scale
2	HN	Numeric	8	0		None	None	8	Right	Nominal
3	AN	String	8	0		None	None	8	Left	Nominal
4	AGE	Numeric	8	0		None	None	8	Right	Scale
5	AGE_1	Numeric	5	0		None	None	5	Right	Scale
6	AGE_2	Numeric	5	0		None	None	5	Right	Scale
7	DOB	Date	10	0		None	None	8	Right	Nominal
8	GENDER	Numeric	8	0	{1, Male}...	None	None	8	Right	Nominal
9	WEIGHT	Numeric	8	2		None	None	8	Right	Nominal
10	HEIGHT	Numeric	8	2		None	None	8	Right	Nominal
11	ADD	String	33	0		None	None	11	Left	Nominal
12	STATUS	Numeric	8	0	{0, Alive}...	None	None	8	Right	Nominal
13	DATESTAT...	Date	10	0		None	None	8	Right	Nominal
14	disc	Date	10	0		None	None	8	Right	Nominal
15	IntrahosMotal	Numeric	5	0		None	None	5	Right	Scale
16	CAUSEOF...	String	29	0		None	None	8	Left	Nominal
17	DISEASE_NO	Numeric	8	0	{0, No}...	None	None	9	Right	Nominal
18	HIV	Numeric	8	0	{0, No}...	None	None	8	Right	Nominal
19	DM	Numeric	8	0	{0, No}...	None	None	8	Right	Nominal
20	CKD	Numeric	8	0	{0, No}...	None	None	8	Right	Nominal
21	ESRD	Numeric	8	0	{0, No}...	None	None	8	Right	Nominal
22	CKDandES...	Numeric	8	2		None	None	9	Right	Nominal
23	RENALT...	Numeric	8	0	{0, No}...	None	None	8	Right	Nominal
24	CANCER	Numeric	8	0	{0, No}...	None	None	8	Right	Nominal

1. Name: กำหนดชื่อตัวแปร
2. Type: กำหนดชนิดของตัวแปร

Variable Type

☒ Numeric
☐ Comma
☐ Dot
☐ Scientific notation
☐ Dollar
☐ Custom currency
☐ String
☐ Restricted Numeric (integer with leading zeros)

Width: 8
 Decimal Places: 0

The Numeric type honors the digit grouping setting, while the Restricted Numeric never uses digit grouping.

OK Cancel Help

ลักษณะของชนิดตัวแปร

Numeric	ใช้กับข้อมูลที่เป็นตัวเลข
Comma	ใช้กับข้อมูลที่เป็นตัวเลขและต้องมี comma เช่น 5,000.00
Dot	ใช้กับข้อมูลที่เป็นตัวเลข (สลับกับการใช้ Comma) เช่น 5.000,00
Scientific notation	ใช้กับข้อมูลที่เป็นตัวเลขและสัญลักษณ์ เช่น 347E+4
Date	ใช้กับข้อมูลที่เป็นวัน เดือน ปี
Dollar	ใช้กับข้อมูลที่เป็นตัวเลข (ตัวเงิน) และเครื่องหมาย \$
Custom currency	ใช้กับข้อมูลที่เป็น Custom currency
String	ใช้กับข้อมูลที่เป็นตัวเลขและหรือตัวอักษร
Restricted Numeric	ใช้กับข้อมูลที่เป็นตัวเลข ที่มี 0 นำหน้า

3. Width กำหนดความกว้างของค่าตัวแปร
4. Decimals กำหนดทศนิยม
5. Label อธิบายความหมายของค่าตัวแปร

6. Values กำหนด code ตัวแปรที่แปลงจากตัวแปรกลุ่มเป็น ตัวเลข

1. Click

2. กำหนด Value เป็นตัวเลข และ Label เป็นข้อความ

3. Click

4. Click

7. Missing กำหนดค่าสูญหาย

1. Click

No missing = ไม่กำหนดค่า missing

Discrete missing = กำหนดค่า missing ได้ไม่เกิน 3 ค่า

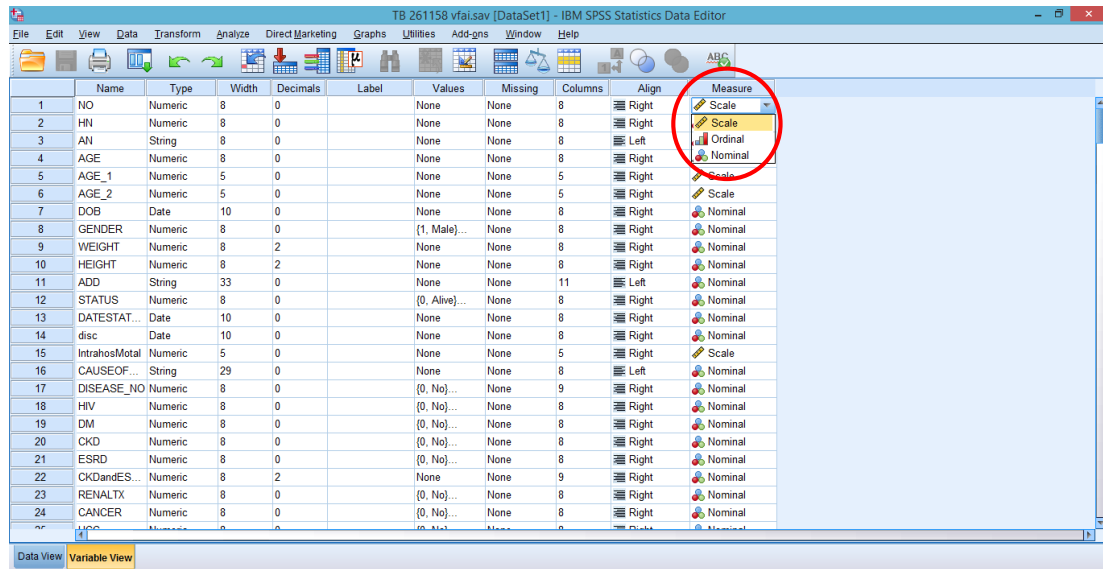
Range of missing = กำหนดค่า missing เป็นช่วง

8. Columns กำหนดความกว้างของคอลัมน์เฉพาะในหน้าจอ Data view

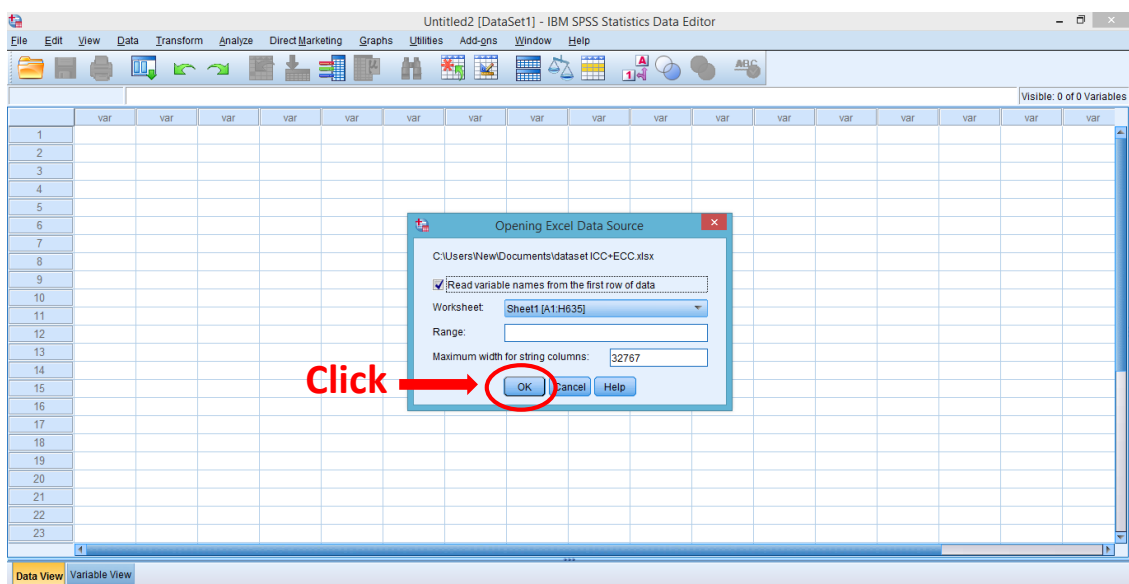
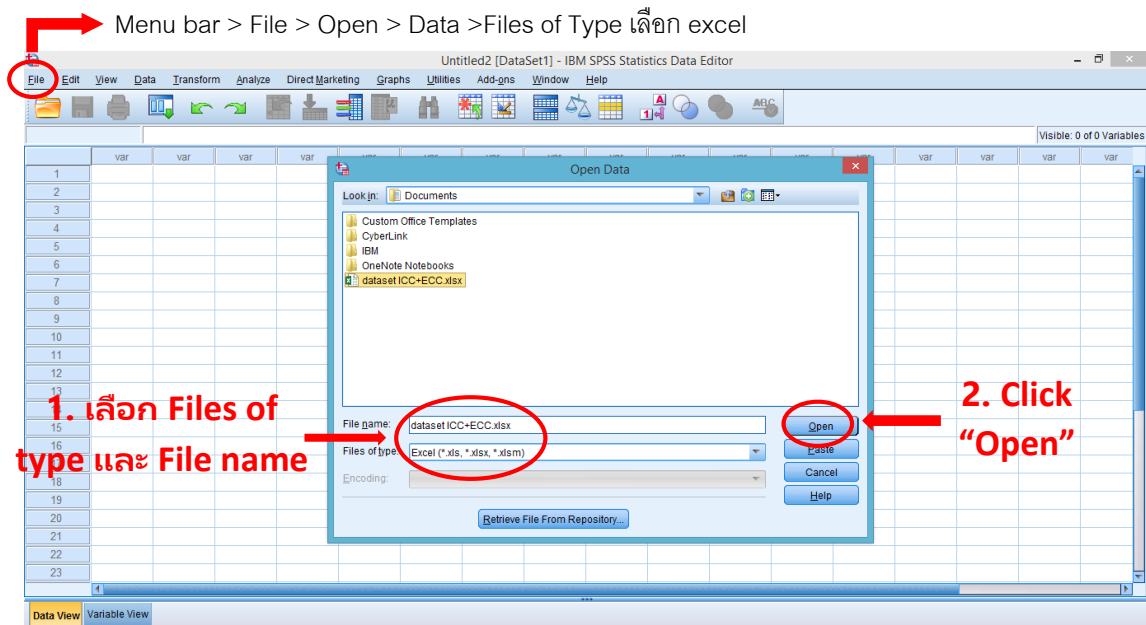
9. Align กำหนดตำแหน่งการแสดงค่าข้อมูลในหน้าจอ Data view

Align

10. Measure ใช้กำหนดมาตรวัดข้อมูลว่าข้อมูลนั้นเป็นข้อมูลแบบ Scale (คือ continuous variable), Ordinal หรือ Nominal (คือ categorical variable)

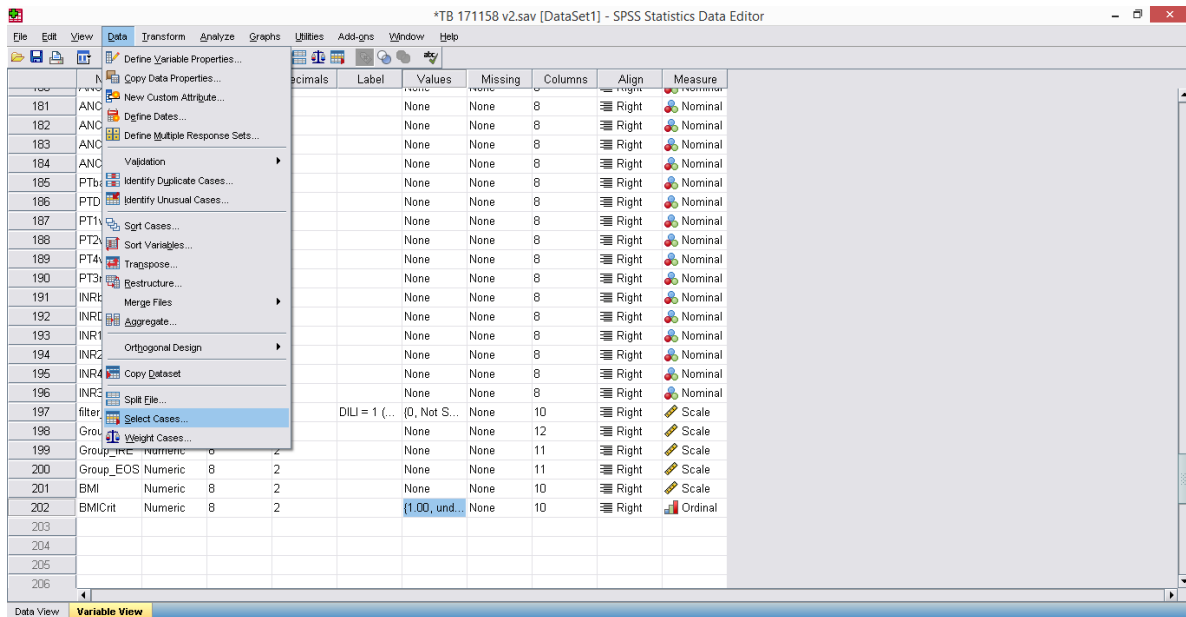


3. การนำเข้าข้อมูลที่เป็น Excel

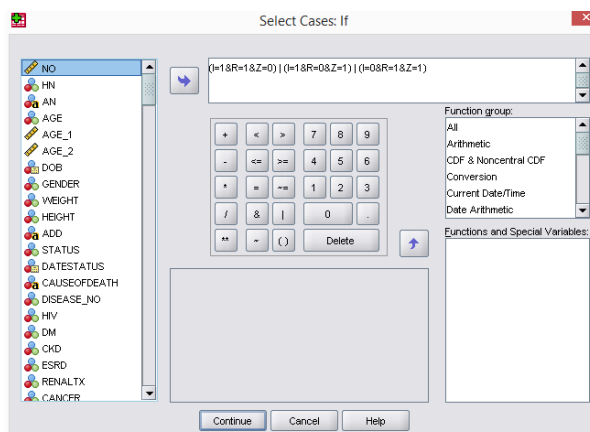
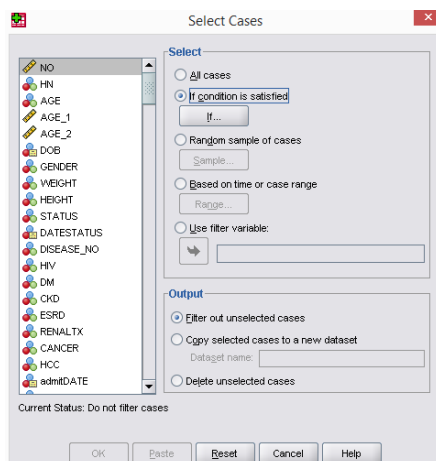


วิธี Select Case เพื่อวิเคราะห์ข้อมูลแบบ subgroup analysis

Data > Select cases



ตัวอย่าง: เลือก เฉพาะ case ที่ใช้ยา 2 ตัว จาก ยา I, R, Z



การสร้างตัวแปรใหม่จากตัวแปรที่มีอยู่แล้วโดยใช้คำสั่ง Compute variable

☞ Transform > Compute variable เพื่อคำนวณค่าใหม่จากตัวแปรเดิม

The screenshot shows the 'Compute Variable' menu under 'Transform' in SPSS. The menu options include: Count Values within Cases..., Shift Values..., Recode into Same Variables..., Recode into Different Variables..., Automatic Recode..., Visual Binning..., Optimal Binning..., Rank Cases..., Date and Time Wizard..., Create Time Series..., Replace Missing Values..., Random Number Generators..., and Run Pending Transformations. Below the menu is a data table with columns: AGE_1, AGE_2, DOB, GENDER, WEIGHT, HEIGHT, ADD, STATUS, DATESTATU, CAUSEOFD, DISEASE_I.

☞ ตัวอย่าง: คำนวณ BMI โดยใช้ $BMI = Weight / (Height * Height / 10000)$

The 'Compute Variable' dialog box shows 'Target Variable' as BMI and 'Numeric Expression' as $WEIGHT / (HEIGHT * HEIGHT / 10000)$. The left pane lists variables: AGE_1, AGE_2, DOB, GENDER, WEIGHT, HEIGHT, ADD, STATUS, DATESTATU, CAUSEOFDEATH, DISEASE_NO, HIV, DM, CKD, ESRD, RENALT, CANCER, HCC. The right pane shows function groups: Search, Significance, Statistical, String, Time Duration Creation, Time Duration Extraction, and Functions and Special Variables.

** หมายถึง ยกกำลังสอง

& หมายถึง และ

| หมายถึง หรือ

~ หมายถึง ปฏิเสธ

~= หมายถึง ≠

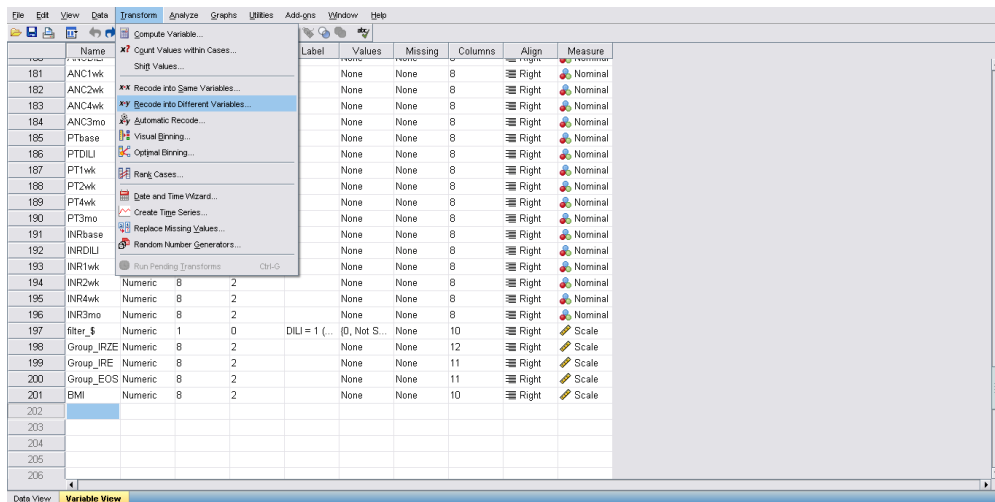
☞ ตัวอย่าง: คำนวณ Date calculation for Length of stay: Date Arithmetic > DATEDIFF

The 'Compute Variable' dialog box shows 'Target Variable' as LOS1 and 'Numeric Expression' as $DATEDIFF(?,?)$. The left pane lists variables: NO, HN, AN, AGE, AGE_1, AGE_2, DOB, GENDER, WEIGHT, HEIGHT, ADD, STATUS, DATESTATUS, CAUSEOFDEATH, DISEASE_NO, HIV, DM, CKD, ESRD, RENALT, CANCER, HCC. The right pane shows function groups: Conversion, Current Date/Time, Date Arithmetic, Date Creation, Date Extraction, Inverse DF, and Functions and Special Variables.

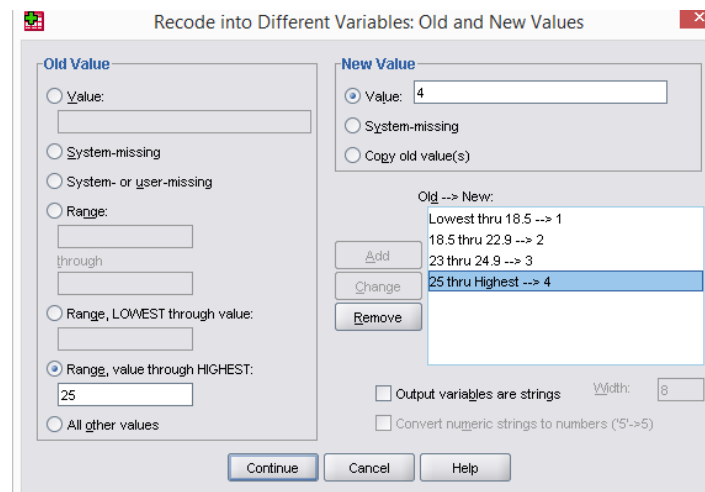
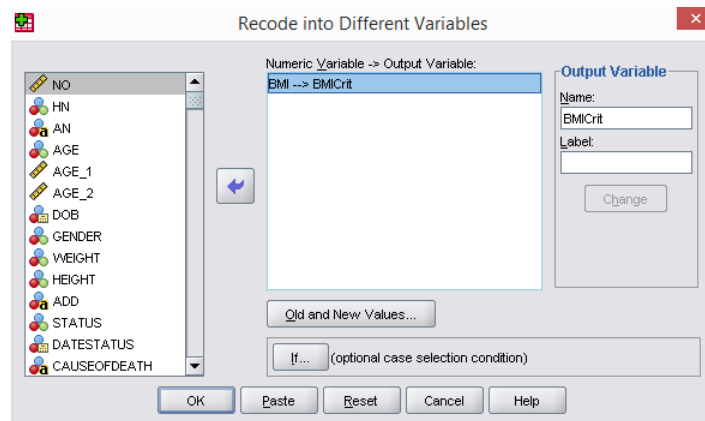
The 'Compute Variable' dialog box shows 'Target Variable' as LOS1 and 'Numeric Expression' as $DATEDIFF(discDATE, admDATE, "days")$. The left pane lists variables: DISEASE_OTHER, NameDisease, MED, admDATE, discDATE, PDiag, LOS_JCU, LOS_ward, LOS, DateDxCir, Childscore, Childclass, MELD, HD, HBV, HCV, ALCOHOL, Autoimmune. The right pane shows function groups: Conversion, Current Date/Time, Date Arithmetic, Date Creation, Date Extraction, Inverse DF, and Functions and Special Variables. A red box highlights the 'Units' field in the 'Date Arithmetic' group, with text: 'DATEDIFF(date1,date2,"units")' and 'Units ที่เลือกใส่ได้ ใส่ " " และ s ด้วย'.

Recode

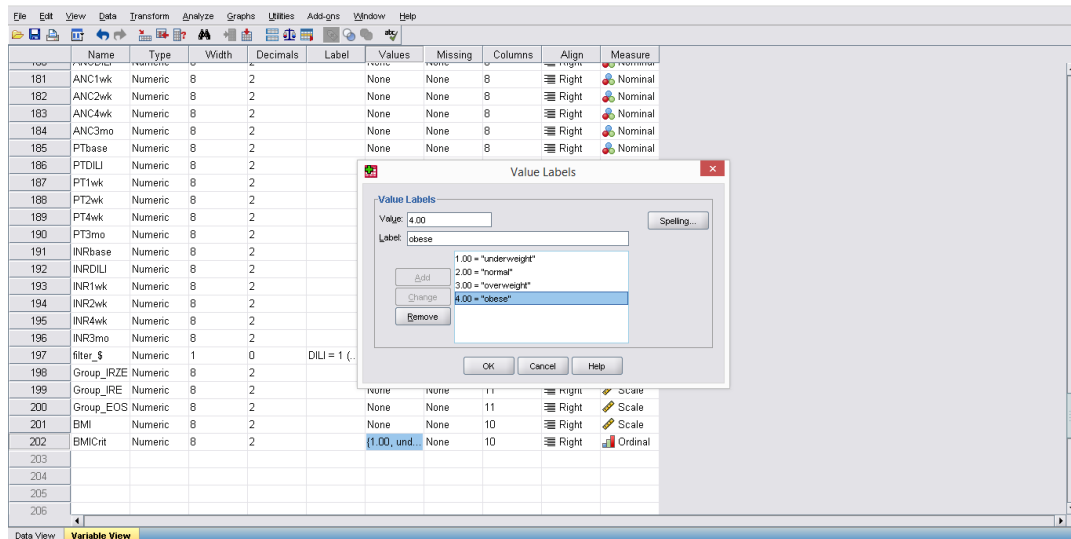
☞ Transform > Recode into Same (แทนที่ตัวแปรเดิม)/ Different Variable (สร้างเป็นตัวแปรใหม่)



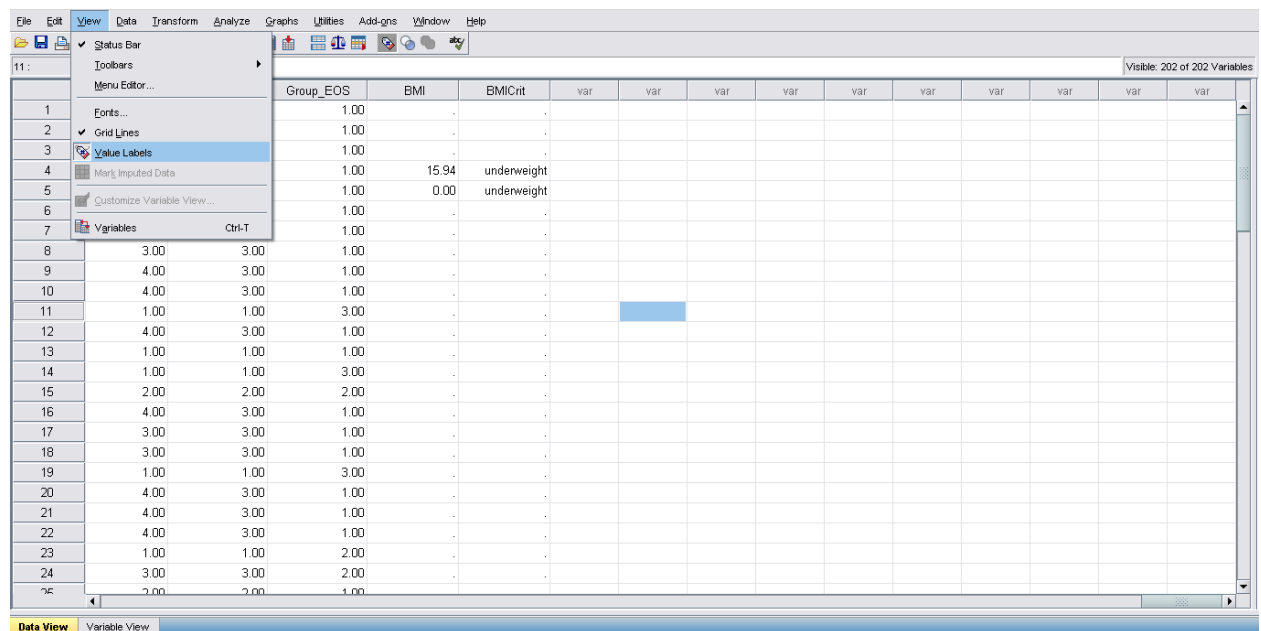
☞ ตัวอย่าง: สร้างตัวแปร BMICrit เพื่อแบ่ง BMI ตาม criteria



กำหนดชื่อให้แต่ละตัวเลข



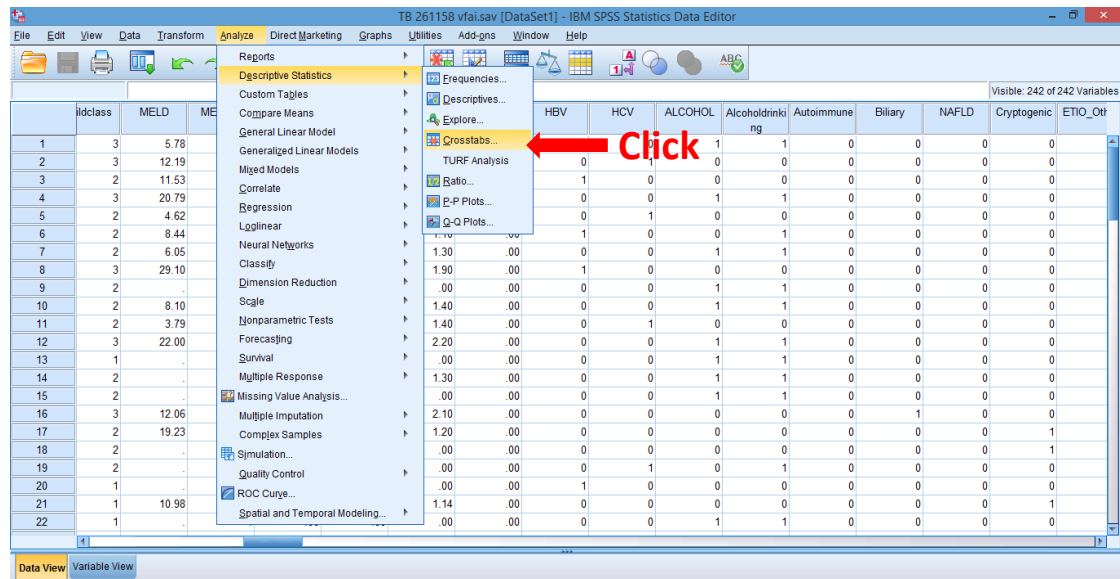
View > Value label เพื่อแสดงค่าตามที่เรากำหนด



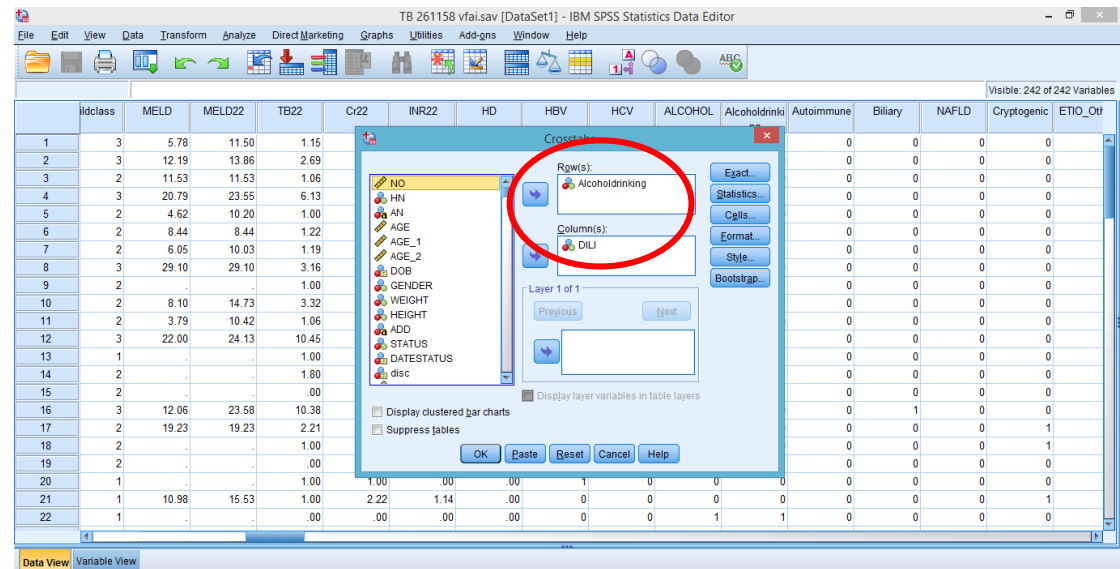
Chi-square test

ใช้สำหรับเปรียบเทียบ categorical variable ระหว่างสองกลุ่ม

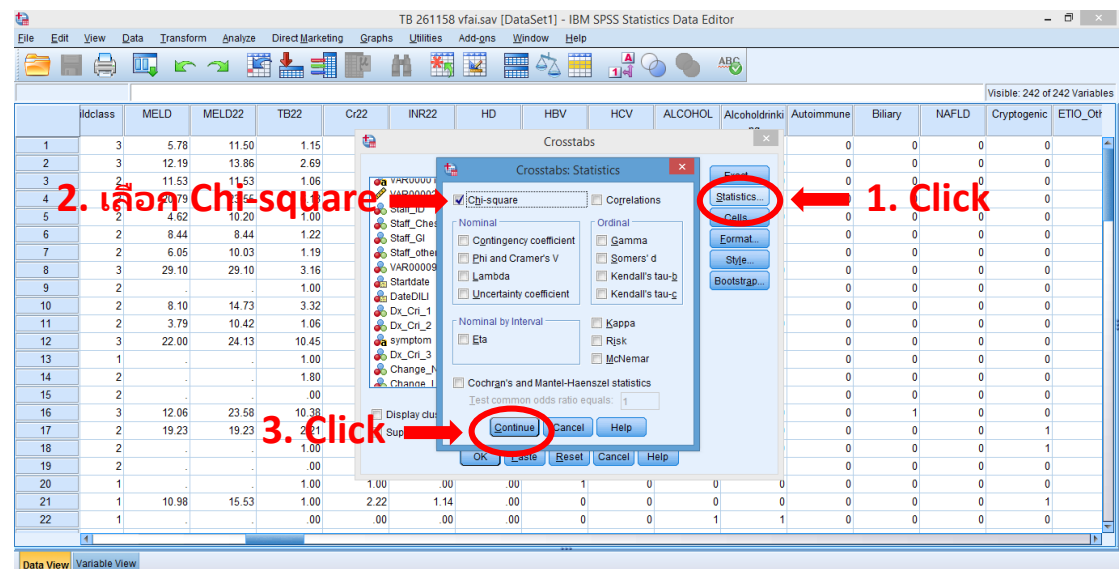
☛ Analyze > Descriptive Statistics > Crosstabs



☛ ใส่ตัวแปรในช่อง Row และ Column > Click OK



☛ กำหนด Statistics ที่ต้องการใช้วิเคราะห์ (ในที่นี้จะใช้ Chi-square)



กำหนดให้แสดงค่า Percentages ในแต่ละ cell

1. Click

2. เลือกแสดง percentages ใน Row และ Column

3. Click

Click

Alcoholdrinking * DILI Crosstabulation

		DILI		Total
		No	Yes	
Alcoholdrinking	No	Count 26	11	37
		% within Alcoholdrinking 70.3%	29.7%	100.0%
		% within DILI 65.0%	33.3%	50.7%
Yes	Count 14	22	36	
	% within Alcoholdrinking 38.9%	61.1%	100.0%	
	% within DILI 35.0%	66.7%	49.3%	
Total	Count 40	33	73	
	% within Alcoholdrinking 54.8%	45.2%	100.0%	
	% within DILI 100.0%	100.0%	100.0%	

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square ^a	7.254 ^a	1	.007	.010	
Continuity Correction ^b	6.043	1	.014		
Likelihood Ratio	7.380	1	.007		
Fisher's Exact Test					
Linear-by-Linear Association	7.155	1	.007		
N of Valid Cases	73				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 16.27.
b. Computed only for a 2x2 table

1. ดูว่ามีเซลล์ที่มี expected count <5 หรือไม่

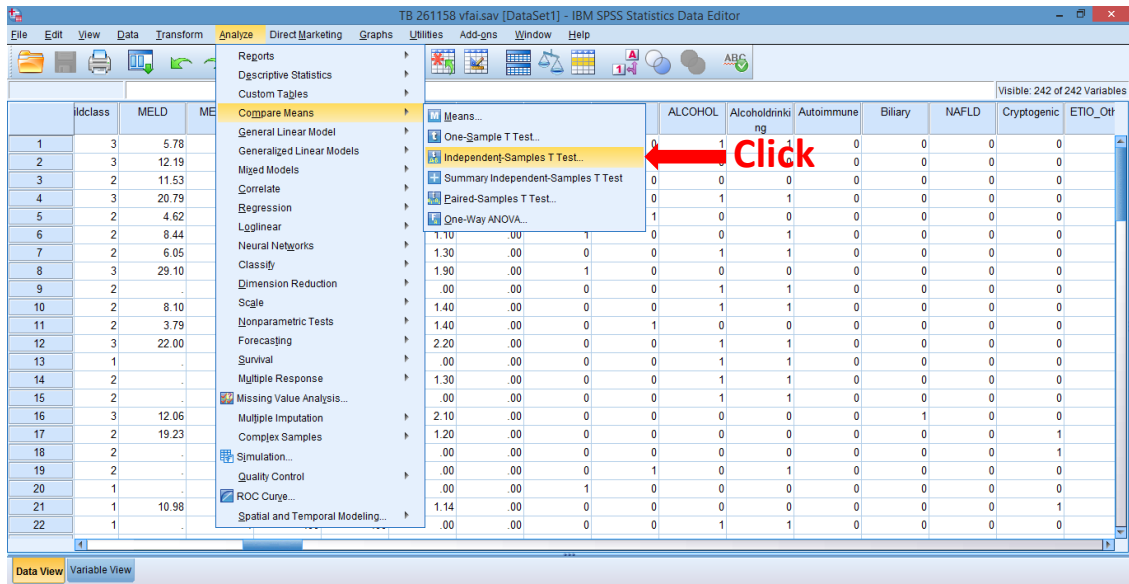
2. กรณีจำนวนเซลล์ที่มี expected count <5 น้อยกว่า 20% ของจำนวนเซลล์ทั้งหมด ใช้ค่า P value ของ Pearson Chi-Square

3. ในกรณีที่ค่า count ในแต่ละเซลล์น้อย หรือมีเซลล์ที่มี expected count <5 มากกว่า 20% ของจำนวนเซลล์ทั้งหมด ใช้ค่า P value ของ Fisher's Exact Test P value= 0.010

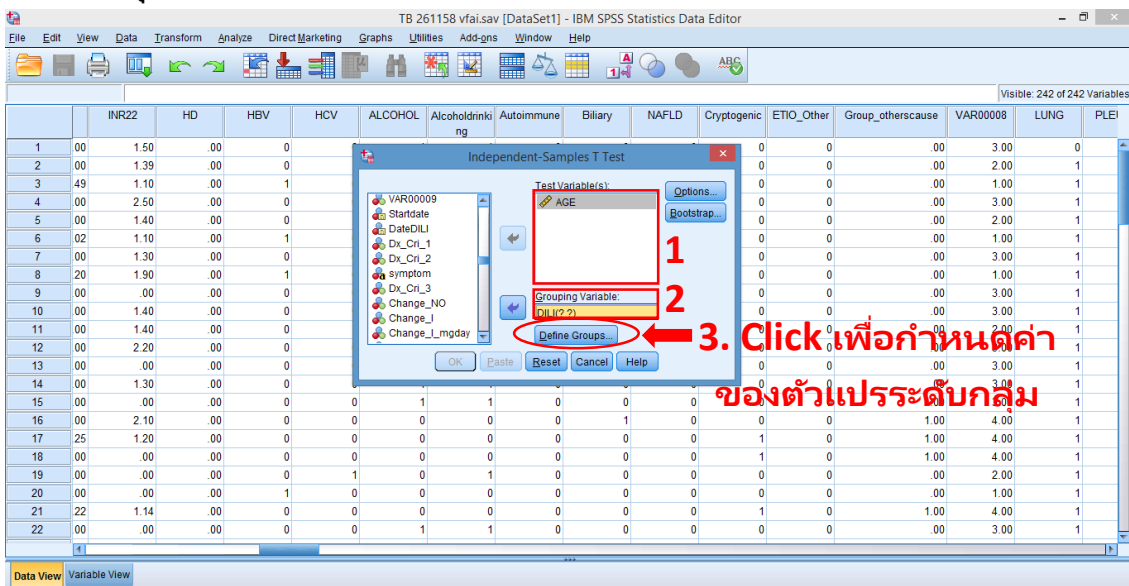
T-test

ใช้สำหรับเปรียบเทียบ continuous variable ระหว่างสองกลุ่ม

☞ Analyze > Compare Means > Independent-Samples T Test



- ☞ 1. ใส่ตัวแปรชนิด Continuous ในช่อง Test Variable
2. ใส่ตัวแปรชนิด Categorical ในช่อง Grouping Variable
3. กำหนดกลุ่มที่จะเปรียบเทียบ ในช่อง Define groups



IBM SPSS Statistics Data Editor

Visible: 242 of 242 Variables

INR22 HD HBV HCV ALCOHOL Alcoholdrinking Autoimmune Biliary NAFLD Cryptogenic ETIO_Other Group_otherscause VAR00008 LUNG PLEI

1 กำหนดค่าของตัวแปรแต่ละกลุ่ม (ในที่นี้ 0 คือ ไม่เกิด DILI, 1=เกิด DILI)

2. Click

Define Groups

Use specified values

Group 1: 0

Group 2: 1

Continue

IBM SPSS Statistics Data Editor

Visible: 242 of 242 Variables

INR22 HD HBV HCV ALCOHOL Alcoholdrinking Autoimmune Biliary NAFLD Cryptogenic ETIO_Other Group_otherscause VAR00008 LUNG PLEI

Click

เพื่อบริหารข้อมูล

Independent-Samples T Test

Test Variable(s): AGE

Grouping Variable: DILI(0 1)

OK

IBM SPSS Statistics Viewer

*Output3 [Document3] - IBM SPSS Statistics Viewer

T-TEST GROUPS=DILI (0 1)

/MISSING=ANALYSIS

/VARIABLES=AGE

/CRITERIA=CI(.95).

T-Test

Group Statistics

	DILI	N	Mean	Std. Deviation	Std. Error Mean
AGE	No	40	55.08	14.811	2.342
	Yes	33	53.21	13.474	2.346

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	
AGE	Equal variances assumed	.049	.825	.557	71	.579	1.863	3.345	-4.807	8.532
	Equal variances not assumed			.562	70.287	.576	1.863	3.314	-4.747	8.473

1. เปรียบเทียบค่าความแปรปรวนของข้อมูลในแต่ละกลุ่ม ว่าแตกต่างกันหรือไม่ โดยดูที่ P ของ Levene's test

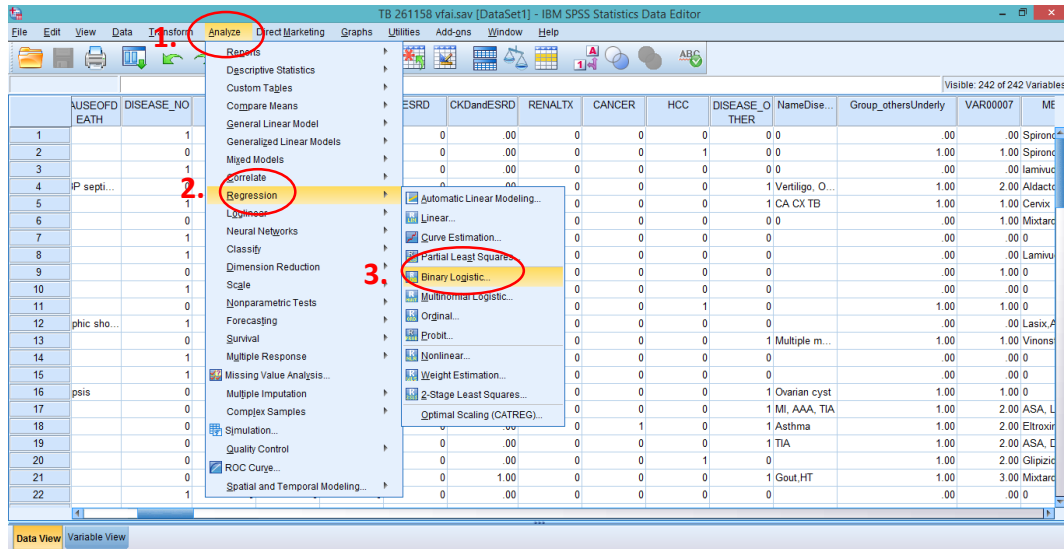
2. หาก P ของ Levene's test ≥ 0.05 แปลว่าค่าความแปรปรวนไม่แตกต่างกัน ให้ใช้ค่า P value for T test = 0.579

3. หาก P ของ Levene's test < 0.05 แปลว่าค่าความแปรปรวนแตกต่างกัน ให้ใช้ค่า P value for T test = 0.576

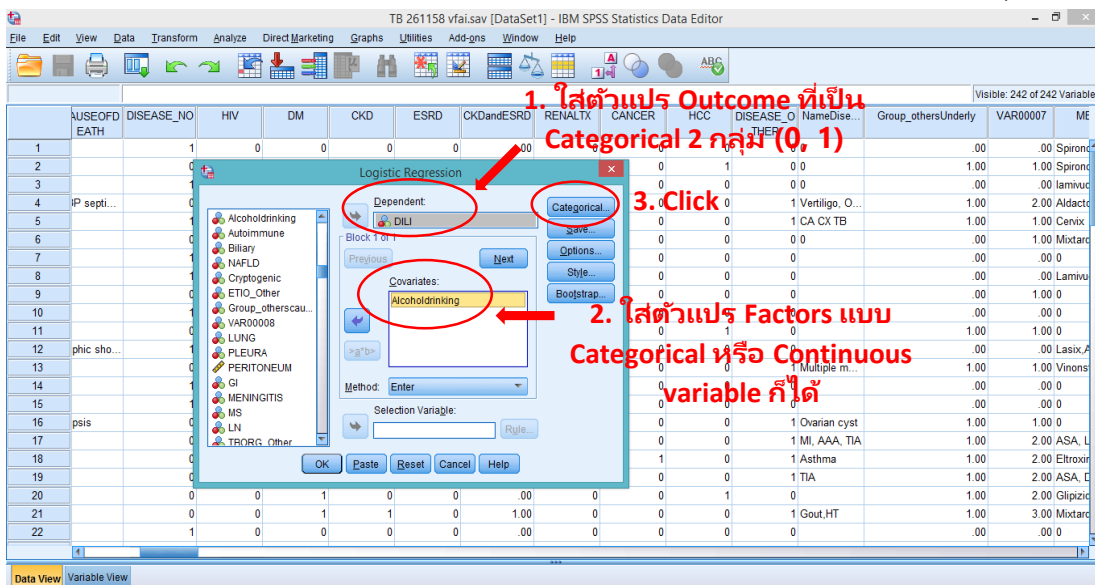
ดังนั้น P value= 0.579 แสดงว่าอายุของผู้ป่วยในกลุ่มที่เกิดและไม่เกิด DILI ไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ

Logistic regression (Odds Ratio)

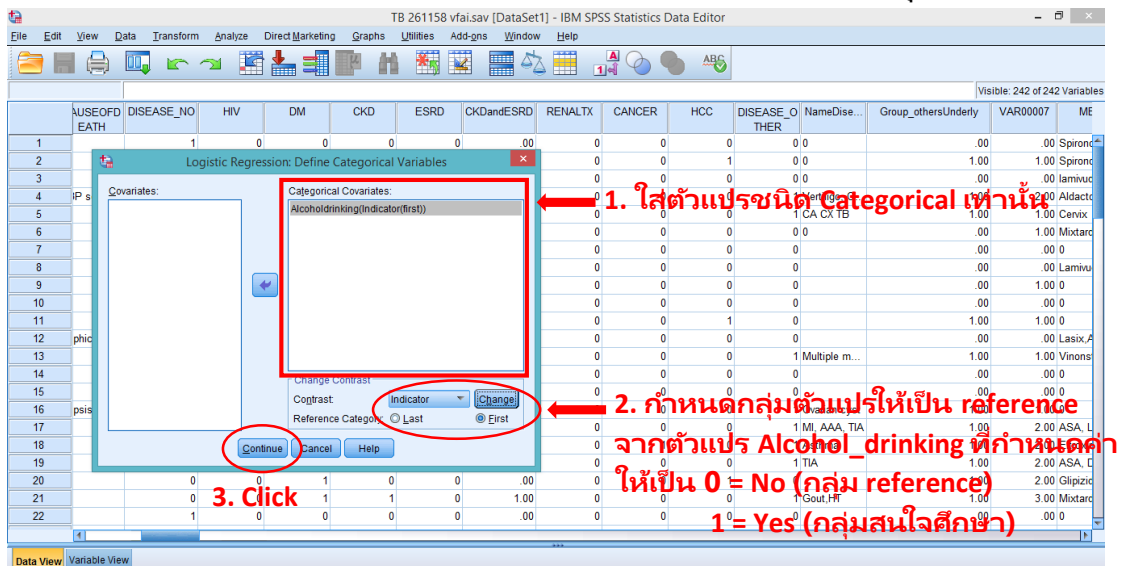
☞ Analyze > Regression > Binary Logistic (ใช้สำหรับ Outcome ที่เป็น categorical variable ที่มีกลุ่มตัวแปรแค่ 2 Category)



☞ ใส่ตัวแปร Outcome ที่เป็นตัวแปรชนิด Categorical ในช่อง Dependent และใส่ Factors อื่นๆ ในช่อง Covariates



☞ นำเฉพาะตัวแปรที่เป็น Categorical ใส่ในช่อง Categorical Covariates เพื่อกำหนดกลุ่ม reference



กำหนด Option ให้แสดงค่าที่ต้องการ

Logistic Regression

Dependent: DILI

Covariates: Alcoholdrinking(Cat)

Method: Enter

Logistic Regression: Options

Statistics and Plots

☒ Classification plots

☒ Hosmer-Lemeshow goodness-of-fit

☒ Casewise listing of residuals

☒ CI for exp(B) 95 %

☒ Include constant in model

3. Continue

Output ที่ได้จากการวิเคราะห์

		Chi-square	df	Sig.
Step 1	Step	7.380	1	.007
	Block	7.380	1	.007
	Model	7.380	1	.007

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	93.147 ^a	.096	.129

a. Estimation terminated at iteration number 3 because parameter estimates changed by less than .001.

Classification Table^a

		Predicted		
		DILI		Percentage Correct
Observed		No	Yes	
	DILI No	26	14	65.0
	Yes	11	22	66.7
Overall Percentage				65.8

a. The cut value is .500

ดังนั้น การดื่มแอลกอฮอล์มีความสัมพันธ์กับการเกิด DILI โดยผู้ที่ดื่มแอลกอฮอล์มีโอกาสเกิด DILI มากกว่าผู้ที่ไม่ได้ดื่มแอลกอฮอล์ 3.714 เท่า อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

Odds Ratio = 3.714

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
Step 1 ^a	Alcoholdrinking(1)	1.312	.496	6.992	1	.008	3.714	Lower	Upper
	Constant	-.860	.360	5.720	1	.017	.423	1.404	9.824

a. Variable(s) entered on step 1: Alcoholdrinking.

Sig. = 0.008