

## Ordered Logistic Regression Workshop #3-GZLM

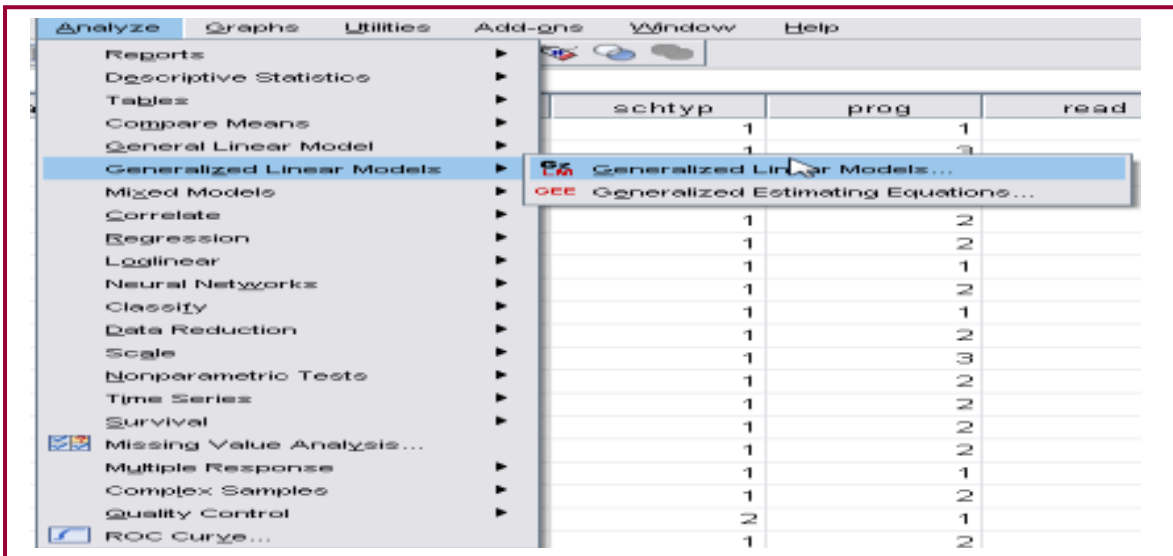
ตารางด้านล่างแสดงข้อมูลที่เราต้องการวิเคราะห์โดยใช้ Ordered logistic regression (OLOGIT) โดยตัวแปรตาม

คือ ses(social economic status) ส่วนตัวแปรอิสระประกอบด้วย female (0=male,1=female) science (คะแนนสาขา

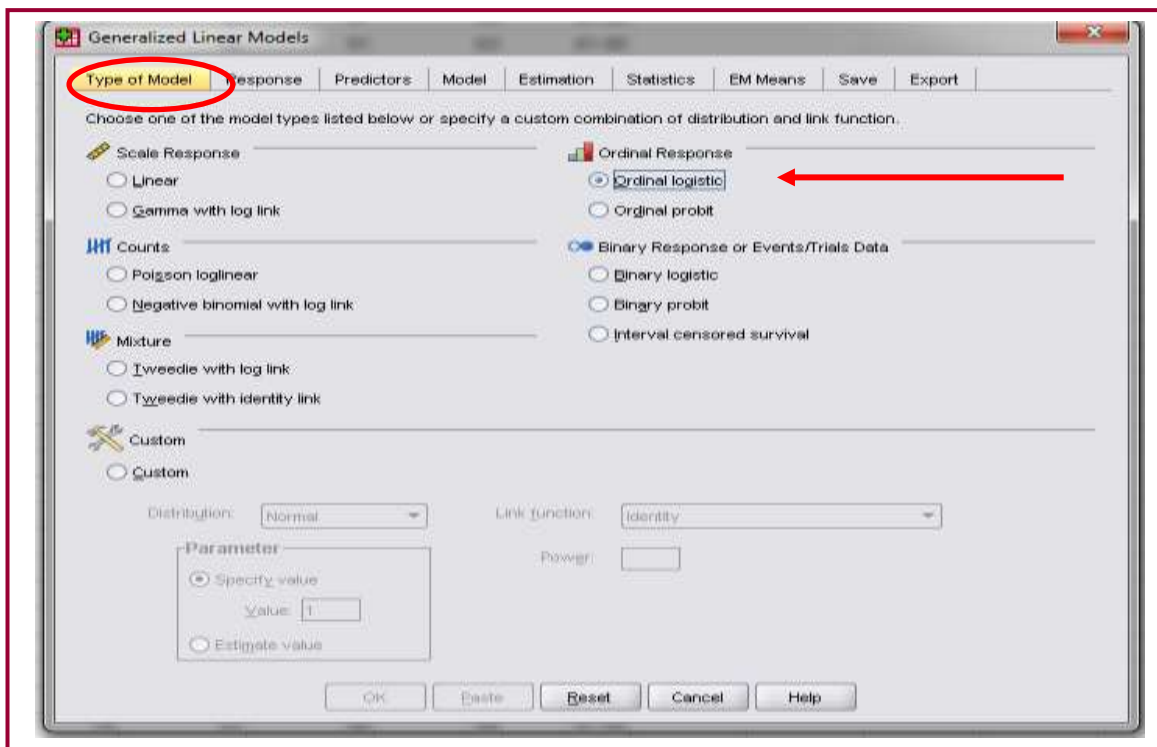
วิทยาศาสตร์) socst (คะแนนวิชาสังคม)

	id	female	race	ses	schtyp	prog	read	write	math	science	socst
1	70	0	4	1	1	1	57	52	41	47	57.00
2	121	1	4	2	1	3	68	59	53	63	61.00
3	85	0	4	3	1	1	44	33	54	58	31.00
4	141	0	4	3	1	3	63	44	47	53	56.00
5	172	0	4	2	1	2	47	52	57	53	61.00
6	113	0	4	2	1	2	44	52	51	63	61.00
7	50	0	3	2	1	1	50	59	42	53	61.00
8	11	0	1	2	1	2	34	46	45	39	36.00
9	84	0	4	2	1	1	63	57	54	58	51.00
10	48	0	3	2	1	2	57	55	52	50	51.00
11	75	0	4	2	1	3	60	46	51	53	61.00
12	60	0	4	2	1	2	57	65	51	63	61.00
13	95	0	4	3	1	2	73	60	71	61	71.00
14	104	0	4	3	1	2	54	63	57	55	46.00
15	38	0	3	1	1	2	45	57	50	31	56.00
16	115	0	4	1	1	1	42	49	43	50	56.00
17	76	0	4	3	1	2	47	52	51	50	56.00
18	195	0	4	2	2	1	57	57	60	58	56.00
19	114	0	4	3	1	2	68	65	62	55	61.00
20	85	0	4	2	1	1	55	39	57	53	46.00
21	167	0	4	2	1	1	63	49	35	66	41.00
22	143	0	4	2	1	3	63	63	75	72	66.00
23	41	0	3	2	1	2	50	40	45	55	56.00
24	20	0	1	3	1	2	60	52	57	61	61.00
25	12	0	1	2	1	3	37	44	45	39	46.00
26	53	0	3	2	1	3	34	37	46	39	31.00
27	154	0	4	3	1	2	65	65	66	61	66.00
28	178	0	4	2	2	3	47	57	57	58	46.00
29	196	0	4	3	2	2	44	38	49	39	46.00
30	29	0	2	1	1	1	52	44	49	55	41.00
31	126	0	4	2	1	1	42	31	57	47	51.00
32	103	0	4	3	1	2	76	52	64	64	61.00
33	192	0	4	3	2	2	65	67	63	66	71.00
34	150	0	4	2	1	3	42	41	57	72	31.00
35	199	0	4	3	2	2	52	59	50	61	61.00
36	144	0	4	3	1	1	60	65	58	61	66.00
37	200	0	4	2	2	2	68	54	75	66	66.00
38	80	0	4	3	1	2	65	62	68	66	66.00

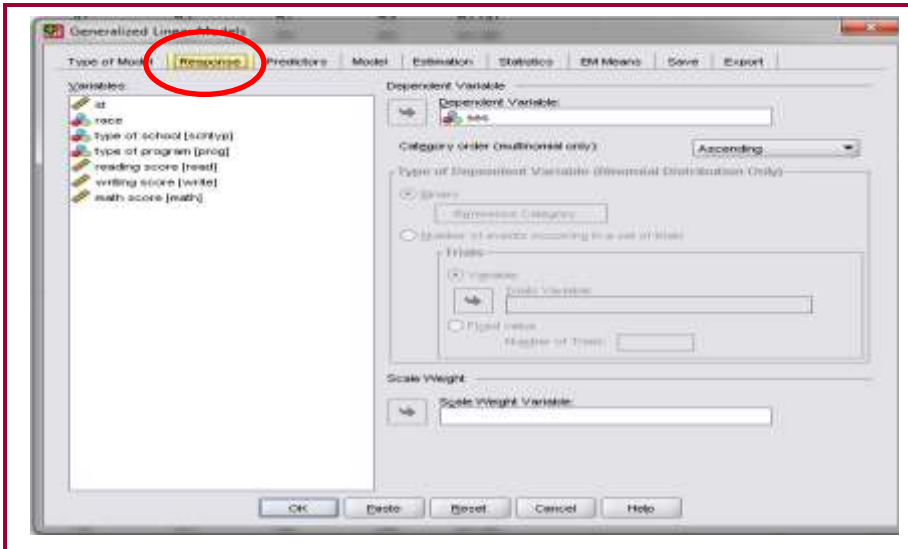
- คลิก Analyze>Generalized Linear Models> Generalized Linear Models ดังแสดงได้ดังภาพต่อไปนี้



จะปรากฏแสดงให้เห็น submenu Generalized Linear Models ในส่วนของ Type of Model ให้คลิกเลือก Ordinal logistic แสดงได้ดังภาพด้านล่าง

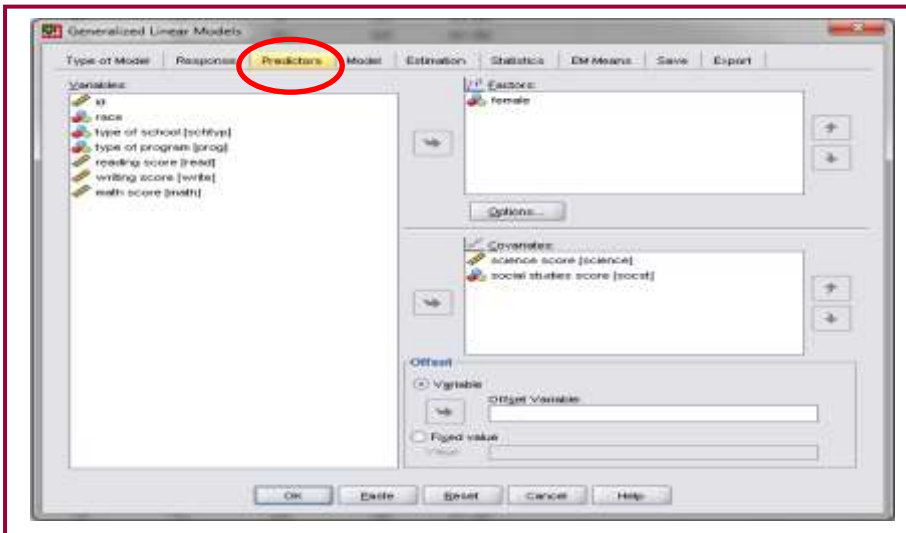


- คลิกที่แถบ Response ของ submenu เพื่อระบุชื่อของตัวแปรตาม ( ในที่นี้ ses)



•คลิกเลือก ses ให้ไปอยู่ในช่อง  
Dependent Variable

•คลิกที่แถบ Predictors เพื่อเลือกตัวแปรอิสระที่เป็นตัวแปรนามบัญญัติ และตัวแปรที่มีค่าต่อเนื่อง



•คลิกเลือก female ให้ไปอยู่ในช่อง Factors และคลิกเลือก science และ socst ให้ไปอยู่ในช่อง Covariates

•คลิกที่แถบ Model เพื่อระบุตัวแปรอิสระที่ต้องการศึกษาว่ามีผลต่อตัวแปรตามหรือไม่



•คลิกเลือกตัวแปรที่เป็นนามบัญญัติและตัวแปรที่มีค่าต่อเนื่องในช่อง Factors and Covariates ด้านซ้าย ให้ไปอยู่ในช่อง Model ด้านขวามือ

- คลิกที่แถบ Statistics เพื่อให้แสดงผลที่ต้องการศึกษา



ในช่อง Chi Square Statistics ให้เลือก Likelihood ratio  
 ในส่วน Print เลือก Case processing summary , Descriptive statistics, Model information , Goodness of fit statistics , Model summary statistics และ Parameter estimates

- คลิก OK

Computer output ที่ได้มีแสดงตามลำดับดังต่อไปนี้

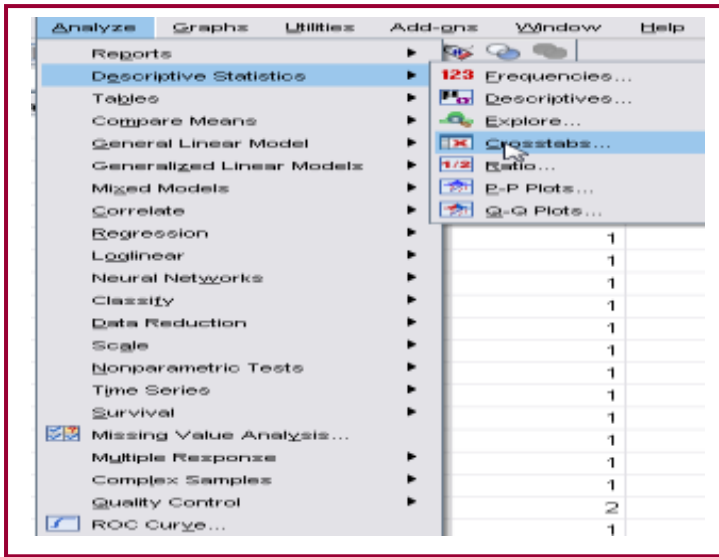
	N	Percent
Included	200	100.0%
Excluded	0	.0%
Total	200	100.0%

ขนาดของกลุ่มตัวอย่างที่ใช้เท่ากับ 200

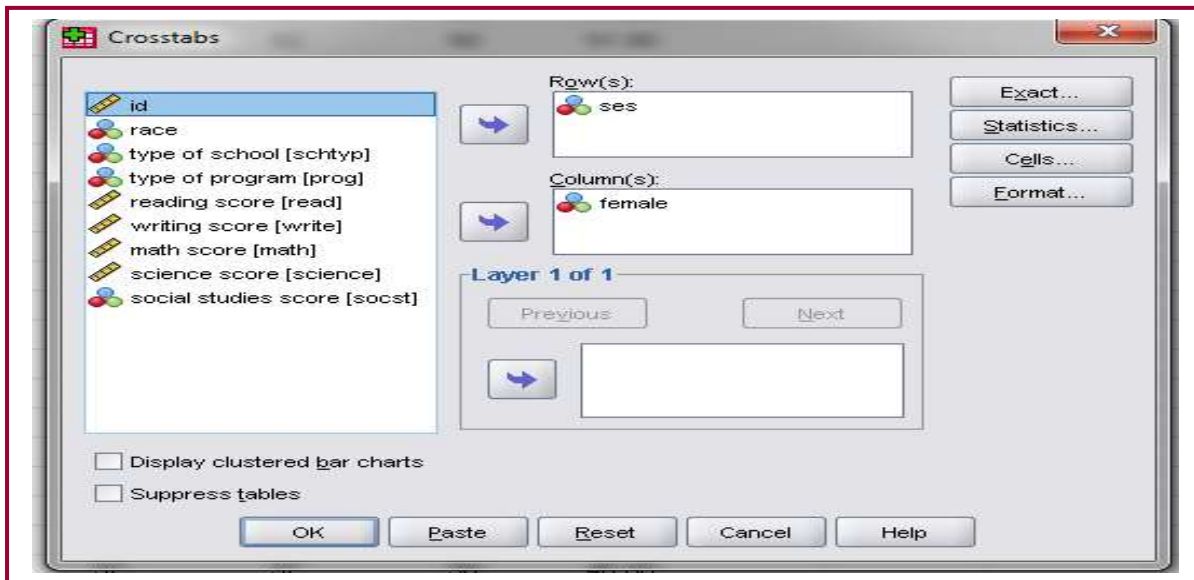
			N	Percent
Dependent Variable	ses	low	47	23.5%
		middle	95	47.5%
		high	58	29.0%
		Total	200	100.0%
Factor	female	male	91	45.5%
		female	109	54.5%
		Total	200	100.0%

ตารางด้านซ้ายแสดงให้เห็นค่าความถี่ของตัวแปรตามและตัวแปรที่เป็น factor อย่างไรก็ตามนี่ไม่ใช่ factor space ที่จะใช้ตรวจสอบว่าแต่ละ cell มีค่าความถี่เพียงพอหรือไม่ หากไม่เพียงพอ เราจะไม่สามารถเชื่อมั่นในผลที่ได้จาก goodness of fit test

- คลิก Analyze>Descriptive Statistics>Crosstabs



- Crosstabs submenu จะปรากฏขึ้น คลิกตัวแปรตามให้ไปอยู่ในช่อง Row คลิกตัวแปรที่เป็น factor (ในที่นี้ female) ให้ไปอยู่ในช่อง column



- คลิก OK เพื่อดูการแจกแจงความถี่ในแต่ละ cell

### ses \* female Crosstabulation

Count		female		Total
		male	female	
ses	low	15	32	47
	middle	47	48	95
	high	29	29	58
Total		91	109	200

ตารางด้านซ้ายแสดงให้เห็นการกระจายของค่าความถี่ใน factor space ในที่นี้แต่ละ cell มีความถี่ที่ไม่ต่ำกว่า 5 ดังนั้นปัญหาความกังวลเรื่อง adequacy of cell size จะหมดไป

### Continuous Variable Information

		N	Minimum	Maximum	Mean	Std. Deviation
Covariate	science score	200	26	74	51.85	9.901
	social studies score	200	26.00	71.00	52.4050	10.73579

ตารางด้านซ้ายแสดงให้เห็นค่าสถิติของ covariates

### Goodness of Fit<sup>a</sup>

	Value	df	Value/df
Deviance	289.148	281	1.029
Scaled Deviance	289.148	281	
Pearson Chi-Square	290.757	281	1.035
Scaled Pearson Chi-Square	290.757	281	
Log Likelihood <sup>b</sup>	-194.802		
Akaike's Information Criterion (AIC)	399.605		
Finite Sample Corrected AIC (AICC)	399.914		
Bayesian Information Criterion (BIC)	416.096		
Consistent AIC (CAIC)	421.096		

Dependent Variable: ses

Model: (Threshold), female, science, socst

a. The full log likelihood function is displayed and used in computing information criteria.

b. Information criteria are in small-is-better form.

คำอธิบาย : ตารางด้านบนแสดงให้เห็น Goodness of fit table ในที่นี้สถิติที่เกี่ยวข้องมีสองตัวคือ

- 1) ratio ของ deviance กับ degree of freedom ควรมีค่าใกล้เคียง 1 สำหรับ model ที่ fit
- 2) ratio ของ Pearson Chi-square กับ degree of freedom ควรมีค่าใกล้เคียง 1 สำหรับ model ที่ fit

หากสูงกว่าแสดงว่า model ที่ใช้ไม่ถูกต้อง (miss-specified) หรือมีค่าที่อยู่นอกกลุ่ม (outliers)

ในที่นี้ ratios ทั้งสองมีค่าใกล้เคียง 1 แสดงให้เห็นว่าการวิเคราะห์เหมาะสม

**Omnibus Test<sup>a</sup>**

Likelihood Ratio Chi-Square	df	Sig.
31.580	3	.000

Dependent Variable: ses  
Model: (Threshold), female, science, socst

a. Compares the fitted model against the thresholds-only model.

Omnibus test แสดงให้เห็นว่า full model โดยใช้ intercept และตัวแปรให้ผลแตกต่างอย่างมีนัยสำคัญจาก intercept-only หรือ null model

**Tests of Model Effects**

Source	Type III		
	Likelihood Ratio Chi-Square	df	Sig.
female	2.998	1	.083
science	3.308	1	.069
socst	12.696	1	.000

Dependent Variable: ses  
Model: (Threshold), female, science, socst

Test of model effects แสดงให้เห็นว่าเพศและคะแนนวิชาวิทยาศาสตร์ไม่มีผลต่อตัวแปรตาม แต่คะแนนวิชาสังคม (socst) มีผลอย่างมีนัยสำคัญ

**Parameter Estimates**

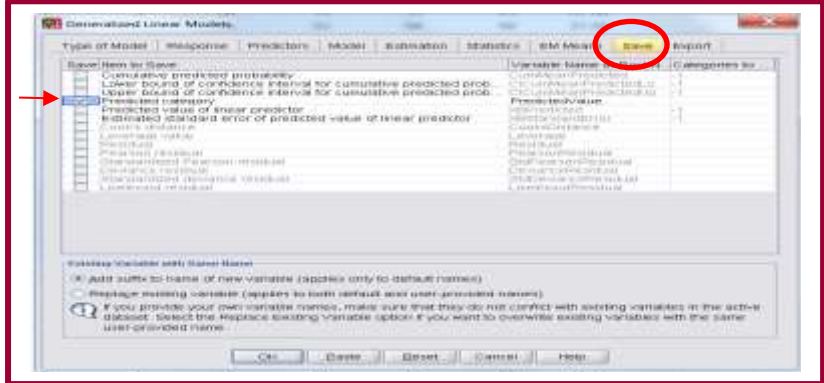
Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.
Threshold [ses=1]	3.237	.8571	1.557	4.917	14.264	1	.000
[ses=2]	5.588	.9250	3.775	7.401	36.492	1	.000
[female=0]	.482	.2797	-.066	1.031	2.975	1	.085
[female=1]	0 <sup>a</sup>						
science	.030	.0166	-.002	.063	3.276	1	.070
socst	.053	.0153	.023	.083	12.128	1	.000
(Scale)	1 <sup>b</sup>						

Dependent Variable: ses  
Model: (Threshold), female, science, socst

a. Set to zero because this parameter is redundant.  
b. Fixed at the displayed value.

Parameter Estimates แสดงให้เห็นผลของการใช้ Generalized Linear Models ในการแก้ปัญหา Ordered Logistic Regression จะสังเกตเห็นว่า ค่าสถิติที่ได้มีความใกล้เคียงกับผลที่ได้จากการวิเคราะห์โดยใช้ drop-down menu (Statistics Talks# 34 ) หรือ PLUM syntax command ( Statistics Talks# 35 )

๑) คลิกเลือก SAVE ใน submenu และระบุ Predicted category

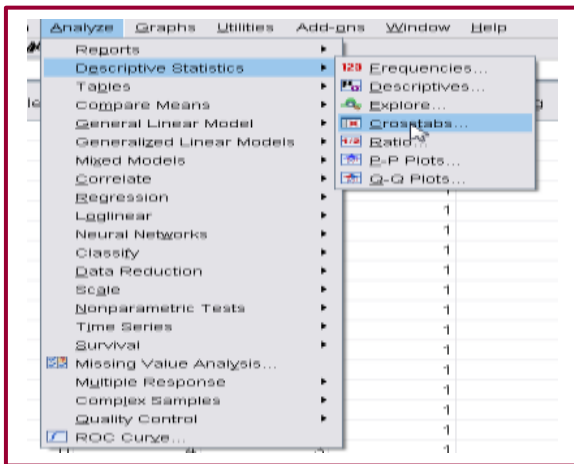


ตาราง data view จะมีข้อมูลเพิ่มขึ้นหนึ่ง column แสดง predicted value ดังภาพต่อไปนี้

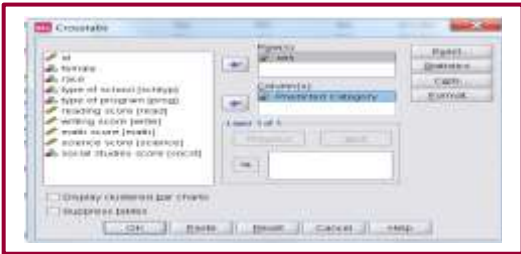
Ses ที่ได้จากการพยากรณ์

	id	female	race	ses	schtyp	prog	read	writ	math	science	socst	Predicted
1	73	0	4	1	1	1	57	52	61	47	57.00	2
2	121	1	4	2	1	3	65	59	53	63	61.00	2
3	86	0	4	3	1	1	44	33	54	58	31.00	2
4	141	0	4	1	1	3	63	44	47	53	50.00	2
5	172	0	4	2	1	2	47	52	57	53	61.00	2
6	113	0	4	2	1	2	44	52	51	63	61.00	3
7	50	0	3	2	1	1	50	58	42	53	61.00	2
8	11	0	1	2	1	2	34	46	45	39	36.00	2
9	84	0	4	2	1	1	63	57	54	59	51.00	2
10	46	0	3	2	1	2	57	56	52	50	51.00	2
11	75	0	4	2	1	3	60	46	51	53	61.00	2
12	60	0	4	2	1	3	57	65	51	63	61.00	3
13	95	0	4	3	1	2	73	60	71	61	71.00	3
14	104	0	4	3	1	2	54	60	57	55	46.00	2
15	38	0	3	1	1	2	45	57	50	31	56.00	2
16	115	0	4	1	1	1	42	49	43	50	50.00	2
17	76	0	4	1	1	2	47	52	51	50	50.00	2
18	125	0	4	2	2	1	57	57	60	58	56.00	2
19	114	0	4	3	1	2	68	66	62	55	61.00	3
20	85	0	4	2	1	1	55	38	57	53	40.00	2
21	167	0	4	2	1	1	63	48	35	66	41.00	2
22	143	0	4	2	1	3	63	63	75	72	66.00	3
23	41	0	3	2	1	2	50	48	45	55	56.00	2
24	30	0	1	3	1	2	60	52	57	61	61.00	3
25	12	0	1	2	1	3	57	44	45	39	46.00	2
26	53	0	5	2	1	3	54	37	46	39	31.00	1
27	154	0	4	3	1	2	65	66	66	61	66.00	3
28	178	0	4	2	2	3	47	57	57	58	46.00	2
29	196	0	4	3	2	2	44	38	49	39	46.00	2
30	29	0	2	1	1	1	52	44	49	55	41.00	2

● คลิก Analyze>Descriptives>Crosstabs ตามภาพด้านล่าง



Submenu crosstabs จะปรากฏขึ้น ให้คลิก ses ไปอยู่ในช่อง row และ predicted category ไปอยู่ในช่อง column คลิก OK



ภาพถัดไปแสดงตารางไขว้ ( cross tabulation )

Count		Predicted Category Value			Total
		low	middle	high	
ses	low	11	35	1	47
	middle	8	74	13	95
	high	1	39	18	58
Total		20	148	32	200

ความถูกต้องในการพยากรณ์  
 $= (11+74+18) \times 100 / 200$   
 $= 51.5\%$

จะเห็นได้ว่า ผลของการพยากรณ์ไม่แตกต่างไปจากผลที่ได้ตาม drop-down menu ( Statistics Talks#34) และ

PLUM syntax command ( Statistics Talks#35)

Contribution this issue : ดร. ดนัย ปัตตพงษ์

อยากเรียนรู้การนำสถิติข้างต้นนี้ไปใช้ในการวิจัยระดับสารนิพนธ์ (independent study)  
วิทยานิพนธ์ (thesis) ดุษฎีนิพนธ์(dissertation) ปรึกษาได้ที่ dpattaphongse@gmail.com

- \* ผู้แต่ง MBA's Made Easy (160+ issues) เอกสารวิชาการด้านศาสตร์การบริหารธุรกิจที่ช่วยให้ธุรกิจสามารถยืนหยัดและอยู่รอดได้ในภาวะที่โลกเปลี่ยนแปลงอยู่ตลอดเวลา
- \* ผู้พัฒนา FINALYSIS... a dedicated software สำหรับให้บริการนักธุรกิจที่ต้องการวิเคราะห์ความเป็นไปได้ทางการเงินของโครงการพัฒนาอสังหาริมทรัพย์ (บ้านจัดสรร/จัดสรรที่ดินเพื่อการอุตสาหกรรม/อาคารชุด/อาคารสำนักงานให้เช่า) โรงแรม โรงพยาบาลเอกชน ห้างสรรพสินค้า โรงงานน้ำตาล โรงงานกระดาษ โรงไฟฟ้าชีวมวล ฯลฯ ได้เห็นตัวเลขก่อนโครงการเกิด หลีกเลี่ยงความผิดพลาดเป็นร้อยเป็นพันล้านหากเกิดการลงทุนจริง(กำหนด DEBUT 1 เมษายน 2569)
- \* ผู้แต่งหนังสือ"การวิเคราะห์ความเป็นไปได้ทางการเงินและการจัดวงเงินเครดิตของโครงการลงทุน"ประกอบด้วยตัวอย่างของธุรกิจจริงที่ไม่เปิดเผยชื่อนับ 100 บริษัท ครอบคลุมอุตสาหกรรม 24 อุตสาหกรรม
- \* Co-developer ซอฟต์แวร์ en@gex@cel<sup>®</sup> สำหรับใช้ทดสอบ/เรียนรู้ศัพท์(ประกอบด้วยแบบฝึกหัดและเฉลยกว่า 90 บทครอบคลุมศัพท์ระดับ SAT/IELTS/TOEFL กว่า 12,000 คำ) และไวยากรณ์อังกฤษ (ประกอบด้วยแบบฝึกหัดและเฉลยกว่า 160 บทหรือกว่า 10,000 ข้อครอบคลุมเนื้อหาระดับอุดมศึกษาและTOEFL) มาพร้อมกับไฟล์เสียง/ไฟล์ข้อมูล/ ฯลฯ อีกมาก(กำหนด DEBUT 1 เมษายน 2569)