

Canonical Correlation Workshop#1-Manova syntax command

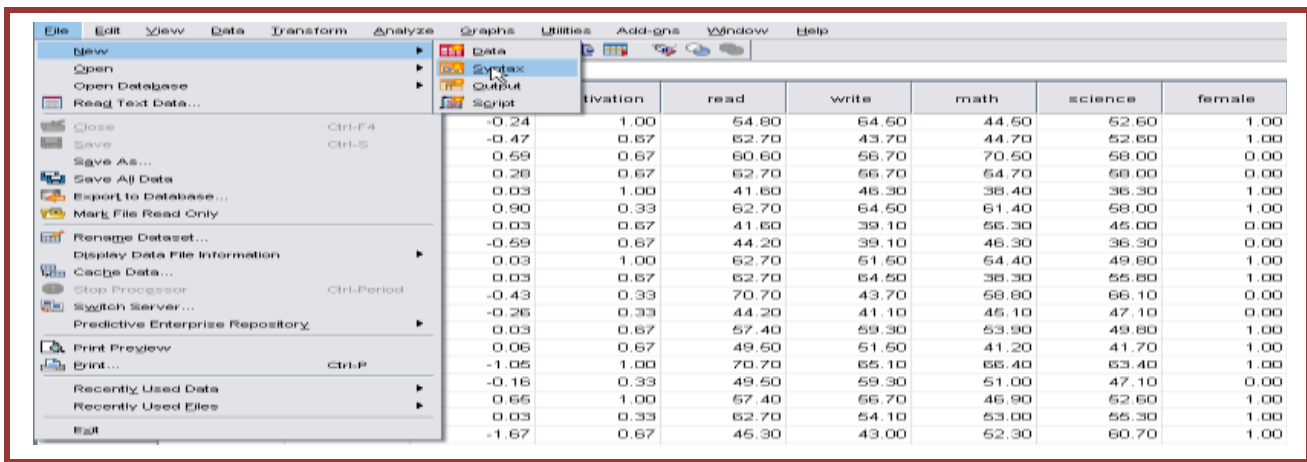
ตารางด้านล่างแสดงข้อมูลที่ต้องวิเคราะห์โดยใช้ Canonical correlation ในที่นี้ข้อมูลจะแบ่งออกเป็นสองกลุ่ม โดยกลุ่มแรกจะประกอบด้วยข้อมูลทางจิตวิทยา(psychological factors) ประกอบด้วยตัวแปรสามตัว ได้แก่ locus_of_control self_concept และ motivation ส่วนกลุ่มที่สองจะประกอบด้วยตัวแปรห้าตัวที่แสดงคะแนนผลการทดสอบทางวิชาการ ได้แก่ read write math science และ female ขนาดของกลุ่มตัวอย่างทั้งหมดเท่ากับ 600 ราย

	id	locus_of_control	self_concept	motivation	read	write	math	science	female
1	303.00	-0.84	-0.24	1.00	54.80	64.50	44.50	52.60	1.00
2	404.00	-0.38	-0.47	0.67	62.70	43.70	44.70	52.60	1.00
3	225.00	0.89	0.59	0.67	60.60	56.70	70.50	58.00	0.00
4	553.00	0.71	0.28	0.67	62.70	56.70	54.70	58.00	0.00
5	433.00	-0.64	0.03	1.00	41.60	46.30	38.40	36.30	1.00
6	189.00	1.11	0.90	0.33	62.70	64.50	61.40	58.00	1.00
7	336.00	0.06	0.03	0.67	41.60	39.10	56.30	45.00	0.00
8	321.00	-0.91	-0.59	0.67	44.20	39.10	46.30	36.30	0.00
9	358.00	0.45	0.03	1.00	62.70	51.50	54.40	49.60	1.00
10	422.00	0.00	0.03	0.67	62.70	64.50	38.30	55.80	1.00
11	461.00	0.24	-0.43	0.33	70.70	43.70	58.80	66.10	0.00
12	593.00	-1.09	-0.26	0.33	44.20	41.10	45.10	47.10	0.00
13	346.00	0.46	0.03	0.67	57.40	59.30	53.90	49.60	1.00
14	531.00	0.68	0.06	0.67	49.50	51.50	41.20	41.70	1.00
15	492.00	-0.14	-1.05	1.00	70.70	65.10	66.40	63.40	1.00
16	582.00	0.10	-0.16	0.33	49.50	59.30	51.00	47.10	0.00
17	102.00	0.45	0.65	1.00	57.40	56.70	46.90	52.60	1.00
18	249.00	-1.56	0.03	0.33	62.70	54.10	53.00	55.30	1.00
19	389.00	-1.02	-1.67	0.67	45.30	43.00	52.30	60.70	1.00
20	324.00	0.94	0.03	1.00	68.00	56.70	59.60	58.00	1.00
21	401.00	-0.40	0.03	1.00	44.20	54.10	59.30	58.00	1.00
22	550.00	-0.19	-1.73	1.00	44.70	47.60	50.30	39.00	1.00
23	25.00	0.26	0.03	0.33	48.40	48.90	52.20	43.90	0.00
24	18.00	0.21	0.94	0.00	38.90	48.90	41.80	58.50	0.00
25	160.00	0.70	0.87	1.00	46.90	54.10	46.40	49.80	1.00
26	470.00	1.36	0.94	1.00	57.40	64.50	49.90	55.30	1.00
27	322.00	-0.84	-0.57	0.33	33.60	33.30	41.00	36.30	0.00
28	396.00	-0.68	0.03	0.33	38.90	38.50	41.20	39.00	1.00
29	84.00	-0.17	0.31	1.00	65.40	67.10	75.50	71.50	0.00
30	473.00	0.06	0.28	1.00	60.10	59.30	56.20	60.70	1.00
31	491.00	-0.14	0.56	1.00	54.80	61.90	54.60	47.10	1.00
32	544.00	0.96	0.03	0.67	44.20	51.50	43.20	44.40	1.00
33	476.00	0.23	-0.26	0.00	52.10	61.90	56.30	58.00	1.00
34	163.00	0.66	0.00	1.00	52.10	48.90	48.50	44.40	1.00
35	171.00	0.08	0.59	1.00	57.40	56.70	53.90	60.70	1.00
36	29.00	-0.90	0.03	0.67	35.30	44.30	36.10	33.60	1.00
37	233.00	0.06	0.03	0.67	62.70	61.90	58.80	60.70	1.00

เนื่องจากในโปรแกรมทางสถิติที่ใช้วิเคราะห์เวอร์ชันนี้ไม่มี menu ที่สามารถคลิกเลือกเหมือนการวิเคราะห์ทางสถิติอื่นๆ

จำเป็นต้องอาศัยการสั่งงานโปรแกรมผ่านทางsyntax (ภาษาที่ใช้สั่งการ-command language) โดยในที่นี้จะใช้ manova command ในการวิเคราะห์

- คลิกที่ปุ่ม File>New จะปรากฏ Menu ดังภาพล่าง



- คลิกที่ Syntax จะปรากฏ syntax editor ให้พิมพ์ข้อความต่อไปนี้

```
manova locus_of_control self_concept motivation with read write math science female  
/discrim all apha(1)  
/print =sig(eigen dim).
```

หมายเหตุ:- manova เป็นคำสั่งcommandของโปรแกรมสำเร็จรูป ตัวแปรที่ตามหลังสามตัวจะเป็นตัวแปรกลุ่มที่หนึ่ง

ตัวแปรที่ตามหลัง with ห้าตัวจะเป็นตัวแปรกลุ่มที่สอง

ตัวแปรกลุ่มที่หนึ่งเป็นกลุ่มตัวแปรตาม(dependent /criterion variable) ตัวแปรกลุ่มที่สองเป็นกลุ่มตัวแปร

อิสระ(independent / predictor variables) หากสลับกัน ผู้วิเคราะห์จะเกิดความสับสนเอง เพราะ format

ของ manova output จะเริ่มจากกลุ่ม dependent variables ตามด้วย covariates ซึ่งก็คือ independent variables

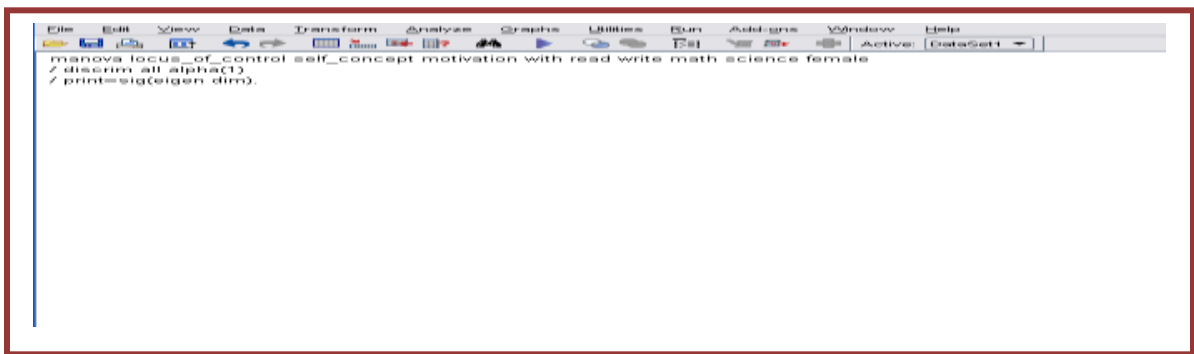
Discrim เป็น subcommand ในที่นี้กำหนดให้แสดงค่า significance level=1

Print เป็น subroutine สั่งให้พิมพ์ eigenvalue

/ เป็นเครื่องหมายกั้นระหว่างบรรทัด

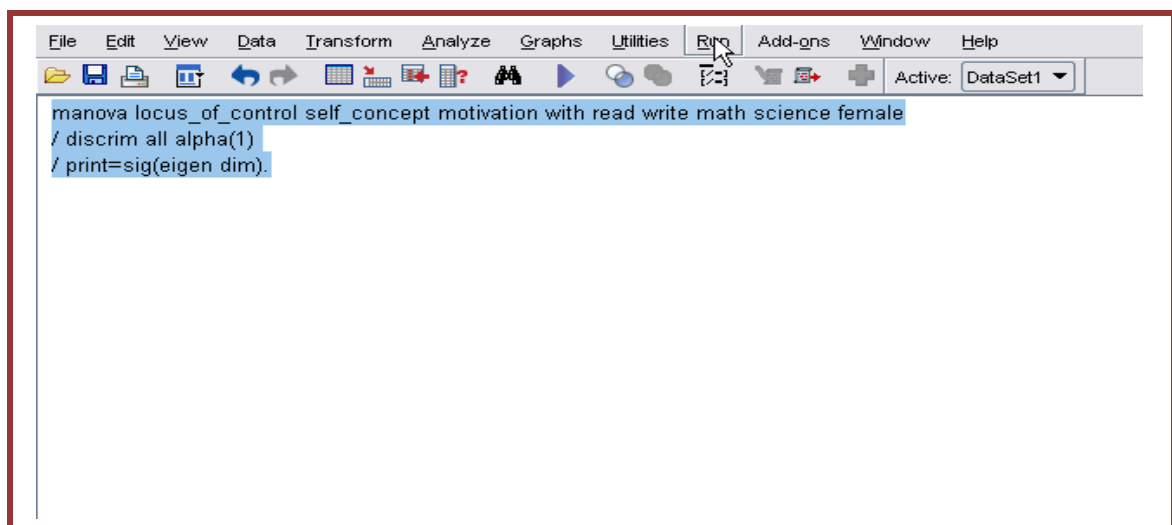
เมื่อจบคำสั่งให้ใช้เครื่องหมาย .

ถ้าทำตามนี้ syntax editor ที่ท่านกรอกข้อมูล จะปรากฏดังภาพต่อไปนี้



```
manova locus_of_control self_concept motivation with read write math science female
/ discrim all alpha(1)
/ print=sig(eigen dim).
```

จากนั้นคลิก Run>All จะปรากฏแถบงานข้อความทั้งหมดใน syntax editor ตามด้วย computer output



```
manova locus_of_control self_concept motivation with read write math science female
/ discrim all alpha(1)
/ print=sig(eigen dim).
```

***** Analysis of Variance -- Design 1 *****

EFFECT .. WITHIN CELLS Regression

Multivariate Tests of Significance (S = 3, M = 1/2, N = 295)

Test Name	Value	Approx. F	Hypoth. DF	Error DF	Sig. of F
Pillais	.25425	11.00057	15.00	1782.00	.000
Hotellings	.31430	12.37633	15.00	1772.00	.000
Wilks	.75436	11.71573	15.00	1634.65	.000
Roys	.21538				

Eigenvalues and Canonical Correlations

Root No.	Eigenvalue	Pct.	Cum. Pct.	Canon Cor.	Sq. Cor
1	.27450	87.33628	87.33628	.46409	.21538
2	.02887	9.18537	96.52164	.16751	.02806
3	.01093	3.47836	100.00000	.10399	.01081

Dimension Reduction Analysis

Roots	Wilks L.	F	Hypoth. DF	Error DF	Sig. of F
1 TO 3	.75436	11.71573	15.00	1634.65	.000
2 TO 3	.96143	2.94446	8.00	1186.00	.003
3 TO 3	.98919	2.16461	3.00	594.00	.091

EFFECT .. WITHIN CELLS Regression (Cont.)

Univariate F-tests with (5,594) D. F.

Variable	Sq. Mul. R	Adj. R-sq.	Hypoth. MS	Error MS	F	Sig. of F
locus_of	.18062	.17372	9.72160	.37123	26.18789	.000
self_con	.01957	.01131	1.16669	.49212	2.37076	.038
motivati	.07874	.07098	1.10799	.10913	10.15338	.000

• Multivariate tests หรือ Omnibus tests ใช้ในการทดสอบสมมติฐานว่า canonical correlation ทุกตัวมีค่าเป็นศูนย์หรือไม่ โดยใช้ Pillais/Hotellings/Wilks พบว่าปฏิเสธ null hypothesis ที่ว่า canonical correlation ทุกตัวมีค่าไม่แตกต่างไปจากศูนย์ที่ระดับนัยสำคัญ $\alpha=0.05$

• Wilk's lamda แสดง Variance ที่แบบจำลองไม่สามารถอธิบายได้ (model สามารถอธิบาย variance ได้เพียงร้อยละ 24.56) $[(1-.75436) \times 100]$

Variate คู่แรกมี canonical correlation² = .215

Variate คู่ที่สองมี canonical correlation² = .028

Variate คู่ที่สามมี canonical correlation² = .011

Dimension Reduction Analysis แสดงให้เห็นผลจากการทดสอบสมมติฐานว่า root หนึ่งๆ กับที่เหลือมีค่าเป็นศูนย์หรือไม่พบว่า root (function) แรกเท่านั้นที่แตกต่างจากศูนย์อย่างมีนัยสำคัญทางสถิติ โดย

$F(15, 1634.65) = 11.71573, P < 0.05$

และ root แรกเท่านั้นที่สมควรได้รับการพิจารณา (variance ของ variate คู่แรกมีร่วมกัน 21.5%)

Raw canonical coefficients for DEPENDENT variables
Function No.

Variable	1	2	3
locus_of	1.25383	-.62148	.66169
self_con	-.35135	-1.18769	-.82672
motivati	1.26242	2.02726	-2.00023

Standardized canonical coefficients for DEPENDENT variables
Function No.

Variable	1	2	3
locus_of	.84042	-.41656	.44352
self_con	-.24788	-.83793	-.58326
motivati	.43267	.69480	-.68554

ในส่วนนี้จะให้ข้อมูลเกี่ยวกับ canonical weights (ความหมายคล้ายคลึงกับ Beta coefficient ในการวิเคราะห์การถดถอยพหุคูณ-multiple regression analysis) ซึ่งเป็นข้อมูลที่จำเป็นหากต้องการคำนวณหา canonical scores

Raw canonical coefficients for COVARIATES
Function No.

COVARIATE	1	2	3
read	.04462	-.00491	-.02138
write	.03588	.04207	-.09131
math	.02342	.00423	-.00940
science	.00503	-.08516	.10984
female	.63212	1.08464	1.79465

Standardized canonical coefficients for COVARIATES
CAN. VAR.

COVARIATE	1	2	3
read	.45080	-.04961	-.21601
write	.34896	.40921	-.88810
math	.22047	.03982	-.08848
science	.04878	-.82660	1.06608
female	.31504	.54057	.89443

นักสถิติมักจะดเว้นไม่ใช้ข้อมูลส่วนนี้ เนื่องด้วยหากมี multicollinearity ในระหว่างตัวแปรจะมีผลทำให้ canonical coefficient มีค่าไม่คงที่ (unstable) และค่าที่ได้จะมีความเพี้ยน (distortion) เกิดขึ้น หรือหากมีกลุ่มตัวอย่างเพิ่มขึ้นค่า canonical coefficient ก็จะไปเปลี่ยนไป

ข้อมูลก่อนนให้ข้อมูลเกี่ยวกับ canonical weights ของ dependent canonical variate ส่วนข้อมูลท่อนล่างให้ข้อมูลเกี่ยวกับ canonical weights ของ covariate/independent canonical variate

Correlations between DEPENDENT and canonical variables
Function No.

Variable	1	2	3
locus_of	.90405	-.38969	.17562
self_con	.02084	-.70874	-.70516
motivati	.56715	.35089	-.74513

ในส่วนข้อมูลของตัวแปรตาม หากใช้เกณฑ์ Hair et al.'s factor loadings threshold ที่ $n=350$ loadings 0.3 หรือมากกว่าจะมีนัยสำคัญทางสถิติ ในที่นี้ Locus_of_control motivation มีส่วนสำคัญในการกำหนด root ที่ 1

Variance in dependent variables explained by canonical variables

CAN. VAR.	Pct Var DEP	Cum Pct DEP	Pct Var COV	Cum Pct COV
1	37.97982	37.97982	8.17994	8.17994
2	25.90966	63.88948	.72701	8.90694
3	36.11052	100.00000	.39050	9.29745

%ของ variance ระหว่างตัวแปรตามที่สามารถอธิบายได้โดย dependent canonical variate และ cov. canonical variate เท่ากับ 37.98% และ 8.18% ตามลำดับ

Correlations between COVARIATES and canonical variables
CAN. VAR.

Covariate	1	2	3
read	.84045	-.35883	-.13536
write	.87654	.06484	-.25456
math	.76395	-.29795	-.14776
science	.65841	-.67680	.23036
female	.36411	.75493	.54340

Variance in covariates explained by canonical variables

CAN. VAR.	Pot Var DEP	Cum Pot DEP	Pot Var COV	Cum Pot COV
1	11.30458	11.30458	52.48769	52.48769
2	.70132	12.00590	24.99409	77.48177
3	.09804	12.10394	9.06617	86.54795

ในส่วนของข้อมูลของตัวแปรอิสระ(covariates) write read math science และ femaleมีส่วนสำคัญในการกำหนด root ที่ 1

%ของ variance ระหว่าง covariates ที่สามารถอธิบายได้โดย dep. canonical variate และcov. canonical variateเท่ากับ 11.30% และ 52.49% ตามลำดับ

Canonical correlation analysis แสดงให้เห็นว่า locus_of_control motivation มีความเกี่ยวข้องกับ read write math science และ female

Communality

เราสามารถคำนวณหา communality ได้โดยอาศัย structure correlation coefficients(Str.) มาสร้างตารางต่อไปนี้

Variable	Str.1	Sq.Str.1	Communality
Locus_of_control	.90405*	.817	.817
Self_concept	.02084	.000	.000
Motivation	.56715*	.322	.322
read	.84045*	.706	.706
write	.87654*	.768	.768
math	.76395*	.584	.584
science	.65841*	.434	.434
female	.36411*	.133	.133

Communalities แสดงให้เห็นตัวแปรเกือบทุกตัวยกเว้น self_concept และ femaleมีค่าสูง บ่งบอกให้เห็นว่า variance ของตัวแปรเกือบทั้งหมดสามารถอธิบายได้จากcanonical function ที่ 1

Variate adequacy coefficient

Root ที่หนึ่งสามารถอธิบาย variance ในระหว่างตัวแปรตามได้ 37.98% และ variance ในระหว่างตัวแปรอิสระได้ 52.49%

Redundancy coefficient

8.18% ของ variance ในระหว่างตัวแปรตามสามารถอธิบายได้ด้วยตัวแปรอิสระตาม root ที่หนึ่ง ในขณะที่เดียวกัน 11.30% ของ variance ในระหว่างตัวแปรอิสระสามารถอธิบายได้ด้วยตัวแปรตามตาม root ที่หนึ่ง อื่นๆ pooled redundancy coefficients ของตัวแปรตามเป็น 9.30% ในขณะที่ของตัวแปรอิสระเป็น 12.10%

ผลการวิเคราะห์ Canonical correlation สามารถสรุปได้ดังตารางต่อไปนี้

Canonical correlation 1						
		CV1-1(8.18%)		CV2-1(11.30%)		
		37.98%		52.49%		
Set1	Loading		$R_c = 0.4641$		Set2	Loading
Locus-of-control	.90405*				Read	.84045*
Self-concept	.02084		$R_c^2 = 0.2154$		Write	.87654*
Motivation	.56715*				Math	.76395*
					Science	.65841*
					Female	.36411*
Wilk's Lambda: .75436						
p<.001						

ประโยชน์ของ Canonical correlation ไม่ใช่เพื่อทดสอบความมีนัยสำคัญทางสถิติหรือหาค่าสถิติที่สำคัญๆ แต่หากเป็นผลที่ได้จากการวิเคราะห์ ซึ่งสามารถสรุปได้ว่า :

Locus_of_control และ motivation มีความสัมพันธ์กับ read write math science และ female หากคะแนน read write math science สูง และเป็นเพศหญิง locus_of_control และ motivation ก็จะถูกจัดอยู่ในระดับสูงเช่นกัน

Contribution this issue: ดร. ดนัย ปัตตพงษ์

อยากเรียนรู้การนำสถิติข้างต้นนี้ไปใช้ในการวิจัยระดับสารนิพนธ์ (independent study) วิทยานิพนธ์ (thesis)

คุณวุฒินิพนธ์(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 หลีกเลี่ยงความผิดพลาดเป็นร้อยเป็นพันล้านบาทเกิดการลงทุนจริง(กำหนด DEBUT 1 เมษายน 2569)
- * ผู้แต่งหนังสือ”การวิเคราะห์ความเป็นไปได้ทางการเงินและการจัดวงเงินเครดิตของโครงการลงทุน”ประกอบด้วยตัวอย่างของธุรกิจจริงที่ไม่เปิดเผยชื่อนับ 100 บริษัท ครอบคลุมอุตสาหกรรม 24 อุตสาหกรรม
- * Co-developer ซอฟต์แวร์ en@gex@cel[®] 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)