



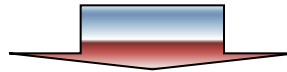
สหสัมพันธ์ (Correlation)



สหสัมพันธ์ (correlation)

เป็นวิธีการวัดค่าความสัมพันธ์ของตัวแปร 2 ตัว
ว่ามีความสัมพันธ์กันมากหรือน้อยเพียงใด

เป็นเทคนิคการวิเคราะห์ข้อมูลที่ใช้กับ



ข้อมูลเชิงปริมาณ (Quantitative Data)

ลักษณะของข้อมูล **Statistics for Research**

สถิติสำหรับงานวิจัย



คนที่	น้ำหนัก	ความสูง
1	y_1	x_1
2	y_2	x_2
.	.	.
.	.	.
.	.	.
n	y_n	x_n

ครั้งที่	ยอดขาย	ค่าใช้จ่ายในการโฆษณา
1	y_1	x_1
2	y_2	x_2
.	.	.
.	.	.
.	.	.
n	y_n	x_n

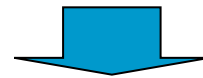
คนที่	เงินเดือน	อายุการทำงาน
1	y_1	x_1
2	y_2	x_2
.	.	.
.	.	.
.	.	.
n	y_n	x_n

ค่าตัวแปรสุ่มได้มาจากหน่วยตัวอย่างเดียวกัน ซึ่งอาจจะมีหน่วยเดียวกัน หรือ หน่วยต่างกันก็ได้



วิธีการวัดความสัมพันธ์

1. เขียนกราฟ



Scatter Plot

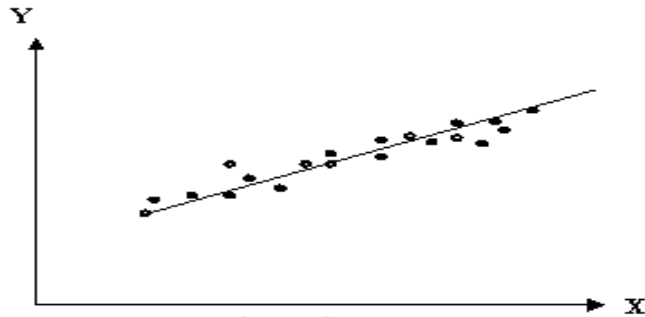
Y



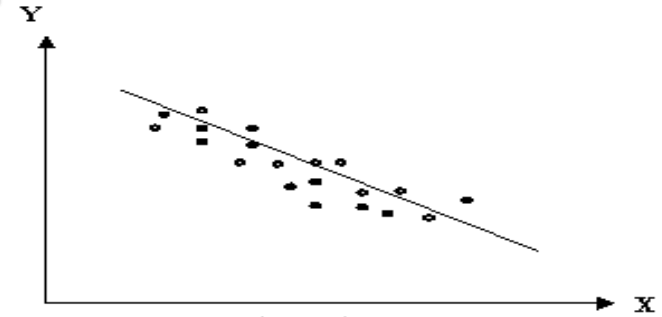
X



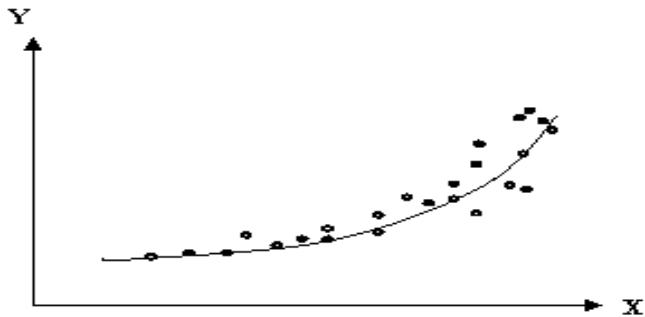
การตรวจสอบความเป็นเชิงเส้นของข้อมูล



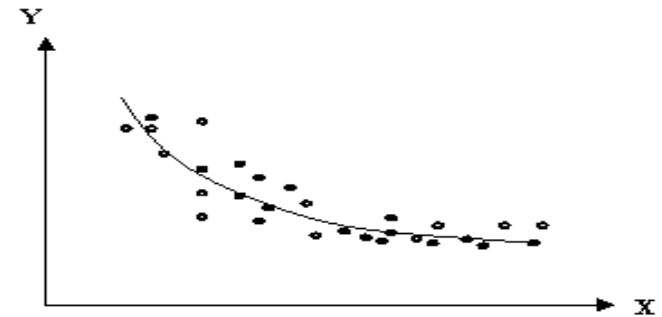
ก) ความสัมพันธ์แบบเส้นตรงทางบวก



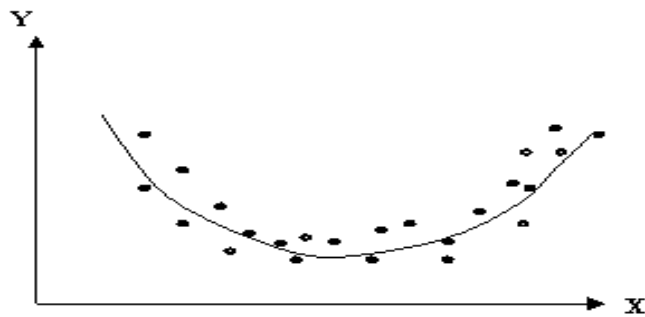
ข) ความสัมพันธ์แบบเส้นตรงทางลบ



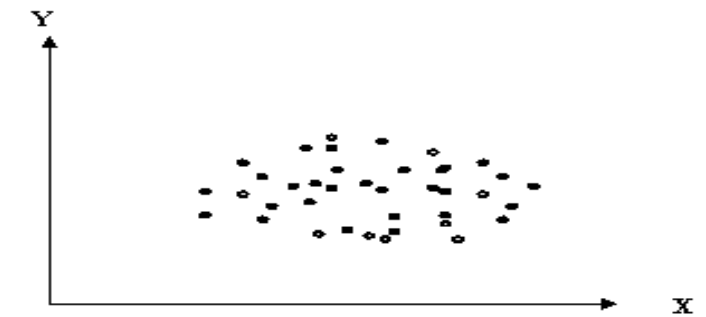
ค) ความสัมพันธ์แบบเอก ไปเนนเจียสทางบวก



ง) ความสัมพันธ์แบบเอก ไปเนนเจียสทางลบ



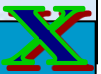
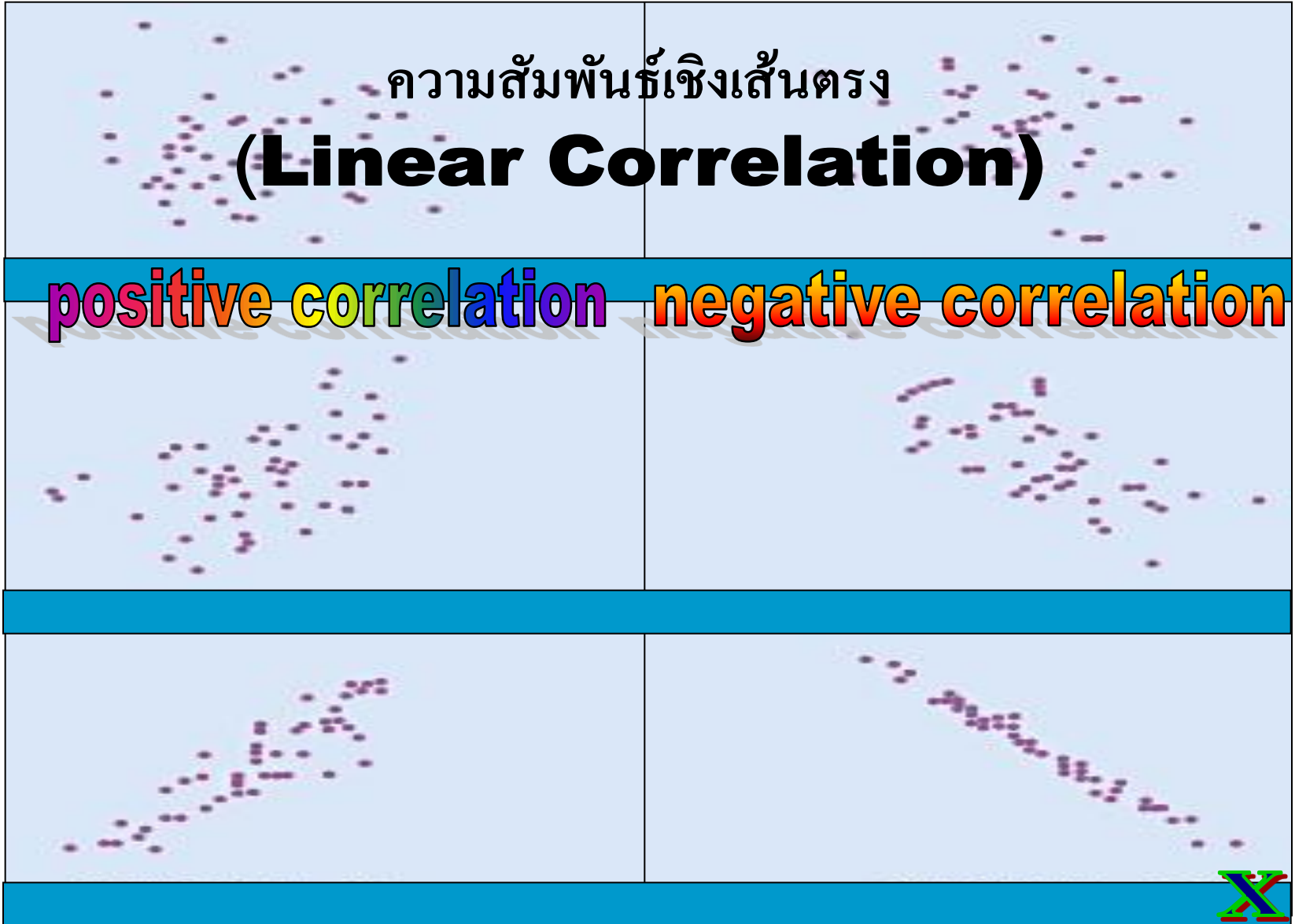
จ) ความสัมพันธ์แบบพาราโบลาโบล่าทางบวก



ฉ) ไม่มีความสัมพันธ์กัน

ความสัมพันธ์เชิงเส้นตรง
(Linear Correlation)

positive correlation negative correlation



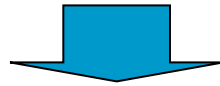
วิธีการวัดความสัมพันธ์

2. หาค่าตัวสถิติที่ใช้วัดความสัมพันธ์



สัมประสิทธิ์สหสัมพันธ์

(Correlation coefficient)



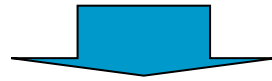
คาร์ล เพียร์สัน **(Karl Pearson)**



Pearson Product moment correlation Coefficient



Pearson's correlation Coefficient



เขียนแทนด้วยสัญลักษณ์

พารามิเตอร์ ρ, ρ_{xy} r, r_{xy} ตัวสถิติ

ข้อตกลงเบื้องต้น

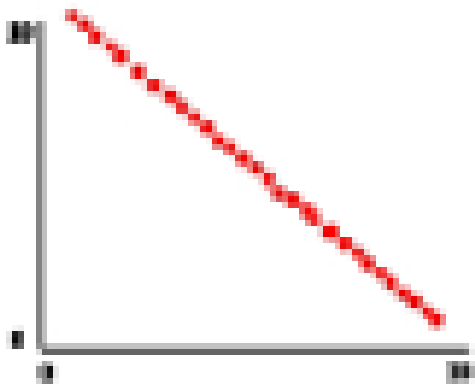
1. ตัวแปร 2 ตัว เป็นตัวแปรสุ่มแบบต่อเนื่อง หรือเป็นข้อมูลเชิงปริมาณ
2. ความสัมพันธ์ระหว่าง 2 ตัวแปร เป็นเส้นตรง (Linear Relationship)

ความหมายของ ρ

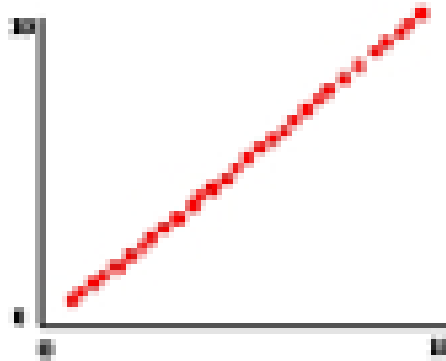
$$-1 \leq \rho \leq 1$$

1. ถ้า ρ เป็น -1 แสดงว่า x และ y มีความสัมพันธ์เชิงลบ
2. ถ้า ρ เป็น +1 แสดงว่า x และ y มีความสัมพันธ์เชิงบวก
3. ถ้า $\rho = 0$ แสดงว่า x และ y ไม่มีความสัมพันธ์เชิงเส้น

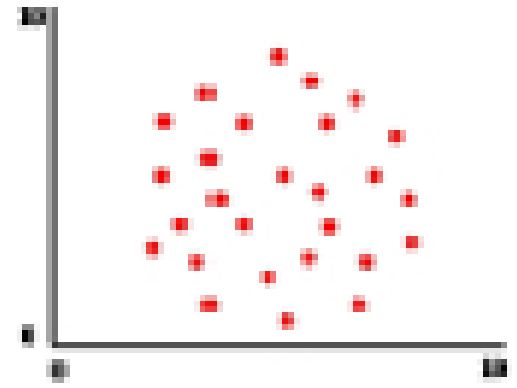
Correlation = -1.0



Correlation = 1.0

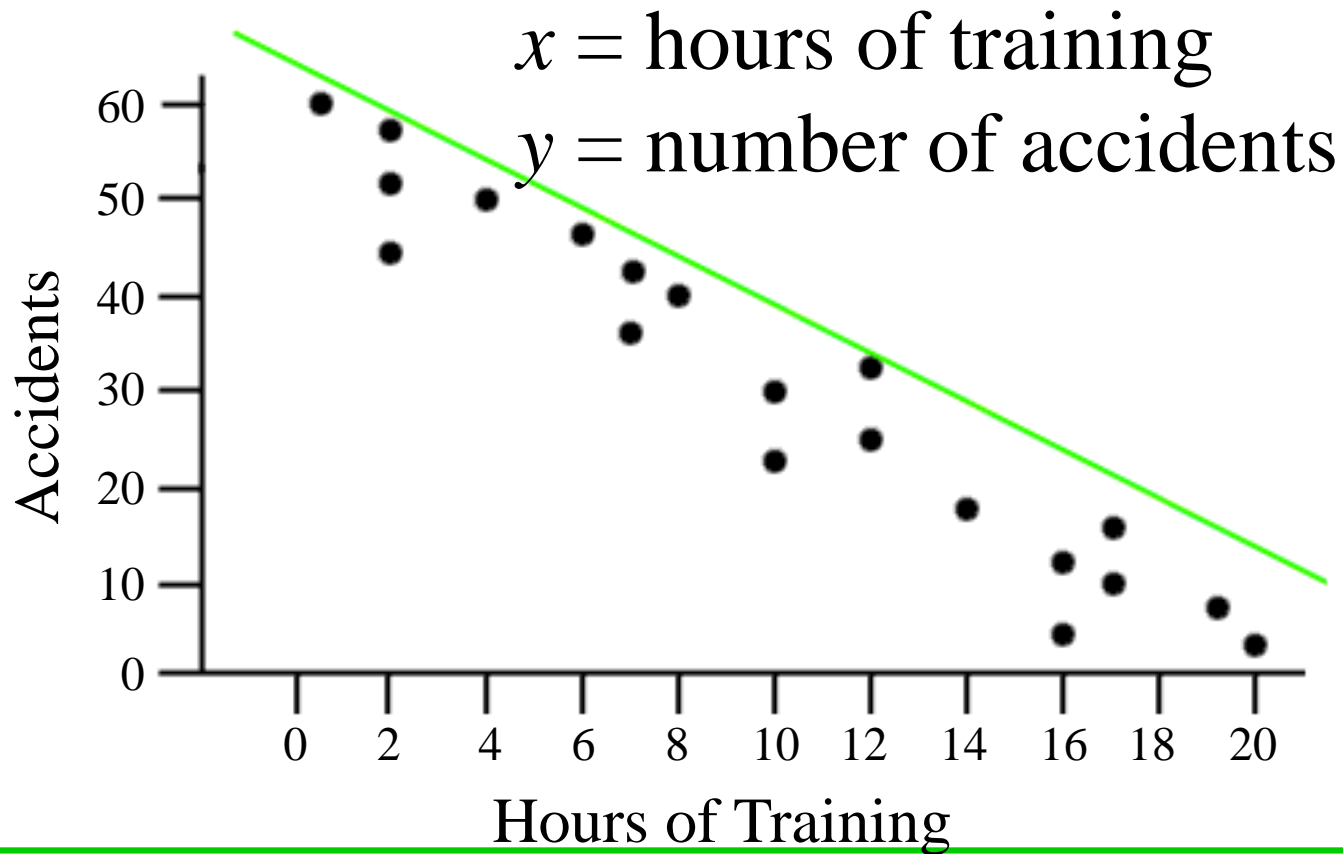


Correlation = 0





Scatter Plots and Types of Correlation

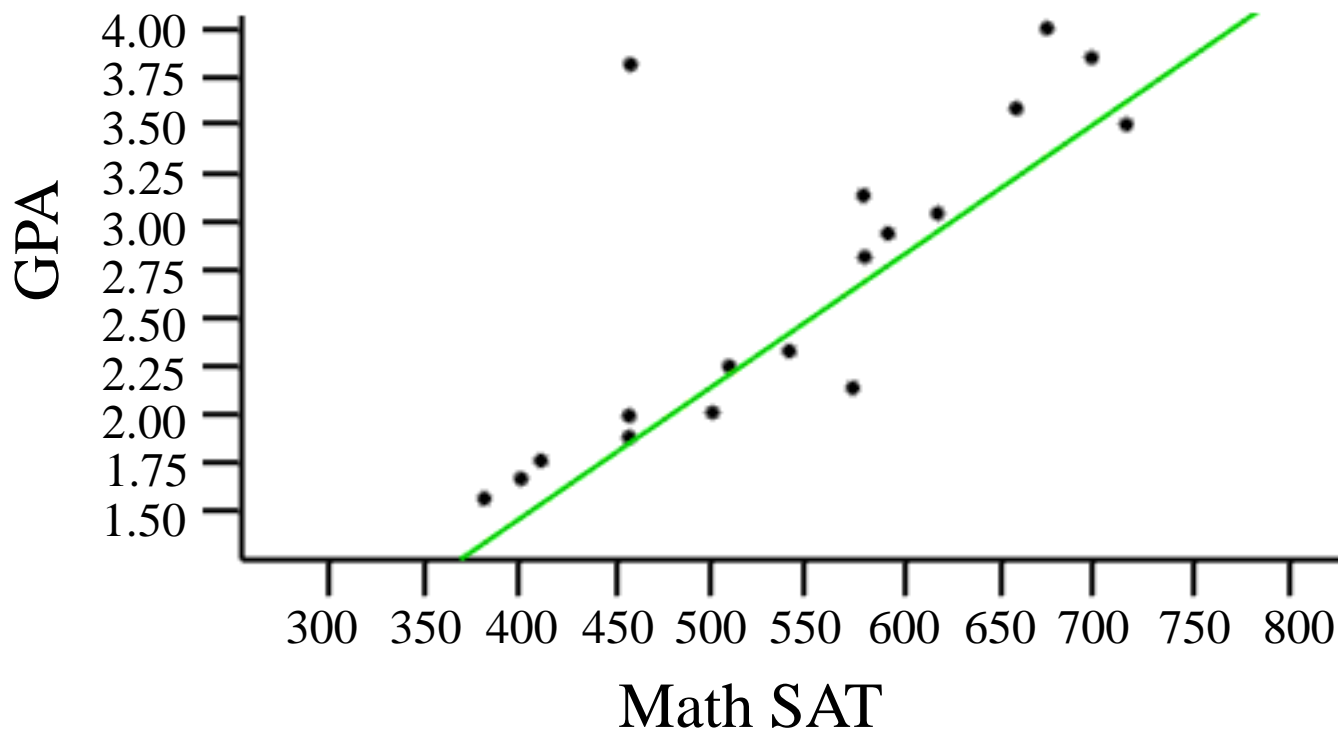


Negative Correlation— x increases, y decreases



$x = \text{SAT score}$

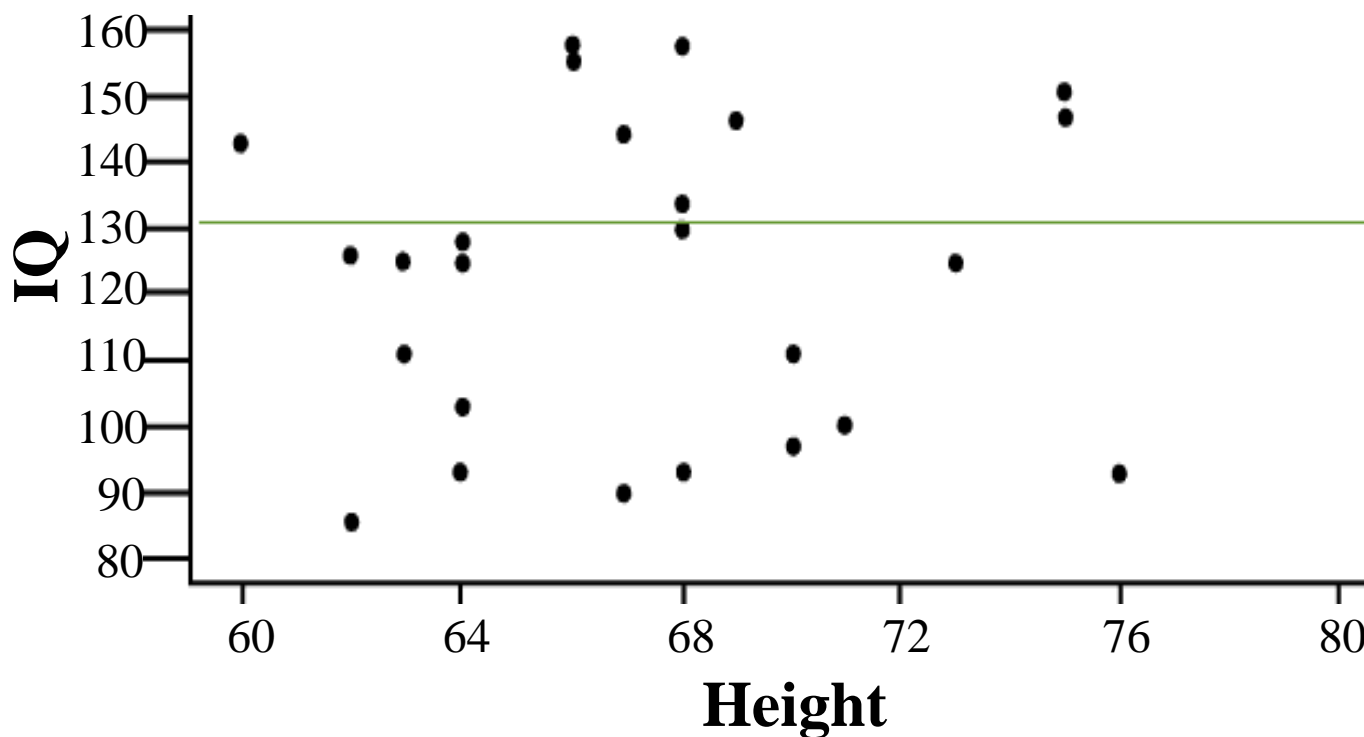
$y = \text{GPA}$



Positive Correlation—as x increases, y increases



$x = \text{height}$ $y = \text{IQ}$



No linear correlation

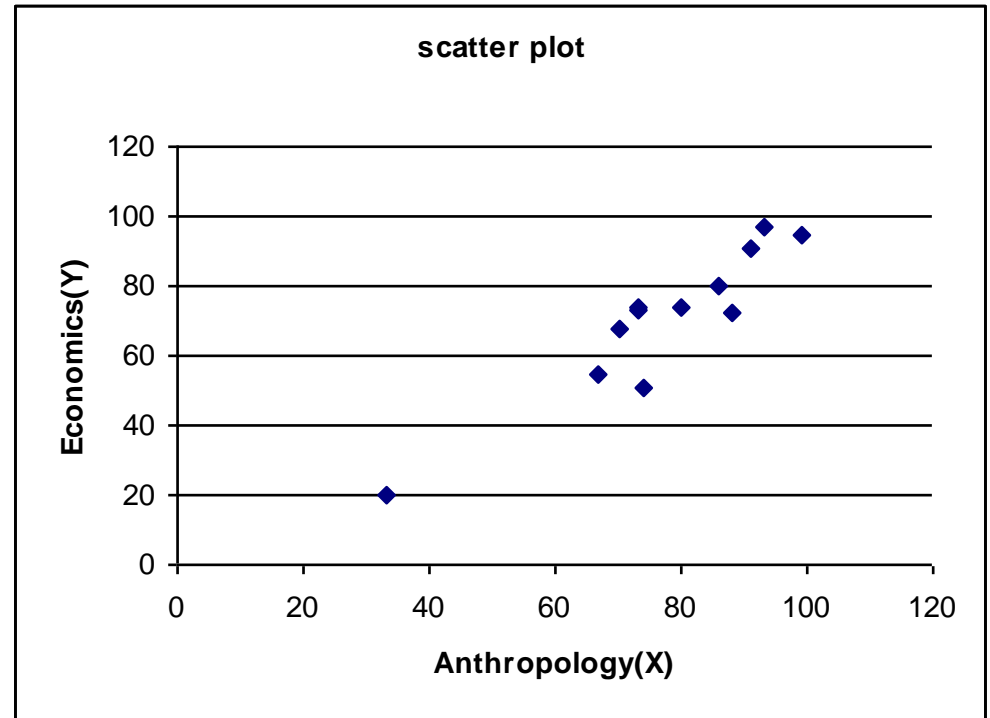


Correlation Coefficient Interpretation

Coefficient Range	Strength of Relationship
0.00 - 0.20	Very Low
0.21 - 0.40	Low
0.41 - 0.60	Moderate
0.61 - 0.80	High Moderate
0.81 - 1.00	Very High

ตัวอย่าง จากคะแนนการสอบของนักศึกษาวิชาเอกบริหารธุรกิจที่สอบวิชาเศรษฐศาสตร์ และมนุษยวิทยา 12 คน เป็นดังตาราง อยากทราบว่าคะแนนสอบทั้ง 2 วิชามีความสัมพันธ์กันเพียงใด

No.	Economics (Y)	Anthropology (X)
1	51	74
2	68	70
3	72	88
4	97	93
5	55	67
6	73	73
7	95	99
8	74	73
9	20	33
10	91	91
11	74	80
12	80	86



การคำนวณหาค่า ρ

$$\rho = \frac{\text{COV}(x, y)}{\sigma_x \sigma_y}$$

หรือ

$$\rho = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

โดยที่ ρ แทนสัมประสิทธิ์สหสัมพันธ์

$\text{COV}(x, y)$ หรือ σ_{xy} แทนความแปรปรวนรวมของตัวแปร x, y

σ_x, σ_y แทนส่วนเบี่ยงเบนมาตรฐานของตัวแปร x, y

$$\rho = \frac{\sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y)}{\sqrt{\sum_{i=1}^N (x_i - \mu_x)^2 \sum_{i=1}^N (y_i - \mu_y)^2}}$$



การคำนวณ

Statistics for Research

NO	Y	X	$(x_i - \mu_x)$	$(y_i - \mu_y)$	$(x_i - \mu_x)(y_i - \mu_y)$	$(x_i - \mu_x)^2$	$(y_i - \mu_y)^2$
1	y_1	x_1					
2	y_2	x_2					
.	.	.					
.	.	.					
.	.	.					
N	y_N	x_N					
Σ	Σy	Σx			$\Sigma(x_i - \mu_x)(y_i - \mu_y)$	$\Sigma(x_i - \mu_x)^2$	$\Sigma(y_i - \mu_y)^2$

$$\rho = \frac{\sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y)}{\sqrt{\sum_{i=1}^N (x_i - \mu_x)^2 \sum_{i=1}^N (y_i - \mu_y)^2}}$$

พารามิเตอร์



ค่าสหสัมพันธ์

$$r = \frac{\text{cov}(x, y)}{S_x S_y}$$

หรือ

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

$$r = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sqrt{\left[n \sum_{i=1}^n (x_i)^2 - \left(\sum_{i=1}^n x_i \right)^2 \right] \left[n \sum_{i=1}^n (y_i)^2 - \left(\sum_{i=1}^n y_i \right)^2 \right]}}$$

r^2 เรียกว่า “สัมประสิทธิ์การตัดสินใจ”
(coefficient of determination)

ใช้วัดอิทธิพลของตัวแปรหนึ่ง
ว่ามีผลต่อตัวแปรอีกตัวหนึ่ง
เพียงใด

$$0 \leq r^2 \leq 1$$

การคำนวณ

สถิติสำหรับงานวิจัย



Statistics for Research

NO	Y	X	$(y_i)^2$	$(x_i)^2$	$x_i y_i$
1	y_1	x_1			
2	y_2	x_2			
.	.	.			
.	.	.			
.	.	.			
N	y_N	x_N			
Σ	Σy	Σx	$\Sigma (y_i)^2$	$\Sigma (x_i)^2$	$\Sigma x_i y_i$

$$r = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sqrt{\left[n \sum_{i=1}^n (x_i)^2 - \left(\sum_{i=1}^n x_i \right)^2 \right] \left[n \sum_{i=1}^n (y_i)^2 - \left(\sum_{i=1}^n y_i \right)^2 \right]}}$$

ตัวอย่าง จากคะแนนการสอบของนักศึกษาวิชาเอกบริหารธุรกิจที่สอบวิชาเศรษฐศาสตร์ และมานุษยวิทยา 12 คน เป็นดังตารางอยากทราบว่าคะแนนสอบทั้ง 2 วิชามีความสัมพันธ์กันเพียงใด

No.	Economics (Y)	Anthropology (X)	y^2	x^2	xy
1	51	74	2601	5476	3774
2	68	70	4624	4900	4760
3	72	88	5184	7744	6336
4	97	93	9409	8649	9021
5	55	67	3025	4489	3685
6	73	73	5329	5329	5329
7	95	99	9025	9801	9405
8	74	73	5476	5329	5402
9	20	33	400	1089	660
10	91	91	8281	8281	8281
11	74	80	5476	6400	5920
12	80	86	6400	7396	6880
	$\sum y = 850$	$\sum x = 927$	65230	74883	$\sum xy = 69453$

No.	Economics (Y)	Anthropology (X)	y^2	x^2	xy
	$\sum y = 850$	$\sum x = 927$	65230	74883	$\sum xy = 69453$

$$r = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sqrt{\left[n \sum_{i=1}^n (x_i)^2 - \left(\sum_{i=1}^n x_i \right)^2 \right] \left[n \sum_{i=1}^n (y_i)^2 - \left(\sum_{i=1}^n y_i \right)^2 \right]}}$$

$$= \frac{12(69453) - (927)(850)}{\sqrt{\left(12(74883) - (927)^2 \right) \left(12(65230) - (850)^2 \right)}}$$

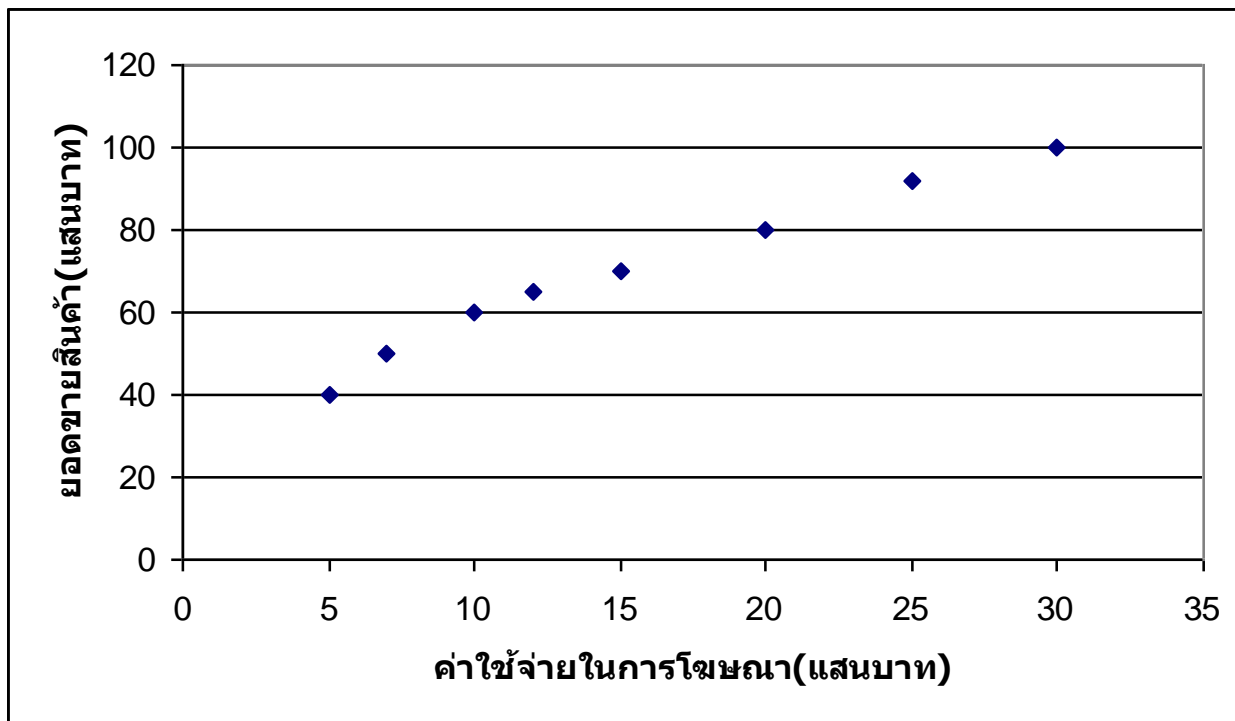
$$= 0.935$$

$$r^2 = 0.874$$

หมายความว่าจากนักศึกษา 12 คน คะแนนสอบวิชามนุษยวิทยามีความสัมพันธ์อย่างมากในเชิงบวก กับคะแนนวิชาเศรษฐศาสตร์ และคะแนนวิชามนุษยวิทยามีผลต่อคะแนนวิชาเศรษฐศาสตร์ 87.4%

ตัวอย่าง นักวิจัยการตลาดของบริษัทแห่งหนึ่งต้องการศึกษาความสัมพันธ์ระหว่างยอดขายสินค้ากับค่าใช้จ่ายในการโฆษณาของบริษัท จึงเก็บรวบรวมข้อมูลที่ต้องการได้ดังตาราง

ครั้งที่	1	2	3	4	5	6	7	8
ค่าใช้จ่ายในการโฆษณา(แสนบาท)	5	7	10	12	15	20	25	30
ยอดขายสินค้า(แสนบาท)	40	50	60	65	70	80	92	100



$$\sum x = 124, \sum y = 557, \sum x^2 = 2468, \sum y^2 = 41689, \sum xy = 9880$$

$$\begin{aligned} r &= \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sqrt{\left[n \sum_{i=1}^n (x_i)^2 - \left(\sum_{i=1}^n x_i \right)^2 \right] \left[n \sum_{i=1}^n (y_i)^2 - \left(\sum_{i=1}^n y_i \right)^2 \right]}} \\ &= \frac{8(9880) - (124)(557)}{\sqrt{\{8(2468) - (124)^2\} \{8(41689) - (557)^2\}}} \\ &= \frac{79040 - 69068}{\sqrt{(4368)(23263)}} \\ &= \frac{9972}{10080.32} \\ r &= 0.9893 \\ r^2 &= (0.9893)^2 \\ &= 0.9787 \end{aligned}$$

หมายความว่าค่าใช้จ่าย
ในการโฆษณาที่เพิ่มขึ้น
หรือลดลงมีผลทำให้
ยอดขายสินค้าเพิ่มขึ้น
หรือลดลงเป็นอย่างมาก
ถึงร้อยละ 97.87 ส่วนอีก
ร้อยละ 2.13 เป็นผลจาก
สาเหตุอื่น

การทดสอบสมมติฐานเกี่ยวกับค่าสัมประสิทธิ์สหสัมพันธ์ ρ_{xy}

เป็นการทดสอบว่าตัวแปร X และ Y มีความสัมพันธ์เชิงเส้นหรือไม่

สมมติฐานเชิงสถิติ

$$H_0 : \rho = 0$$

หรือ

$$H_1 : \rho \neq 0$$

H_0 : X และ Y ไม่มีความสัมพันธ์เชิงเส้นตรง

H_1 : X และ Y มีความสัมพันธ์เชิงเส้นตรง

ตัวสถิติทดสอบ

$$T = \frac{r - 0}{\sqrt{(1 - r^2) / (n - 2)}}$$

หรือ

$$T = \frac{r\sqrt{n - 2}}{\sqrt{1 - r^2}}$$

ค่าวิกฤติ

$$-t_{\frac{\alpha}{2}, n-2}, \quad t_{\frac{\alpha}{2}, n-2}$$

การทดสอบสมมติฐานว่า ส.ป.ส สหสัมพันธ์มีค่าเท่ากับค่าใดค่าหนึ่งหรือไม่

$$H_0: \rho = \rho_0$$

$$H_1: \rho \neq \rho_0$$

เมื่อ

$$\rho_0 \neq 0$$

ตัวสถิติทดสอบ

$$Z = \frac{Z_r - Z_{\rho_0}}{\sqrt{1/(n-3)}}$$

โดยที่

$$Z_r = \frac{1}{2} \ln \left(\frac{1+r}{1-r} \right)$$

$$Z_{\rho_0} = \frac{1}{2} \ln \left(\frac{1+\rho_0}{1-\rho_0} \right)$$

ค่าวิกฤติ

$$-Z \leq -\frac{\alpha}{2}, \quad Z \leq \frac{\alpha}{2}$$

ตัวอย่าง จากตัวอย่างคะแนนการสอบของนักศึกษาวิชาเอกบริหารธุรกิจที่สอบวิชา เศรษฐศาสตร์ และมนุษยวิทยา 12 คน อยากทราบว่าคะแนนทั้งสองวิชา มีความสัมพันธ์เชิงเส้นตรงหรือไม่ ที่ระดับนัยสำคัญ 0.05

H_0 : คะแนนวิชาเศรษฐศาสตร์กับมนุษยวิทยาไม่มีความสัมพันธ์เชิงเส้นตรง

H_1 : คะแนนวิชาเศรษฐศาสตร์กับมนุษยวิทยามีความสัมพันธ์เชิงเส้นตรง

$$H_0 : \rho = 0$$

$$H_1 : \rho \neq 0$$

ตัวสถิติทดสอบ

$$\begin{aligned} T &= \frac{r - 0}{\sqrt{\frac{(1 - r^2)}{n - 2}}} \\ &= \frac{0.935}{\sqrt{\frac{(1 - 0.935^2)}{12 - 2}}} \\ &= \frac{0.935}{\sqrt{\frac{0.126}{10}}} \\ &= 8.33 \end{aligned}$$

ค่าวิกฤต

$$\pm t_{\frac{\alpha}{2}, n-2} = \pm t_{0.025, 10} = \pm 2.228$$

เนื่องจาก $T = 8.33$ อยู่ในบริเวณ ปฏิเสธ H_0 หมายความว่าคะแนน วิชา เศรษฐศาสตร์ และ วิชา มนุษยวิทยามีความสัมพันธ์เชิง เส้นตรง ที่ระดับนัยสำคัญ 0.05

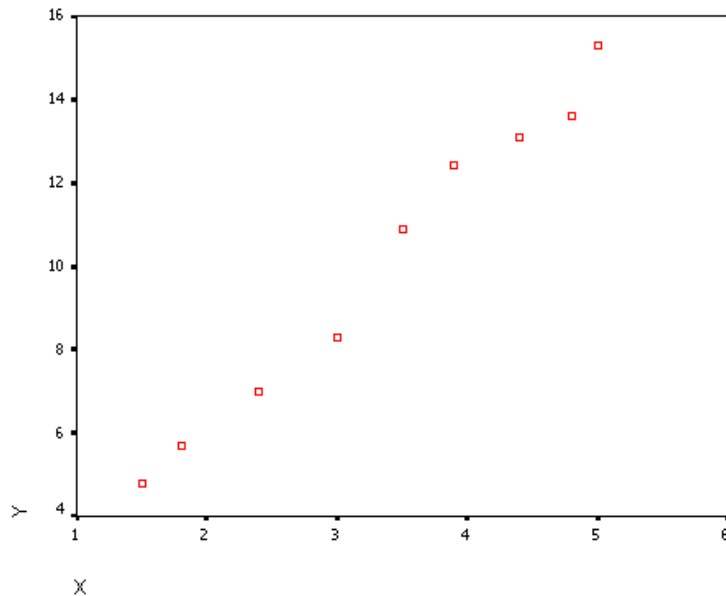
การวิเคราะห์สหสัมพันธ์ด้วย SPSS

คำสั่ง

Graphs... → Scatter...

คำสั่ง

Analyze → Correlate → Bivariate...



ตัวอย่าง file example16.sav

สมมติฐาน: x กับ y มีความสัมพันธ์กัน

Correlations

		X	Y
X	Pearson Correlation	1	.991*
	Sig. (2-tailed)	.	.000
	N	9	9
Y	Pearson Correlation	.991*	1
	Sig. (2-tailed)	.000	.
	N	9	9

สัมประสิทธิ์สหสัมพันธ์

$r=0.991$

P-value

** . Correlation is significant at the 0.01 level

H_0 : X ไม่มีความสัมพันธ์กับ Y

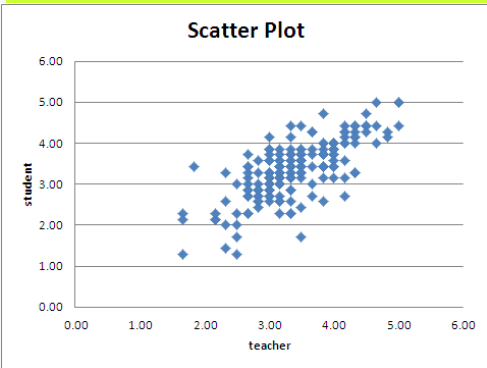
H_1 : X มีความสัมพันธ์กับ Y

จากตารางสหสัมพันธ์พบว่า X และ Y มีความสัมพันธ์กันอย่างมากในระดับ 0.991 อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 (sig<0.05) Y ได้รับอิทธิพลจาก X ร้อยละ 98.2

Home Work

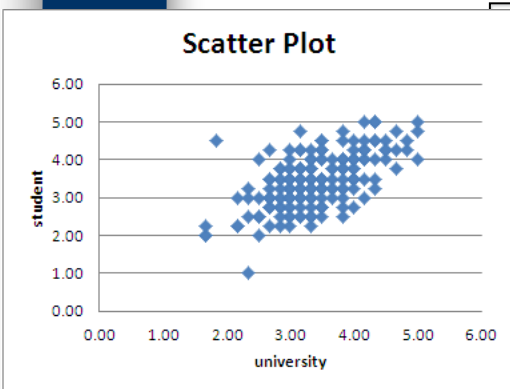
ที่ระดับนัยสำคัญ 0.05

1. ทดสอบสมมติฐานที่ว่าระดับความสำคัญของปัญหาและอุปสรรคในการทำงานที่ได้รับมอบหมายด้านอาจารย์และด้านนักศึกษามีความสัมพันธ์กัน



	ระดับของปัญหาและอุปสรรคในการปฏิบัติงานที่ได้รับมอบหมายด้าน	
	นักศึกษา	อาจารย์
นักศึกษา		
อาจารย์		

2. ทดสอบสมมติฐานที่ว่าระดับความสำคัญของปัญหาและอุปสรรคในการทำงานที่ได้รับมอบหมายด้านมหาวิทยาลัยและด้านนักศึกษามีความสัมพันธ์กัน



	ระดับของปัญหาและอุปสรรคในการปฏิบัติงานที่ได้รับมอบหมายด้าน	
	นักศึกษา	มหาวิทยาลัย
นักศึกษา		
มหาวิทยาลัย		