



๑๐๐ ปี คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
100th Anniversary of The Faculty of Science, CU

เอกสารเผยแพร่เพื่อการศึกษา

การวิเคราะห์ข้อมูลทางสถิติ

SPSS for Windows Version 16

รองศาสตราจารย์ ดำรงค์ ทิพย์โยธา

ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์

คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

การวิเคราะห์ข้อมูลทางสถิติ

SPSS for Windows

Version 16



รองศาสตราจารย์ ดำรงค์ ทิพย์โยธา
ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์
คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

<http://www.math.sc.chula.ac.th>
<http://www.math.sc.chula.ac.th/~tdumrong/homepage>
<http://pioneer.netserv.chula.ac.th/~tdumrong/homepage>

คำนำ

หนังสือ การวิเคราะห์ข้อมูลทางสถิติด้วย SPSS for Windows version 16 เป็นคู่มือในการใช้งานโปรแกรม SPSS for Windows version 16 เนื้อหาภายในเล่มประกอบด้วย การสร้างแฟ้มข้อมูล การแก้ไขแฟ้มข้อมูล การวิเคราะห์ข้อมูลทางสถิติ การคำนวณค่าสถิติเบื้องต้น เช่น การหาค่าเฉลี่ย ค่ามัธยฐาน ค่าความแปรปรวน ส่วนเบี่ยงเบนมาตรฐาน ฯลฯ การนำเสนอข้อมูลในรูปแบบการแจกแจงความถี่แบบ 1 ทาง และการแจกแจงความถี่แบบ 2 ทาง การนำเสนอข้อมูลในรูปแบบตารางที่สวยงาม การนำเสนอในรูปแบบกราฟ

ในส่วนของการวิเคราะห์ข้อมูลโปรแกรม SPSS for Windows version 12 สามารถคำนวณค่าช่วงความเชื่อมั่นของค่าพารามิเตอร์ สามารถทำการทดสอบสมมติฐานแบบต่างๆ เช่น การทดสอบสมมติฐานว่า $\mu = \mu_0$, $\mu_1 = \mu_2$, $\sigma^2 = \sigma_0^2$, $\sigma_1^2 = \sigma_2^2$ การทดสอบภาวะสารูปสนิทธิ การทดสอบว่าข้อมูลมีการแจกแจงปกติจริงหรือไม่ การทดสอบสมมติฐานว่าข้อมูลเป็นอิสระต่อกัน การหาสมการเส้นถดถอยเชิงเส้นเชิงเดียวและ สหสัมพันธ์ การหาสมการเส้นถดถอยแบบไม่เป็นเส้นตรง เช่นการหาความสัมพันธ์ในรูปแบบสมการ log exponential การคำนวณค่าสัมประสิทธิ์สหสัมพันธ์และสมการถดถอยพหุคูณ การวิเคราะห์ความแปรปรวนแบบจำแนกทางเดียว การวิเคราะห์ความแปรปรวนแบบจำแนกสองทาง การทดสอบสมมติฐานแบบนอนพาราเมตริก

เนื้อหาภายในเล่มจะแสดงขั้นตอนการสั่งงานอย่างละเอียดพร้อม คำอธิบายและภาพประกอบทุกขั้นตอน มีเหตุผลทางทฤษฎีความน่าจะเป็นและสถิติประกอบการทำงาน และแสดงสูตรทางคณิตศาสตร์ซึ่งเป็นที่มาของค่าสถิติที่ SPSS คำนวณมาให้ สรุปผล วิเคราะห์และแปลความหมายทางสถิติเพื่อนำไปสรุปผลของข้อมูลได้ นอกจากนี้เพื่อสะดวกในการเชื่อมโยงข้อมูลกับ Microsoft Word และ Excel จึงได้เพิ่มภาคผนวกเรื่องการเชื่อมโยงข้อมูล SPSS for Windows กับ Microsoft Word และ การเชื่อมโยงข้อมูล SPSS for Windows กับ Excel นอกจากนี้ยังมีเนื้อเกี่ยวกับโปรแกรมภาษาของ SPSS ชนิดที่เป็น Syntax ในภาคผนวกที่ 4 ซึ่งจะประโยชน์ในการวิเคราะห์ข้อมูลมากขึ้น

เนื่องจากรูปแบบการใช้งานของ SPSS for Windows version 10 - 16 มีลักษณะคล้ายกันหนังสือเล่มนี้จึงสามารถประยุกต์ใช้งานได้กับ SPSS for Windows version 10 - 16

เพื่อความสะดวกและรวดเร็วในการศึกษาผู้อ่านทุกท่านสามารถ Download ข้อมูลตัวอย่าง โดยเข้าไปที่

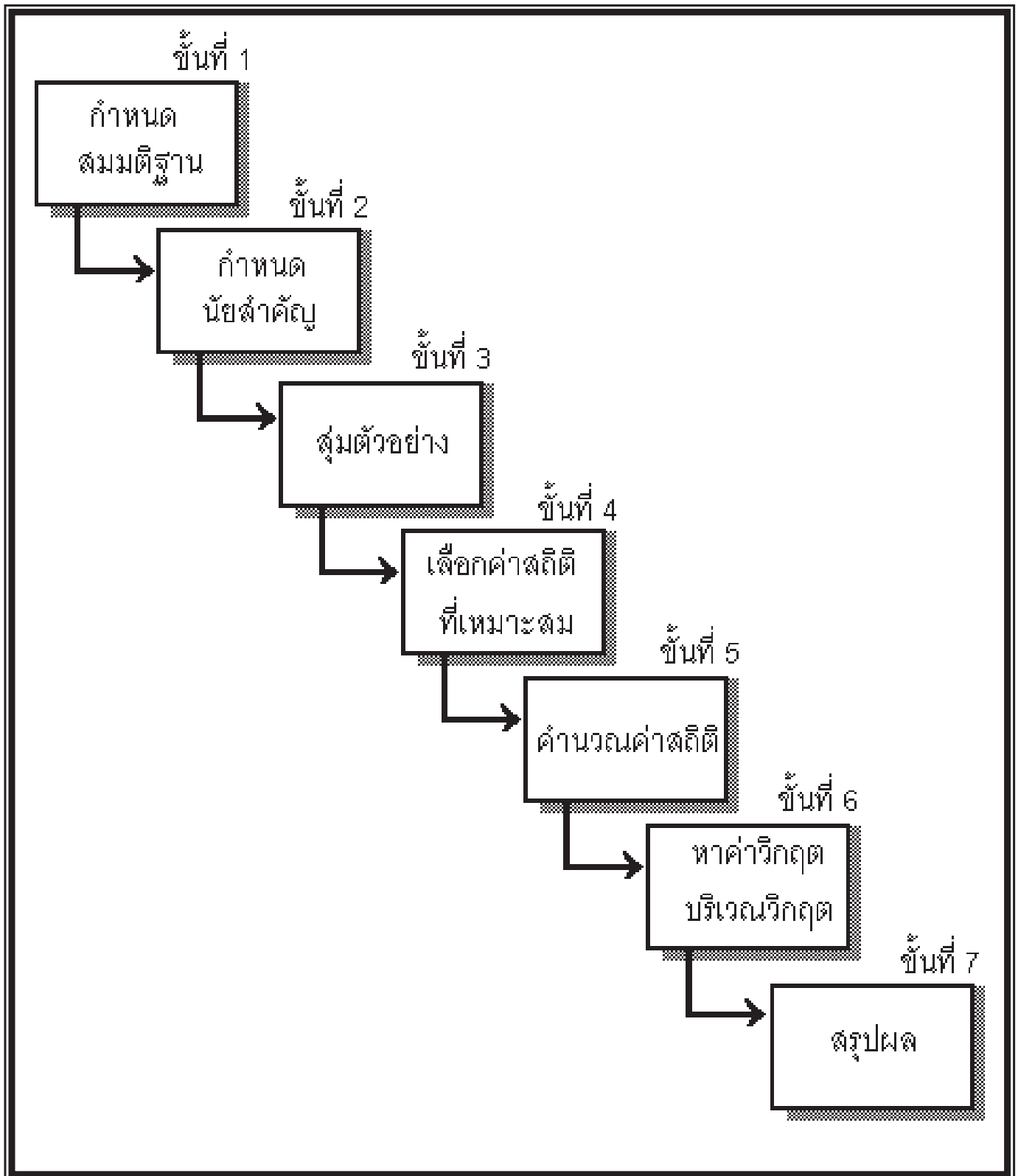
<http://pioneer.netserv.chula.ac.th/~tdumrong/2301286data>

หรือ <http://www.math.sc.chula.ac.th/~tdumrong/2301286data>

ผู้เขียนหวังว่าหนังสือเล่มนี้จะช่วยให้ผู้อ่านทุกท่านสามารถนำไปใช้วิเคราะห์ข้อมูลได้เป็นอย่างดี และขอขอบคุณผู้อ่านทุกท่านที่ได้ติดตามผลงานของผู้เขียนมาโดยตลอด

รองศาสตราจารย์ ดำรงค์ ทิพย์โยธา

การทดสอบสมมติฐาน



หมายเหตุ ในขั้นที่ 4 ค่าสถิติที่เหมาะสมและนิยมใช้กันมากคือค่า Z, t, F และ χ^2

ในขั้นที่ 7 การสรุปผลทำได้ 2 แบบคือ นำค่าสถิติจากตัวอย่างเปรียบเทียบกับค่าวิกฤตและบริเวณวิกฤต หรือ เปรียบเทียบค่า Significant ที่คำนวณได้จากตัวอย่าง กับค่านัยสำคัญของการทดสอบ α

สารบัญ

	หน้าที่
บทที่ 1 ความรู้เบื้องต้นเกี่ยวกับ SPSS for Windows	1 – 22
1.1 คอมพิวเตอร์ที่สามารถทำงานกับโปรแกรม SPSS for Windows	2
1.2 ความสามารถของโปรแกรม SPSS for Windows	2
1.3 การเข้าสู่การทำงานของโปรแกรม SPSS for Windows	4
1.4 WINDOW ของการทำงานแบบต่างๆ ของ SPSS for Windows	6
1.5 สรุปเนื้อหาของคำสั่งและขั้นตอนการทำงานโดยย่อของ SPSS for Windows	9
1.6 Icon บนเมนูบาร์กับการทำงานของ SPSS for Windows	19
บทที่ 2 การสร้างแฟ้มข้อมูล	21 – 40
2.1 การสร้างแฟ้มข้อมูลใน SPSS Data Editor	25
2.2 การบันทึกแฟ้มข้อมูล	34
2.3 การเปิดแฟ้มข้อมูล	35
2.4 การดูรายละเอียดตัวแปร	36
2.5 การสั่งให้ SPSS Data Editor แสดง Value Labels	37
บทที่ 3 การวิเคราะห์ข้อมูลด้วยคำสั่ง Analyze / Descriptive Statistics	39 – 62
3.1 การคำนวณค่าสถิติเบื้องต้นด้วยคำสั่ง Analyze / Descriptive Statistics / Descriptives	39
3.2 การเปลี่ยนรูปแบบของตารางในการแสดงผลของ SPSS Viewer	40
3.3 การกำหนดตำแหน่งทศนิยมของการคำนวณในตารางของ SPSS Viewer	42
3.4 การคำนวณค่าสถิติอื่นๆ ด้วยคำสั่ง Descriptives	43
3.5 สูตรของค่าสถิติและเปรียบเทียบการคำนวณ MATHCAD กับ SPSS	45
3.6 การคำนวณค่าสถิติเบื้องต้นด้วยคำสั่ง Analyze / Descriptive Statistics / Frequencies	49
3.7 การคำนวณค่าสถิติเบื้องต้นด้วยคำสั่ง Analyze / Descriptive Statistics / Explore	54
3.8 การคำนวณค่าสถิติเบื้องต้นด้วยคำสั่ง Analyze / Descriptive Statistics / Crosstabs	58
3.9 การคำนวณค่าสถิติเบื้องต้นด้วยคำสั่ง Analyze / Descriptive Statistics / Ratio	61
บทที่ 4 การวิเคราะห์ข้อมูลด้วยคำสั่ง Analyze / Reports และ Tables	63 – 76
4.1 การคำนวณค่าสถิติเบื้องต้นด้วยคำสั่ง Analyze / Reports / OLAP Cubes	63
4.2 การคำนวณค่าสถิติเบื้องต้นด้วยคำสั่ง Analyze / Reports / Case Summaries	66
4.3 การคำนวณค่าสถิติเบื้องต้นด้วยคำสั่ง Analyze / Reports / Report Summaries in Rows Analyze / Reports / Report Summaries in Columns	69
4.4 การนำเสนอข้อมูลด้วยคำสั่ง Analyze / Tables	73

บทที่ 5 การปรับปรุงเพิ่มข้อมูลด้วยคำสั่ง Edit, Data และ Transform	77 – 94
5.1 การเพิ่มตัวแปร การลดตัวแปร การแทรกตัวแปร	78
5.2 การลบคำสั่งเกิด	81
5.3 การรวมเพิ่มข้อมูลแบบเพิ่มตัวแปร	81
5.4 การรวมเพิ่มข้อมูลแบบเพิ่มคำสั่งเกิด	82
5.5 การเรียงลำดับข้อมูล	84
5.6 การกำหนดตัวแปรน้ำหนัก	85
5.7 การนำค่าจากตัวแปรเก่าไปสร้างเป็นค่าของตัวแปรใหม่	86
5.8 การปรับเปลี่ยนค่าของตัวแปรด้วยคำสั่ง Transform	89
บทที่ 6 การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่าพารามิเตอร์	95 – 122
6.1 การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่าเฉลี่ย μ	96
6.2 การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของผลต่างค่าเฉลี่ย $\mu_1 - \mu_2$ กรณีประชากร 2 ชุดเป็นอิสระต่อกัน	102
6.3 การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของผลต่างค่าเฉลี่ย μ_D กรณีประชากร 2 ชุดไม่เป็นอิสระต่อกัน	111
6.4 การวิเคราะห์ข้อมูลด้วยคำสั่ง Analyze / Compare Means / Means	116
6.5 การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่าเฉลี่ย μ ด้วยคำสั่ง Analyze / Compare Means / One-Way ANOVA	120
บทที่ 7 การทดสอบสมมติฐาน	123 – 164
7.1 การทดสอบสมมติฐาน $H_0 : \mu = \mu_0$	124
7.2 การทดสอบสมมติฐาน $H_0 : \mu_1 = \mu_2$ กรณีประชากร 2 ชุดเป็นอิสระต่อกัน	128
7.3 การทดสอบสมมติฐาน $H_0 : \mu_D = d_0$ กรณีประชากร 2 ชุดไม่เป็นอิสระต่อกัน	138
7.4 การทดสอบสมมติฐาน $H_0 : \sigma^2 = \sigma_0^2$	145
7.5 การทดสอบสมมติฐาน $H_0 : \sigma_1^2 = \sigma_2^2$	146
7.6 การทดสอบภาวะสภาวะรูปสัณฐาน	150
7.7 การทดสอบสมมติฐานว่าข้อมูลเป็นอิสระต่อกันหรือไม่	155
7.7 การทดสอบสมมติฐานว่าข้อมูลมีการแจกแจงปกติจริงหรือไม่	161
บทที่ 8 สหสัมพันธ์และการถดถอยเชิงเส้น	165 – 200
8.1 การหาสมการเส้นถดถอยเชิงเส้นเชิงเดี่ยว(Simple Linear Regression) และ สหสัมพันธ์ (Correlation)	166
8.2 การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของสัมประสิทธิ์การถดถอย β และระยะตัดแกน α	178
8.3 การทดสอบสมมติฐาน $H_0 : \rho = 0$	182
8.4 การทดสอบสมมติฐาน $H_0 : \beta = \beta_0$	184

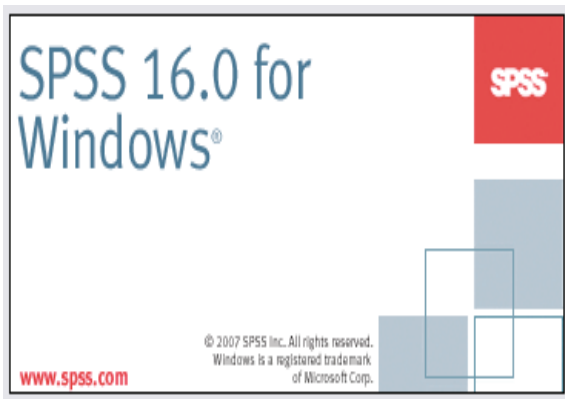
8.5	การเลือกรูปแบบความสัมพันธ์ที่เหมาะสมกับข้อมูล	190
8.6	การคำนวณค่าสัมประสิทธิ์สหสัมพันธ์และสมการถดถอยพหุคูณ	196
บทที่ 9	การวิเคราะห์ความแปรปรวน	201 – 228
9.1	การวิเคราะห์ความแปรปรวนแบบจำแนกทางเดียว	201
9.2	การวิเคราะห์ความแปรปรวนแบบที่มีการสุ่มอย่างสมบูรณ์ในแต่ละกลุ่ม	209
บทที่ 10	การทดสอบสมมติฐานแบบนอนพาราเมตริก	229 – 250
10.1	การทดสอบว่าตัวอย่างที่เราเลือกมาเป็นไปโดยสุ่มหรือไม่	229
10.2	การทดสอบว่าประชากรมีการแจกแจงตามที่เราคาดไว้หรือไม่	233
10.3	การทดสอบว่าประชากร 2 กลุ่มมีค่าเฉลี่ยเท่ากันหรือไม่	235
10.4	การทดสอบว่าประชากร k กลุ่มมีค่าเฉลี่ยเท่ากันหรือไม่	242
10.5	การหาสัมประสิทธิ์สหสัมพันธ์ตำแหน่งที่ (Rank Correlation Coefficient)	248
ภาคผนวกที่ 1	การคำนวณค่า Significant ของค่าสถิติ	251 – 254
ภาคผนวกที่ 2	การเชื่อมโยงข้อมูล SPSS for Windows กับ Microsoft Word	255 – 260
ภาคผนวกที่ 3	การเชื่อมโยงข้อมูล SPSS for Windows กับ Excel	261 – 266
ภาคผนวกที่ 4	SPSS Syntax Editor กับ โปรแกรมภาษา SPSS	267 – 271
บรรณานุกรม		272

สารบัญของการทำงานพื้นฐานในการวิเคราะห์ข้อมูลทางสถิติ

การทำงานเกี่ยวกับเพิ่มข้อมูล และ การวิเคราะห์ข้อมูล	หน้า
การสร้างเพิ่มข้อมูล	25
การบันทึกเพิ่มข้อมูล	34
การเปิดเพิ่มข้อมูล	35
การดูรายละเอียดตัวแปร	36
การเพิ่มตัวแปร การลดตัวแปร การแทรกตัวแปร	78
การลบค่าสังเกต	81
การรวมเพิ่มข้อมูลแบบเพิ่มตัวแปร และ การรวมเพิ่มข้อมูลแบบเพิ่มค่าสังเกต	81
การเรียงลำดับข้อมูล	84
การนำค่าจากตัวแปรเก่าไปสร้างเป็นค่าของตัวแปรใหม่	86
การปรับเปลี่ยนค่าของตัวแปรด้วยคำสั่ง	89
การคำนวณค่าสถิติเบื้องต้นด้วยคำสั่ง Analyze / Descriptive Statistics / Descriptives	39
การเปลี่ยนรูปแบบของตารางในการแสดงผลของ SPSS Viewer	40
การกำหนดตำแหน่งทศนิยมของการคำนวณในตารางของ SPSS Viewer	42
การแจกแจงความถี่ของข้อมูลด้วยคำสั่ง Analyze / Descriptive Statistics / Frequencies	51
การคำนวณค่าสถิติเบื้องต้นด้วยคำสั่ง Analyze / Descriptive Statistics / Explore	56
การแจกแจงความถี่ของข้อมูลด้วยคำสั่ง Analyze / Descriptive Statistics / Crosstabs	60
การนำเสนอข้อมูลด้วยคำสั่ง Analyze / Tables / Custom Tables	73
การหาช่วงความเชื่อมั่นของค่าเฉลี่ย μ	96
การหาช่วงความเชื่อมั่นของผลต่างค่าเฉลี่ย $\mu_1 - \mu_2$ กรณีประชากร 2 ชุดเป็นอิสระต่อกัน	102
การหาช่วงความเชื่อมั่นของผลต่างค่าเฉลี่ย $\mu_1 - \mu_2$ กรณีประชากร 2 ชุดไม่เป็นอิสระต่อกัน	111
การทดสอบสมมติฐาน $H_0 : \mu = \mu_0$	124
การทดสอบสมมติฐาน $H_0 : \mu_1 = \mu_2$ กรณีประชากร 2 ชุดเป็นอิสระต่อกัน	128
การทดสอบสมมติฐาน $H_0 : \mu_1 = \mu_2$ กรณีประชากร 2 ชุดไม่เป็นอิสระต่อกัน	138
การทดสอบสมมติฐานว่าข้อมูลเป็นอิสระต่อกันหรือไม่	155
การทดสอบสมมติฐานว่าข้อมูลมีการแจกแจงปกติจริงหรือไม่	161
การหาสัมประสิทธิ์การถดถอยและสหสัมพันธ์	166
การเลือกรูปแบบความสัมพันธ์ที่เหมาะสมกับข้อมูล	190
การวิเคราะห์ความแปรปรวนแบบจำแนกทางเดียว	201
การวิเคราะห์ความแปรปรวนแบบที่มีการสุ่มอย่างสมบูรณ์ในแต่ละกลุ่ม	209

บทที่ 1

ความรู้เบื้องต้นเกี่ยวกับ SPSS for Windows



โปรแกรมสำเร็จรูปที่มีความสามารถในการวิเคราะห์ข้อมูลทางสถิติในปัจจุบันมีหลายโปรแกรมเช่น SAS, MINITAB, SPSS for Windows แต่โปรแกรมที่นิยมใช้กันมากคงจะเป็นโปรแกรม SPSS for Windows โปรแกรม SPSS (Statistical Package for the Social Sciences) มีการใช้งานมานาน เริ่มตั้งแต่การใช้งานบนเครื่องคอมพิวเตอร์ขนาดใหญ่ (Main Frame) จนถึงรุ่นที่ใช้กับเครื่องไมโครคอมพิวเตอร์ได้เช่น

SPSS/PC version 3.0	สำหรับระบบปฏิบัติการ DOS
SPSS for Windows version 6.0	สำหรับระบบปฏิบัติการ Windows 3.0
SPSS for Windows version 7.5, 8.0, 9.0	สำหรับระบบปฏิบัติการ Windows 95, 98
SPSS for Windows version 10, 11	สำหรับระบบปฏิบัติการ Windows 2000
SPSS for Windows version 12.0 – 16.0	สำหรับระบบปฏิบัติการ Windows XP

โปรแกรม SPSS for Windows version 16.0 สามารถนำข้อมูล SPSS หรือ ข้อมูลโปรแกรมเดิมที่สร้างมาจาก SPSS version 3.0 – 15.0 ทั้งในระบบ DOS และ ระบบปฏิบัติการ Windows กลับมาใช้ได้ และสามารถรับข้อมูลที่สร้างจากโปรแกรมประเภทต่างๆ ได้เช่น Excel Mathcad Microsoft Word ฯลฯ นอกจากนี้ โปรแกรม SPSS for Windows ยังสามารถบันทึกคำสั่งที่เกิดจากขั้นตอนการทำงานตามลำดับต่างๆ จากการใช้เมาส์เลือกเมนูของโปรแกรมที่มีอยู่มาบันทึกเป็น ชุดคำสั่ง (Command Language) เพื่อประโยชน์ในการเรียกคำสั่งเหล่านี้มาใช้ได้อีกในครั้งต่อไป ภายหลังจาก ผู้ที่เคยใช้โปรแกรมอื่นๆ ที่ทำงานบน Window สามารถเรียนรู้การใช้งานโปรแกรม SPSS for Windows version 16.0 ได้อย่างรวดเร็ว และสามารถนำคุณสมบัติของ Window มาใช้ได้เช่น copy cut paste การย้าย การคัดลอก การพิมพ์ การแลกเปลี่ยนข้อมูลระหว่างโปรแกรม ฯลฯ

1.1 คอมพิวเตอร์ที่สามารถทำงานกับโปรแกรม SPSS for Windows

ความต้องการของเครื่องคอมพิวเตอร์ฮาร์ดแวร์และซอฟต์แวร์ที่สามารถนำโปรแกรม SPSS for Windows ไปใช้ได้ควรมีคุณสมบัติดังต่อไปนี้

- เครื่องคอมพิวเตอร์ IBM PC หรือ IBM Compatible ที่ใช้ Windows XP
- หน่วยความจำ RAM ในคอมพิวเตอร์ส่วนใหญ่ควรมี 512 Mb หรือมากกว่า ซึ่งเพียงพอใช้งาน
- Hard disk ควรมีที่ว่างมากพอสมควร
- จอภาพ (Monitor) ต้องสามารถแสดงผลทางด้านกราฟฟิกได้
- โปรแกรมระบบปฏิบัติการ Microsoft Windows XP
- โปรแกรม SPSS for Windows version 16.0

เพื่อความสะดวกในการทำงานและการเชื่อมโยงข้อมูลควรจะต้องมี Excel, Microsoft Word, Mathcad

1.2 ความสามารถของโปรแกรม SPSS for Windows

1.2.1 ความสามารถในการวิเคราะห์ข้อมูล

SPSS for Windows มีความสามารถที่จะทำการวิเคราะห์ข้อมูลทางสถิติ ดังต่อไปนี้

1. การคำนวณค่าสถิติเบื้องต้น (Descriptive Statistics) สามารถคำนวณค่าสถิติพื้นฐานต่างๆ ไป เช่น ค่าเฉลี่ย (Mean) มัชยฐาน (Median) ฐานนิยม (Mode) พิสัย (Range) ความแปรปรวน (Variance) ส่วนเบี่ยงเบนมาตรฐาน (Standard deviation) ฯลฯ
2. การแจกแจงความถี่ (Frequency Distributions) สามารถแจกแจงค่าของตัวแปรตามจำนวนที่นับได้ทั้งแบบทางเดียวและแบบหลายทาง (Crosstabs) พร้อมทั้งแสดงค่าสถิติที่เกี่ยวข้อง เช่น ค่าเฉลี่ย (Mean) มัชยฐาน (Median) ฐานนิยม (Mode) พิสัย (Range) ความแปรปรวน (Variance) ส่วนเบี่ยงเบนมาตรฐาน (Standard deviation) เปอร์เซ็นไทล์ (Percentiles) กราฟแท่งหรือค่าสถิติที่เกี่ยวข้องกับการทดสอบทางสถิติ เช่น Chi-Squares, Phi
3. การเปรียบเทียบค่าเฉลี่ย (Mean Groups Comparison) สามารถเปรียบเทียบและทดสอบค่าเฉลี่ยระหว่างกลุ่ม 2 กลุ่มตัวอย่างโดยค่าสถิติ t (Student's t) และสำหรับหลายกลุ่มตัวอย่างโดยค่าสถิติ F ด้วยการวิเคราะห์ความแปรปรวน (Analysis of Variance : ANOVA) ทั้งแบบทางเดียวและแบบหลายทาง
4. การหาความสัมพันธ์ระหว่างตัวแปร (Correlation) สามารถคำนวณหาค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรแบบต่างๆ เช่น Pearson, Kendall, Spearman
5. การวิเคราะห์การถดถอย (Regression Analysis) สามารถหาความสัมพันธ์เพื่อการพยากรณ์แบบการถดถอยเชิงเส้น (Linear Regression Analysis) ทั้งชนิด 1 ตัวแปรอิสระ และ ตัวแปรอิสระมากกว่า 1 ตัว และสามารถรูปแบบความสัมพันธ์ในลักษณะอื่นที่ไม่ใช่เส้นตรง เช่น Linear, Quadratic, Logarithmic ฯลฯ
6. การทดสอบแบบนอนพาราเมตริก (Nonparametric Test) สามารถวิเคราะห์ข้อมูลโดยวิธีของนอนพาราเมตริกสำหรับการทดสอบแบบต่างๆ เช่น Sign Test, Wilcoxon, Friedman, Kolmogorov – Smirnov ฯลฯ
7. การวิเคราะห์ข้อมูลสำหรับคำตอบแบบหลายคำตอบ (Multiple Response Analysis) สามารถวิเคราะห์ข้อมูลจากแบบสอบถามที่มีตัวเลือกมาให้และผู้ตอบสามารถตอบได้มากกว่า 1 คำตอบ

1.2.2 ความสามารถในการนำเสนอข้อมูลด้วยกราฟ

โปรแกรม SPSS for Windows สามารถนำเสนอข้อมูลในรูปของกราฟหรือตารางแบบต่างๆ เช่น กราฟแท่ง (Bar, Histogram) กราฟเส้น (Line) กราฟวงกลม (Pie) และกราฟชนิดอื่น ๆ

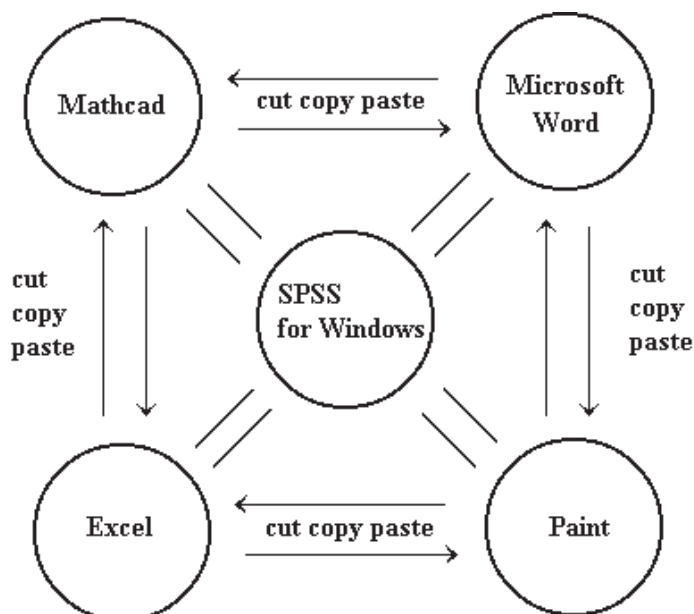
1.2.3 ความสามารถในการทำงานด้านอื่น ๆ

ในการใช้งานโปรแกรม SPSS นอกจากจะทำการวิเคราะห์ข้อมูลด้วยวิธีการทางสถิติแล้วผู้ใช้อาจจะมีการดำเนินการกับข้อมูลในลักษณะต่างๆ เช่น สร้างตัวแปรเพิ่ม เรียงลำดับข้อมูล คัดเลือกข้อมูลมาทำการวิเคราะห์ ฯลฯ ซึ่งสามารถแบ่งเป็นประเภทต่างๆ ได้ดังนี้

1. การเปลี่ยนรูปแบบข้อมูล (Data Transformation) โดยการเปลี่ยนค่าใหม่ จัดค่าใหม่ หรือสร้างตัวแปรใหม่ด้วยฟังก์ชันพิเศษต่างๆ ทางคณิตศาสตร์ที่มีในโปรแกรม SPSS
2. การจัดกลุ่มตัวแปร (Define Set of Variables) โดยการเลือกตัวแปร หรือจัดกลุ่มตัวแปรไว้เป็นชุดต่างๆ เพื่อนำมาวิเคราะห์เป็นชุดๆ ในภายหลัง
3. การเลือกข้อมูล (Select Case) โดยการกำหนดเงื่อนไขต่างๆ หรือการเลือกข้อมูลแบบสุ่มตัวอย่าง
4. การสร้างข้อมูลแบบอนุกรมเวลา (Create Time Series) โดยการสร้างข้อมูลที่เกิดขึ้นตามเวลา เช่น วัน เดือน ไตรมาส ฯลฯ สำหรับการวิเคราะห์แบบอนุกรมเวลา
5. การดำเนินการกับข้อมูลในลักษณะอื่น ๆ โดยการเรียงลำดับข้อมูล การให้นำหนักหรือความสำคัญแก่ชุดข้อมูล การสลับที่ข้อมูลระหว่างแถวและคอลัมน์
6. การจัดการกับแฟ้มข้อมูล โดยการรวมแฟ้มข้อมูลตั้งแต่ 2 แฟ้มเช่น รวมตัวแปร รวมชุดข้อมูล ฯลฯ

1.2.4. ความสามารถในการเชื่อมโยงข้อมูลกับโปรแกรมอื่น ๆ

การทำงานของโปรแกรม SPSS for Windows version 12 เป็นการทำงานภายใต้ระบบปฏิบัติการ Windows ดังนั้นเราสามารถใช้ความสามารถขั้นพื้นฐาน เช่น การเลือกบริเวณเพื่อ copy cut paste ฯลฯ แล้วนำข้อมูลนั้นไปใช้กับโปรแกรมอื่นๆ เช่น Excel, Microsoft Word, Mathcad หรือนำข้อมูลจาก Excel, Microsoft Word, Mathcad มาใช้กับ SPSS for Windows ตัวอย่างเช่น ข้อมูลในรูปแบบ column สามารถนำมาเป็นข้อมูลในรูปแบบตัวแปรของ SPSS for Windows ได้ หรือข้อมูลที่วิเคราะห์ได้จาก SPSS for Windows สามารถ copy รูปแบบตารางไปเป็นตารางของ Microsoft Word ได้ทันที



หมายเหตุ การคัดลอกข้อมูล ตารางแสดงผล และอื่น ๆ ระหว่าง SPSS, Excel, Microsoft Word ดูได้ที่ภาคผนวกที่ 2. และ ภาคผนวกที่ 3.

1.3 การเข้าสู่การทำงานของโปรแกรม SPSS for Windows

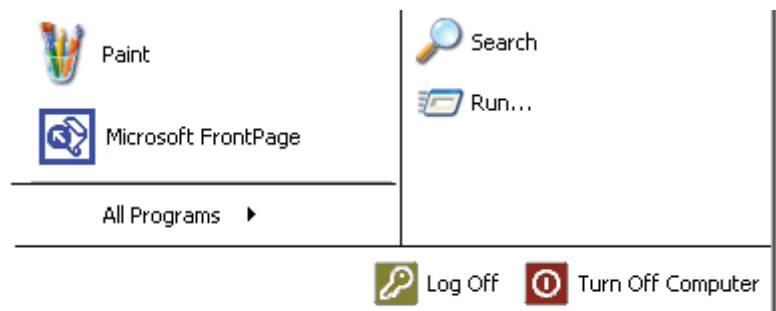
สำหรับคอมพิวเตอร์ที่ติดตั้งโปรแกรม SPSS for Windows เสร็จเรียบร้อยแล้ว การเข้าสู่การทำงานมีขั้นตอนดังนี้

ขั้นที่ 1. เปิดเครื่องคอมพิวเตอร์

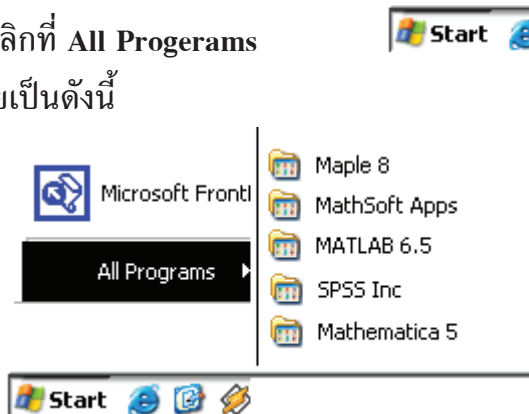
รอนจนจอภาพขึ้น Icon ของโปรแกรมต่างๆ ที่มีในคอมพิวเตอร์ขณะนั้น



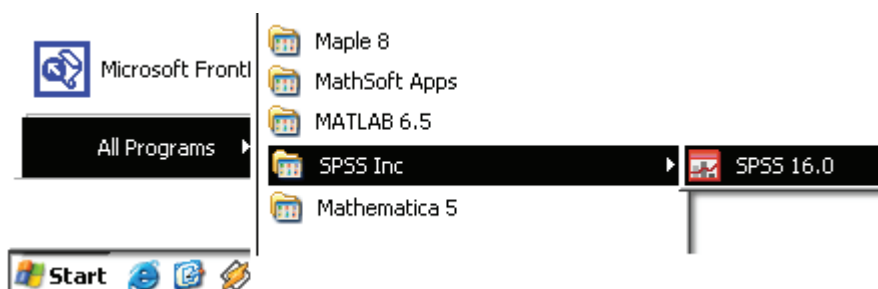
ขั้นที่ 2. เลื่อนเมาส์ไปคลิกที่ Start จะได้เมนูย่อยเป็นดังนี้



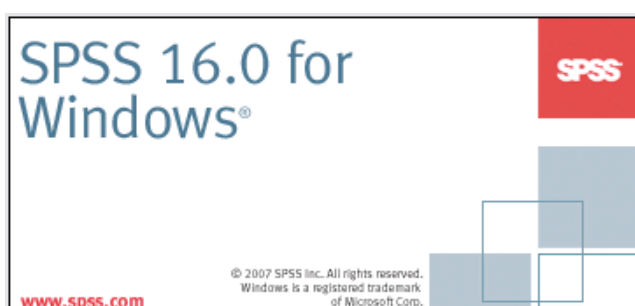
ขั้นที่ 3. คลิกที่ All Programs จะได้เมนูย่อยเป็นดังนี้



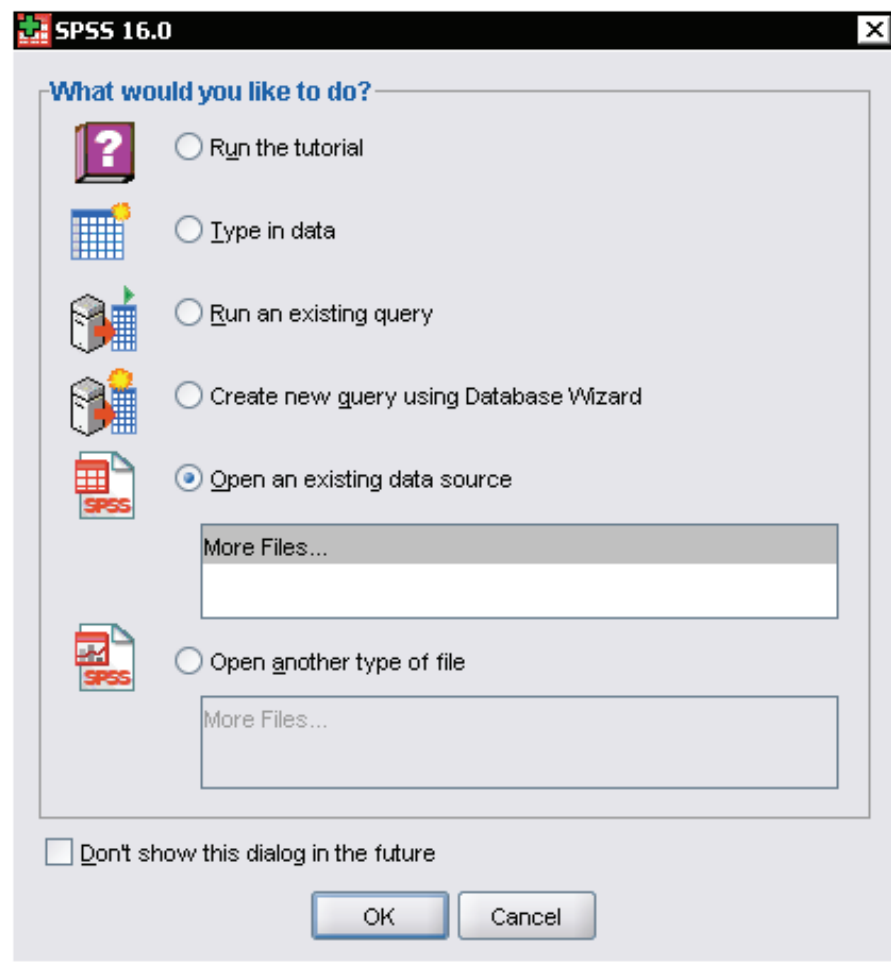
คลิกที่ SPSS Inc จะได้เมนูย่อยเป็นดังนี้



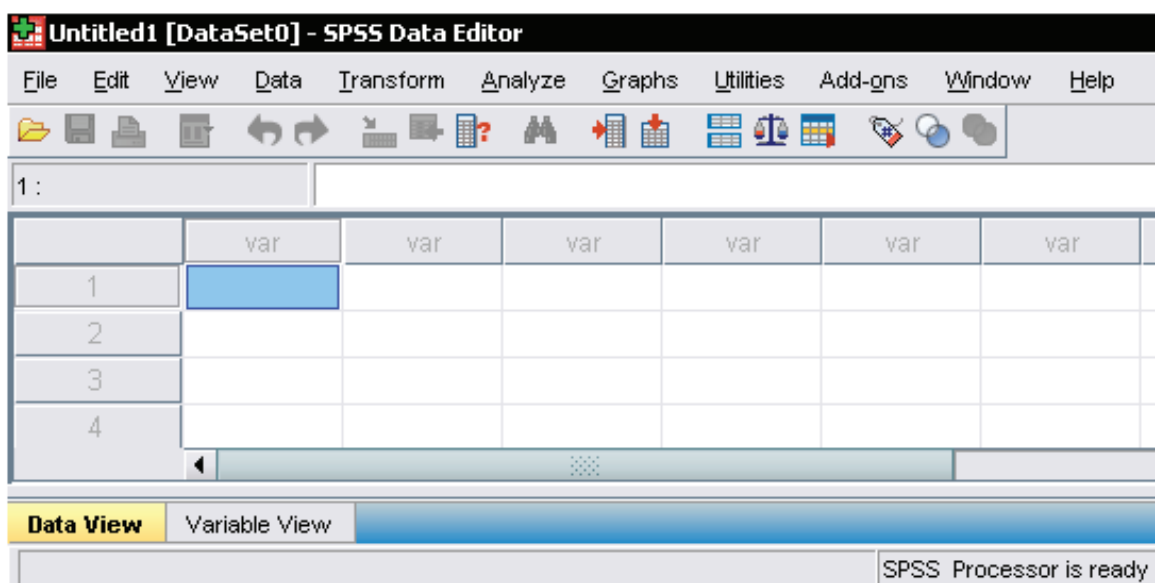
ขั้นที่ 4. คลิกที่ SPSS 16.0 จะได้ Logo ของ SPSS 16.0 for Windows



ต่อจากภาพของ LOGO จะมีเมนูเริ่มต้นให้เราเลือกทำงานตามความเหมาะสมเช่น Run the tutorial เปิดแฟ้มตามที่กำหนด พิมพ์ข้อมูล ฯลฯ ขณะนี้เพื่อความสะดวกและเข้าใจได้ง่ายขอให้คลิก **Cancel**



จะเข้าสู่การทำงานของ SPSS Data Editor



ขณะนี้เราพร้อมที่จะทำงานกับ SPSS 16.0 for Windows แล้ว

1.4 WINDOW ของการทำงานแบบต่าง ๆ ของ SPSS for Windows

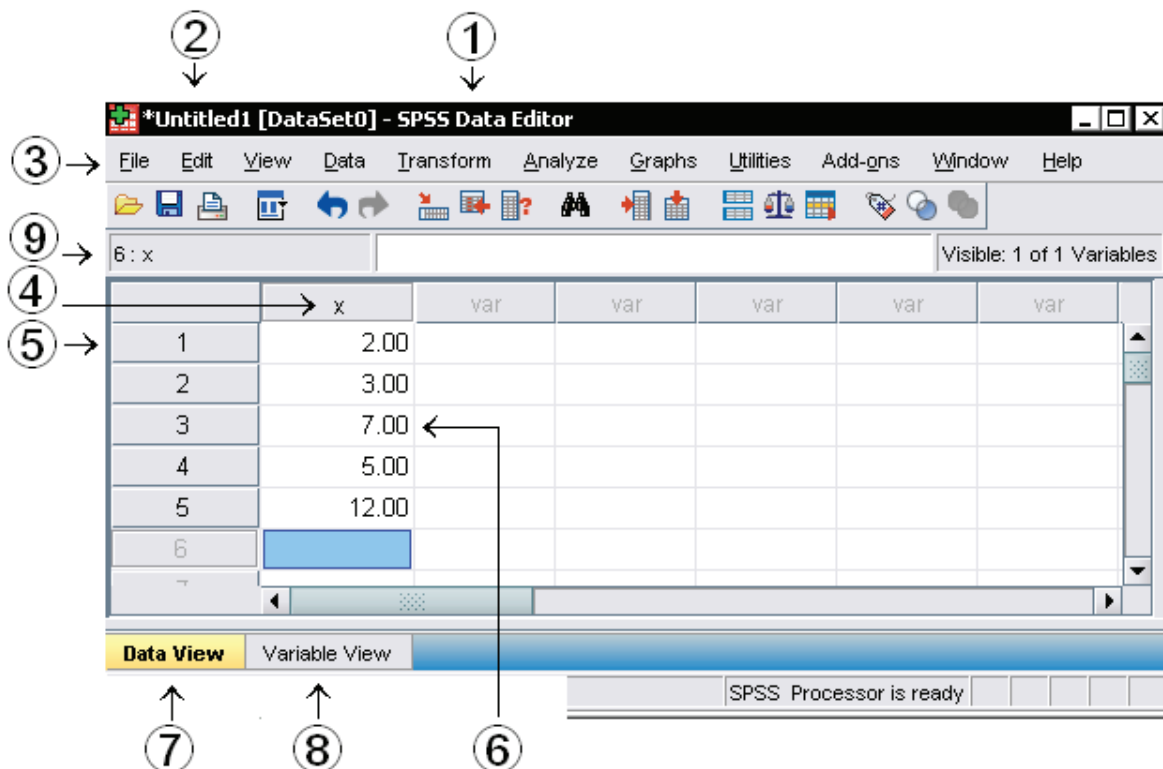
การทำงานของโปรแกรม SPSS มีการจำแนกส่วนของ WINDOW ที่สำคัญดังนี้

1. SPSS Data Editor

SPSS Data Editor เป็น Window สำหรับทำงานกับแฟ้มข้อมูลที่จะนำมาวิเคราะห์ด้วยโปรแกรม SPSS ซึ่งผู้ใช้อาจจะสร้างแฟ้มข้อมูลใหม่ หรือนำข้อมูลที่สร้างจากโปรแกรมอื่นๆ เรียกเข้ามาไว้ใน SPSS Data Editor แล้วใช้งานต่อไป ใน SPSS Data Editor แบ่งทำงานสองส่วนคือ Data View และ Variable View

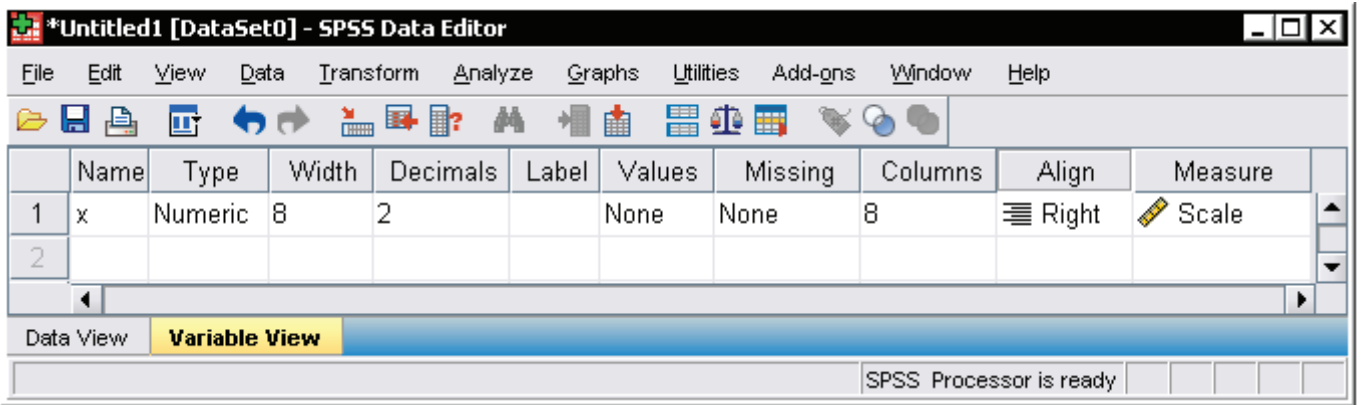
SPSS Data Editor สำหรับ SPSS 16.0 for Windows จะเปิด SPSS Data Editor ได้หลาย Window พร้อมกัน

ข้อควรทราบเกี่ยวกับ SPSS Data Editor ในส่วนการทำงานของ Data View



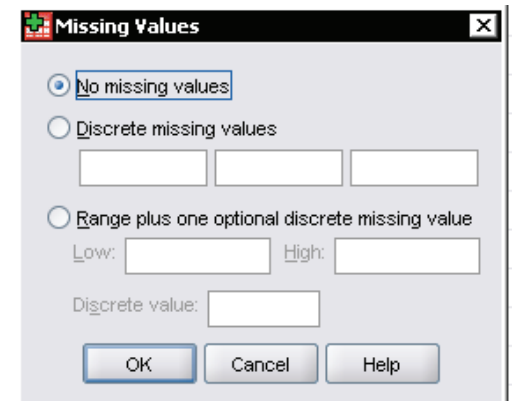
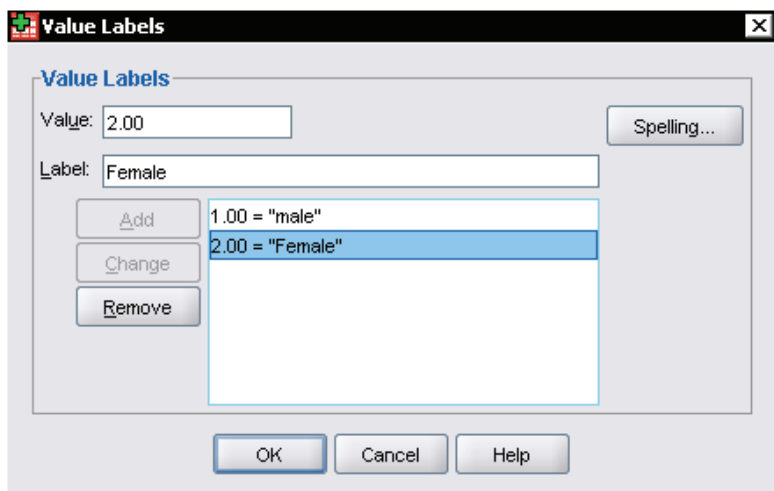
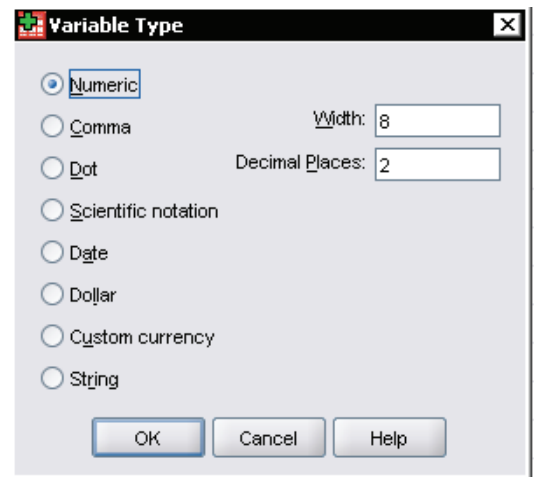
- หมายเลข 1 ชื่อชนิดของ Window ใน SPSS ขณะนี้คือ SPSS Data Editor
- หมายเลข 2 ชื่อแฟ้มข้อมูลที่กำลังใช้งาน หากยังไม่ได้ตั้งชื่อ SPSS จะใช้ชื่อว่า Untitled1
- หมายเลข 3 แถบเมนูของ SPSS Data Editor
- หมายเลข 4 ชื่อตัวแปร x ของข้อมูล
- หมายเลข 5 ลำดับที่ของค่าสังเกตในแฟ้มข้อมูล
- หมายเลข 6 ค่าของข้อมูล ค่าสังเกตตัวที่ 3 ของตัวแปร x
- หมายเลข 7 เป็นการเลือกทำงานในส่วนของ Data View
Data View เป็นส่วนทำงานเกี่ยวกับข้อมูลเช่นการวิเคราะห์ข้อมูลเช่น การใส่ค่าของข้อมูล
- หมายเลข 8 เป็นการเลือกทำงานในส่วนของ Variable View
Variable View เป็น Window ที่ทำงานเกี่ยวกับการกำหนดค่าต่าง ๆ ให้กับตัวแปร
- หมายเลข 9 แสดงตำแหน่งของ ค่าสังเกต และ ตัวแปรที่ cell pointer กำลังทำงาน

ข้อควรทราบเกี่ยวกับ SPSS Data Editor ในส่วนของการทำงาน Variable View



ความหมายของแต่ละ columns

- Name** กำหนดชื่อตัวแปร
- Type** กำหนดค่าตัวแปรเป็น Numeric, String
- Width** กำหนดจำนวนหลักของตัวเลข หรือจำนวน character ในการแสดงผล
- Decimals** กำหนดตำแหน่งทศนิยม
- Label** กำหนดคำอธิบายค่าตัวแปร
- Values** กำหนดค่าให้กับ Value Label



Missing กำหนดค่าสำหรับข้อมูลที่ไม่สมบูรณ์

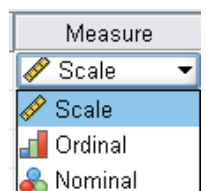
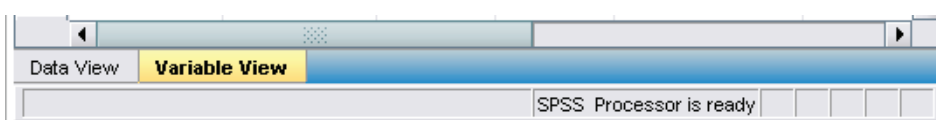
Columns กำหนดความกว้างของ columns ในการแสดงผลของ Data View

Align กำหนดการแสดงผลว่าต้องการ ชิดซ้าย ชิดขวา หรือ กึ่งกลาง



Measure กำหนดลักษณะข้อมูลว่าเป็นข้อมูลเชิงปริมาณ (Scale) หรือ ข้อมูลเชิงคุณภาพ (Ordinal)

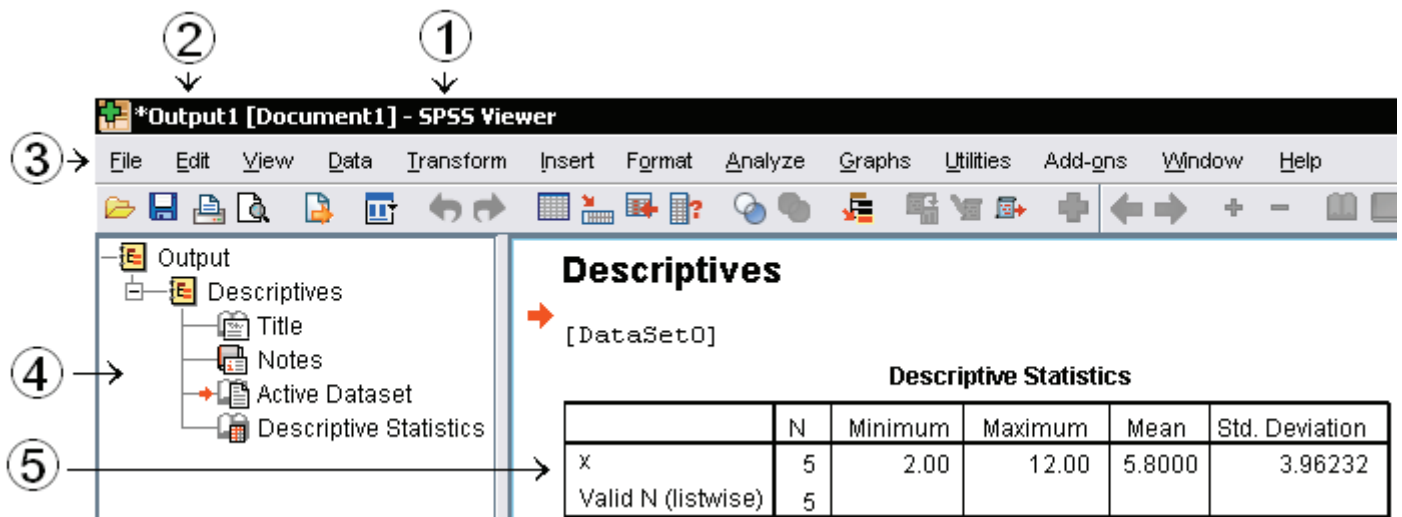
ข้อควรจำ การเลือกทำงานกับ Data View หรือ Variable View ให้คลิกที่มุมล่างด้านซ้าย



2. SPSS Viewer

SPSS Viewer เป็น Window สำหรับเก็บบันทึกผลลัพธ์ของการวิเคราะห์ข้อมูล ที่เกิดขึ้นจากการใช้งานโปรแกรม SPSS โดยจะบันทึกผลลัพธ์ที่เกิดขึ้นทุกครั้งที่มีการวิเคราะห์ข้อมูล และผลลัพธ์จะถูกบันทึกอย่างต่อเนื่องจนกว่าจะมีการสั่งให้บันทึกผลลัพธ์ใน SPSS Viewer อื่น ผู้ใช้สามารถเปิด SPSS Viewer ได้มากกว่า 1 SPSS Viewer ถ้ามีการเปิด SPSS Viewer มากกว่า 1 Window จะต้องมีการกำหนด SPSS Viewer ให้ทำหน้าที่เก็บผลลัพธ์ที่เกิดจากการประมวลผล

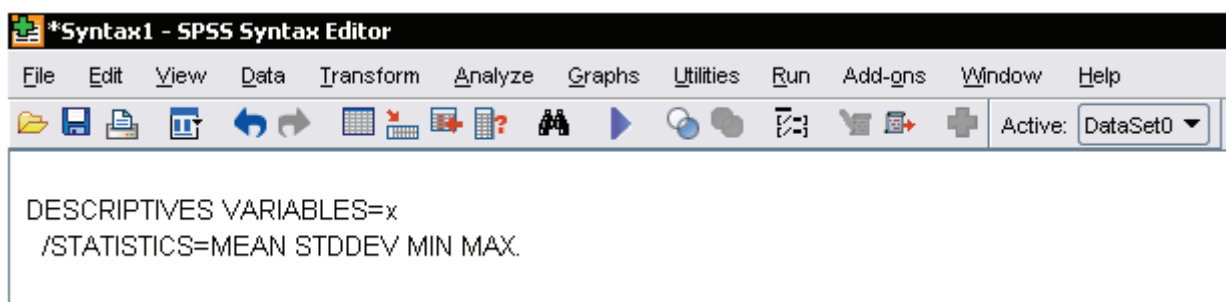
ข้อควรทราบเกี่ยวกับ SPSS Viewer



- หมายเลข 1 ชื่อชนิดของ Window ใน SPSS ขณะนี้คือ SPSS Viewer
- หมายเลข 2 ชื่อแฟ้ม Output ที่กำลังใช้งาน หากยังไม่ได้ตั้งชื่อจะใช้ชื่อว่า Output1
- หมายเลข 3 แถบเมนูของ SPSS Viewer
- หมายเลข 4 แผนภูมิต้นไม้แสดงลำดับและตำแหน่งของการแสดงผล
- หมายเลข 5 ผลของการวิเคราะห์ข้อมูล

3. SPSS Syntax Editor

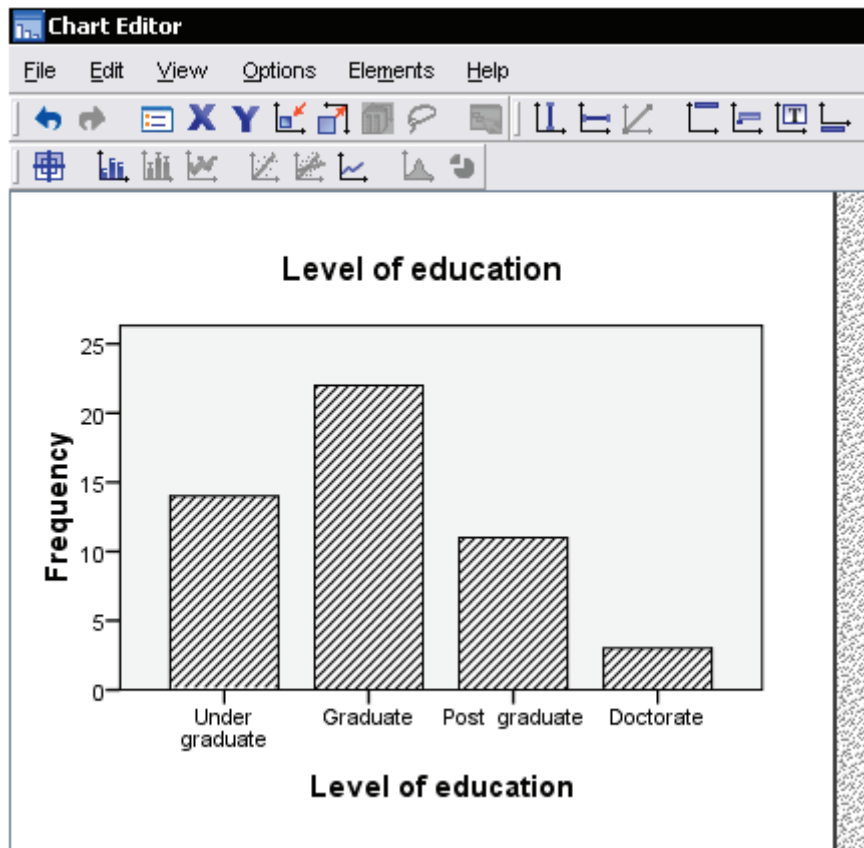
SPSS Syntax Editor เป็น Window สำหรับเก็บบันทึกคำสั่งที่ได้จากการใช้งานโปรแกรม SPSS ตามขั้นตอนต่างๆ ที่ทำของผู้ใช้ขณะนั้น (โดยการคลิกที่ Paste) ให้ผู้ใช้นำคำสั่งที่เกิดขึ้นนี้มาใช้ได้อีกโดยไม่ต้องสั่งการทำงานแบบเก่าซ้ำอีก หรือผู้ใช้สามารถเปลี่ยนแปลงแก้ไขใหม่ได้



หมายเหตุ อ่านเนื้อหาเพิ่มเติมเกี่ยวกับ SPSS Syntax ได้ที่ภาคผนวกที่ 4.

4. SPSS Chart Editor

SPSS Chart Editor เป็น Window ของการสร้าง หรือแก้ไขกราฟ เพื่อให้ผู้ใช้เปลี่ยนแปลง แก้ไข กราฟที่สร้างขึ้นมา เช่น เปลี่ยนรูปแบบตัวอักษร เปลี่ยนสี ฯลฯ



1.5 สรุปเนื้อหาของคำสั่งและขั้นตอนการทำงานโดยย่อของ SPSS for Windows

1. ประเภทของ Windows ในโปรแกรม SPSS for Windows

1.1 SPSS Data Editor

เป็น Window ที่เก็บแฟ้มข้อมูลที่จะนำมาวิเคราะห์ด้วยโปรแกรม SPSS

1.2 SPSS Viewer

เป็น Window ที่เก็บบันทึกรวบรวมผลลัพธ์ที่เกิดขึ้นจากการใช้งานโปรแกรม SPSS สามารถเปิดได้ครั้งละหลายๆ Window พร้อมๆ กัน

1.3 SPSS Syntax Editor

เป็น Window ที่เก็บบันทึกคำสั่งที่ได้จากการใช้งานโปรแกรม SPSS ตามขั้นตอนต่างๆ มารวบรวมไว้ เพื่อประโยชน์ในการนำคำสั่งมาใช้ภายหลัง

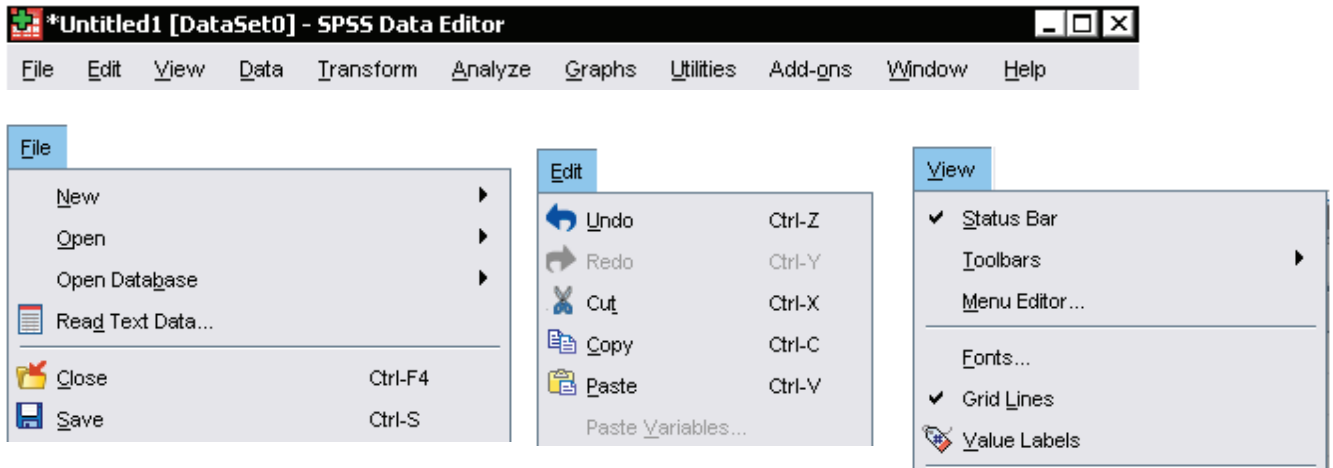
1.4 SPSS Chart Editor

เป็น Window ที่เก็บบันทึกรวบรวมกราฟ ต่างๆ ทั้งหมดที่เกิดขึ้นจากการโปรแกรม SPSS และเป็น Window ของกราฟ มีเมนูสำหรับให้ผู้ใช้เปลี่ยนแปลง แก้ไขรายละเอียดต่างๆ

1.5 Help Window

เป็น Window ที่เก็บข้อมูลรายละเอียด คำสั่ง คำอธิบาย ตัวอย่างการใช้งานต่างๆ ของโปรแกรม SPSS

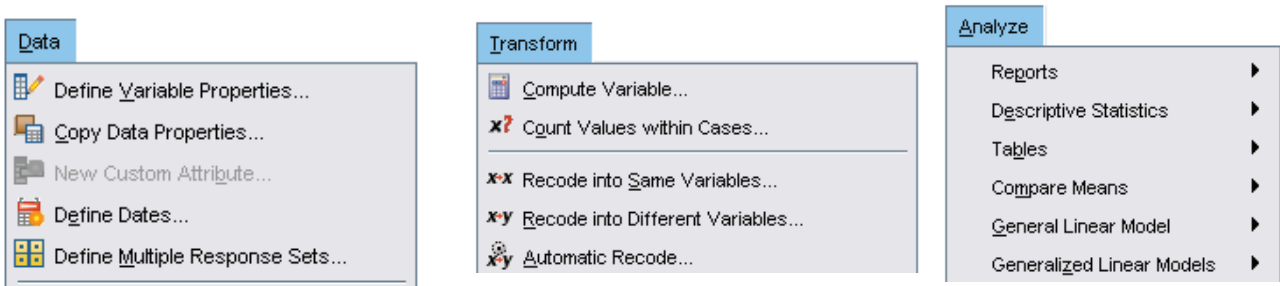
2. Menu ของโปรแกรม SPSS Data Editor



File ใช้เปิดแฟ้มข้อมูล บันทึกข้อมูล พิมพ์ข้อมูล ฯลฯ

Edit ใช้ย้ายข้อมูล คัดลอกข้อมูล ค้นหาข้อมูล ลบข้อมูล

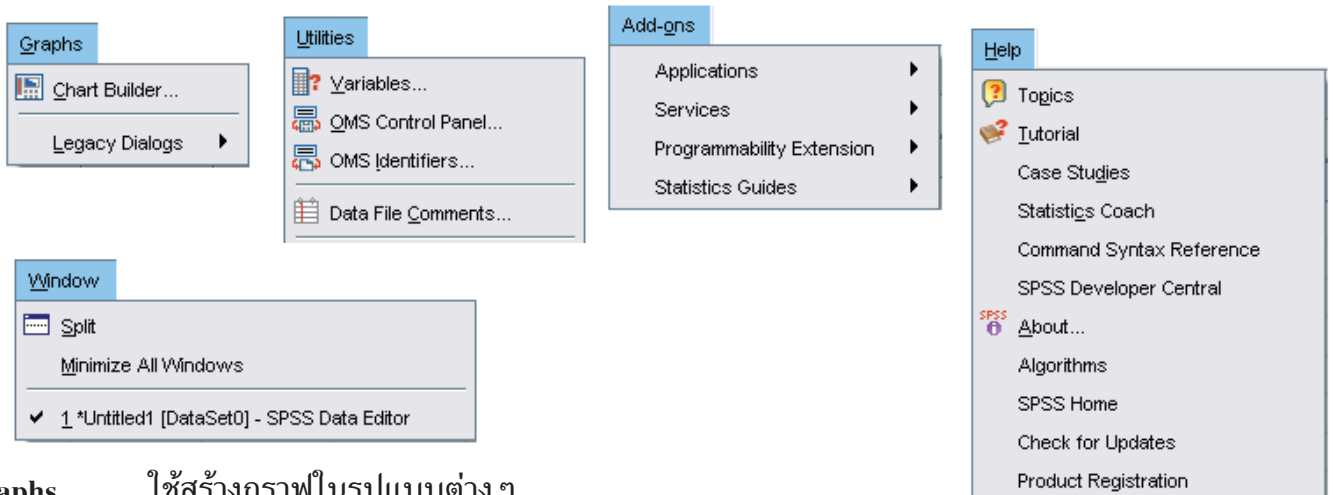
View ปรับรูปแบบและขนาดตัวอักษร แสดง Value Labels, Toolbars, เลือก Data View, Variable View



Data ใช้จัดการกับข้อมูลเช่น สร้างตัวแปร แก้ไข เรียงลำดับข้อมูล รวมแฟ้ม แทรกตัวแปร

Transform ใช้สร้างตัวแปรเพิ่ม หรือ จัดค่าตัวแปรใหม่ สร้างตัวแปรใหม่จากตัวแปรเก่า

Analyze ใช้เรียกคำสั่งเกี่ยวกับการวิเคราะห์ข้อมูลทางสถิติ



Graphs ใช้สร้างกราฟในรูปแบบต่างๆ

Utilities ใช้แสดงรายละเอียดตัวแปร กำหนดกลุ่มตัวแปร กำหนดรูปแบบเมนู

Add-ons เพิ่มเติมความสามารถของการวิเคราะห์ข้อมูล

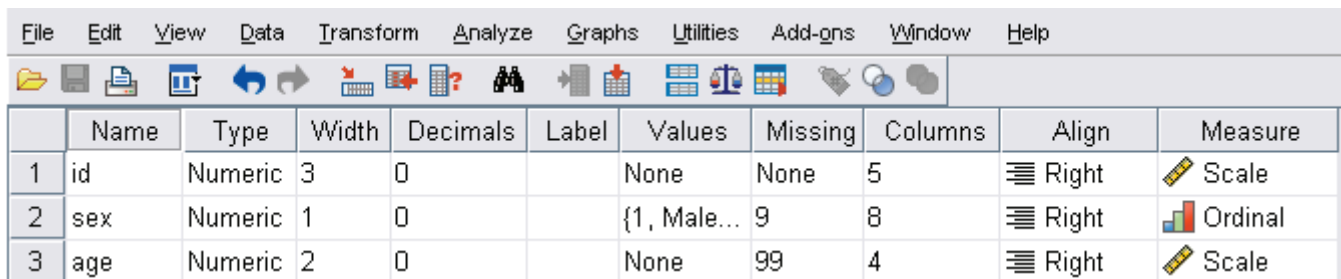
Window ใช้จัดเรียง Windows หรือเลือก Window ของ SPSS ขึ้นมาใช้งาน

Help ใช้แสดงคำอธิบาย ในการใช้โปรแกรม SPSS for Windows

3. การจัดเตรียมข้อมูลโดย SPSS Data Editor

3.1 กำหนดชื่อตัวแปรและรายละเอียดของตัวแปร

- เลือก SPSS Data Editor และ คลิกที่ Variable View หรือเลือกคำสั่ง View / Variables



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	id	Numeric	3	0		None	None	5	Right	Scale
2	sex	Numeric	1	0		{1, Male...	9	8	Right	Ordinal
3	age	Numeric	2	0		None	99	4	Right	Scale

- กำหนดชื่อตัวแปรในช่อง **Name**
- ถ้าต้องการเปลี่ยนแปลงรายละเอียดต่าง ๆ ของตัวแปรให้กำหนดในแต่ละช่อง ซึ่งมีความหมายดังนี้
 - Type** กำหนดค่าตัวแปรเป็น Numeric, String
 - Width** กำหนดจำนวนหลักของตัวเลข หรือจำนวน character ในการแสดงผล
 - Decimals** กำหนดตำแหน่งทศนิยม
 - Label** กำหนดคำอธิบายค่าตัวแปร
 - Values** กำหนดค่าให้กับ Value Label
 - Missing** กำหนดค่าสำหรับข้อมูลที่ไม่สมบูรณ์
 - Columns** กำหนดความกว้างของ columns ในการแสดงผลของ Data View
 - Align** กำหนดการแสดงผลว่าต้องการ ชิดซ้าย ชิดขวา หรือ กึ่งกลาง
 - Measure** กำหนดลักษณะข้อมูลว่าเป็นข้อมูลเชิงปริมาณ (Scale) หรือข้อมูลเชิงคุณภาพ (Ordinal)

3.2 การพิมพ์ข้อมูล

- ใช้แป้น **Enter** สำหรับการพิมพ์ข้อมูลครั้งละ 1 ตัวแปร
- ใช้แป้น **↓ → ← ↑** สำหรับการพิมพ์ข้อมูลแล้วเลื่อนไปเซลล์ถัดไป
- ใช้แป้น **Tab** สำหรับการพิมพ์ข้อมูลครั้งละ 1 ชุด (แถว)

3.3 การบันทึกข้อมูล

- File / Save Data** สำหรับการบันทึกภายใต้ชื่อเดิม
- File / Save As** สำหรับการบันทึกภายใต้ชื่อใหม่

3.4 การเรียกใช้ข้อมูลที่บันทึกไว้แล้ว

- File / Open / Data** เลือกหรือพิมพ์ชื่อแฟ้มที่ต้องการ

3.5 การพิมพ์ ข้อมูล คำสั่ง หรือ ผลลัพธ์ออกเครื่องพิมพ์

- เลือก Window ที่ต้องการ (Data Editor, SPSS Viewer, SPSS Syntax)
- File / Print / OK**

4. การทำงานที่สำคัญใน SPSS Data Editor

4.1 การค้นหาชุดข้อมูลและตัวแปร

4.1.1 การค้นหาชุดข้อมูล

- Edit / Go to Case เสร็จแล้วพิมพ์ตำแหน่งของชุดข้อมูลที่ต้องการค้นหา

4.1.2 การค้นหาตัวแปร

- Edit / Go to Variables และ เลือกตัวแปรที่ต้องการ

4.2 การคัดลอก หรือ ย้ายข้อมูล

- เลือกข้อมูลที่ต้องการคัดลอก หรือ ย้ายข้อมูล เสร็จแล้วเลือกเมนู Edit / Copy หรือ Edit / Cut
- เลือกเซลล์ซึ่งเป็นตำแหน่งที่ต้องการคัดลอกข้อมูลมาไว้
- Edit / Paste

4.3 การแทรก หรือ ลบชุดข้อมูล

4.3.1 การแทรกชุดข้อมูล

- คลิกที่หัวแถวที่ต้องการแทรกไว้ (จะแทรกไว้เหนือแถวที่เลือก)
- Edit / Insert Cases

4.3.2 การลบชุดข้อมูล

- คลิกที่หัวแถวหรือกลุ่มของหัวแถว (drag ตามแถว)
- กด Delete

4.4 การแทรก หรือ ลบตัวแปร

4.4.1 การแทรกตัวแปร

- คลิกที่ชื่อตัวแปรที่ต้องการแทรก (จะแทรกไว้ข้างหน้าตัวแปรที่เลือก)
- Edit / Insert Variable

4.4.2 การลบตัวแปร

- คลิกที่ชื่อตัวแปร หรือกลุ่มของตัวแปร
- Edit / Clear (หรือกดแป้น Del)

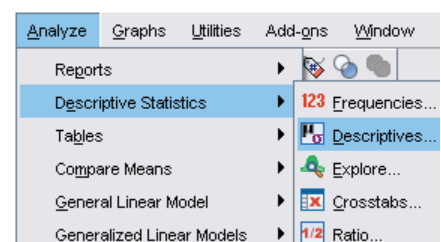
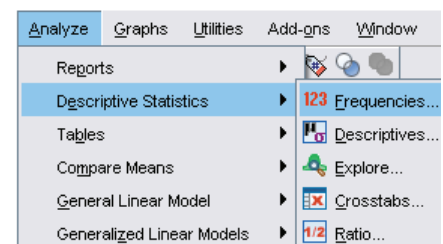
5. การวิเคราะห์ข้อมูลด้วยเมนู Analyze / Descriptive Statistics

5.1 การแจกแจงความถี่แบบทางเดียว

- Analyze / Descriptive Statistics / Frequencies
- เลือกตัวแปรไว้ในกรอบของ Variable(s)
- กำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Statistics, Chart หรือ Format
- คลิกปุ่ม OK

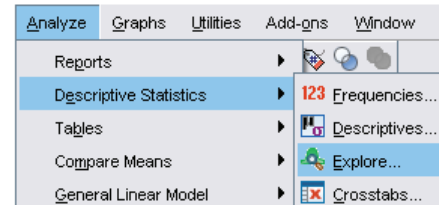
5.2 การคำนวณค่าสถิติเบื้องต้น

- Analyze / Descriptive Statistics / Descriptives
- เลือกตัวแปรไว้ในกรอบของ Variable(s)
- กำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Options เสร็จแล้วคลิก OK



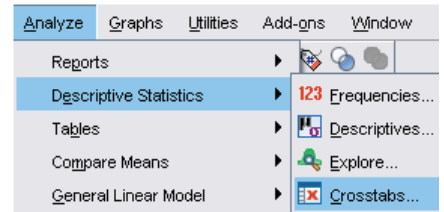
5.3 การตรวจสอบข้อมูล

- Analyze / Descriptive Statistics / Explore
- เลือกตัวแปรมาไว้ในกรอบของ Dependent List
- กำหนดรายละเอียดเพิ่มเติมที่ Statistics, Plots, Options
- คลิกปุ่ม OK



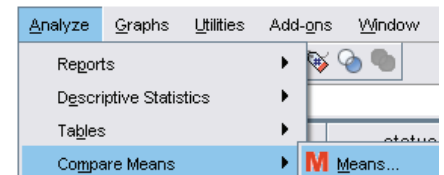
5.4 การแจกแจงความถี่ตั้งแต่ 2 ทาง

- Analyze / Descriptive Statistics / Crosstabs
- เลือกตัวแปรอย่างน้อย 1 ตัว ที่ต้องการ ให้อยู่ด้านแถวไว้ในกรอบของ Row[s]
- เลือกตัวแปรอย่างน้อย 1 ตัว ที่ต้องการ ให้อยู่ด้านหลังไว้ในกรอบของ Column[s]
- กำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Statistics, Cell หรือ Format เสร็จแล้วคลิกปุ่ม OK



5.5 การทดสอบค่าสัดส่วน Analyze / Descriptive Statistics / Ratio

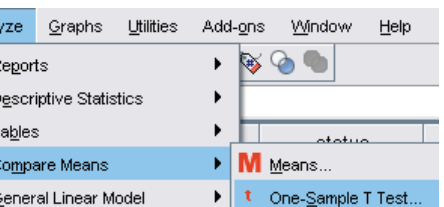
ทำการทดสอบเกี่ยวกับค่าอัตราส่วน



6. การวิเคราะห์ข้อมูลด้วยเมนู Analyze / Compare Means

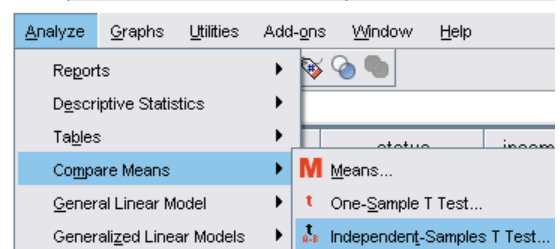
6.1 การคำนวณค่าสถิติเบื้องต้นจำแนกตามกลุ่ม

- Analyze / Compare Means / Means
- เลือกตัวแปรไว้ในกรอบของ Dependent List และ Independent List
- กำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Options
- คลิกปุ่ม OK



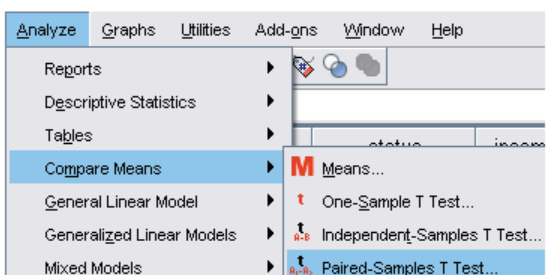
6.2 การทดสอบค่าเฉลี่ย 1 กลุ่ม

- Analyze / Compare Means / One-Sample T Test
- เลือกตัวแปรไว้ในกรอบของ Test Variable[s]
- กำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Options
- คลิกปุ่ม OK



6.3 การทดสอบค่าเฉลี่ย 2 กลุ่มที่เป็นอิสระต่อกัน

- Analyze / Compare Means / Independent-Samples T Test
- เลือกตัวแปรไว้ในกรอบของ Test Variable[s] และ Grouping Variables
- กำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Options
- คลิกปุ่ม OK

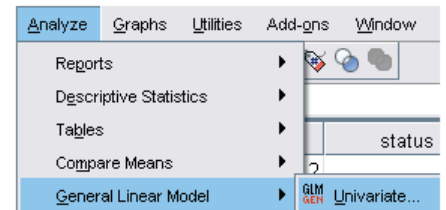


6.4 การทดสอบค่าเฉลี่ย 2 กลุ่มที่มีความสัมพันธ์กัน

- Analyze / Compare Means / Paired-Samples T Test
- เลือกตัวแปรมาไว้ในกรอบของ Paired Variables
- กำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Options เสร็จแล้วคลิกปุ่ม OK

6.5 การวิเคราะห์ความแปรปรวนแบบจำแนกทางเดียว

- Analyze / Compare Means / One-Way ANOVA
- เลือกตัวแปรอย่างน้อย 2 ตัวไว้ในกรอบของ Dependent List และ Factor(s)
- กำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Contrasts, Options
- การทดสอบหาคู่ที่ค่าเฉลี่ยต่างกัน Post Hoc เสร็จแล้วคลิก OK



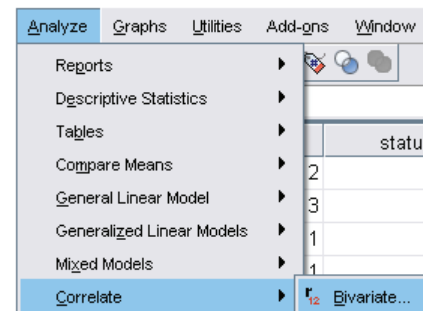
6.6 การวิเคราะห์ความแปรปรวนแบบจำแนกหลายทาง

- Analyze / General Linear Model / Univariate
- เลือกตัวแปรอย่างน้อย 2 ตัวไว้ในกรอบของ Dependent Variables และ Fixed Factor
- กำหนดรายละเอียดเพิ่มเติมที่กรอบของ Model, Covariate[s] หรือ Options
- กำหนดการทดสอบหาคู่ที่ค่าเฉลี่ยต่างกัน Post Hoc เสร็จแล้วคลิกปุ่ม OK

7. การวิเคราะห์ข้อมูลด้วยเมนู Analyze / Correlate หรือ Regression

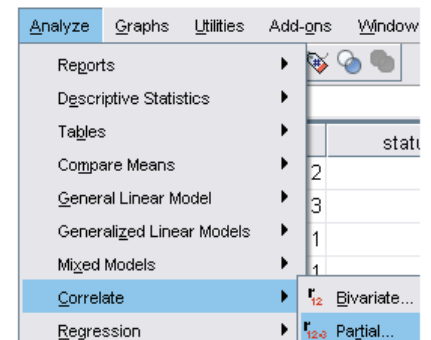
7.1 การหาความสัมพันธ์ของข้อมูลเชิงปริมาณ

- Analyze / Correlate / Bivariate
- เลือกตัวแปรไว้ในกรอบของ Variables
- เลือกวิธีการวิเคราะห์ที่จะใช้ในส่วนของ Correlation Coefficients
- เลือกวิธีการทดสอบในส่วนของ Test of Significance
- กำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Options เสร็จแล้วคลิกปุ่ม OK



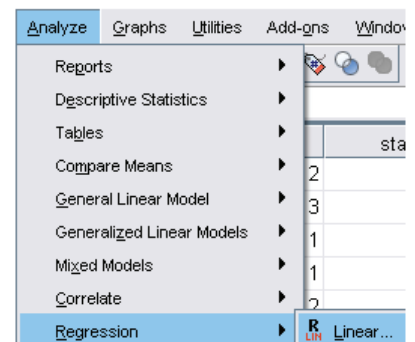
7.2 การหาความสัมพันธ์บางส่วนข้อมูลเชิงปริมาณ

- Analyze / Correlate / Partial
- เลือกตัวแปรไว้ในกรอบของ Variables และ Controlling for
- เลือกวิธีการทดสอบในส่วนของ Test of Significance
- กำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Options
- คลิกปุ่ม OK



7.3 การพยากรณ์โดยวิธีวิเคราะห์การถดถอย

- Analyze / Regression / Linear
- เลือกตัวแปรตามไว้ในกรอบของ Dependent
- เลือกตัวแปรอิสระอย่างน้อย 1 ตัวไว้ในกรอบของ Independent[s]
- กำหนดรายละเอียดเพิ่มเติมที่ปุ่ม WLS, Statistics, Plot, Save, Options
- เสร็จแล้วคลิกปุ่ม OK



7.4 การเลือกรูปแบบของการพยากรณ์

- Analyze / Regression / Curve Estimation
- เลือกตัวแปรตามไว้ในกรอบของ Dependent
- เลือกตัวแปรอิสระไว้ในกรอบของ Independent
- กำหนดรายละเอียดเพิ่มเติมที่ปุ่ม WLS, Statistics, Plot, Save, Options เสร็จแล้วคลิกปุ่ม OK



8. การวิเคราะห์ข้อมูลด้วยเมนู Analyze / Nonparametric Tests

8.1 การทดสอบอัตราส่วน

- Analyze / Nonparametric Tests / Chi-Square
- เลือกตัวแปรที่ต้องการทดสอบไว้ในกรอบของ Test Variable List
- กำหนดค่าความถี่ใหม่ที่ต้องการไว้ในส่วนของ Expected Values
- กำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Options
- คลิกปุ่ม OK

8.2 การทดสอบสัดส่วน

- Analyze / Nonparametric Tests / Binomial
- เลือกตัวแปรที่ต้องการทดสอบไว้ในกรอบของ Test Variables List
- กำหนดค่าสัดส่วนใหม่ที่ต้องการไว้ในส่วนของ Test Proportion
- กำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Options
- คลิกปุ่ม OK

8.3 การทดสอบความเป็นตัวอย่างสุ่ม

- Analyze / Nonparametric Tests / Runs
- เลือกตัวแปรที่ต้องการทดสอบไว้ในกรอบของ Test Variable List
- เลือกวิธีการแบ่งกลุ่มข้อมูลเพิ่มอีกในส่วนของ Cut Point
- กำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Options
- คลิกปุ่ม OK

8.4 การทดสอบรูปแบบการแจกแจงของข้อมูล

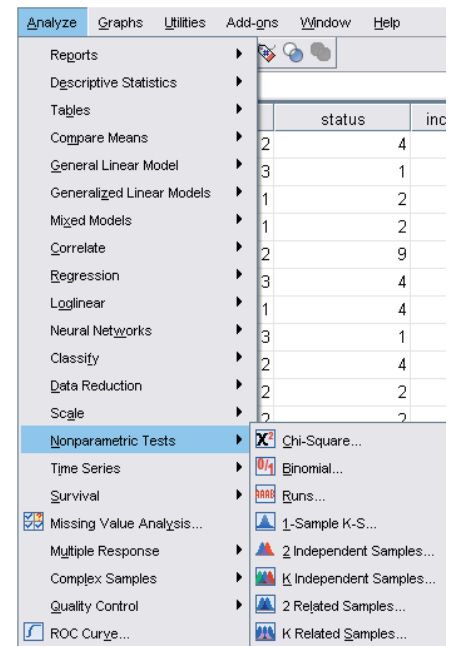
- Analyze / Nonparametric Tests / 1-Sample K-S
- เลือกตัวแปรที่ต้องการทดสอบไว้ในกรอบของ Test Variable List
- เลือกวิธีการแบ่งกลุ่มข้อมูลเพิ่มอีกในส่วนของ Cut Point
- กำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Options
- คลิกปุ่ม OK

8.5 การทดสอบสำหรับข้อมูล 2 กลุ่มที่เป็นอิสระต่อกัน

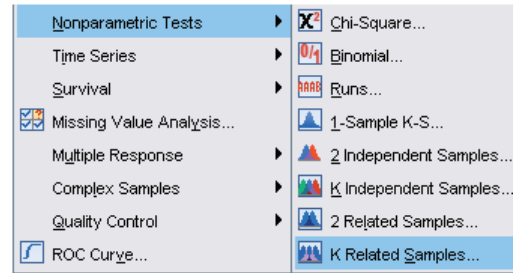
- Analyze / Nonparametric Tests / 2 Independents Samples
- เลือกตัวแปรที่ต้องการทดสอบไว้ในกรอบของ Test Variable List
- เลือกตัวแปรที่ต้องการเป็นตัวแบ่งกลุ่มไว้ในกรอบของ Grouping Variable
- เลือกวิธีทางสถิติที่จะใช้ทดสอบในส่วนของ Test Type
- กำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Options
- คลิกปุ่ม OK

8.6 การทดสอบสำหรับข้อมูล k กลุ่มที่เป็นอิสระต่อกัน

- Analyze / Nonparametric Tests / K Independent Samples



- เลือกตัวแปรที่ต้องการทดสอบไว้ในกรอบของ Test Variables List
- เลือกตัวแปรที่ต้องการเป็นตัวแบ่งกลุ่มไว้ในกรอบของ Grouping Variable
- เลือกวิธีการสถิติที่จะใช้ทดสอบในส่วนของ Test Type
- กำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Options
- คลิกปุ่ม OK



8.7 การทดสอบสำหรับข้อมูล 2 กลุ่มที่มีความสัมพันธ์

- Analyze / Nonparametric Tests / 2 Related Samples
- เลือกตัวแปรที่ต้องการทดสอบ 2 ตัวไว้ในกรอบของ Test Variable List
- เลือกวิธีการสถิติที่จะใช้ทดสอบในส่วนของ Test Type
- กำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Options เสร็จแล้วคลิกปุ่ม OK

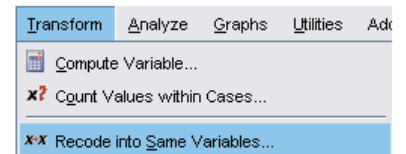
8.8 การทดสอบสำหรับข้อมูล k กลุ่มที่มีความสัมพันธ์

- Analyze / Nonparametric Tests / K Related Samples
- เลือกตัวแปรอย่างน้อย 2 ตัวแปรไว้ในกรอบของ Test Variable List
- เลือกวิธีการสถิติที่จะใช้ทดสอบในส่วนของ Test Type
- กำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Statistics
- คลิกปุ่ม OK

9. การปรับปรุงข้อมูลด้วยเมนู Transform

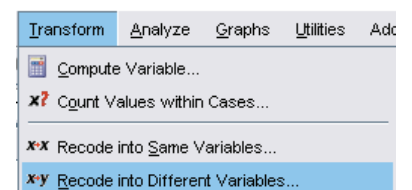
9.1 การเปลี่ยนค่าตัวแปรไว้ในตัวแปรเดิม

- Transform / Recode into Same Variables
- เลือกตัวแปรที่ต้องการเปลี่ยนค่าไว้ในกรอบของ Variables
- คลิกปุ่ม Old and New Values
 - ◆ กำหนดค่าที่ต้องการเปลี่ยนในกรอบของ Old Value
 - ◆ กำหนดค่าใหม่ที่จะแทนค่าเดิมในกรอบของ New Value
- คลิกปุ่ม Continue
- ถ้าต้องเปลี่ยนข้อมูลบางชุดให้เลือกที่ปุ่ม If
- คลิกปุ่ม OK



9.2 การเปลี่ยนค่าตัวแปรไว้ในตัวแปรใหม่

- Transform / Recode into Different Variables
- เลือกตัวแปรที่ต้องการเปลี่ยนค่าไว้ในกรอบของ Variables
- กำหนดชื่อตัวแปรใหม่ในกรอบของ Output Variable
- กำหนดข้อความขยายชื่อตัวแปรไว้ในกรอบของ Label / เลือก Change
- คลิกปุ่ม Old and New Values
 - ◆ กำหนดค่าที่ต้องการเปลี่ยนในกรอบของ Old Value

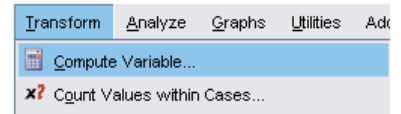


◆ กำหนดค่าใหม่ที่จะแทนค่าเดิมในกรอบของ New Value

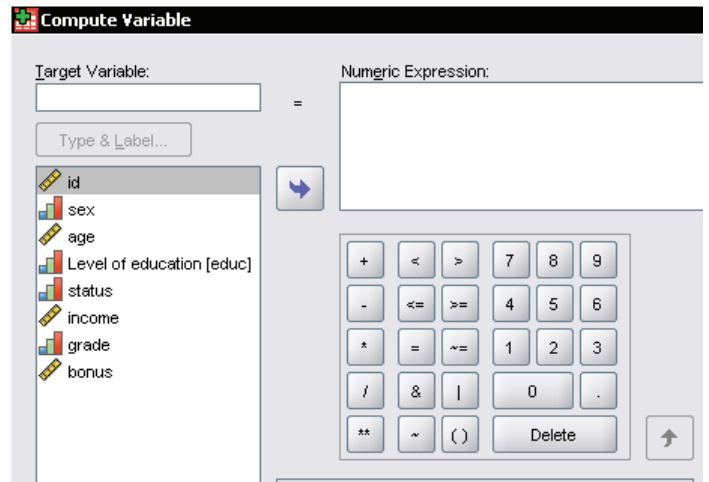
- คลิกปุ่ม Continue
- ถ้าต้องเปลี่ยนข้อมูลบางชุดให้เลือกที่ปุ่ม If
- คลิกปุ่ม OK

9.3 การสร้างตัวแปรใหม่จากการคำนวณและเงื่อนไข

- Transform / Compute Variable
- กำหนดชื่อตัวแปรใหม่หรือใช้ตัวแปรเก่าในกรอบของ Target Variable
- กำหนดสูตรทางคณิตศาสตร์ของตัวแปรไว้ในกรอบของ Numeric Expression



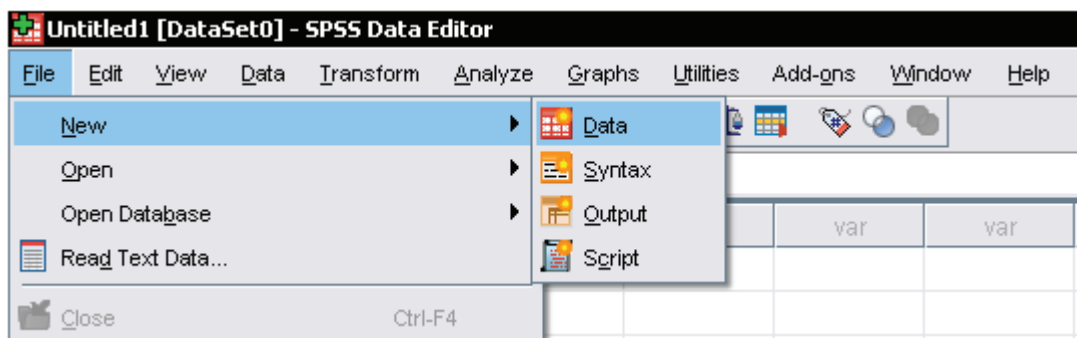
- ถ้าต้องการสร้างตัวแปรแบบมีเงื่อนไขให้เลือกที่ปุ่ม If
- คลิกปุ่ม OK



10. การเปิด Windows หลายแบบพร้อมกัน

เมื่อเริ่มใช้โปรแกรม SPSS ครั้งแรกของการเรียกโปรแกรมขึ้นมาจะปรากฏ Window SPSS Data Editor เมื่อทำการวิเคราะห์ข้อมูลจะเกิด Window SPSS Viewer ถ้าต้องการเปิด Window อื่นๆ เพิ่มเติมสามารถทำได้ดังนี้

คลิกที่เมนู File และเลือกรายการ New หรือ Open จะมีชนิดของ Window ให้เลือก 4 ชนิดคือ



- ◆ Data สำหรับเปิดแฟ้มข้อมูลของ SPSS Data Editor
- ◆ Syntax สำหรับเปิดโปรแกรม Syntax ของ SPSS Syntax Editor
- ◆ Output สำหรับเปิดผลการวิเคราะห์ข้อมูลของ SPSS Viewer
- ◆ Script สำหรับเปิด SPSS for Windows ที่จัดการเกี่ยวกับโปรแกรม
- คลิกที่ชนิดที่ต้องการ

11. การบันทึกข้อมูลใน Windows

ผู้ใช้สามารถบันทึกข้อมูลใน Window ที่ถูกเปิดขึ้นมาใช้งาน โดยบันทึกไว้ในรูปของแฟ้ม ซึ่งโปรแกรม SPSS ได้จัดแบ่งประเภทของแฟ้มดังนี้

ชนิดแฟ้มของ Window	ส่วนขยายของแฟ้ม	ชื่อที่ SPSS ตั้งให้ชั่วคราว
SPSS Data Editor	*.sav	Untitled1
SPSS Viewer	*.spv	Output1
SPSS Syntax Editor	*.sps	Syntax1

ข้อสังเกต โปรแกรม Output1 ของ SPSS version 12.0, version 15.0 จะมีนามสกุลเป็น spo

การบันทึกข้อมูลที่อยู่ใน Window ใดๆ ทำได้ดังนี้

- เลือก Window ที่จะบันทึกข้อมูลโดยการใช้เมาส์คลิกบริเวณใดๆ ใน Window ที่ต้องการจะปรากฏแถบแสงที่ชื่อ Window นั้น เปิดเมนู File และเลือกรายการใดรายการหนึ่ง
 - ◆ Save ชื่อ และ ชนิดของแฟ้มข้อมูลสำหรับบันทึกภายใต้ชื่อแฟ้มเดิมที่เคยบันทึกไว้แล้ว
 - ◆ Save as สำหรับการบันทึกภายใต้ชื่อแฟ้มใหม่
- กำหนดชื่อ ตำแหน่งไดรฟ์ และ ประเภทของแฟ้ม
- คลิกปุ่ม Save

12. การเปิดแฟ้มข้อมูล

แฟ้มข้อมูลของ Window ที่ถูกบันทึกไว้แล้วเมื่อต้องการนำมาใช้ต้องทำดังนี้

- คลิกเมนู File และเลือกรายการ Open จะปรากฏรายการให้เลือกตามชนิดของ Window ต่างๆ ในความหมายของต่อไปนี้
 - ◆ Data สำหรับเปิด SPSS Data Editor
 - ◆ Syntax สำหรับเปิด SPSS Syntax Editor
 - ◆ Output สำหรับเปิด SPSS Viewer
 - ◆ Script สำหรับเปิด SPSS for Windows ที่จัดการเกี่ยวกับโปรแกรม
- พิมพ์ชื่อที่ต้องการ แล้วคลิก Open

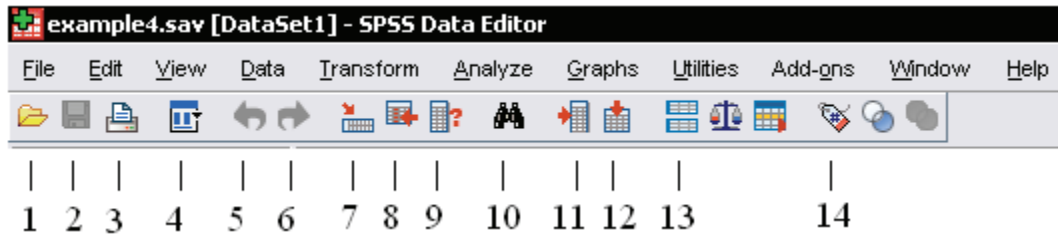
13. การบันทึกข้อมูล

การบันทึกข้อมูลใน Window มีขั้นตอนที่สำคัญดังนี้

- เลือกชนิด Window ที่ต้องการบันทึกข้อมูลเช่น Data, Output, Syntax, ...
- เลือกเมนู File
- เลือกรายการ Save หรือ Save as
- กำหนดชื่อแฟ้ม และตำแหน่งที่จะบันทึกตามความต้องการ
- คลิกปุ่ม Save

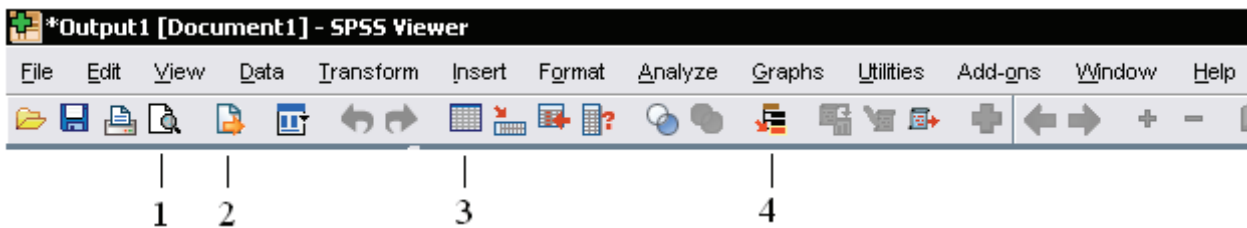
1.6 Icon บนเมนูบาร์กับการทำงานของ SPSS for Windows

SPSS Data Editor



- | | |
|----------------------------|--|
| 1. เปิดแฟ้มข้อมูล | 2. Save ข้อมูล |
| 3. พิมพ์ข้อมูล | 4. ดับเบิลคลิกคำสั่งล่าสุดที่วิเคราะห์ข้อมูล |
| 5. Undo | 6. Redo |
| 7. ไปหาคำสั่งที่ต้องการ | 8. ไปหาค่าตัวแปรที่ต้องการ |
| 9. แสดงรายละเอียดของตัวแปร | 10. ค้นหาข้อมูล |
| 11. แทรกตัวแปร | 12. แทรกตัวแปร |
| 13. แยกแฟ้มเป็น 2 ส่วน | 14. แสดงผลเป็น Value Label หรือค่าตัวเลข |

SPSS Viewer



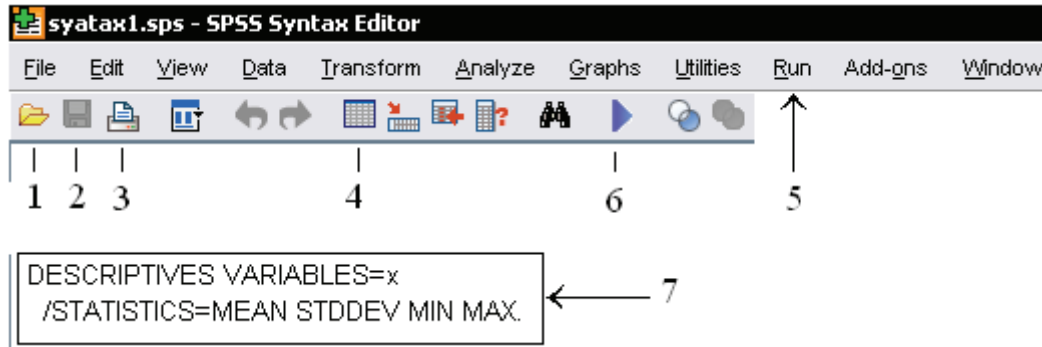
- | | |
|-------------------------------|-----------------------|
| 1. พิมพ์ Output ดูแบบ Preview | 2. Export Output |
| 3. กลับไปที่ SPSS Data Editor | 4. Select last output |

SPSS Chart Editor



- | | |
|-----------------------------|-----------------------------|
| 1. Show Properties Windows | 2. คำสั่งในการ Format แกน X |
| 3. คำสั่งในการ Format แกน Y | 4. แทรก Title |
| 5. แทรกข้อความในกราฟ | 6. แทรก Footnote |
| 7. Show / Hide grid line | 8. หมุนกราฟ 90 องศา |

SPSS Syntax Editor



1. เปิดแฟ้มชนิด Syntax
2. บันทึกแฟ้มชนิด Syntax
3. สั่งพิมพ์แฟ้มข้อมูลใน SPSS Syntax Editor
4. ไป SPSS Data Editor
5. สั่งให้โปรแกรมใน Syntax Editor ทำงานด้วยคำสั่ง Run
6. สั่งให้โปรแกรม Syntax ขณะนั้นเริ่มต้นทำงานที่บรรทัดที่ Cursor อยู่
7. ตัวโปรแกรม Syntax

หมายเหตุ



เปิดแฟ้มตามชนิดของ Windows ขณะนั้น



พิมพ์ข้อมูลของ Windows ขณะนั้น



บันทึกข้อมูลตามชนิดของ Windows ขณะนั้น



ค้นหาข้อความ



ปุ่มสลับการแสดงค่าตัวแปร หรือ Value Label ของตัวแปร



ปุ่มแสดง Variable Information

หนังสือ ความน่าจะเป็นและสถิติ สรุปเนื้อหา โจทย์แบบฝึกหัดและเฉลย

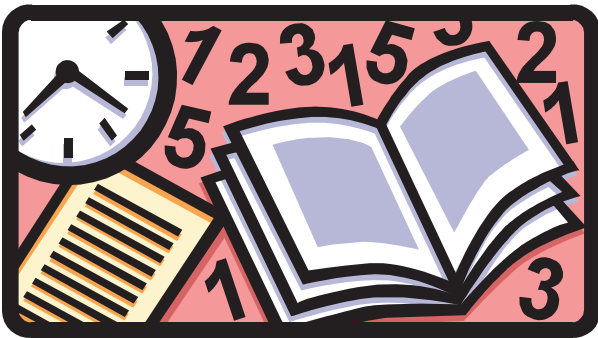
เขียนโดย รองศาสตราจารย์ ดำรงค์ ทิพย์โยธา

ภายในเล่มประกอบด้วย การนับ ความน่าจะเป็น ความน่าจะเป็นแบบมีเงื่อนไข ตัวแปรสุ่มต่อเนื่อง และ ไม่ต่อเนื่อง ค่าคาดหวัง ความแปรปรวน การแจกแจง ทวินาม ไฮเพอร์จีโอเมตริก ปัวส์ซอง ปกติ ไคสแควร์ ที และ เอฟ การแจกแจงค่าสถิติของตัวอย่าง การหาช่วงความเชื่อมั่น การทดสอบสมมติฐาน สหสัมพันธ์ การถดถอยเชิงเส้น การวิเคราะห์ความแปรปรวน (ANOVA) การทดสอบแบบไม่อิงพารามิเตอร์ มีโจทย์ 392 ข้อพร้อมเฉลย และ วิเคราะห์ข้อมูลด้วย SPSS for Windows, Mathcad เหมาะสำหรับ นิสิต และ นักศึกษาทุกคนที่ต้องเรียนสถิติ ภายในเล่มมีสรุปเนื้อหา

จัดจำหน่ายโดย ศูนย์หนังสือจุฬาลงกรณ์มหาวิทยาลัย

บทที่ 2

การสร้างแฟ้มข้อมูล



สิ่งที่สำคัญของผู้ที่ต้องการวิเคราะห์ข้อมูลต้องทำคือ การวางแผนเก็บข้อมูล การสร้างแบบสอบถาม การแปลความหมายแบบสอบถามเพื่อเป็นข้อมูลของการวิเคราะห์ด้วยโปรแกรม SPSS for Windows ตัวอย่างเช่น บริษัทแห่งหนึ่งต้องการเก็บรวบรวมข้อมูลและวิเคราะห์ข้อมูลพนักงานเกี่ยวกับ เพศ อายุ ระดับการศึกษา สถานะภาพการแต่งงาน เงินเดือน ระดับคะแนนความสามารถ และ เงินตอบแทนประจำปี จึงทำการสอบถามข้อมูลด้วยแบบสอบถามดังนี้

แบบสอบถามข้อมูลพนักงาน

สำหรับเจ้าหน้าที่กรอกข้อมูล

1. เลขประจำตัว.....

2. เพศ ชาย หญิง

3. อายุ ปี

4. ระดับการศึกษา

ต่ำกว่าระดับปริญญาตรี

จบระดับปริญญาตรี

จบระดับปริญญาโท

จบระดับปริญญาเอก

5. สถานะภาพ

โสด

แต่งงานแล้ว

เป็นหม้าย

หย่าร้าง

6. เงินเดือนบาท

7. ระดับคะแนนความสามารถ.....

8. เงินตอบแทนประจำปี.....บาท

ข้อกำหนดในการสร้างแฟ้มข้อมูล

จากแบบสอบถามที่ผู้ที่ต้องการวิเคราะห์ข้อมูล เมื่อต้องการจะทำเป็นข้อมูลสำหรับ SPSS for Windows ต้องทำการกำหนดค่าต่างๆ เช่น ชื่อแฟ้ม (file name) ชื่อตัวแปร (variable name) ชนิดของค่าตัวแปร (variable type) กำหนดค่าข้อมูลที่ไม่สมบูรณ์ (missing value) คำอธิบายความหมายของชื่อตัวแปร (variable label) คำอธิบายความหมายของค่าตัวแปร (value label)

ข้อกำหนดของแฟ้มข้อมูลที่เราต้องการเป็นดังนี้ กำหนดชื่อแฟ้มข้อมูล example4.sav

1. เลขประจำตัว	กำหนดชื่อตัวแปร	id
	กำหนดชนิดของข้อมูล	จำนวนเต็ม 3 หลัก
	ค่าที่กำหนดให้สำหรับข้อมูลที่ไม่สมบูรณ์	ไม่มี
	คำอธิบายความหมายของชื่อตัวแปร	ไม่มี
	คำอธิบายความหมายของค่าตัวแปร	ไม่มี
	กำหนดความกว้าง Columns การแสดงผลเป็น	5

หมายเหตุ เนื่องจาก id เป็นเลขจำนวนเต็ม 3 หลัก เราจึงใช้ความกว้าง Columns เป็น 5

2. เพศ	กำหนดชื่อตัวแปร	sex
	กำหนดชนิดของข้อมูล	จำนวนเต็ม 1 หลัก
	ค่าที่กำหนดให้สำหรับข้อมูลที่ไม่สมบูรณ์	9
	คำอธิบายความหมายของชื่อตัวแปร	ไม่มี
	คำอธิบายความหมายของค่าตัวแปร	1. Male 2. Female
	กำหนดความกว้าง Columns การแสดงผลเป็น	8

หมายเหตุ เนื่องจากคำอธิบายค่าตัวแปร Female กว้าง 6 Characters เราจึงใช้ความกว้าง Columns เป็น 8

3. อายุ	กำหนดชื่อตัวแปร	age
	กำหนดชนิดของข้อมูล	จำนวนเต็ม 2 หลัก
	ค่าที่กำหนดให้สำหรับข้อมูลที่ไม่สมบูรณ์	99
	คำอธิบายความหมายของชื่อตัวแปร	ไม่มี
	คำอธิบายความหมายของค่าตัวแปร	ไม่มี

หมายเหตุ เนื่องจาก age เป็นเลขจำนวนเต็ม 2 หลัก เราจึงใช้ความกว้าง Columns เป็น 4

4. ระดับการศึกษา	กำหนดชื่อตัวแปร	educ
	กำหนดชนิดของข้อมูล	จำนวนเต็ม 1 หลัก
	ค่าที่กำหนดให้สำหรับข้อมูลที่ไม่สมบูรณ์	9
	คำอธิบายความหมายของชื่อตัวแปร	Level of education
	คำอธิบายความหมายของค่าตัวแปร	1. Under graduate 2. Graduate 3. Post graduate 4. Doctorate

หมายเหตุ เนื่องจากคำอธิบายค่าตัวแปร Under graduate กว้าง 14 Characters เราจึงใช้ความกว้าง Columns เป็น 16

5. สถานะภาพ	กำหนดชื่อตัวแปร	status
	กำหนดชนิดของข้อมูล	จำนวนเต็ม 1 หลัก
	ค่าที่กำหนดให้สำหรับข้อมูลที่ไม่สมบูรณ์	9
	คำอธิบายความหมายของชื่อตัวแปร	ไม่มี
	คำอธิบายความหมายของค่าตัวแปร	1. Single 2. Married 3. Widowhood 4. Divorce

หมายเหตุ เนื่องจากคำอธิบายค่าตัวแปร Widowhood กว้าง 9 Characters เราจึงใช้ความกว้าง Columns เป็น 11

6. เงินเดือน	กำหนดชื่อตัวแปร	income
	กำหนดชนิดของข้อมูล	จำนวนเต็ม 4 หลัก
	ค่าที่กำหนดให้สำหรับข้อมูลที่ไม่สมบูรณ์	9999
	คำอธิบายความหมายของชื่อตัวแปร	ไม่มี
	คำอธิบายความหมายของค่าตัวแปร	ไม่มี

หมายเหตุ เนื่องจาก income เป็นเลขจำนวนเต็ม 4 หลัก เราจึงใช้ความกว้าง Columns เป็น 6

7. ระดับคะแนน	กำหนดชื่อตัวแปร	grade
	กำหนดชนิดของข้อมูล	จำนวนจริง xxxx.xx
	ค่าที่กำหนดให้สำหรับข้อมูลที่ไม่สมบูรณ์	9.99
	คำอธิบายความหมายของชื่อตัวแปร	ไม่มี
	คำอธิบายความหมายของค่าตัวแปร	ไม่มี

หมายเหตุ เนื่องจาก grade เป็นเลขจำนวนจริง xxxx.xx เราจึงใช้ความกว้าง Columns เป็น 8

8. เงินตอบแทนประจำปี	กำหนดชื่อตัวแปร	bonus
	กำหนดชนิดของข้อมูล	จำนวนจริง xxxxx.xx
	ค่าที่กำหนดให้สำหรับข้อมูลที่ไม่สมบูรณ์	9.99
	คำอธิบายความหมายของชื่อตัวแปร	ไม่มี
	คำอธิบายความหมายของค่าตัวแปร	ไม่มี

หมายเหตุ เนื่องจาก bonus เป็นเลขจำนวนจริง xxxxx.xx เราจึงใช้ความกว้าง Columns เป็น 10

คำแนะนำสำหรับเจ้าหน้าที่พิมพ์ข้อมูล

- | | | |
|----------------|-------------------------------|-----------------------------------|
| 1. เลขประจำตัว | | พิมพ์ข้อมูลตามค่าจริงจากแบบสอบถาม |
| 2. เพศ | <input type="checkbox"/> ชาย | พิมพ์ข้อมูลเป็นเลข 1 |
| | <input type="checkbox"/> หญิง | พิมพ์ข้อมูลเป็นเลข 2 |

หมายเหตุ ไม่ตอบ หรือ ข้อมูลไม่สมบูรณ์ให้พิมพ์ข้อมูลเป็นเลข 9

- | | | |
|---------|--|-----------------------------------|
| 3. อายุ | | พิมพ์ข้อมูลตามค่าจริงจากแบบสอบถาม |
|---------|--|-----------------------------------|
- หมายเหตุ ไม่ตอบ หรือ ข้อมูลไม่สมบูรณ์ให้พิมพ์ข้อมูลเป็นเลข 99

4. ระดับการศึกษา ต่ำกว่าระดับปริญญาตรี พิมพ์ข้อมูลเป็นเลข 1
 จบระดับปริญญาตรี พิมพ์ข้อมูลเป็นเลข 2
 จบระดับปริญญาโท พิมพ์ข้อมูลเป็นเลข 3
 จบระดับปริญญาเอก พิมพ์ข้อมูลเป็นเลข 4

หมายเหตุ ไม่ตอบ หรือ ข้อมูลไม่สมบูรณ์ให้พิมพ์ข้อมูลเป็นเลข 9

5. สถานะภาพ โสด พิมพ์ข้อมูลเป็นเลข 1
 แต่งงานแล้ว พิมพ์ข้อมูลเป็นเลข 2
 เป็นหม้าย พิมพ์ข้อมูลเป็นเลข 3
 หย่าร้าง พิมพ์ข้อมูลเป็นเลข 4

หมายเหตุ ไม่ตอบ หรือ ข้อมูลไม่สมบูรณ์ให้พิมพ์ข้อมูลเป็นเลข 9

6. เงินเดือน พิมพ์ข้อมูลตามค่าจริงจากแบบสอบถาม

หมายเหตุ ไม่ตอบ หรือ ข้อมูลไม่สมบูรณ์ให้พิมพ์ข้อมูลเป็นเลข 9999

7. ระดับคะแนน พิมพ์ข้อมูลตามค่าจริงจากแบบสอบถาม

หมายเหตุ ไม่ตอบ หรือ ข้อมูลไม่สมบูรณ์ให้พิมพ์ข้อมูลเป็นเลข 9.99

8. เงินตอบแทนประจำปี พิมพ์ข้อมูลตามค่าจริงจากแบบสอบถาม

หมายเหตุ ไม่ตอบ หรือ ข้อมูลไม่สมบูรณ์ให้พิมพ์ข้อมูลเป็นเลข 9.99

ตัวอย่างแบบสอบถามข้อมูลพนักงานที่กรอกแล้ว

สำหรับเจ้าหน้าที่กรอกข้อมูล

1. เลขประจำตัว 1

2. เพศ ชาย หญิง

3. อายุ 37 ปี

4. ระดับการศึกษา

- ต่ำกว่าระดับปริญญาตรี จบระดับปริญญาตรี
 จบระดับปริญญาโท จบระดับปริญญาเอก

5. สถานะภาพ

- โสด แต่งงานแล้ว
 เป็นหม้าย หย่าร้าง

6. เงินเดือน 5500 บาท

7. ระดับคะแนน 3.78

8. เงินตอบแทนประจำปี 11000.00 บาท

1

2

4

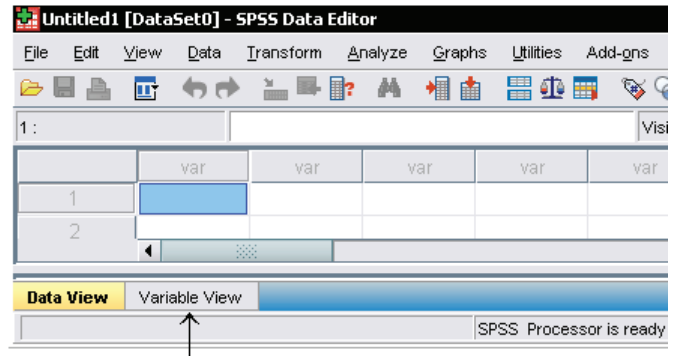
2.1 การสร้างแฟ้มข้อมูลใน SPSS Data Editor

เริ่มต้นการสร้างแฟ้มข้อมูลที่ SPSS Data Editor

คลิกที่ Variable View

SPSS Data Editor จะเปลี่ยนไปทำงาน

ในส่วนของการกำหนดตัวแปร



ความหมายของแต่ละ Column ของ Variable View

Name กำหนดชื่อตัวแปร

Type กำหนดชนิดของตัวแปรเช่น

ตัวเลข (Numeric) ตัวอักษร (String)

Width กำหนดความกว้างสำหรับเก็บค่าของตัวแปร

Decimals กำหนดตำแหน่งทศนิยมของข้อมูลตัวเลข

Label กำหนดคำอธิบายชื่อของตัวแปร

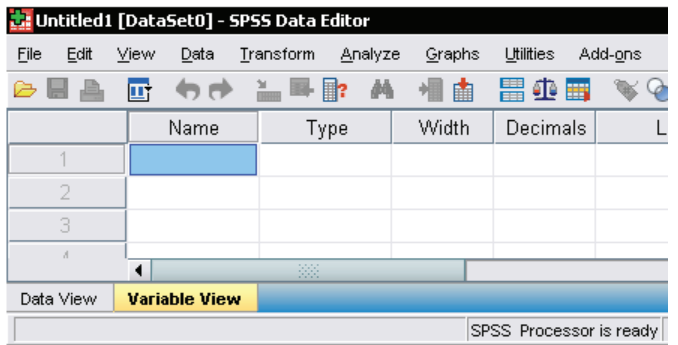
Values กำหนดความหมายให้กับค่าตัวเลข เช่น 1 หมายถึง ชาย 2 หมายถึง หญิง

Missing กำหนดค่าของข้อมูลไม่สมบูรณ์เช่น กรอกตัวเลขอายุผิด ไม่ตอบค่าที่ต้องการ

Columns กำหนดความกว้างของ Columns ใน Data View

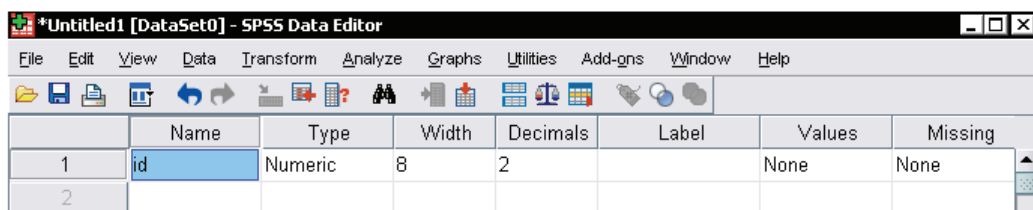
Align กำหนดการแสดงผลใน Column เป็น ชิดซ้าย ชิดขวา หรือ กึ่งกลาง

Measure กำหนดชนิดข้อมูลเป็น Scale, Ordinal, Nominal

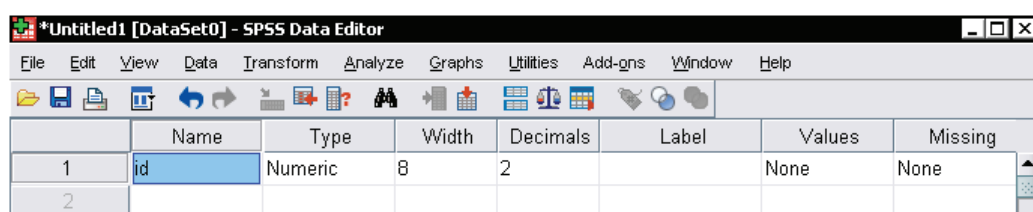


การกำหนดค่าต่าง ๆ ของตัวแปร id

ขั้นที่ 1. พิมพ์ชื่อตัวแปรแรกคือ id ในช่อง Name ของตัวแปรตัวที่ 1



เมื่อกด Enter โปรแกรมจะนำค่า Default ของ SPSS เกี่ยวกับตัวแปรมาเติมให้ ซึ่งจะได้ผลบนจอภาพเป็นดังนี้



จากข้อกำหนดของการสร้างแฟ้มของตัวแปร id

1. เลขประจำตัว กำหนดชื่อตัวแปร
กำหนดชนิดของข้อมูล
ค่าที่กำหนดให้สำหรับข้อมูลที่ไม่สมบูรณ์
คำอธิบายความหมายของชื่อตัวแปร
คำอธิบายความหมายของค่าตัวแปร

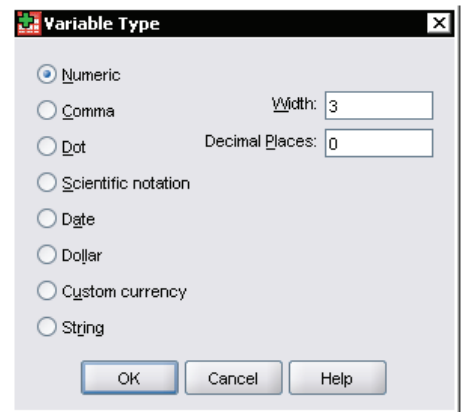
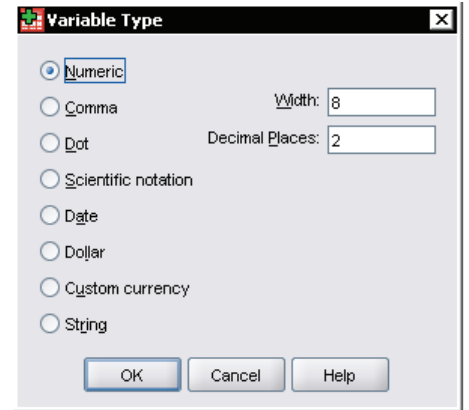
id
จำนวนเต็ม 3 หลัก
ไม่มี
ไม่มี
ไม่มี

การเปลี่ยนข้อกำหนดต่างๆ ของตัวแปร id

1. พิมพ์ค่าที่ต้องการในแต่ละ Column ของตัวแปร id

หรือ 2. คลิกที่ **Numeric** ... จะได้เมนูย่อยเป็น

- ในเมนูย่อยเราสามารถเลือกชนิดของตัวแปรเป็น Numeric, String, ...
- เปลี่ยน Width เป็น 3
- กำหนดการแสดงผลทศนิยมเป็น 0 ตำแหน่ง
- กำหนดความกว้าง columns เป็น 5
- จะได้ผลของเมนูย่อย Variable Type บนจอภาพเป็นดังนี้



เสร็จแล้วคลิก OK จอภาพ SPSS Data Editor ในส่วนของ Variable View จะมีผลดังนี้

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	id	Numeric	3	0		None	None	5	Right	Scale
2										

การกำหนดค่าต่างๆ ของตัวแปร sex

- ขั้นที่ 1. พิมพ์ชื่อตัวแปร sex ในช่อง Name ของตัวแปรตัวที่ 2

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	id	Numeric	3	0		None	None	5	Right	Scale
2	sex									

เมื่อกด Enter โปรแกรมจะนำค่า Default ของ SPSS เกี่ยวกับตัวแปรมาเติมให้ ซึ่งจะได้ผลบนจอภาพเป็นดังนี้

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	id	Numeric	3	0		None	None	5	Right	Scale
2	sex	Numeric	8	2		None	None	8	Right	Scale
3										

จากข้อกำหนดของการสร้างแฟ้มข้อมูลของตัวแปร sex

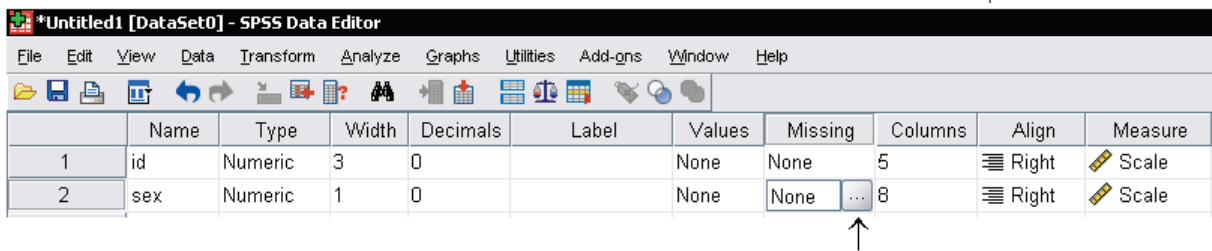
2. เพศ	กำหนดชื่อตัวแปร	sex
	กำหนดชนิดของข้อมูล	จำนวนเต็ม 1 หลัก
	ค่าที่กำหนดให้สำหรับข้อมูลที่ไม่สมบูรณ์	9
	คำอธิบายความหมายของชื่อตัวแปร	ไม่มี
	คำอธิบายความหมายของค่าตัวแปร	1. Male 2. Female

ให้เปลี่ยนค่า Decimals จาก 2 เป็น 0

ให้เปลี่ยนค่า Width จาก 8 เป็น 1

หมายเหตุ ในกรณีที่ Width มีค่าต่ำกว่า Decimals ต้องกำหนดค่า Decimals ก่อน

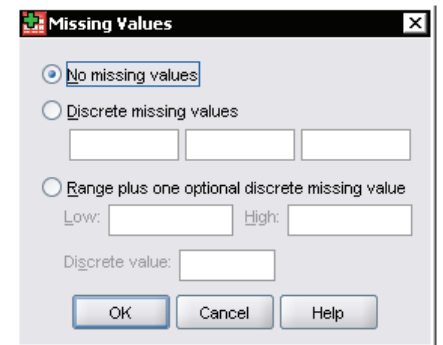
ขั้นที่ 2. การกำหนดค่า Missing Values ให้กับตัวแปร sex ให้คลิกเมาส์ที่ 



จะได้เมนูย่อยของการกำหนด Missing Values เป็นดังนี้

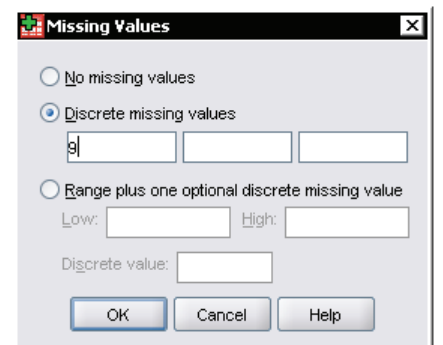
หมายเหตุ ความหมายของ Options ในเมนูย่อย

- No missing values ตัวแปรนี้ไม่มีค่า Missing values
- Discrete missing values ตัวแปรนี้มีค่า Missing เป็นตัว ๆ
- Range plus one optional discrete missing value
ตัวแปรนี้มีค่า Missing แบบผสมทั้งชนิดเป็นช่วง และเป็นค่าแบบ discrete อีก 1 ค่า



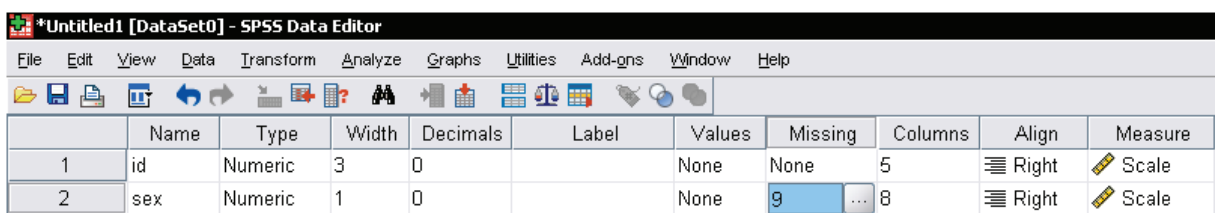
สำหรับตัวแปร sex เลือกชนิด Missing เป็น Discrete missing values

และพิมพ์ค่าในช่องเป็นเลข 9



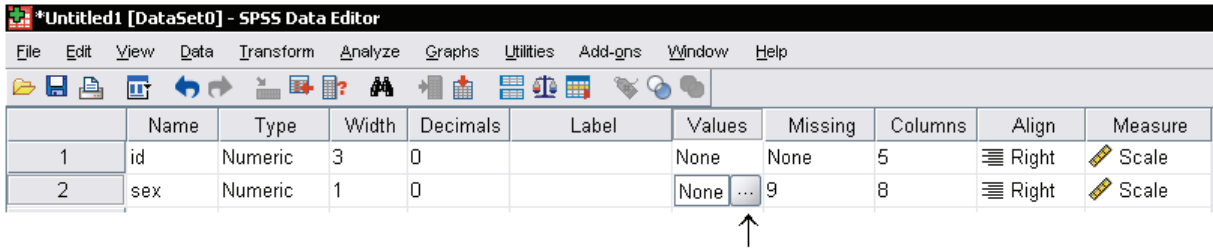
เสร็จแล้วคลิก OK

บนจอภาพจะกลับไป SPSS Data Editor ในส่วน Variable View



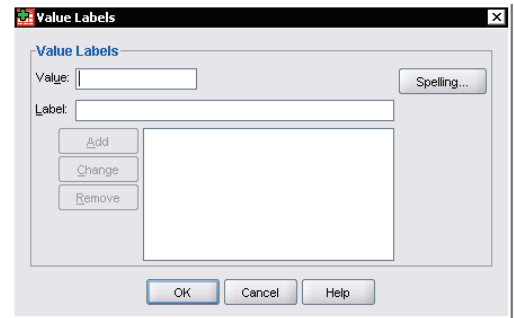
ขั้นที่ 3. การกำหนด Values และ Value Labels ให้กับตัวแปร sex

ให้คลิกที่ช่อง ของตัวแปร sex บนจอภาพที่ช่อง Values ของตัวแปร sex จะเปลี่ยนเป็น



ให้คลิกที่ตำแหน่งลูกศรชี้

จะได้เมนูย่อย Value Labels ของการกำหนดค่าเป็นดังนี้



ขั้นที่ 3.1 ไปที่ช่อง Value พิมพ์ค่า 1 เสร็จแล้วกด

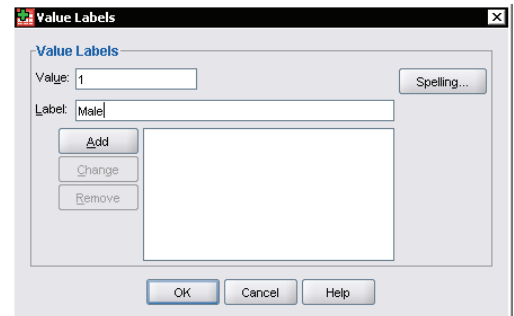
Tab เพื่อไปที่ช่อง Label

ขั้นที่ 3.2 พิมพ์ความหมายของค่า Label

เป็น Male

ขั้นที่ 3.3 คลิก Add จะได้ผลบนเมนูย่อยของ

Value Labels เป็นดังนี้



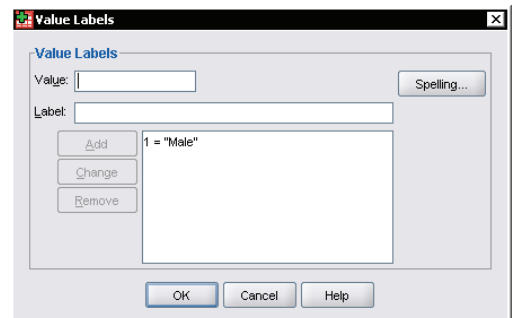
ในทำนองเดียวกัน

ขั้นที่ 3.4 ไปที่ช่อง Value พิมพ์ค่า 2 เสร็จแล้วกด

Tab เพื่อไปที่ช่อง Label

ขั้นที่ 3.5 พิมพ์ความหมายของค่าเป็น Female

เสร็จแล้วคลิกที่ Add

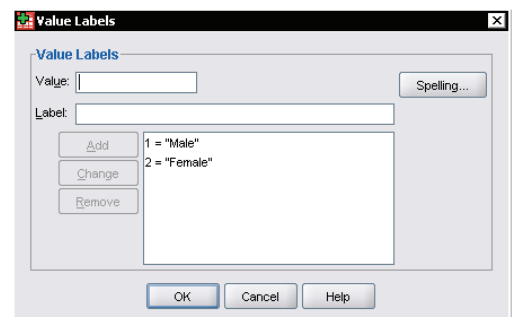


หมายเหตุ ในแถบเมนูย่อย

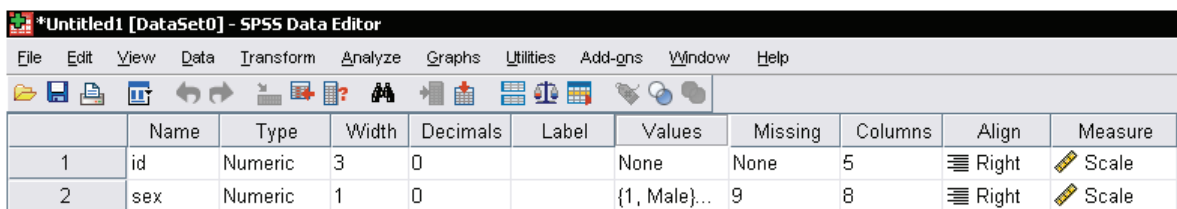
Change เป็นการเลือกที่จะเปลี่ยนค่า Value และ Label

Remove เป็นการยกเลิกค่า Value และ Label

จะได้ผลบนเมนูย่อยของ Value Labels เป็นดังนี้



เสร็จแล้วคลิก OK จอภาพจะกลับมาที่ SPSS Data Editor ในส่วนของ Variable View ดังนี้



ข้อสังเกต ในช่อง values ของตัวแปร sex มีคำอธิบายความหมายบางส่วนของ Value Label ปรากฏ
ขณะนี้เรากำหนดค่าต่าง ๆ เกี่ยวกับตัวแปร sex เสร็จแล้ว

ในทำนองเดียวกันการกำหนดค่าเกี่ยวกับตัวแปรอื่น ๆ สามารถทำได้ตามขั้นตอนโดยย่อดังนี้

การกำหนดค่าต่าง ๆ ของตัวแปร age

ขั้นที่ 1. พิมพ์ชื่อตัวแปร age ในช่อง Name ของตัวแปรตัวที่ 3

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	id	Numeric	3	0		None	None	5	Right	Scale
2	sex	Numeric	1	0		{1, Male}...	9	8	Right	Scale
3	age									

เมื่อกด Enter โปรแกรมจะนำค่า Default ของ SPSS เกี่ยวกับตัวแปรมาเติมให้ ซึ่งจะได้ผลบนจอภาพเป็นดังนี้

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	id	Numeric	3	0		None	None	5	Right	Scale
2	sex	Numeric	1	0		{1, Male}...	9	8	Right	Scale
3	age	Numeric	8	2		None	None	8	Right	Scale

ขั้นที่ 2. กำหนดชนิดของตัวแปร age เลือกชนิดเป็น Numeric

เปลี่ยน Decimals จากเดิม 2 เป็น 0

เปลี่ยน Width จากเดิม 8 เป็น 2

เปลี่ยนค่า Columns จากเดิม 8 เป็น 4

ขั้นที่ 3. กำหนด Missing Value โดยเลือกชนิดเป็น Discrete missing values

และพิมพ์ค่าในช่องเป็นเลข 99

จะได้ผลบนจอภาพเป็น

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	id	Numeric	3	0		None	None	5	Right	Scale
2	sex	Numeric	1	0		{1, Male}...	9	8	Right	Scale
3	age	Numeric	2	0		None	99	4	Right	Scale

การกำหนดค่าต่าง ๆ ของตัวแปร educ

ขั้นที่ 1. พิมพ์ชื่อตัวแปร educ ในช่อง Name ของตัวแปรตัวที่ 4

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	id	Numeric	3	0		None	None	5	Right	Scale
2	sex	Numeric	1	0		{1, Male}...	9	8	Right	Scale
3	age	Numeric	2	0		None	99	4	Right	Scale
4	educ									

เมื่อกด Enter โปรแกรมจะนำค่า Default ของ SPSS เกี่ยวกับตัวแปรมาเติมให้ ซึ่งจะได้ผลบนจอภาพเป็นดังนี้

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	id	Numeric	3	0		None	None	5	Right	Scale
2	sex	Numeric	1	0		{1, Male}...	9	8	Right	Scale
3	age	Numeric	2	0		None	99	4	Right	Scale
4	educ	Numeric	8	2		None	None	8	Right	Scale

ขั้นที่ 2. กำหนดชนิดของตัวแปร educ เลือกชนิดเป็น Numeric

เปลี่ยน Decimals จากเดิม 2 เป็น 0

เปลี่ยน Width จากเดิม 8 เป็น 1

เปลี่ยนค่า Columns จากเดิม 8 เป็น 16

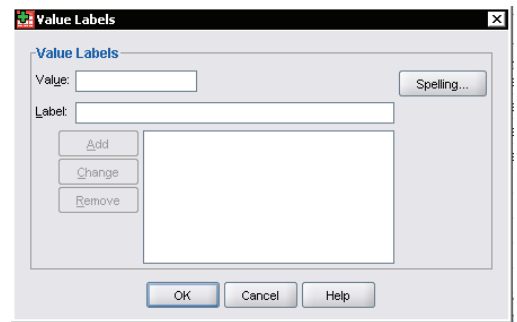
ขั้นที่ 3. กำหนด Missing Value โดยเลือกชนิดเป็น Discrete missing values และพิมพ์ค่าในช่องเป็นเลข 9

ขั้นที่ 4. ไปที่ช่อง Label พิมพ์ Level of education

ขั้นที่ 5. การกำหนด Value Labels ให้กับตัวแปร educ ให้คลิกที่ช่อง ของตัวแปร educ บนจอ

ภาพที่ช่อง Values ของตัวแปร educ จะเปลี่ยนเป็น ...

ให้คลิกที่ตำแหน่งลูกศรชี้ ... จะได้เมนูย่อยเป็นดังนี้



ขั้นที่ 5.1 ไปที่ช่อง Value พิมพ์ค่า 1

เสร็จแล้วกด Tab เพื่อไปที่ช่อง Label

พิมพ์ความหมายของค่าเป็น Under graduate

เสร็จแล้วคลิกที่ Add

ขั้นที่ 5.2 ไปที่ช่อง Value พิมพ์ค่า 2 เสร็จแล้วกด Tab เพื่อไปที่ช่อง Label

พิมพ์ความหมายของค่าเป็น Graduate เสร็จแล้วคลิกที่ Add

ขั้นที่ 5.3 ไปที่ช่อง Value พิมพ์ค่า 3 เสร็จแล้วกด Tab เพื่อไปที่ช่อง Label

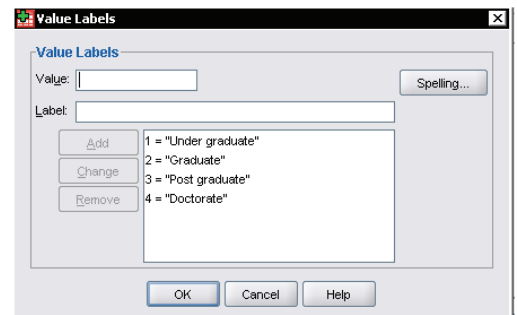
พิมพ์ความหมายของค่าเป็น Post graduate เสร็จแล้วคลิกที่ Add

ขั้นที่ 5.4 ไปที่ช่อง Value พิมพ์ค่า 4

เสร็จแล้วกด Tab เพื่อไปที่ช่อง Label

พิมพ์ความหมายของค่าเป็น Doctorate

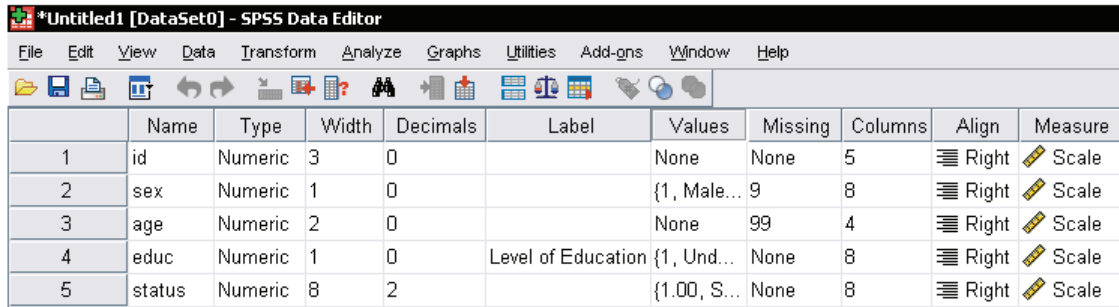
เสร็จแล้วคลิกที่ Add



ขั้นที่ 6. คลิกที่ OK จะได้ผลบนจอภาพเป็นดังนี้

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	id	Numeric	3	0		None	None	5	Right	Scale
2	sex	Numeric	1	0		{1, Male}...	9	8	Right	Scale
3	age	Numeric	2	0		None	99	4	Right	Scale
4	educ	Numeric	1	0	Level of Education	{1, Und...	None	8	Right	Scale

ขั้นที่ 6. คลิกที่ OK จะได้ผลบนจอภาพเป็นดังนี้



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	id	Numeric	3	0		None	None	5	Right	Scale
2	sex	Numeric	1	0		{1, Male...	9	8	Right	Scale
3	age	Numeric	2	0		None	99	4	Right	Scale
4	educ	Numeric	1	0	Level of Education	{1, Und...	None	8	Right	Scale
5	status	Numeric	8	2		{1.00, S...	None	8	Right	Scale

การกำหนดค่าต่าง ๆ ของตัวแปร income

พิมพ์ชื่อตัวแปร income ในช่อง Name ของตัวแปรตัวที่ 6

กำหนดชนิดของตัวแปรเป็น Numeric

เปลี่ยน Decimals จากของเดิม 2 เป็น 0

เปลี่ยน Width จากของเดิม 8 เป็น 4

กำหนด Missing Value เป็นเลข 9999

กำหนด Columns จากของเดิม 8 เป็น 6

การกำหนดค่าต่าง ๆ ของตัวแปร grade

พิมพ์ชื่อตัวแปร grade ในช่อง Name ของตัวแปรตัวที่ 7

กำหนดชนิดของตัวแปรเป็น Numeric

เปลี่ยน Width จากของเดิม 8 เป็น 6

กำหนด Missing Value เป็นเลข 9.99

การกำหนดค่าต่าง ๆ ของตัวแปร bonus

พิมพ์ชื่อตัวแปร bonus ในช่อง Name ของตัวแปรตัวที่ 8

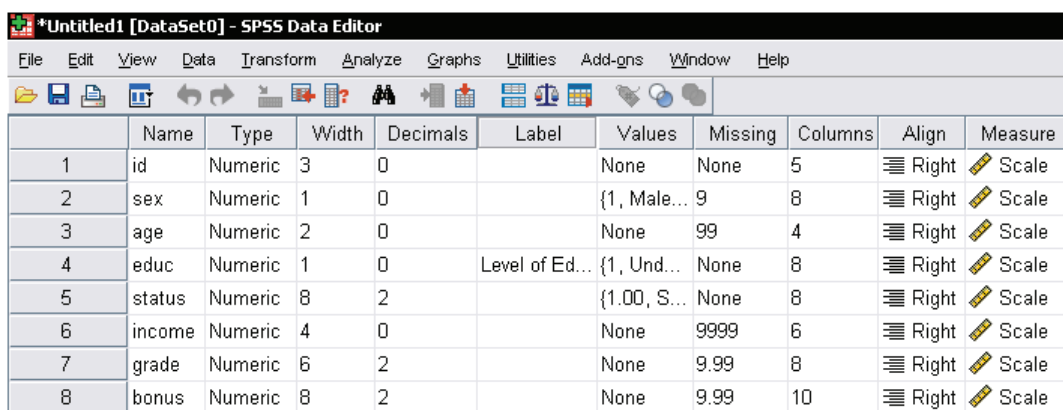
กำหนดชนิดของตัวแปรเป็น Numeric

เปลี่ยน Width จากของเดิม 8 เป็น 9

กำหนด Missing Value เป็นเลข 9.99

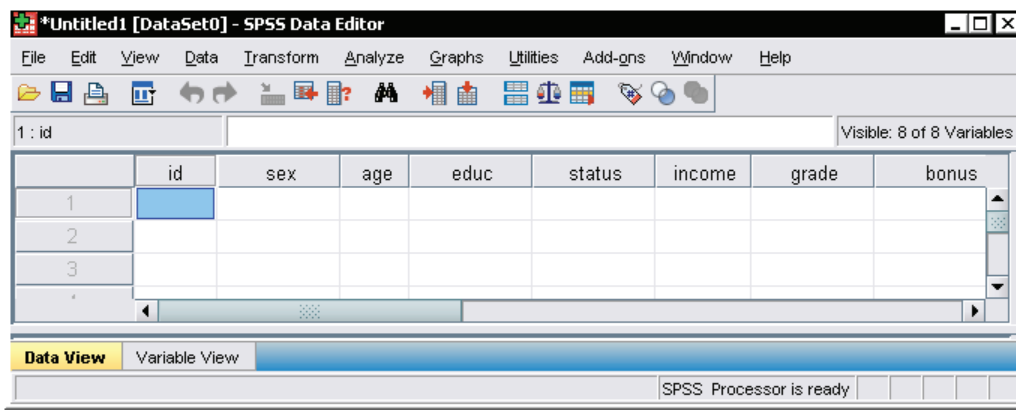
กำหนด Columns จากของเดิม 8 เป็น 10

ผลบนจอภาพเมื่อกำหนดค่าต่าง ๆ เสร็จแล้วคือ



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	id	Numeric	3	0		None	None	5	Right	Scale
2	sex	Numeric	1	0		{1, Male...	9	8	Right	Scale
3	age	Numeric	2	0		None	99	4	Right	Scale
4	educ	Numeric	1	0	Level of Ed...	{1, Und...	None	8	Right	Scale
5	status	Numeric	8	2		{1.00, S...	None	8	Right	Scale
6	income	Numeric	4	0		None	9999	6	Right	Scale
7	grade	Numeric	6	2		None	9.99	8	Right	Scale
8	bonus	Numeric	8	2		None	9.99	10	Right	Scale

คลิกที่ Data View จะกลับไปส่วนที่ทำงานเกี่ยวกับการใส่ข้อมูล ผลบนจอภาพจะเป็นดังนี้



จากข้อมูลที่เก็บมาได้ของพนักงาน 50 คน ได้ข้อมูลดังนี้

id	sex	age	educ	status	income	grade	bonus
1	1	37	2	4	5500	3.78	11000.00
2	2	29	3	1	4100	3.89	12300.00
3	2	48	1	2	5400	3.67	21600.00
4	1	99	1	2	9999	2.78	19998.00
5	2	33	2	9	9999	3.00	29997.00
6	2	45	3	4	8300	3.45	16600.00
7	2	38	1	4	7700	3.89	7700.00
8	2	23	3	1	3900	3.67	11700.00
9	1	34	2	4	4500	2.56	9000.00
10	1	50	2	2	6700	2.69	6700.00
11	2	43	2	2	4700	3.56	18800.00
12	2	37	3	2	3900	3.00	3900.00
13	1	24	2	1	3300	2.45	9900.00
14	1	46	2	2	4900	2.45	14700.00
15	1	32	1	1	4000	3.87	8000.00
16	1	42	2	3	6600	3.67	13200.00
17	1	38	4	2	8000	3.23	32000.00
18	2	41	2	3	7000	3.45	21000.00
19	2	99	1	9	2000	3.21	2000.00
20	1	54	2	2	7400	3.00	22200.00
21	2	32	3	9	6200	2.56	24800.00
22	1	43	1	2	4700	2.45	18800.00
23	2	22	1	1	3400	3.78	3400.00
24	1	40	2	2	5900	2.67	17700.00
25	1	37	4	9	7500	3.45	22500.00
26	1	28	1	1	3100	2.78	9300.00
27	1	44	3	2	6800	2.56	13600.00
28	1	56	2	2	6400	2.78	19200.00
29	1	35	3	1	5800	3.33	5800.00
30	2	42	1	2	3900	2.56	11700.00
31	1	21	2	1	4700	2.67	14100.00
32	1	39	2	2	5900	2.89	17700.00
33	1	45	1	2	4900	2.56	4900.00
34	1	31	1	2	3100	3.23	9300.00

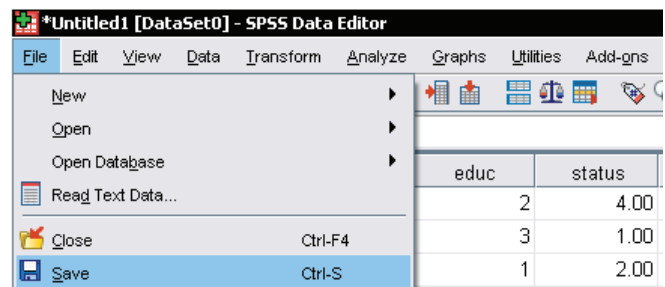
35	1	51	2	3	5400	3.01	5400.00
36	1	23	3	1	6300	2.77	12600.00
37	1	40	3	2	7100	2.89	21300.00
38	1	47	2	3	6600	2.77	19800.00
36	1	53	2	2	7200	2.31	21600.00
40	2	27	2	1	1700	2.67	5100.00
41	1	29	4	1	5000	2.89	15000.00
42	1	40	3	2	6000	3.67	18000.00
43	2	30	1	1	3000	2.56	12000.00
44	2	53	2	2	4700	3.00	9400.00
45	1	31	1	1	2800	2.74	5600.00
46	1	45	2	2	5700	2.67	22800.00
47	1	22	2	4	4300	3.07	4300.00
48	2	34	1	1	3900	2.56	7800.00
49	2	33	3	2	6700	2.12	20100.00
50	1	54	2	2	4800	2.66	19200.00

เมื่อพิมพ์ข้อมูลเสร็จแล้วจะได้ผลเป็น

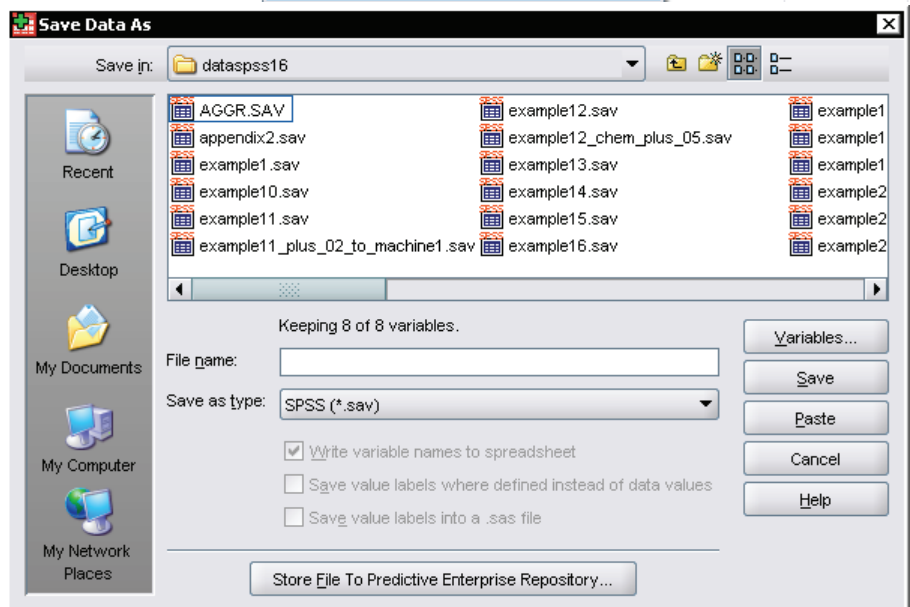
	id	sex	age	educ	status	income	grade	bonus
1	1	1	37	2	4.00	5500	3.78	11000.00
2	2	2	29	3	1.00	4100	3.89	12300.00
3	3	2	48	1	2.00	5400	3.67	21600.00
4	4	1	99	1	2.00	9999	2.78	19998.00

2.2 การบันทึกแฟ้มข้อมูล

ขั้นที่ 1. คลิกคำสั่ง File / Save



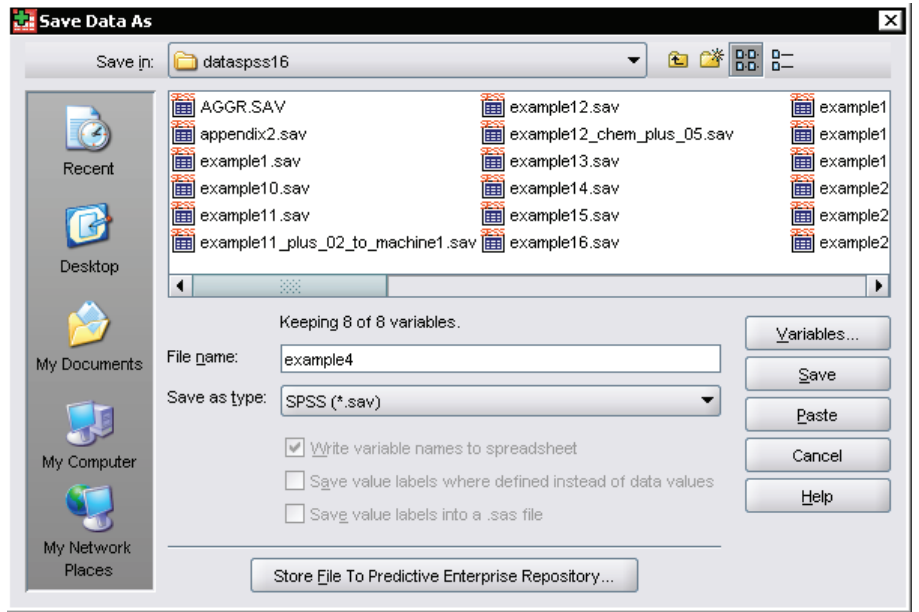
จะได้เมนูย่อย Save Data As



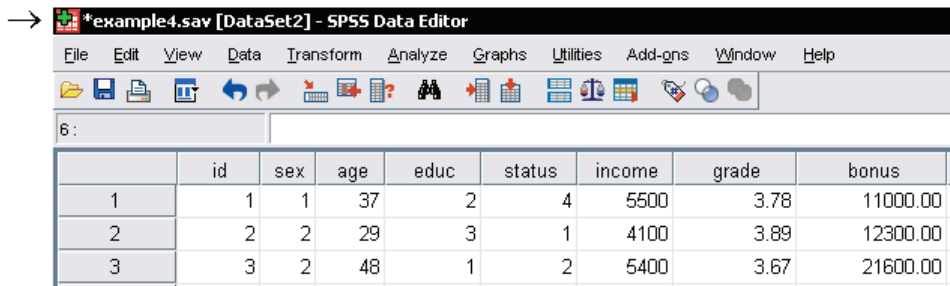
ขั้นที่ 2. บันทึกเป็นแฟ้มข้อมูล

ในช่อง File name

โดยพิมพ์ชื่อ example4



ขั้นที่ 3. คลิก Save จะเห็นได้ว่า Untitled1 เปลี่ยนเป็น example4.sav แล้ว

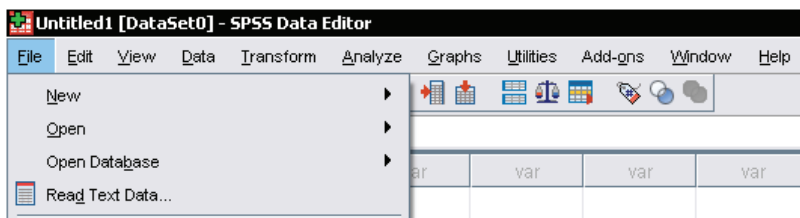


ขณะนี้ถือว่าการสร้างแฟ้มข้อมูลเสร็จและบันทึกไว้ที่แฟ้มข้อมูลชื่อ example4.sav เรียบร้อยแล้ว

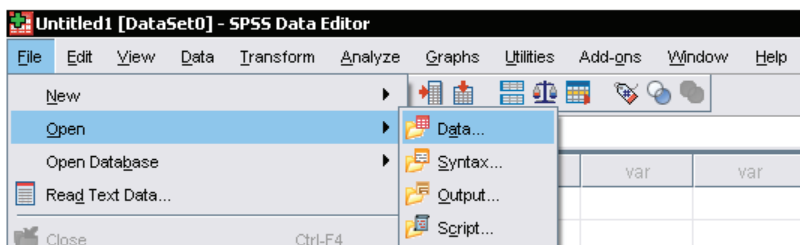
2.3 การเปิดแฟ้มข้อมูล

ขณะนี้ขอสมมติว่าได้สร้างแฟ้มข้อมูลชื่อ example4.sav บันทึกไว้แล้วประกอบด้วยตัวแปร 8 ตัว และมีค่าสังเกต 50 ค่า (หมายเหตุ หากไม่ต้องการพิมพ์ข้อมูลของตัวเอง ท่านสามารถ Download ได้จาก Web <http://www.math.sc.chula.ac.th/~tdumrong/2301286data>)

ขั้นที่ 1. เข้าสู่ SPSS Data Editor และ เลือกคำสั่ง File

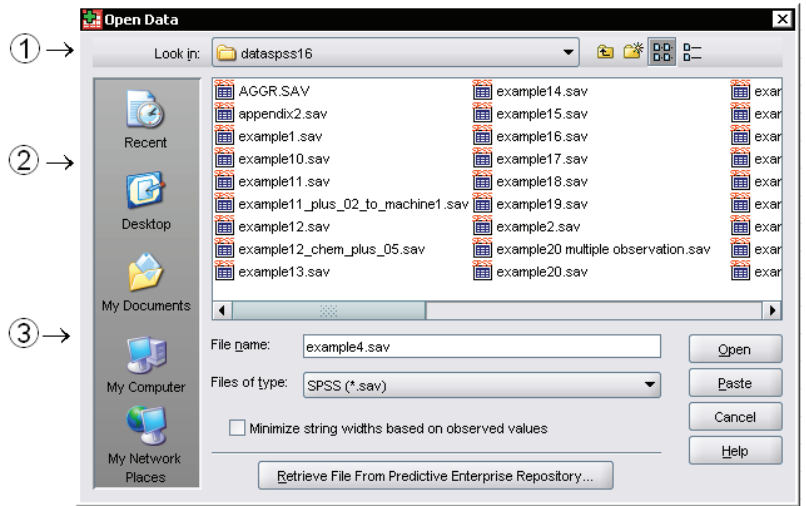


ขั้นที่ 2. เลือกคำสั่ง Open จะมีเมนูย่อยว่าจะเปิดแฟ้มแบบใด โดยเลือกจาก Data, Syntax, Output, ...



ขั้นที่ 3. คลิกที่ Data จะได้เมนูย่อยเป็น

1. เลือก Directory ที่มีแฟ้มข้อมูลอยู่
 2. ดูรายชื่อแฟ้มข้อมูลแล้วเลือกแฟ้มที่ต้องการโดยการคลิกที่ชื่อแฟ้มข้อมูล หรือ
 3. พิมพ์ชื่อแฟ้มข้อมูลที่ต้องการ เช่น example4.sav
- เสร็จแล้วคลิกที่ปุ่ม Open
จะได้ข้อมูลบนจอภาพดังนี้



*example4.sav [DataSet1] - SPSS Data Editor

	id	sex	age	educ	status	income	grade	bonus
1	1	1	37	2	4	5500	3.78	11000.00
2	2	2	29	3	1	4100	3.89	12300.00
3	3	2	48	1	2	5400	3.67	21600.00

ขณะนี้เราเปิดแฟ้มข้อมูล example4.sav เข้าสู่การทำงานของ SPSS Data Editor เรียบร้อยแล้ว

2.4 การดูรายละเอียดของตัวแปร

เราสามารถตรวจสอบรายละเอียดต่างๆ เกี่ยวกับแฟ้มข้อมูลได้ ตามขั้นตอนดังนี้

แบบที่ 1. ดูรายละเอียดและข้อกำหนดต่างของตัวแปรผ่านทาง Variable View

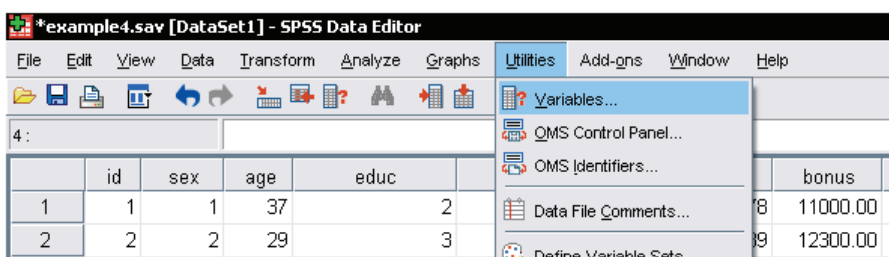
*example4.sav [DataSet1] - SPSS Data Editor

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	id	Numeric	3	0		None	None	3	Right	Scale
2	sex	Numeric	1	0		{1, Male}...	9	5	Right	Ordinal
3	age	Numeric	2	0		None	99	4	Right	Scale
4	educ	Numeric	1	0	Level of education	{1, Under gr...	9	11	Right	Ordinal
5	status	Numeric	1	0		{1, Single}...	9	8	Right	Ordinal
6	income	Numeric	4	0		None	9999	6	Right	Scale
7	grade	Numeric	6	2		None	9.99	6	Right	Ordinal
8	bonus	Numeric	8	2		None	9.99	7	Right	Scale

การดูรายละเอียดของตัวแปรผ่านทาง Variable View เราสามารถแก้ไขรายละเอียดต่างๆ ของตัวแปรได้

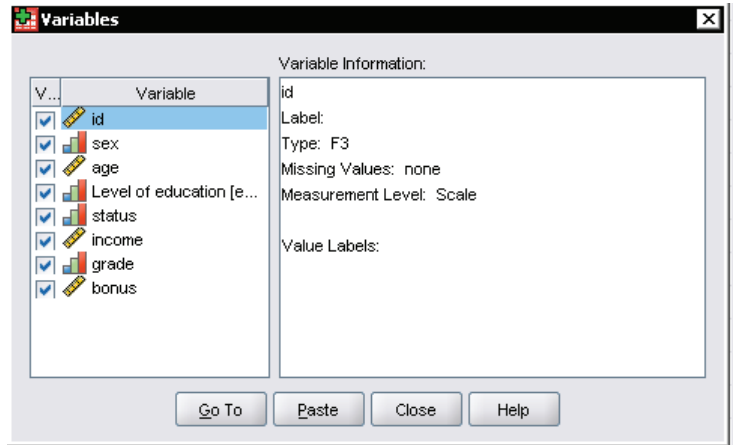
แบบที่ 2. ดูรายละเอียดและข้อกำหนดต่างของตัวแปรด้วยคำสั่ง Utilities

ขั้นที่ 1. คลิกคำสั่ง Utilities



ขั้นที่ 2. คลิกคำสั่ง Variables

จะได้เมนูย่อย Variables



ต้องการดูรายละเอียดของตัวแปรใด

ให้คลิกที่ชื่อของตัวแปรที่ต้องการ

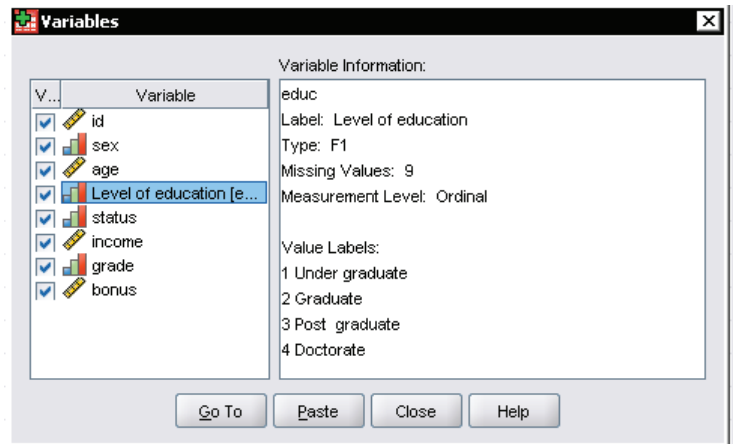
เช่นตัวแปร educ

หมายเหตุ

เพราะว่า ตัวแปร educ มีค่าของ Variable Label

จะแสดงค่าของ Variable Label เป็น

Level of education

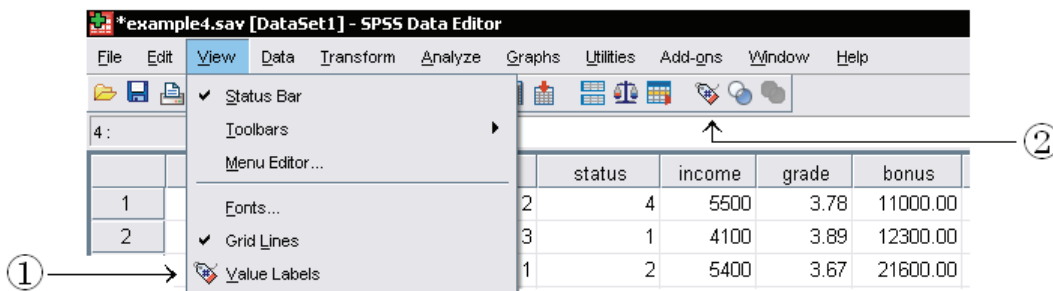


2.5 การสั่งให้ SPSS Data Editor แสดง Value Labels

ข้อมูลที่กำหนด Value Labels ไว้แล้วหากต้องการให้แสดงผลในลักษณะของ Value Labels ต้องทำดังนี้ จากจอภาพของ SPSS Data Editor

	id	sex	age	educ	status	income	grade	bonus
1	1	1	37	2	4	5500	3.78	11000.00
2	2	2	29	3	1	4100	3.89	12300.00

ขั้นที่ 1. คลิกที่ View จะได้แถบเมนูย่อยเป็น

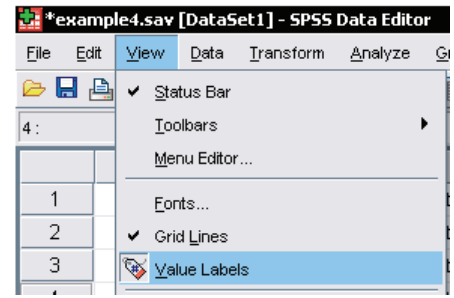


หมายเหตุ มีเครื่องหมายถูกหน้าคำสั่ง Grid Lines แปลว่าให้แสดงเส้นตารางใน SPSS Data Editor

คำสั่ง Fonts ใช้ในการเปลี่ยน Fonts ของตัวอักษรใน SPSS Data Editor

การเลือกให้แสดง Value label กดที่ปุ่ม Value Labels หมายเลข 1 หรือ 2 ก็ได้

ขั้นที่ 2. คลิกที่คำสั่ง Value Labels หน้า Value Labels
 การแสดงผลของตัวแปรจะแสดงค่า Value Labels ตามที่กำหนดไว้

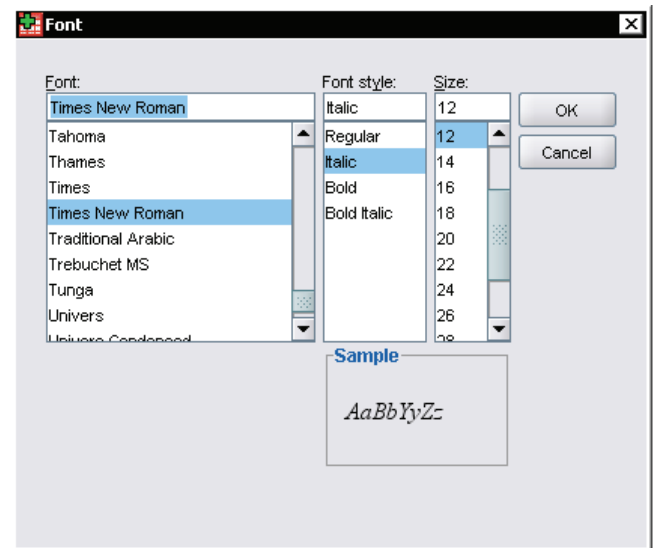
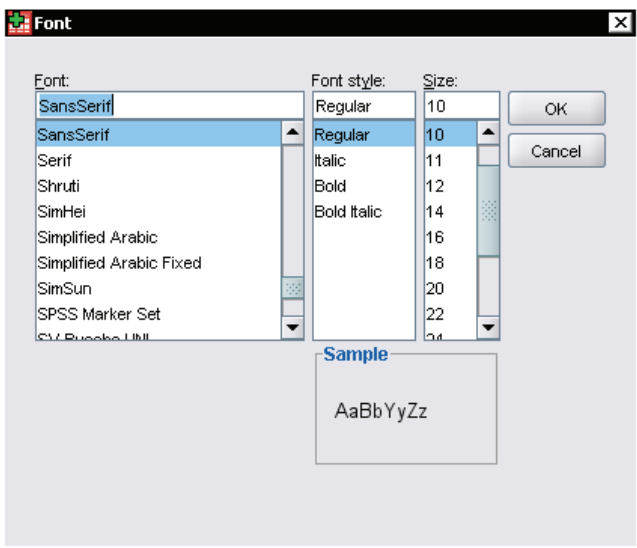


การแสดงผลบนจอภาพ SPSS Data Editor
 จะแสดงค่าของ Value Labels ออกมาแทนค่าตัวเลข ดังนี้

	id	sex	age	educ	status	income	grade	bonus
1	1	Male	37	Graduate	Divorce	5500	3.78	11000.00
2	2	Female	29	Post graduate	Single	4100	3.89	12300.00
3	3	Female	48	Under graduate	Married	5400	3.67	21600.00

จะเห็นว่าการแสดงผลจะเปลี่ยนไปเช่นตัวแปร sex 1 จะแสดงค่าเป็น Male 2 จะแสดงค่าเป็น Female ตามที่กำหนดไว้ตอนที่สร้างแฟ้มข้อมูล

หมายเหตุ การเปลี่ยน Fonts ของข้อมูลใน SPSS Data Editor ให้เลือกคำสั่ง View / Fonts จะได้เมนูย่อย



ตัวอย่างเช่น เลือก Font : Times New Roman, Font style : Italic, Size : 12 เสร็จแล้วคลิก OK
 ผลลัพธ์บนจอภาพจะเป็นดังนี้

	<i>id</i>	<i>sex</i>	<i>age</i>	<i>educ</i>	<i>status</i>	<i>income</i>	<i>grade</i>	<i>bonus</i>
1	1	Male	37	Graduate	Divorce	5500	3.78	11000.00
2	2	Female	29	Post graduate	Single	4100	3.89	12300.00
3	3	Female	48	Under graduate	Married	5400	3.67	21600.00

บทที่ 3

การวิเคราะห์ข้อมูลด้วยคำสั่ง Analyze / Descriptive Statistics



การวิเคราะห์ข้อมูลทางสถิติที่สำคัญคือ การคำนวณค่าสถิติเบื้องต้น การแจกแจงความถี่ของข้อมูลแบบ 1 ทาง การแจกแจงความถี่ของข้อมูลแบบ 2 ทาง การตรวจสอบความถูกต้องของข้อมูล คำสั่งสำคัญในการคำนวณค่าสถิติเบื้องต้นของ SPSS คือ คำสั่ง Analyze / Descriptive Statistics ซึ่งมีคำสั่งย่อยต่างๆ ในการทำงานเช่น

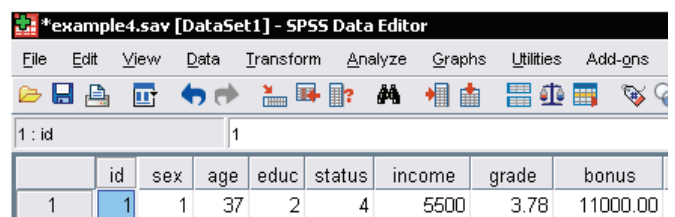
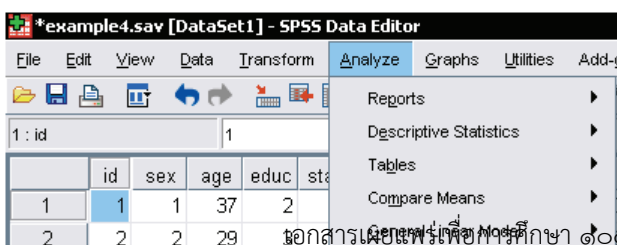
- Analyze / Descriptive Statistics / Frequencies แจกแจงความถี่ คำนวณค่าสถิติเบื้องต้น
- Analyze / Descriptive Statistics / Descriptives คำนวณค่าสถิติเบื้องต้น
- Analyze / Descriptive Statistics / Explore คำนวณค่าสถิติเบื้องต้น
- Analyze / Descriptive Statistics / Crosstabs แจกแจงความถี่ คำนวณค่าสถิติเบื้องต้น
- Analyze / Descriptive Statistics / Ratio คำนวณค่าอัตราส่วน

3.1 การคำนวณค่าสถิติเบื้องต้นด้วยคำสั่ง Analyze / Descriptive Statistics / Descriptives

คำสั่ง Analyze / Descriptive Statistics / Descriptives เป็นคำสั่งใช้ในการหาค่าสถิติเบื้องต้นเช่น ค่าเฉลี่ยเลขคณิต ค่าฐานนิยม ค่ามัธยฐาน ค่าส่วนเบี่ยงเบนมาตรฐาน ค่าสูงสุด ค่าต่ำสุด การใช้คำสั่งนี้ได้ต้องมีข้อมูลใน SPSS Data Editor

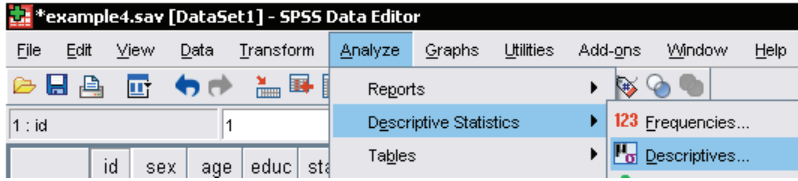
ขั้นที่ 1. เปิดแฟ้มข้อมูล example4.sav

เข้ามาใน SPSS Data Editor



ขั้นที่ 2. คลิกเมาส์ที่ Analyze บนเมนูบาร์
จอภาพจะเป็นดังนี้

ขั้นที่ 3. เลื่อนเมาส์ไปที่ Descriptive Statistics จอภาพจะเปลี่ยนแปลงเป็นเป็นดังนี้



ขั้นที่ 4. คลิก Descriptives จะได้เมนูย่อย

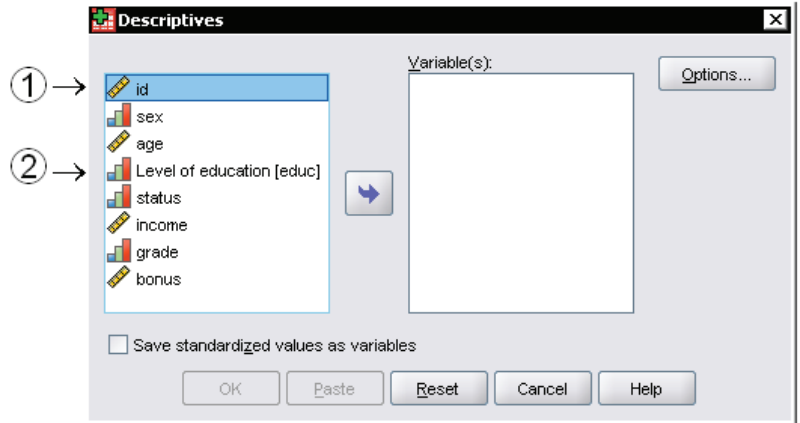
หมายเหตุ 1. เครื่องหมายแสดงชนิด

ตัวแปรว่าเป็นข้อมูลตัวเลข

2. ตัวแปรที่มี Value Labels

จะแสดงค่า Value Label

เช่นตัวแปร educ



ขั้นที่ 5. การเลือกตัวแปร age

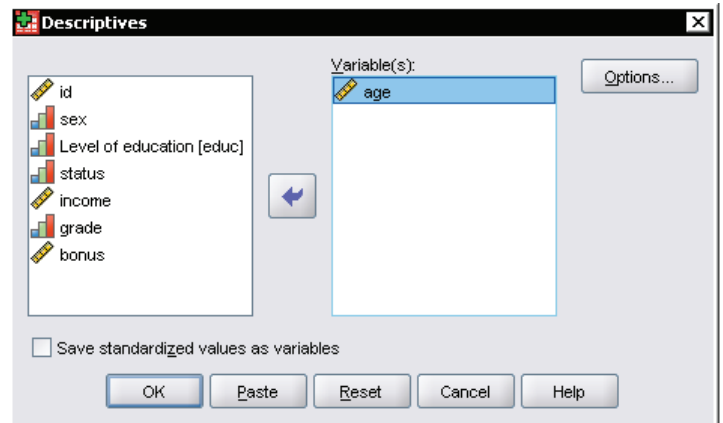
เพื่อทำการคำนวณ

ให้นำเมาส์ไปคลิกที่ตัวแปร age

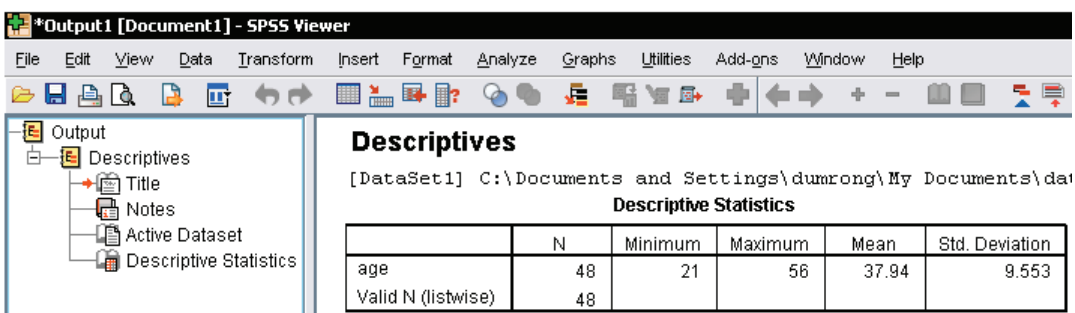
แล้วคลิกที่ปุ่ม  เพื่อย้าย

ตัวแปร age ไปทางขวา

บนจอภาพจะกลายเป็น



ขั้นที่ 6. เสร็จแล้วให้คลิกปุ่ม OK จะได้ผลการคำนวณที่ SPSS Viewer ดังนี้



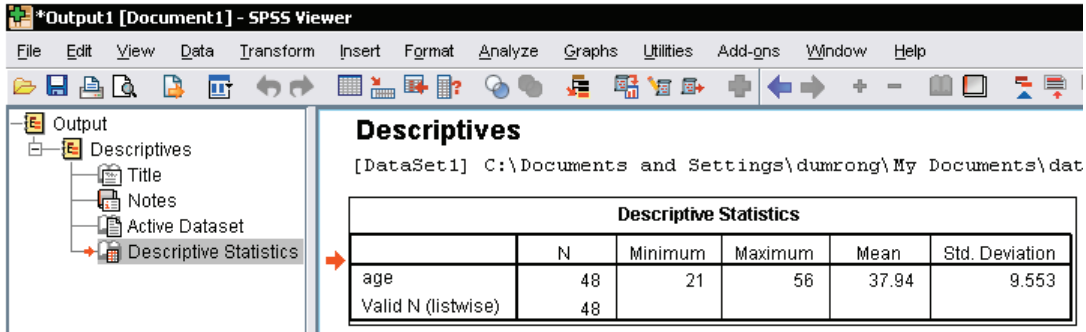
3.2 การเปลี่ยนรูปแบบของตารางในการแสดงผลของ SPSS Viewer

การแสดงผลของการคำนวณหากตารางแนวนอนมีความยาวมากจะทำให้เราไม่มีความสะดวกที่จะเห็นผลการคำนวณทั้งหมดในหน้าจอ

ดังนั้นเราควรจะทำ Transpose ให้ตารางแสดงผลในแนวตั้ง

ขั้นที่ 1. จากจอภาพใน SPSS Viewer ให้เลือกตารางที่ต้องการ

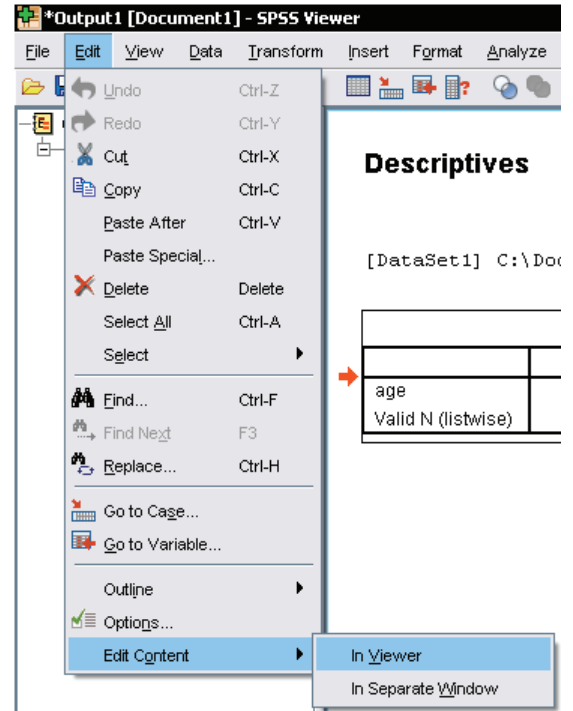
ในที่นี้ขอให้เลือกตารางโดยการคลิกที่ชื่อของตาราง จะเห็นว่าบนจอภาพจะมีลูกศรสีแดงขึ้นที่ขอบของตาราง



ขั้นที่ 2. คลิกที่เมนู Edit และเลื่อนเมาส์ไปเลือก

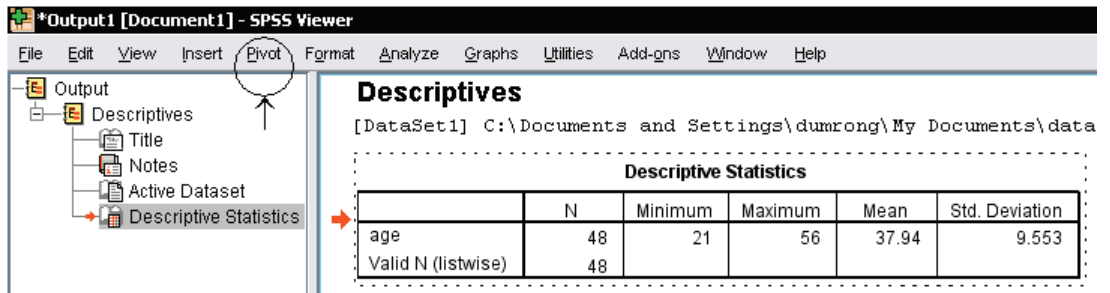
In View หากต้องการแก้ไขใน SPSS Viewer
In Separate Window หากต้องการเปิด SPSS Pivot Table เพื่อการแก้ไข

หมายเหตุ ใน version 12 ให้เลือก SPSS Pivot Table Object / Edit หรือ Open



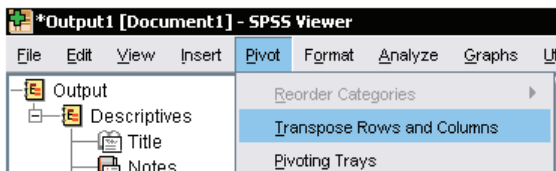
ขั้นที่ 3. ให้คลิกที่ Edit / In Viewer จอภาพจะเปลี่ยนแปลงโดยที่ขอบเส้นทึบ จะกลายเป็นเส้นไขว้ปลา และ เมนูบาร์จะเป็นเมนูของการแก้ไขตาราง

หมายเหตุ ขณะนี้เมนูบาร์มีคำสั่ง Pivot แล้ว



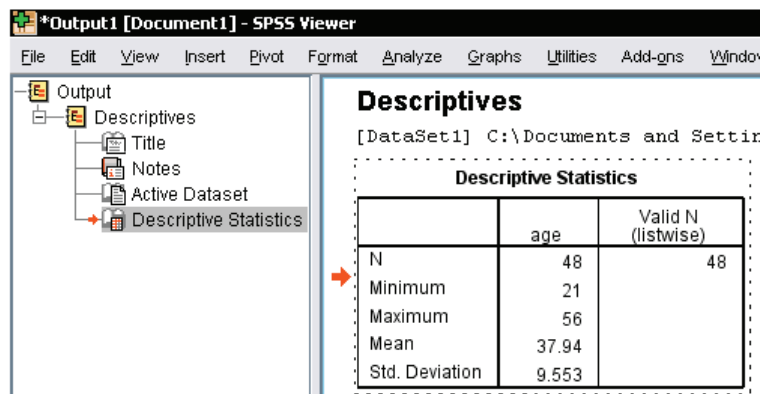
หมายเหตุ การเลื่อน pointer เข้ามาในตารางและกดดับเบิลคลิกจะให้ผลเหมือนกับการทำตาม ขั้นที่ 1 - 3

ขั้นที่ 4. ให้คลิกที่ Pivot จะได้เมนูย่อยของการเปลี่ยนรูปแบบตาราง



ขั้นที่ 5. ให้คลิกที่ Transpose Rows and Columns จะได้ตารางในรูปแบบแนวตั้ง

ขั้นที่ 6. การออกจากการทำงานส่วนนี้ให้เลื่อน Pointer นอกตาราง และคลิกเมาส์นอกบริเวณของตาราง



3.3 การกำหนดตำแหน่งทศนิยมของการคำนวณในตารางของ SPSS Viewer

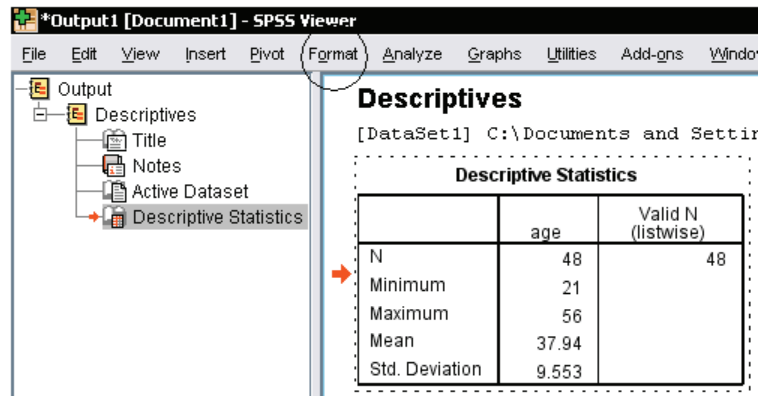
ตารางผลการคำนวณที่ได้เราสามารถกำหนดการแสดงผลว่าต้องการให้แสดงผลเป็นทศนิยม k ตำแหน่งได้ตามความต้องการ สมมติว่าเราต้องการให้แสดงผลการคำนวณของ Mean ให้เป็นทศนิยม 6 ตำแหน่ง มีขั้นตอนการทำงานดังนี้ จากการคำนวณล่าสุดบนจอภาพ

ขั้นที่ 1. จากจอภาพใน SPSS Viewer ให้เลือกตารางที่ต้องการด้วยการเลื่อนเมาส์เข้าไปในบริเวณของตาราง แล้วกดดับเบิลคลิกจะได้ผลเป็นดังนี้ (มีกรอบเส้นไขปลาล้อมรอบตาราง)

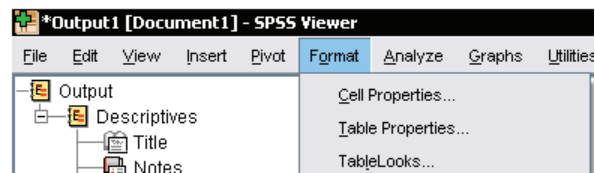
หมายเหตุ แต่ละช่องในตารางแสดงผล มีลักษณะเป็นเซลล์ที่สามารถปรับเปลี่ยนรูปแบบการแสดงผลได้

ขั้นที่ 2. เลื่อนเมาส์ไปคลิกที่ 37.94 จะมีการเปลี่ยนแปลงดังนี้

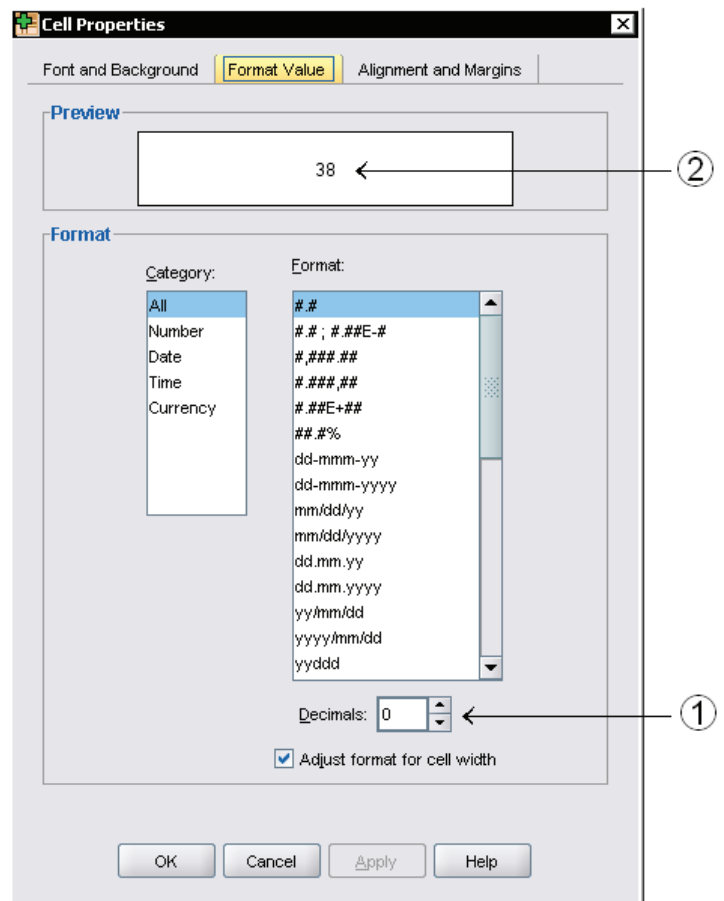
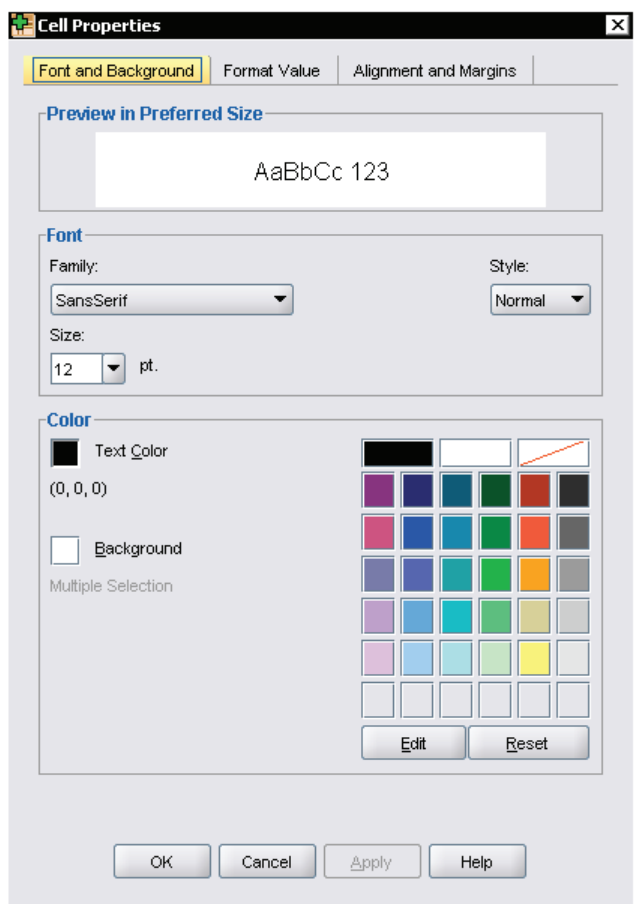
Descriptive Statistics		
	age	Valid N (listwise)
N	48	48
Minimum	21	
Maximum	56	
Mean	37.94	
Std. Deviation	9.553	



ขั้นที่ 3. คลิกคำสั่ง Format ที่เมนูบาร์ของ SPSS Viewer จะได้เมนูย่อยเป็นดังนี้



ขั้นที่ 4. คลิกที่ Cell Properties จะได้เมนูย่อยของ Cell Properties ต่อไปให้คลิกที่ Format Value



ขั้นที่ 5. ขอให้สังเกตที่หมายเลขในรูป

1. ของเดิมเป็นทศนิยม 0 ตำแหน่ง
2. ตัวอย่างของรูปแบบการแสดงผล คือ 38

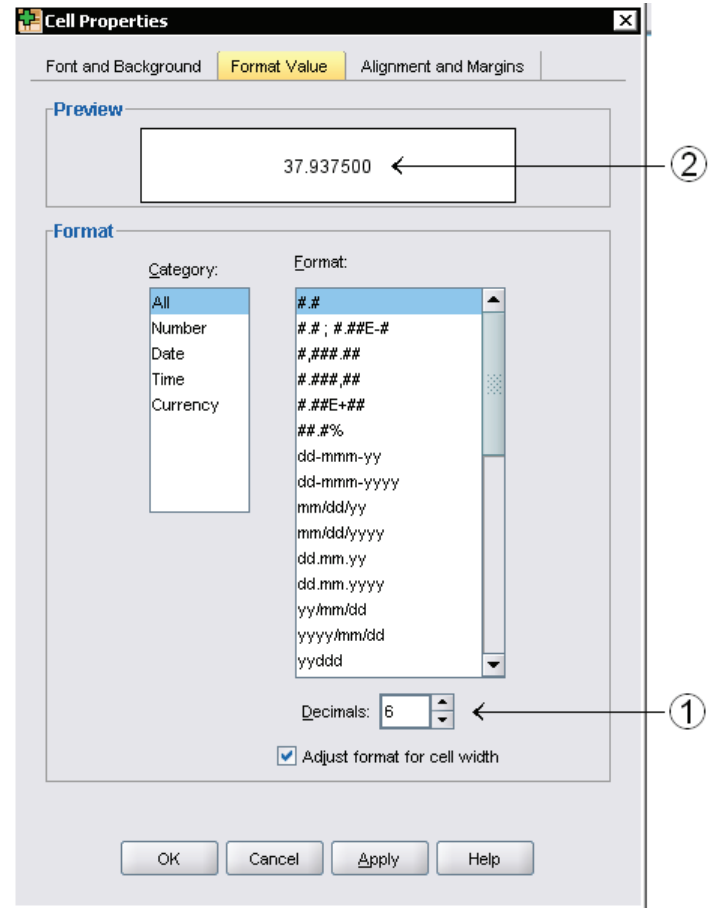
ทำการเพิ่มตำแหน่งทศนิยม

โดยการลบเลข 2 ออกแล้วพิมพ์ของใหม่เป็นเลข 6

1. ของใหม่เป็นทศนิยม 6 ตำแหน่ง
2. ตัวอย่างของรูปแบบการแสดงผล คือ 37.937500

เสร็จแล้วคลิก OK จะได้ผลตามที่ต้องการ

Descriptive Statistics		
	age	Valid N (listwise)
N	48	48
Minimum	21	
Maximum	56	
Mean	37.937500	
Std. Deviation	9.553	



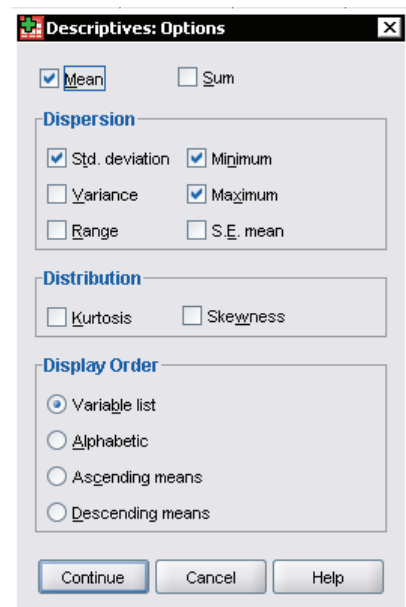
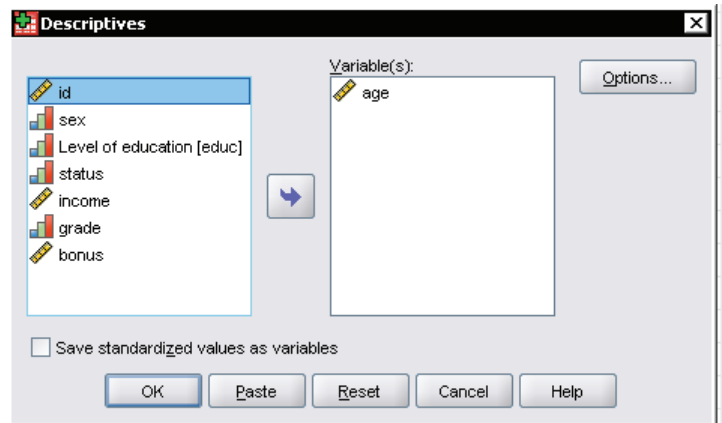
ขั้นสุดท้ายให้นำเมาส์มาคลิกนอกบริเวณตาราง จอภาพจะกลับไป SPSS Viewer

3.4 การคำนวณค่าสถิติเบื้องต้นอื่น ๆ ด้วยคำสั่ง Descriptives

จากขั้นตอนที่เราเลือกตัวแปร age เสร็จแล้ว

หากต้องการคำนวณค่าสถิติอื่นเพิ่มเติม

ให้นำเมาส์ไปคลิกที่ปุ่ม Options



บนจอภาพจะมีเมนูย่อย ให้เลือกค่าสถิติต่างๆ เพิ่มเติมค่าสถิติอื่น ๆ ที่ต้องการคำนวณให้คลิกที่กรอบสี่เหลี่ยมเพื่อให้เกิดเครื่องหมายถูก

- มีเครื่องหมายถูก แสดงว่า ให้คำนวณค่า
 - ไม่มีเครื่องหมายถูก แสดงว่า ไม่ให้คำนวณค่า
 - เลือกอย่างใดอย่างหนึ่งเท่านั้น
- โดยการคลิกให้เกิดจุดสีดำหน้าข้อที่ต้องการ

ในตัวอย่างนี้ขอให้เลือกทุกกรอบสี่เหลี่ยมให้เป็นเครื่องหมายถูก
เสร็จแล้วให้คลิก Continue
และ OK ตามลำดับ

จะได้ผลการคำนวณดังนี้

*Output1 [Document1] - SPSS Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Add-ons W

Output

- Descriptives
- Descriptives
 - Title
 - Notes
 - Active Dataset
 - Descriptive Statistics

Descriptives
[DataSet1] C:\Documents and Settings\dun

	N	Range	Minimum
	Statistic	Statistic	Statistic
age	48	35	21
Valid N (listwise)	48		

Descriptives: Options

Mean Sum

Dispersion

Std. deviation Minimum

Variance Maximum

Range S.E. mean

Distribution

Kurtosis Skewness

Display Order

Variable list

Alphabetic

Ascending means

Descending means

Continue Cancel Help

ทำการ Transpose ตาราง
ได้ผลเป็นดังนี้

*Output1 [Document1] - SPSS Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Add-ons W

Output

- Descriptives
- Descriptives
 - Title
 - Notes
 - Active Dataset
 - Descriptive Statistics

Descriptives
[DataSet1] C:\Documents and Settings\dunr

Descriptive Statistics

		age	Valid N (listwise)
N	Statistic	48	48
Range	Statistic	35	
Minimum	Statistic	21	
Maximum	Statistic	56	

ผลการคำนวณทั้งหมดคือ

Descriptives

[DataSet1] C:\Documents and Settings\dumrong\My Documents\dataspss16\example4.sav

Descriptive Statistics

		age	Valid N (listwise)
N	Statistic	48	48
Range	Statistic	35	
Minimum	Statistic	21	
Maximum	Statistic	56	
Sum	Statistic	1821	
Mean	Statistic	37.94	
	Std. Error	1.379	
Std. Deviation	Statistic	9.553	
Variance	Statistic	91.251	
Skewness	Statistic	.025	
	Std. Error	.343	
Kurtosis	Statistic	-.812	
	Std. Error	.674	

Skewness เป็นค่าที่บอกว่า โค้งความถี่ของข้อมูล มีลักษณะความเบ้ของเส้นโค้งเป็นอย่างไร

- Skewness < 0 โค้งความถี่จะมีลักษณะ เบ้ทางด้านซ้าย หรือ เบ้ทางด้านลบ
- Skewness = 0 โค้งความถี่จะมีลักษณะสมมาตร เป็นรูปประฆังคว่ำ หรือ normal curve
- Skewness > 0 โค้งความถี่จะมีลักษณะ เบ้ทางด้านขวา หรือ เบ้ทางด้านบวก

Kurtosis เป็นค่าที่บอกว่า โค้งความถี่ของข้อมูล มีลักษณะของเส้นโค้งมีการกระจายเป็นอย่างไร

- Kurtosis < 0 ข้อมูลมีการกระจายมาก โค้งความถี่จะมีลักษณะค่อนข้างแบน
- Kurtosis = 0 ข้อมูลมีการกระจายแบบปกติ โค้งความถี่จะมีลักษณะคล้ายการแจกแจงปกติ
- Kurtosis > 0 ข้อมูลมีการกระจายน้อย โค้งความถี่จะมีลักษณะสูงโด่ง

Mean (Std. Error) เป็นค่าที่ได้มาจากสูตร $\frac{\text{Standard Deviation}}{\sqrt{n}}$

การคำนวณด้วยโปรแกรม Mathcad

ORIGIN:= 1

i := 1.. 14

$x_i :=$

$$\sum_{i=1}^{14} x_i = 60$$

$$n := \text{length}(x)$$

$$n = 14$$

$$\min(x) = 1$$

$$\max(x) = 8$$

$$\text{mean}(x) = 4.2857$$

$$\text{median}(x) = 4$$

$$\text{var}(x) = 4.3469$$

$$\text{stdev}(x) = 2.0849$$

$$\text{Var}(x) = 4.6813$$

$$\text{Stdev}(x) = 2.1636$$

3
3
6
4
5
8
1
2
3
8
4
5
2
6

สูตร ความแปรปรวนและส่วนเบี่ยงเบนมาตรฐาน เมื่อกำหนดว่าข้อมูลนั้นคือประชากร

$$\sum_{i=1}^n (x_i - \text{mean}(x))^2$$

$$\text{variance_population_formula} := \frac{\sum_{i=1}^n (x_i - \text{mean}(x))^2}{n}$$

$$\text{variance_population_formula} = 4.3469$$

$$\text{standard_deviation_population_formula} := \sqrt{\frac{\sum_{i=1}^n (x_i - \text{mean}(x))^2}{n}}$$

$$\text{standard_deviation_population_formula} = 2.0849$$

สูตรความแปรปรวนและส่วนเบี่ยงเบนมาตรฐาน เมื่อกำหนดข้อมูลเป็นข้อมูลจากการสุ่มตัวอย่าง

$$\sum_{i=1}^n (x_i - \text{mean}(x))^2$$

$$\text{variance_sample_formula} := \frac{\sum_{i=1}^n (x_i - \text{mean}(x))^2}{n-1}$$

$$\text{variance_sample_formula} = 4.6813$$

$$\text{standard_deviation_sample_formula} := \sqrt{\frac{\sum_{i=1}^n (x_i - \text{mean}(x))^2}{n - 1}}$$

standard_deviation_sample_formula = 2.1636

คำสั่งของ Mathcad ที่ช่วยในการคำนวณค่าสถิติ

var(x) = ค่าความแปรปรวนของประชากร

stdev(x) = ค่าส่วนเบี่ยงเบนมาตรฐานของประชากร

Var(x) = ค่าความแปรปรวนของตัวอย่าง

Stdev(x) = ค่าส่วนเบี่ยงเบนมาตรฐานของตัวอย่าง

mean(x) = ค่าเฉลี่ยเลขคณิต

median(x) = มัธยฐาน

cnorm(z) = พื้นที่ใต้โค้งปกติมาตรฐานบนช่วง $(-\infty, z)$

slope(x, y) = สัมประสิทธิ์การถดถอยเชิงเส้นของ Y เทียบกับ X

intercept(x, y) = ระยะตัดแกน Y ของสมการถดถอย $\hat{y} = a + bx$

corr(x, y) = ค่าสัมประสิทธิ์สหสัมพันธ์ (Correlation)

min(x) = ค่าต่ำสุดของข้อมูล

max(x) = ค่าสูงสุดของข้อมูล

length(x) = จำนวนข้อมูล

ตัวอย่าง 3.5.2 จงสร้างแฟ้มข้อมูล (กำหนดตัวแปรชื่อ score) ที่ประกอบด้วยข้อมูล 3, 3, 6, 4, 5, 1, 2, 3, 8, 4, 5, 6 เสร็จแล้ว Save ไว้ที่ชื่อ example2.sav และคำนวณค่าสถิติเบื้องต้น

วิธีทำ เริ่มต้นที่ SPSS Data Editor สร้างแฟ้มข้อมูล

example2.sav และใช้คำสั่ง

Analyze / Descriptive Statistics / Descriptives

จะได้ผลการคำนวณดังนี้

	score	var	y
1	3.00		
2	3.00		
3	6.00		
4	4.00		
5	5.00		
6	1.00		
7	2.00		
8	3.00		
9	8.00		
10	4.00		
11	5.00		
12	6.00		

Descriptive Statistics

		score	Valid N (listwise)
N	Statistic	12	12
Range	Statistic	7.00	
Minimum	Statistic	1.00	
Maximum	Statistic	8.00	
Mean	Statistic	4.1667	
	Std. Error	.56183	
Std. Deviation	Statistic	1.94625	
Variance	Statistic	3.788	
Skewness	Statistic	.342	
	Std. Error	.637	
Kurtosis	Statistic	.004	
	Std. Error	1.232	

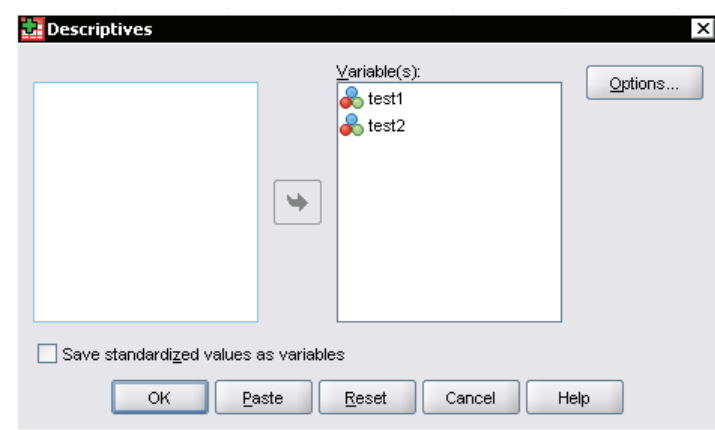
ตัวอย่าง 3.5.3 การสร้างแฟ้มข้อมูล 2 ตัวแปรเช่นข้อมูลคะแนนสอบย่อย 2 ครั้งของนิสิต 10 คน

ครั้งที่ 1	ครั้งที่ 2
76	81
60	52
85	87
58	70
91	86
75	77
82	90
64	63
79	85
88	83

ทำการบันทึกโดยใช้ชื่อแฟ้ม example3.sav และวิเคราะห์ข้อมูลเพื่อหาค่าสถิติเบื้องต้น

วิธีทำ จาก SPSS Data Editor สร้างแฟ้มแล้ว Save ชื่อ example3.sav

ใช้คำสั่ง Analyze / Descriptive Statistics / Descriptives



	test1	test2
1	76.00	81.00
2	60.00	52.00
3	85.00	87.00
4	58.00	70.00
5	91.00	86.00
6	75.00	77.00
7	82.00	90.00
8	64.00	63.00
9	79.00	85.00
10	88.00	83.00

เลือกตัวแปร test1 และ test2 ไปวิเคราะห์จะได้ผลการคำนวณดังนี้

		test1	test2	Valid N (listwise)
N	Statistic	10	10	10
Range	Statistic	33.00	38.00	
Minimum	Statistic	58.00	52.00	
Maximum	Statistic	91.00	90.00	
Sum	Statistic	758.00	774.00	
Mean	Statistic	75.8000	77.4000	
	Std. Error	3.68118	3.85054	
Std. Deviation	Statistic	11.64092	12.17648	
Variance	Statistic	135.511	148.267	
Skewness	Statistic	-.427	-1.200	
	Std. Error	.687	.687	
Kurtosis	Statistic	-1.173	.683	
	Std. Error	1.334	1.334	

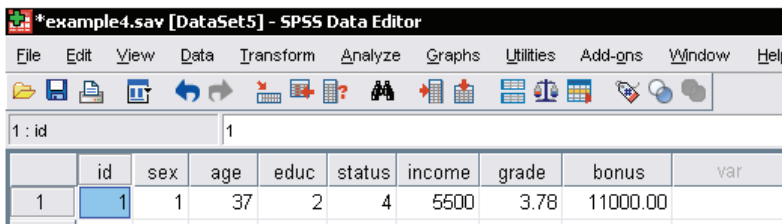
3.6 การคำนวณค่าสถิติเบื้องต้นด้วยคำสั่ง Analyze / Descriptive Statistics / Frequencies

คำสั่ง Analyze / Descriptive Statistics / Frequencies ใช้ในการคำนวณ

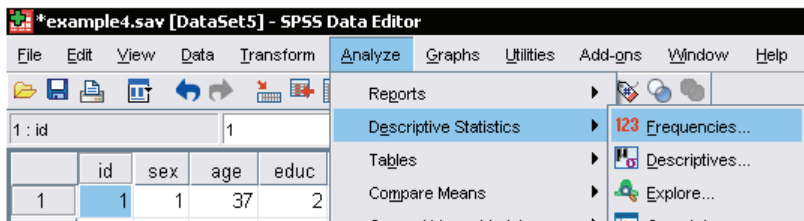
- ความถี่ข้อมูลแบบ 1 ทาง เช่น จำนวนชายและหญิง จำนวนคนที่มี status ต่าง ๆ กัน
- ค่าสถิติเบื้องต้นเช่น ค่าเฉลี่ยเลขคณิต มัธยฐาน เปอร์เซ็นไทล์
- เขียนกราฟความถี่ของข้อมูล

ตัวอย่างการใช้คำสั่ง Analyze / Descriptive Statistics / Frequencies กับตัวแปร age ในแฟ้มข้อมูล example4.sav

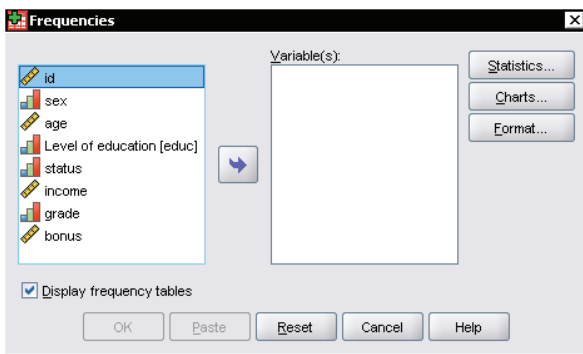
ขั้นที่ 1. นำแฟ้ม example4.sav เข้ามาใน SPSS Data Editor โดยใช้คำสั่ง File / Open



ขั้นที่ 2. เลือกคำสั่ง Analyze / Descriptive Statistics / Frequencies บนจอภาพจะเป็นดังนี้



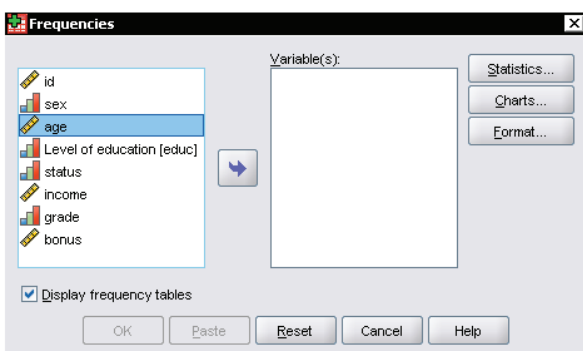
ขั้นที่ 3. คลิกเมาส์ที่ Frequencies จะได้เมนูย่อย Frequencies ของคำสั่งดังนี้



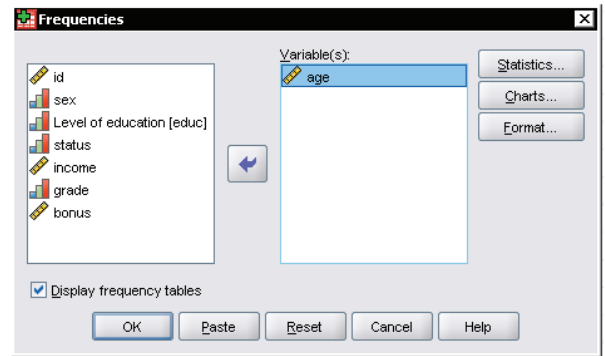
หมายเหตุ

1. ลำดับตัวแปรเรียงตามลำดับของ column
2. เมื่อเข้ามาครั้งแรกตัวแปรตัวแรกจะมีแถบสีเข้มทับอยู่

ขั้นที่ 4. เลือกตัวแปรโดยการเอาเมาส์คลิกที่ตัวแปรที่ต้องการ เช่น ตัวแปร age ขึ้นเป็นแถบสีเข้ม



ขั้นที่ 5. คลิกที่ปุ่ม  ตัวแปร age
ที่เราเลือกไว้จะมาอยู่ทางช่องขวามือ



ขั้นที่ 6. คลิก OK

ผลการคำนวณที่ SPSS Viewer เป็นดังนี้

***Output1 [Document1] - SPSS Viewer**

File Edit View Data Transform Insert Format Analyze Graphs Utilities Add-ons Window Help

Output

- Frequencies
 - Title
 - Notes
 - Active Dataset
 - Statistics
 - age

Frequencies
[DataSet5] C:\Documents and Settings\dumrong\My Document

Statistics

age

N	Valid	48
	Missing	2

age

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	21	1	2.1	2.1

ผลการคำนวณทั้งหมดเป็นดังนี้

Frequencies

[DataSet5] C:\Documents and Settings\dumrong\My Documents\datasps16\example4.sav

Statistics

age

N	Valid	48
	Missing	2

age

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	21	1	2.1	2.1
	22	2	4.2	6.2
	23	2	4.2	10.4
	24	1	2.1	12.5
	27	1	2.1	14.6
	28	1	2.1	16.7
	29	2	4.2	20.8
	30	1	2.1	22.9
	31	2	4.2	27.1
	32	2	4.2	31.2
	33	2	4.2	35.4
	34	2	4.2	39.6
	35	1	2.1	41.7

37	3	6.0	6.2	47.9
38	2	4.0	4.2	52.1
39	1	2.0	2.1	54.2
40	3	6.0	6.2	60.4
41	1	2.0	2.1	62.5
42	2	4.0	4.2	66.7
43	2	4.0	4.2	70.8
44	1	2.0	2.1	72.9
45	3	6.0	6.2	79.2
46	1	2.0	2.1	81.2
47	1	2.0	2.1	83.3
48	1	2.0	2.1	85.4
50	1	2.0	2.1	87.5
51	1	2.0	2.1	89.6
53	2	4.0	4.2	93.8
54	2	4.0	4.2	97.9
56	1	2.0	2.1	100.0
Total	48	96.0	100.0	
Missing	99	2	4.0	
Total	50	100.0		

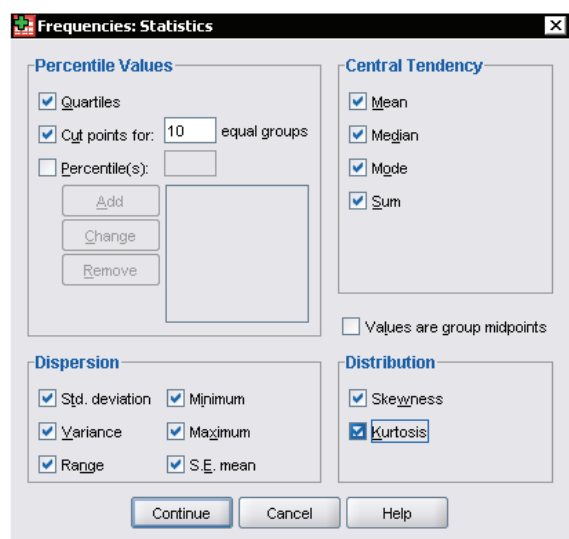
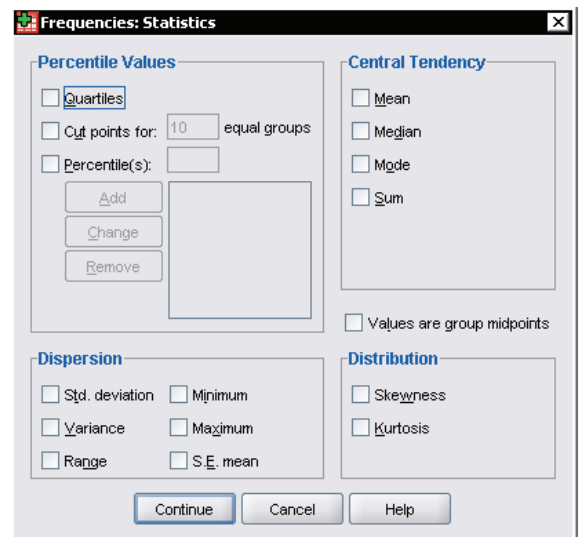
ความสามารถอื่นๆ ของคำสั่ง Frequencies ที่สามารถทำได้เช่น

- หาค่าสถิติเบื้องต้น (เหมือนคำสั่ง Analyze / Descriptive Statistics / Descriptives)
- หาเปอร์เซ็นต์ไทล์ 1, 2, 3, ... , 99
- เขียนกราฟความถี่ แบบ บาร์กราฟ และ แบบฮิสโตแกรม

ขั้นที่ 7. จากขั้นที่ 5.

เมื่อเลือกตัวแปร age ยังไม่ต้องคลิก OK

ให้คลิกที่ปุ่ม Statistics จะได้เมนูย่อย สำหรับเลือกค่านวนค่าสถิติที่ต้องการ



ขั้นที่ 8.

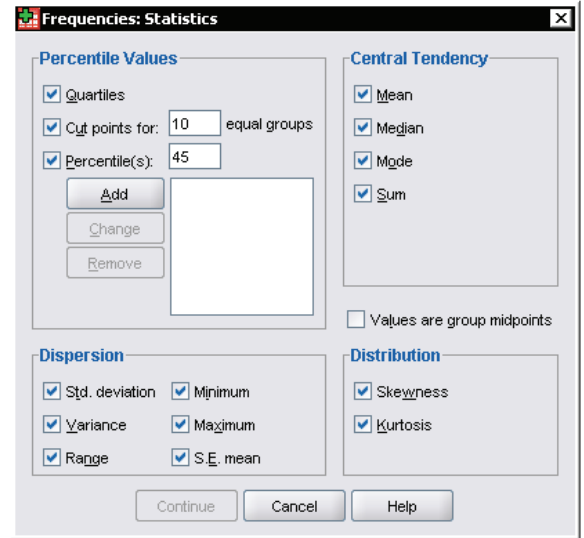
ต้องการค่านวนค่าสถิติใดให้ใส่ เครื่องหมายถูก ในช่องสี่เหลี่ยม (ในที่นี้ขอให้เลือกหมด)

ผลบนจอภาพจะเป็นดังนี้

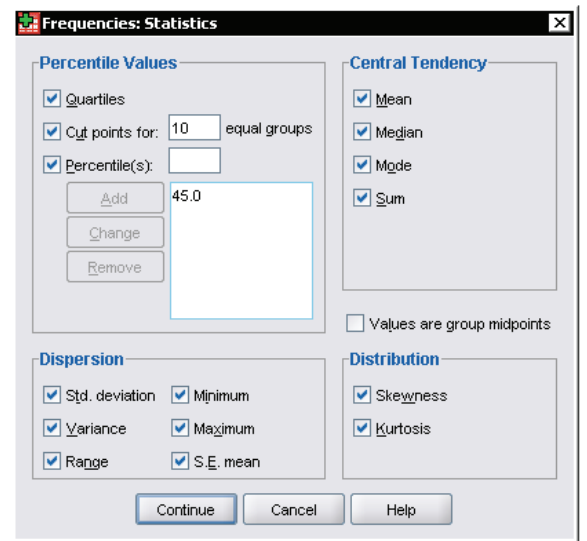
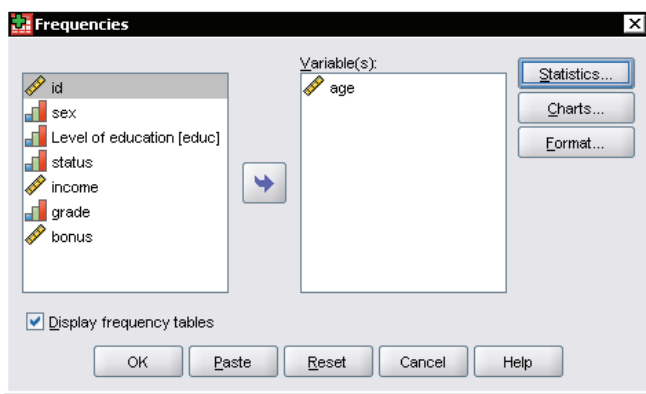
หมายเหตุ การเลือก Percentile ต้องกำหนด รายละเอียดของคำสั่งเพิ่มเติมดังนี้

การเลือกเปอร์เซ็นต์ไทล์

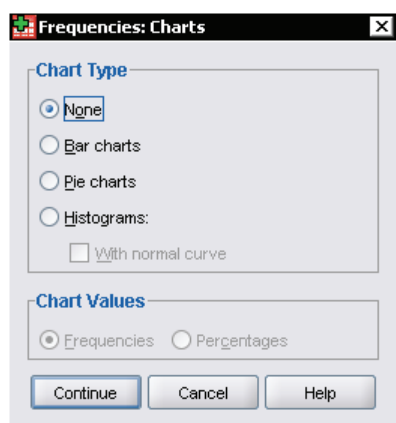
1. คลิกที่ Percentile(s)
2. ที่ช่องหลัง Percentile(s)
พิมพ์ค่าเปอร์เซ็นต์ไทล์ที่ต้องการ
เช่น เปอร์เซ็นต์ไทล์ 45
จะสังเกตเห็นว่าปุ่ม Add จะมีสีเข้มขึ้น
3. คลิก Add
เสร็จแล้วจอภาพจะมีเลข 45
ในกรอบของ Percentile



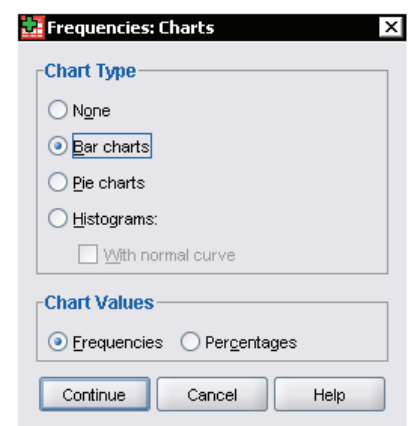
- ขั้นที่ 9. เสร็จแล้วให้คลิกปุ่ม Continue
จอภาพจะกลับไปเมนูย่อยของ
คำสั่ง Frequencies



- ขั้นที่ 10. การสั่งให้เขียนกราฟของการแจกแจงความถี่ให้คลิกที่ปุ่ม Charts จะได้เมนูย่อย



ให้คลิกที่ Bar charts

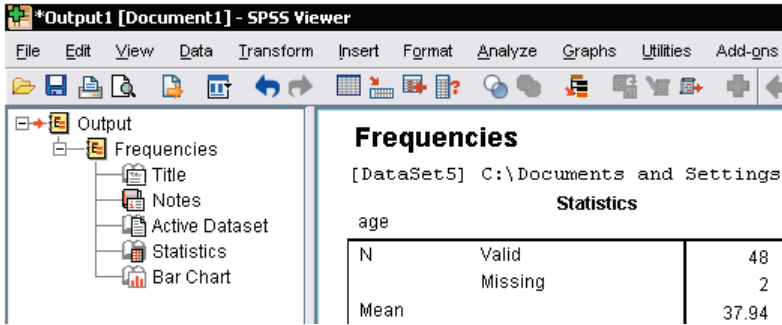


เสร็จแล้วให้คลิกที่ปุ่ม Continue จอภาพจะกลับไปเมนูย่อยของคำสั่ง Frequencies

หมายเหตุ ในกรณีที่ไม่ต้องการตารางแจกแจงความถี่ให้ยกเลิกเครื่องหมายถูกที่หน้า Display frequency tables



ต่อไปให้คลิก OK จะได้ผลการคำนวณค่าสถิติต่างๆ ที่ SPSS Viewer เป็นดังนี้



หมายเหตุ ตารางผลการคำนวณสามารถกำหนดให้แสดงผลตามแนวนอน หรือ แนวตั้งก็ได้เพื่อประหยัดพื้นที่ และสะดวกในการพิมพ์ลงกระดาษเราควรเลือกใช้การแสดงผลตามแนวตั้ง ผลการคำนวณคือ

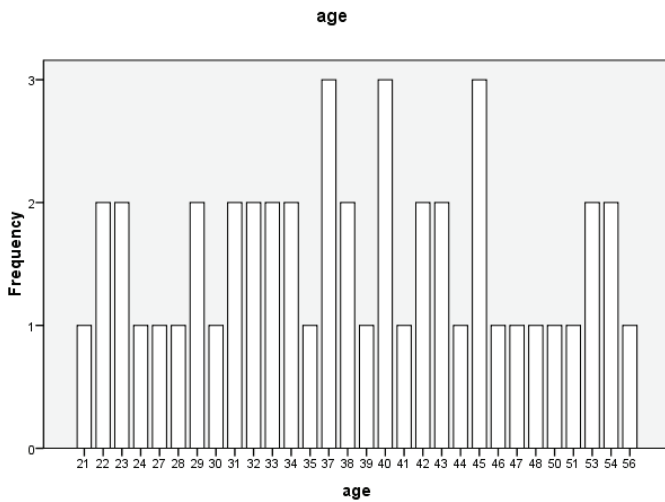
Frequencies

[DataSet5] C:\Documents and Settings\dumrong\My Documents\datasps16\example4.sav

Statistics		
age		
N	Valid	48
	Missing	2
Mean		37.94
Std. Error of Mean		1.379
Median		38.00
Mode		37 ^a
Std. Deviation		9.553
Variance		91.251
Skewness		.025
Std. Error of Skewness		.343
Kurtosis		-.812
Std. Error of Kurtosis		.674
Range		35
Minimum		21
Maximum		56
Sum		1821
Percentiles	10	23.00
	20	29.00
	25	31.00
	30	32.00
	40	34.60
	45	37.00
	50	38.00
	60	40.40
	70	43.30
	75	45.00
	80	46.20
	90	53.00

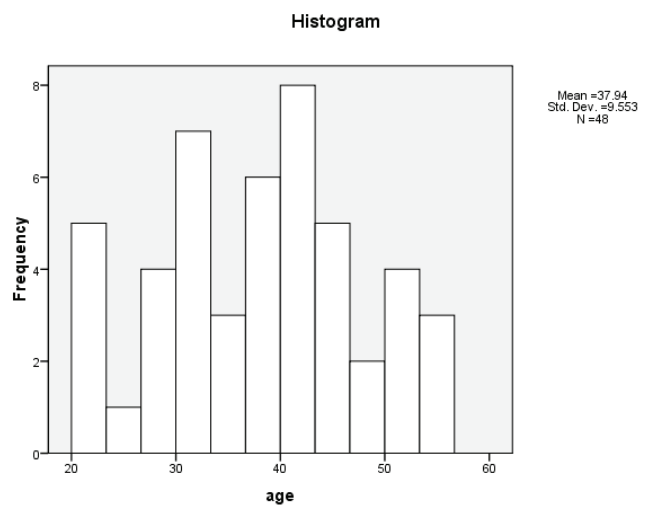
a. Multiple modes exist. The smallest value is shown.

ตัวอย่างกราฟของความถี่แบบ Bar chart



Bar chart

ตัวอย่างกราฟของความถี่แบบ Histogram



Histogram

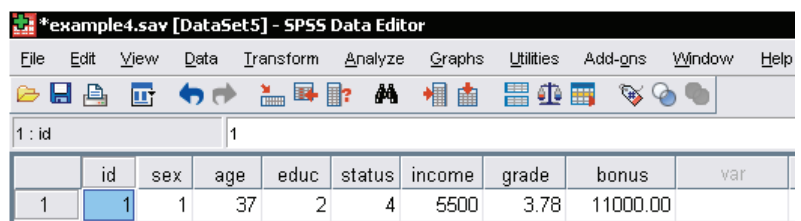
3.7 การคำนวณค่าสถิติเบื้องต้นด้วยคำสั่ง Analyze / Descriptive Statistics / Explore

คำสั่ง Analyze / Descriptive Statistics / Explore สามารถคำนวณค่าสถิติเบื้องต้นต่างๆ ได้เช่น ค่าเฉลี่ย มัธยฐาน ความแปรปรวน ส่วนเบี่ยงเบนมาตรฐาน ฯลฯ และสามารถหาช่วงความเชื่อมั่น 95% ของค่าเฉลี่ย ประชากรได้ ตัวอย่างการใช้คำสั่ง Analyze / Descriptive Statistics / Explore กับตัวแปร age ในแฟ้มข้อมูล example4.sav

ขั้นที่ 1.

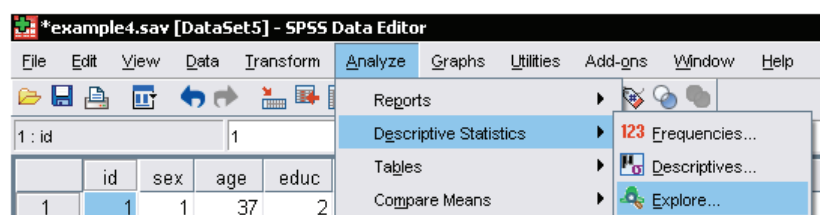
นำข้อมูลเข้า SPSS Data Editor

โดยใช้คำสั่ง File / Open



ขั้นที่ 2. เลือกคำสั่ง

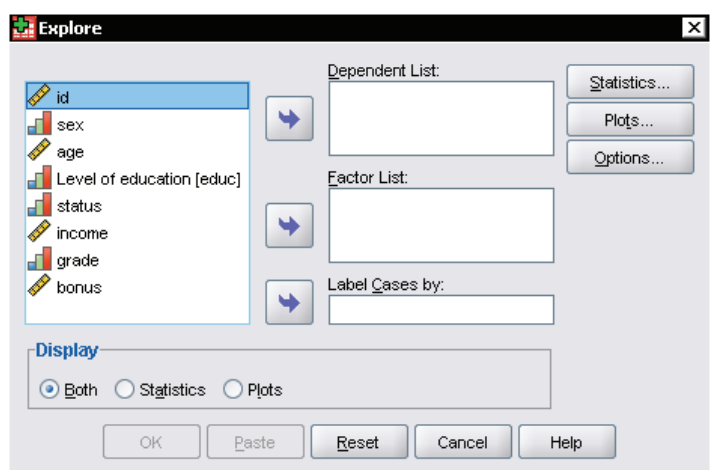
Analyze / Descriptive Statistics / Explore



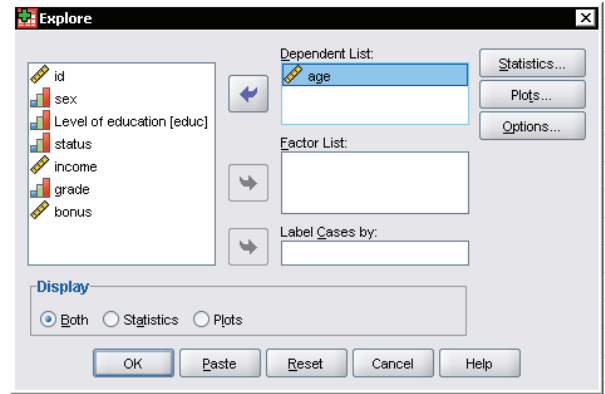
ขั้นที่ 3. คลิกที่ Explore

บนจอภาพจะขึ้นเมนูย่อย

ของคำสั่ง Explore ดังนี้

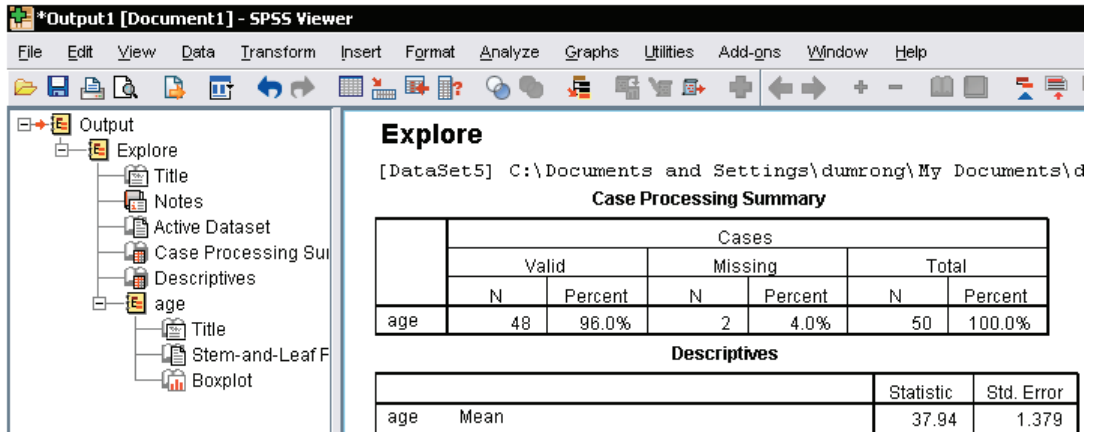


ขั้นที่ 4. เลือกตัวแปร age
มาไว้ที่ Dependent List



ขั้นที่ 5. คลิกที่ OK
จะได้ผลการคำนวณเป็นดังนี้

ผลการคำนวณ
ทั้งหมดคือ



Explore

[DataSet5] C:\Documents and Settings\dumrong\My Documents\datspps16\example4.sav

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
age	48	96.0%	2	4.0%	50	100.0%

หมายเหตุ จำนวนข้อมูลต้องคิดจาก n = 48 ซึ่งตัดค่าที่ไม่สมบูรณ์ออกไป 2 ตัว

Descriptives

		Statistic	Std. Error
age	Mean	37.94	1.379
	95% Confidence Interval for Mean	Lower Bound 35.16	Upper Bound 40.71
	5% Trimmed Mean	37.91	
	Median	38.00	
	Variance	91.251	
	Std. Deviation	9.553	
	Minimum	21	
	Maximum	56	
	Range	35	
	Interquartile Range	14	
	Skewness	.025	.343
	Kurtosis	-.812	.674

age

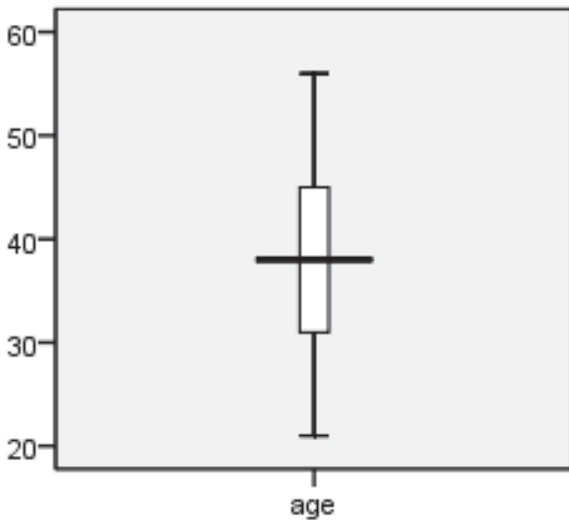
AGE Stem-and-Leaf Plot

Frequency	Stem & Leaf
6.00	2 . 122334
4.00	2 . 7899
9.00	3 . 011223344
7.00	3 . 5777889
9.00	4 . 000122334
6.00	4 . 555678
6.00	5 . 013344
1.00	5 . 6

Stem width: 10

Each leaf: 1 case(s)

กราฟแบบ Stem-and-Leaf Plot ที่ได้



หมายเหตุ Stem width = 10 เป็นตัวเลขบอกขนาดของตัวคูณ

ตัวอย่างเช่น 6.00 2 . 122334

stem = 2 , Leaf = 1 ข้อมูลคือ 2(10) + 1 = 21

เพราะฉะนั้นข้อมูล 6 ตัวคือ 21, 22, 22, 23, 23, 24

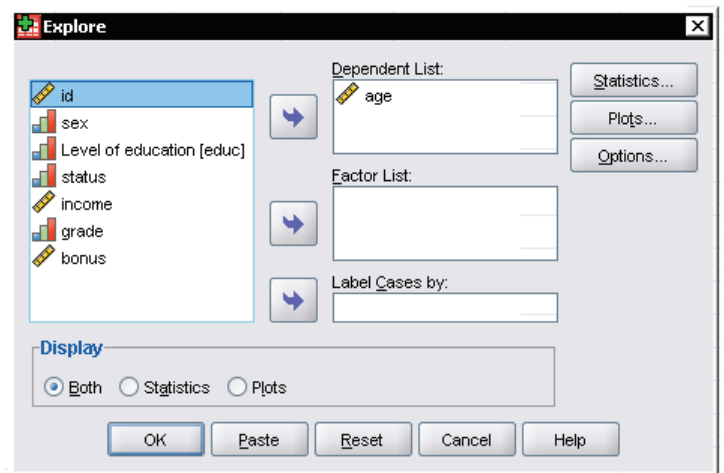
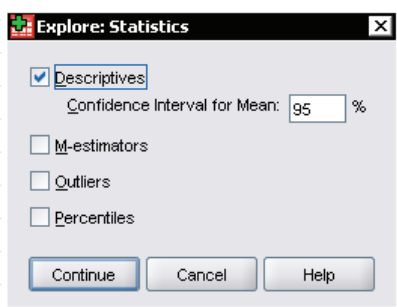
หรือ 4.00 2 . 7899 มีข้อมูล 5 ตัวคือ 27, 28, 29, 29

หมายเหตุ ความสามารถอื่นๆ ของคำสั่ง

Analyze / Descriptive Statistics / Explore

โดยเลือกค่าเพิ่มเติมได้ที่ปุ่ม Statistics

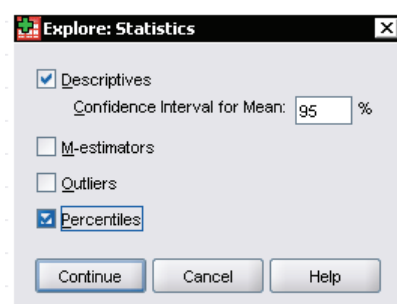
คลิกปุ่ม Statistics จะได้เมนูย่อย



ต้องการตารางเปอร์เซ็นต์คลิกที่ Percentiles

หรือต้องการเปลี่ยนเปอร์เซ็นต์ช่วงความเชื่อมั่น

ให้พิมพ์ค่าใหม่ลงไป



	1
1	37
2	29
3	48
4	33
5	45
6	38
7	23
x = 8	34
9	50
10	43
11	37
12	24
13	46
14	32
15	42
16	38

	1
17	41
18	54
19	32
20	43
21	22
22	40
23	37
x = 24	28
25	44
26	56
27	35
28	42
29	21
30	39
31	45
32	31

	1
33	51
34	23
35	40
36	47
37	53
38	27
39	29
x = 40	40
41	30
42	53
43	31
44	45
45	22
46	34
47	33
48	54

$$n := \text{length}(x) \quad n = 48 \quad \text{Var}(x) = 91.251$$

$$\text{mean}(x) = 37.938 \quad \text{median}(x) = 38$$

$$s := \text{Stdev}(x) \quad s = 9.553$$

$$\text{min}(x) = 21 \quad \text{max}(x) = 56$$

$$\text{Std_Error} := \frac{s}{\sqrt{n}} \quad \text{Std_Error} = 1.379$$

$$\text{Lower} := \text{mean}(x) - 2.012 \left(\frac{s}{\sqrt{n}} \right)$$

$$\text{Lower} = 35.163$$

$$\text{Upper} := \text{mean}(x) + 2.012 \left(\frac{s}{\sqrt{n}} \right)$$

$$\text{Upper} = 40.712$$

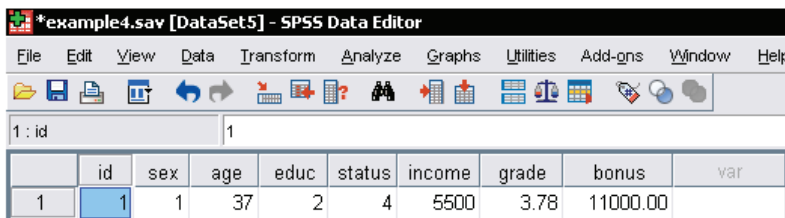
3.7 การคำนวณค่าสถิติเบื้องต้นด้วยคำสั่ง Analyze / Descriptive Statistics / Crosstabs

คำสั่ง Statistics / Descriptive Statistics / Crosstabs ใช้ในการคำนวณเกี่ยวกับ

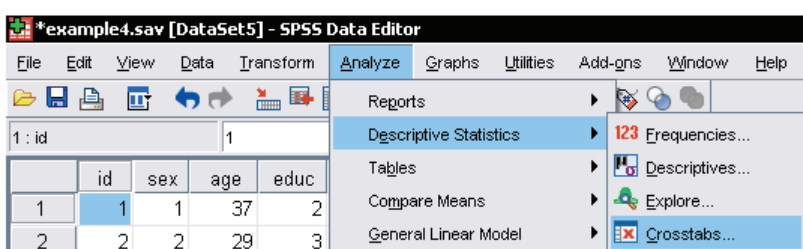
- การแจกแจงความถี่ข้อมูลแบบจำแนก 2 ทาง
- ค่าสถิติเบื้องต้นเช่น ค่าเฉลี่ย มัธยฐาน เปอร์เซ็นไทล์
- เขียนกราฟเปรียบเทียบความถี่ของข้อมูล
- คำนวณค่าสถิติไคสแควร์เพื่อทดสอบความเป็นอิสระของตัวแปร

ตัวอย่างการใช้คำสั่ง Analyze / Descriptive Statistics / Crosstabs กับตัวแปร educ และ sex โดยทำการแจกแจงความถี่จำแนกตาม ระดับการศึกษา และ เพศ ในแฟ้มข้อมูล example4.sav

ขั้นที่ 1. นำข้อมูลเข้า SPSS Data Editor โดยใช้คำสั่ง File / Open



ขั้นที่ 2. เลือกคำสั่ง Analyze / Descriptive Statistics / Crosstabs



ขั้นที่ 3. คลิกที่ Crosstabs

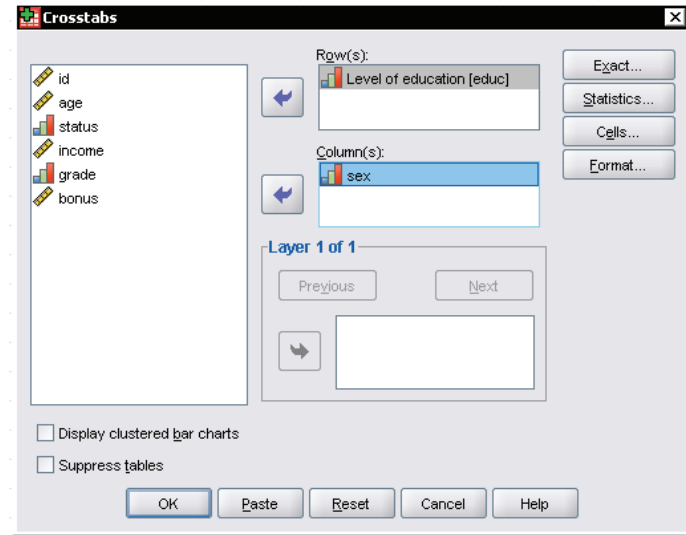
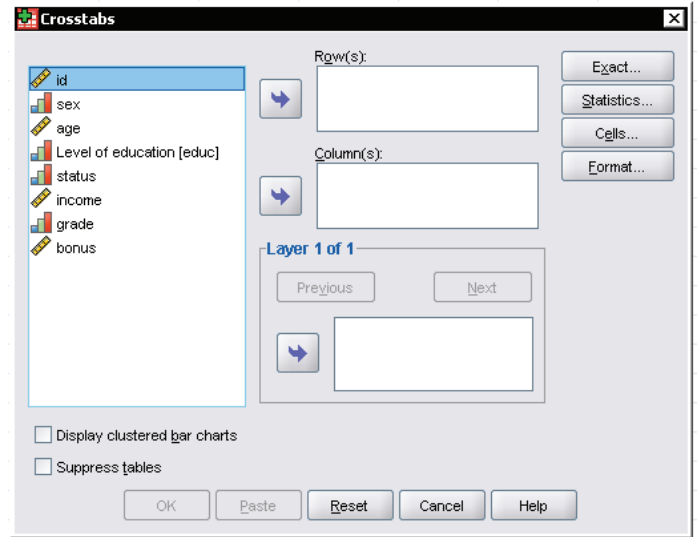
บนจอภาพจะขึ้นเมนูย่อย

ของคำสั่ง Crosstabs

ขั้นที่ 4.

เลือกตัวแปร educ มาไว้ที่ช่อง Row(s)

เลือกตัวแปร sex มาไว้ที่ช่อง Column(s)

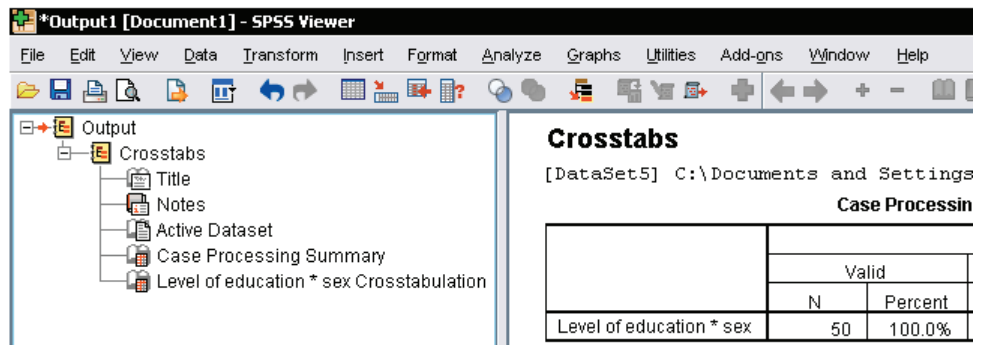


ขั้นที่ 5. คลิกที่ปุ่ม OK

จะได้ผลการคำนวณที่ SPSS Viewer ดังนี้

ผลการคำนวณทั้งหมดคือ

Crosstabs



[DataSet5] C:\Documents and Settings\dumrong\My Documents\dataspss16\example4.sav

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Level of education * sex	50	100.0%	0	.0%	50	100.0%

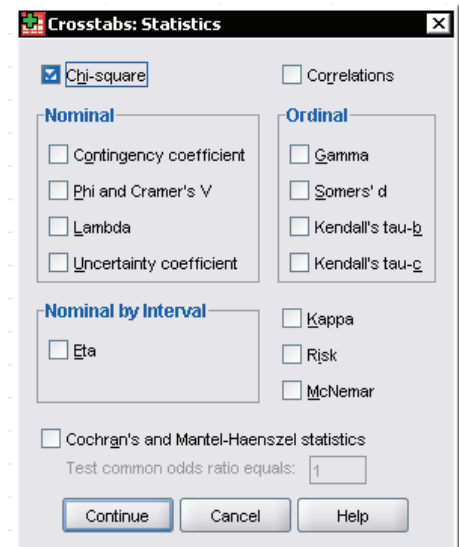
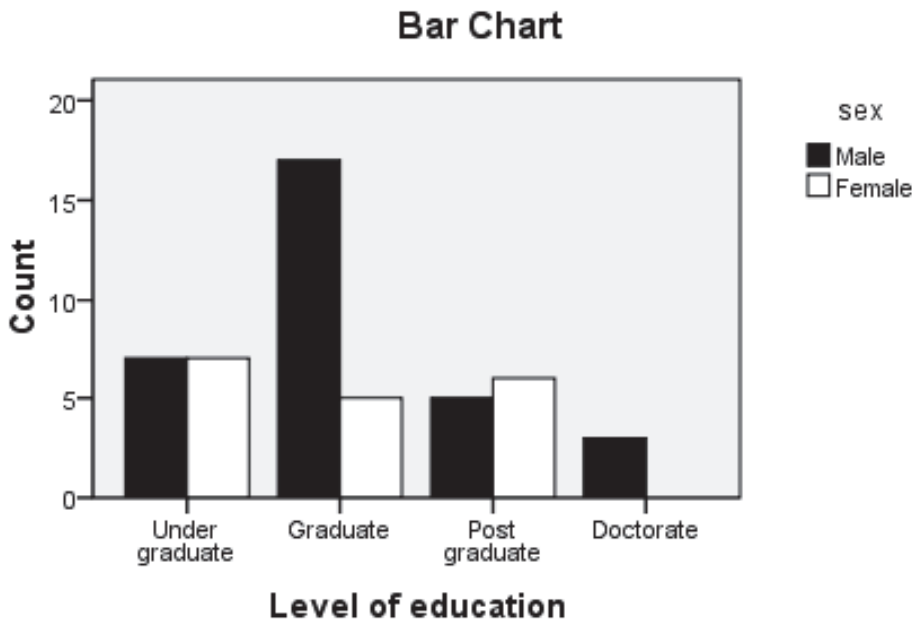
Level of education * sex Crosstabulation

Count		sex		Total
		Male	Female	
Level of education	Under graduate	7	7	14
	Graduate	17	5	22
	Post graduate	5	6	11
	Doctorate	3	0	3
Total		32	18	50

หมายเหตุ ในกรณีที่เรานำเลือก

Display clustered bar charts

จะได้กราฟของการแจกแจงความถี่ดังนี้



ในกรณีที่เรานำเลือกปุ่ม Statistics จะได้เมนูย่อย Crosstabs : Statistics

ของการคำนวณค่าสถิติต่าง ๆ เลือกค่าสถิติ Chi-square

คลิก Continue และ OK ตามลำดับ

จะได้ผลการคำนวณ

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	6.203 ^a	3	.102
Likelihood Ratio	7.193	3	.066
Linear-by-Linear Association	.500	1	.480
N of Valid Cases	50		

a. 3 cells (37.5%) have expected count less than 5. The minimum expected count is 1.08.

หมายเหตุ ค่า Pearson Chi-Square 6.203, df = 3 และ Asymp. Sig. (2-sided) สามารถนำไปสรุปผลได้ว่าระดับการศึกษา กับ เพศ ไม่เป็นอิสระต่อกัน ที่ระดับนัยสำคัญ 0.05 ซึ่งเนื่องจากการทดสอบความเป็นอิสระมีเนื้อหาและขั้นตอนการทำงานอย่างสมบูรณ์ จะกล่าวไว้ในบทที่ 7 หัวข้อ 7.7

ข้อควรระวัง จากการคำนวณทางทฤษฎีพบว่า $P(\chi^2 > 6.203) = 0.102$ เมื่อ df = 3

จากตาราง Chi-Square Tests ค่าแท้จริงของ Asymp. Sig. (2-sided) ควรจะเป็น 0.204

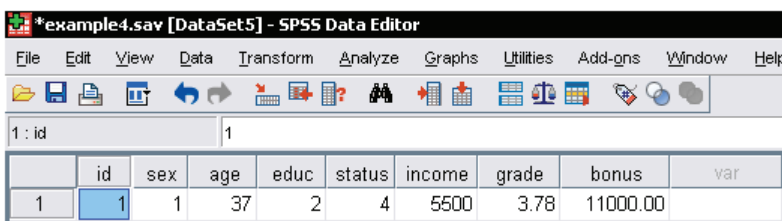
3.8 การคำนวณค่าสถิติเบื้องต้นด้วยคำสั่ง Analyze / Descriptive Statistics / Ratio

คำสั่ง Statistics / Descriptive Statistics / Crosstabs ใช้ในการคำนวณเกี่ยวกับ

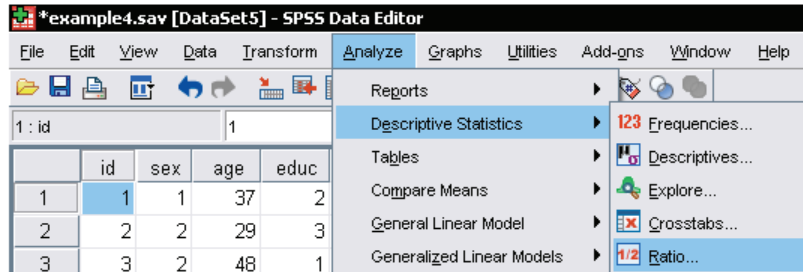
- ความถี่ข้อมูลแบบจำแนก 2 ทาง
- ค่าสถิติเบื้องต้นเช่น ค่าเฉลี่ย มัธยฐาน เปอร์เซ็นไทล์
- เขียนกราฟเปรียบเทียบความถี่ของข้อมูล
- คำนวณค่าสถิติโคสแควร์เพื่อทดสอบความเป็นอิสระ

ตัวอย่างการใช้คำสั่ง Analyze / Descriptive Statistics / Ratio กับตัวแปร sex income และ bonus ในแฟ้มข้อมูล example4.sav

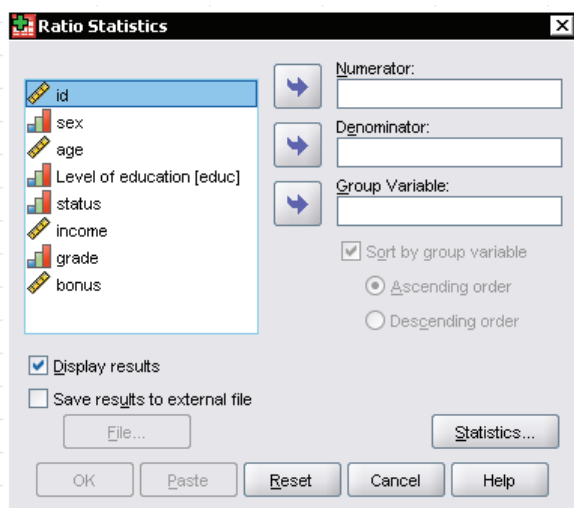
ขั้นที่ 1. นำข้อมูลเข้า SPSS Data Editor โดยใช้คำสั่ง File / Open



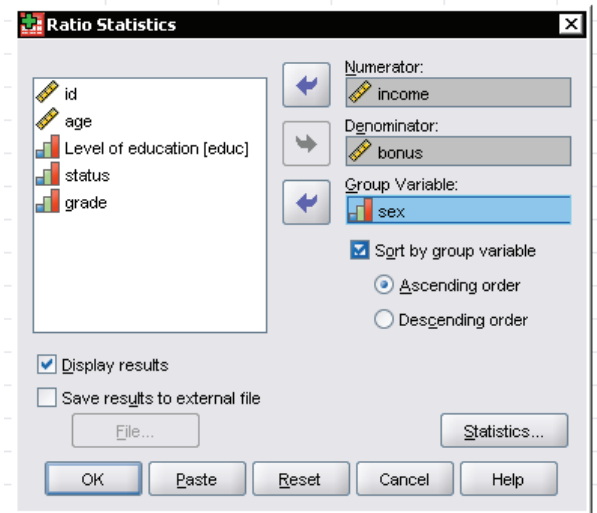
ขั้นที่ 2. เลือกคำสั่ง Analyze / Descriptive Statistics / Ratio



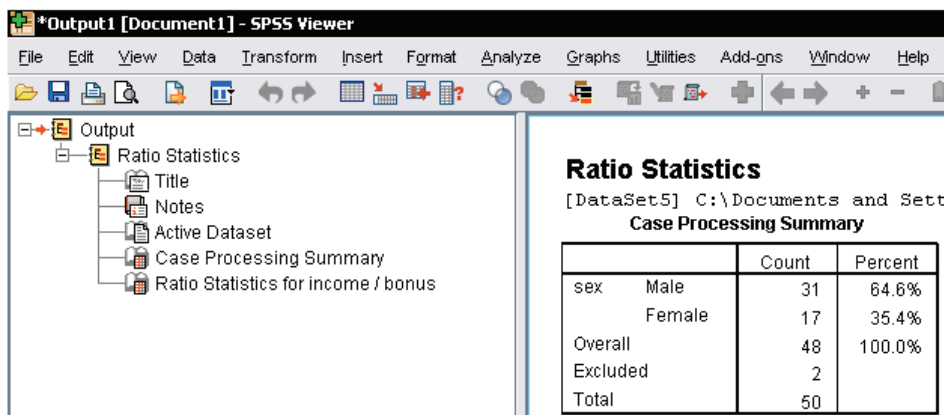
ขั้นที่ 3. คลิกที่ Ratio บนจอภาพจะขึ้นเมนูย่อย



- ขั้นที่ 4. เลือกตัวแปร income ไปไว้ที่ Numerator
เลือกตัวแปร bonus ไปไว้ที่ Denominator
เลือกตัวแปร sex ไปไว้ที่ Group Variable



เสร็จแล้วคลิก OK จะได้ผลการคำนวณดังนี้



ผลการคำนวณทั้งหมดคือ

Ratio Statistics

[DataSet5] C:\Documents and Settings\dumrong\My Documents\dataspss16\example4.sav

Case Processing Summary

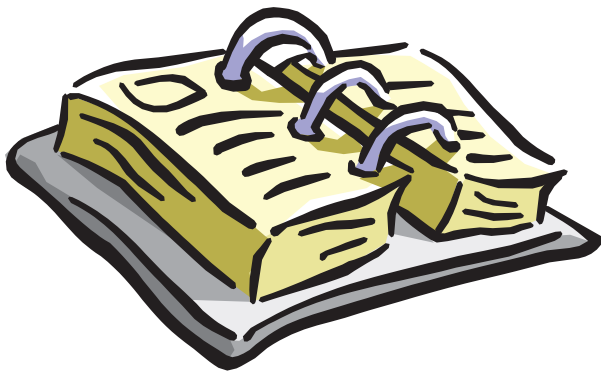
	Count	Percent
sex Male	31	64.6%
Female	17	35.4%
Overall	48	100.0%
Excluded	2	
Total	50	

Ratio Statistics for income / bonus

Group	Price Related Differential	Coefficient of Dispersion	Coefficient of Variation
			Median Centered
Male	1.218	.468	85.6%
Female	1.304	.618	103.1%
Overall	1.249	.521	91.1%

บทที่ 4

การวิเคราะห์ข้อมูลด้วยคำสั่ง Analyze Reports และ Tables



คำสั่ง Analyze / Reports และ Analyze / Tables ของ SPSS for Windows เป็นคำสั่ง รวบรวมข้อมูล นำเสนอข้อมูล และสามารถทำการวิเคราะห์ข้อมูลทางสถิติเบื้องต้นได้ ตัวอย่างคำสั่งต่างๆ เช่น

- Analyze / Reports / OLAP Cubes
- Analyze / Reports / Case Summaries
- Analyze / Reports / Case Summaries in Rows
- Analyze / Reports / Case Summaries in Columns
- Analyze / Tables / Custom Tables

4.1 การคำนวณค่าสถิติเบื้องต้นด้วยคำสั่ง Analyze / Reports / OLAP Cubes

คำสั่ง Analyze / Reports / OLAP Cubes เป็นคำสั่งใช้ในการตรวจสอบข้อมูลเบื้องต้น สามารถหาค่าสถิติเบื้องต้นต่างๆ เช่น ค่าเฉลี่ยเลขคณิต ค่าฐานนิยม ค่ามัธยฐาน ค่าส่วนเบี่ยงเบนมาตรฐาน ค่าสูงสุด ค่าต่ำสุด และสามารถคำนวณจำแนกตามกลุ่มได้

ขั้นที่ 1. เปิดแฟ้มข้อมูล example4.sav
เข้ามาใน SPSS Data Editor

	id	sex	age	educ	status	income	grade	bonus
1	1	1	37	2	4	5500	3.78	11000.00

ขั้นที่ 2. เลือกคำสั่ง

Analyze / Reports / OLAP Cubes

	id	sex	age	educ	st
1	1	1	37	2	

File Edit View Data Transform **Analyze** Graphs Utilities Add-ons Window Help

Reports
Descriptive Statistics
Tables
Compare Means

OLAP Cubes...
Case Summaries...
Report Summaries in Rows...
Report Summaries in Columns...

จะได้เมนูย่อยของคำสั่ง

Analyze / Reports / OLAP Cubes
เป็นดังนี้

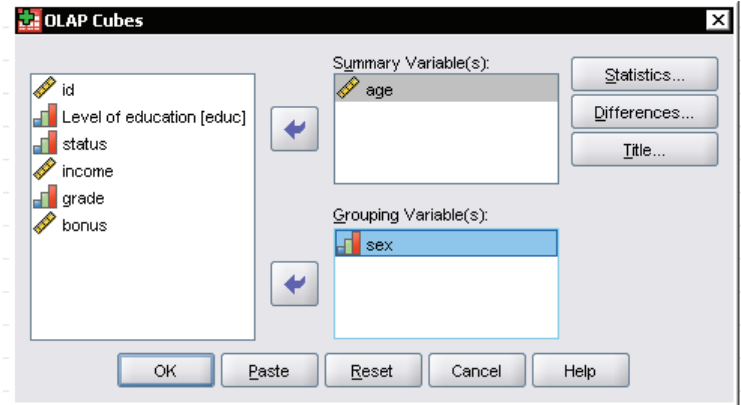
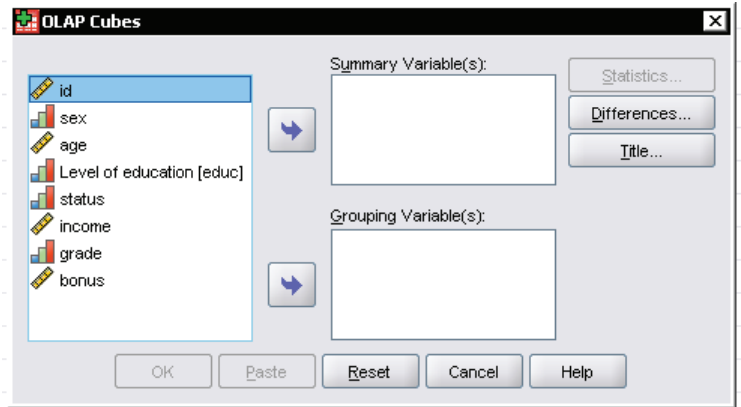
ขั้นที่ 3.

เลือกตัวแปร age

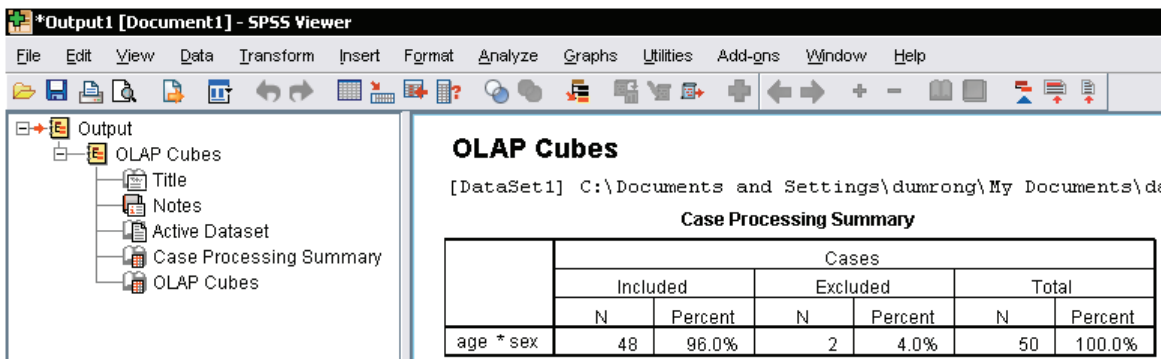
ไว้ที่ช่อง Summary Variable(s)

เลือกตัวแปร sex

ไว้ที่ช่อง Grouping Variable(s)

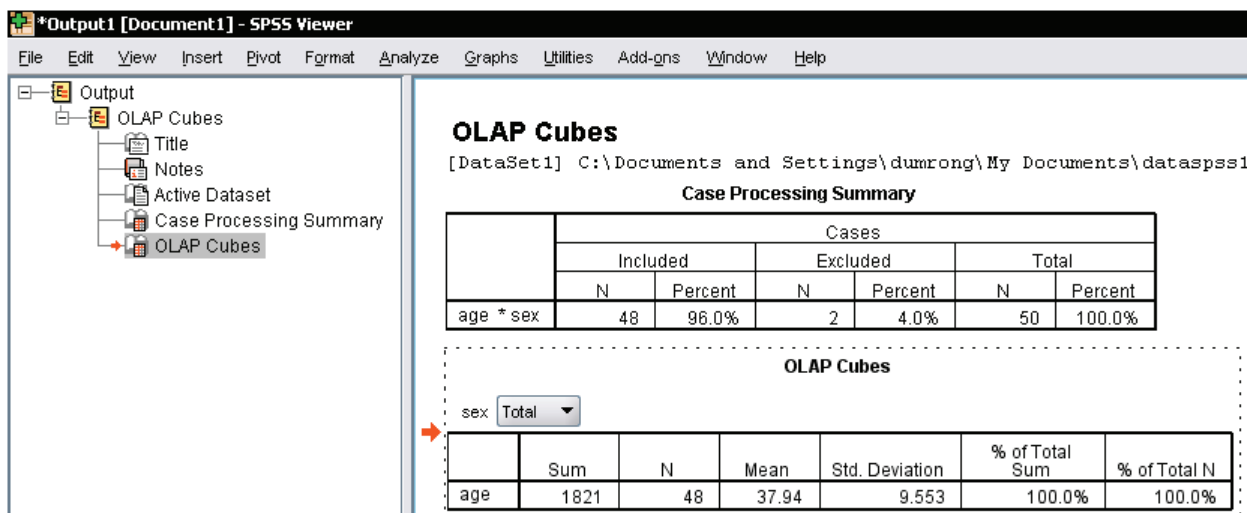


เสร็จแล้วคลิก OK จะได้ผลดังนี้



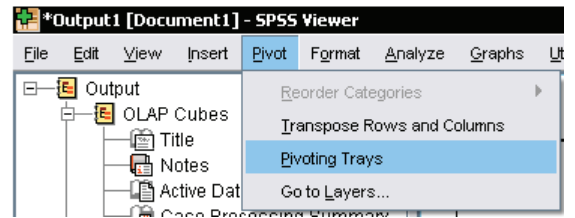
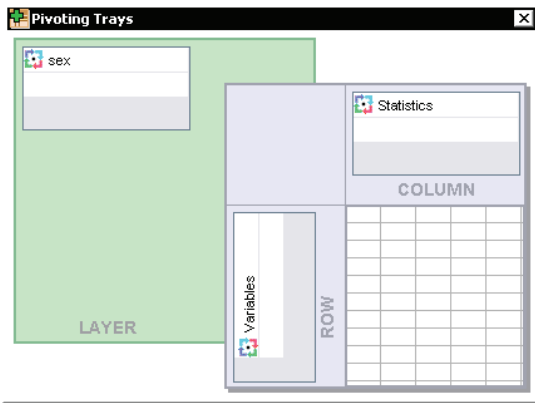
การแสดงผลตารางข้อมูลจำแนกตามกลุ่มชาย และกลุ่มหญิง ให้ทำตามขั้นตอนดังนี้

ขั้นที่ 4. นำเมาส์มากดดับเบิลคลิกที่ตาราง OLAP Cubes จะได้ผลดังนี้

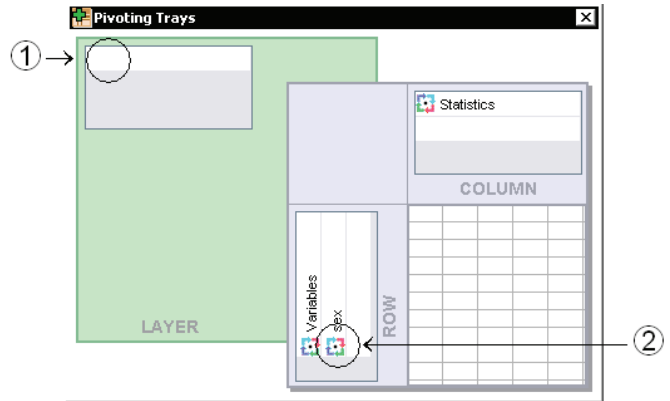


ขั้นที่ 5. เลือกคำสั่ง Pivot / Pivoting Trays

จะได้เมนูย่อย Pivoting Trays



ย้ายตัวแปร Sex ไปที่ช่อง Variables โดยการ drag
เมาส์จากตัวแปร Sex ในวงกลม 1 ไปไว้ที่วงกลม 2
เสร็จแล้วปิดหน้าต่าง Pivoting Trays



ลักษณะของตารางแสดงผลใน SPSS Viewer

จะได้ผลตารางของตัวแปร Sex

จำแนกเป็น 2 กลุ่มดังนี้

OLAP Cubes

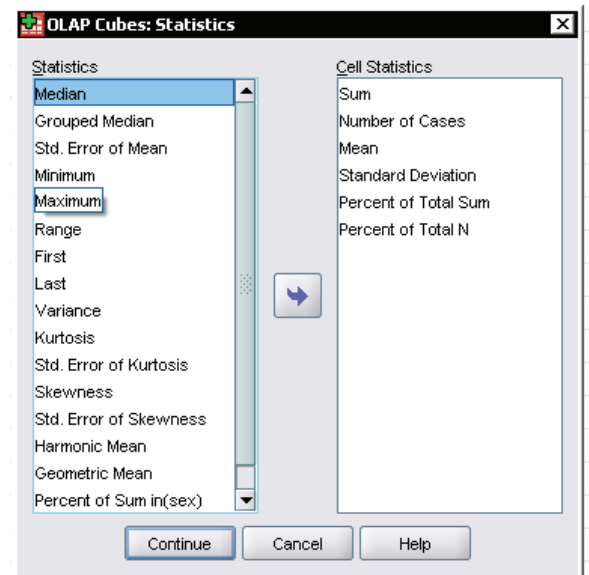
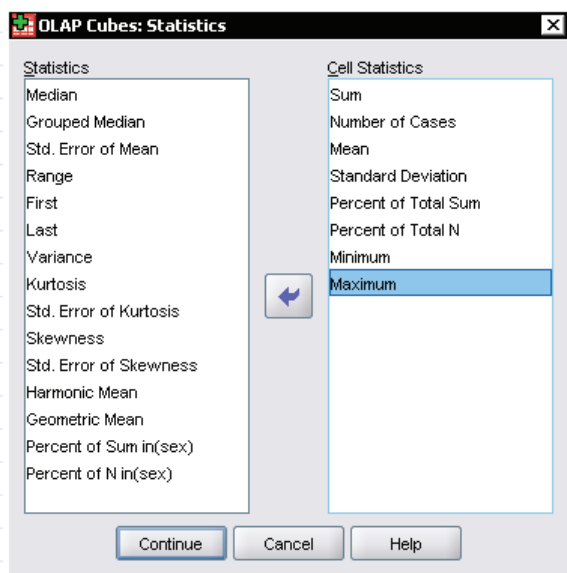
sex		Sum	N	Mean	Std. Deviation	% of Total Sum	% of Total N
age	Male	1211	31	39.06	9.966	66.5%	64.6%
	Female	610	17	35.88	8.652	33.5%	35.4%
	Total	1821	48	37.94	9.553	100.0%	100.0%

การคำนวณค่าสถิติต่างๆ เพิ่มเติม

ขั้นที่ 6. จากขั้นที่ 3 เมื่อเลือกตัวแปรเสร็จแล้ว

ให้คลิกปุ่ม Statistics จะได้เมนูย่อยเป็น

เลือก Minimum, Maximum มาไว้ที่ช่องขวามือ



เสร็จแล้วคลิก Continue และ OK ตามลำดับ และปรับแต่งการแสดงผลตามขั้นที่ 3 และ 4 จะได้ผลดังนี้

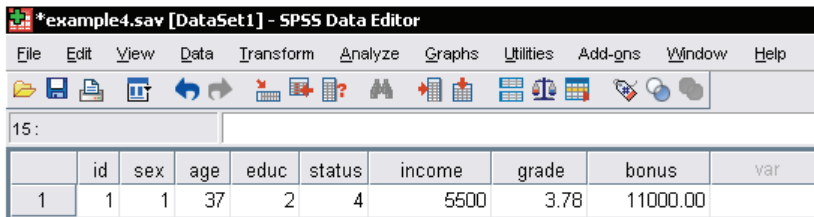
OLAP Cubes

	sex	Sum	N	Mean	Std. Deviation	% of Total Sum	% of Total N	Minimum	Maximum
age	Male	1211	31	39.06	9.966	66.5%	64.6%	21	56
	Female	610	17	35.88	8.652	33.5%	35.4%	22	53
	Total	1821	48	37.94	9.553	100.0%	100.0%	21	56

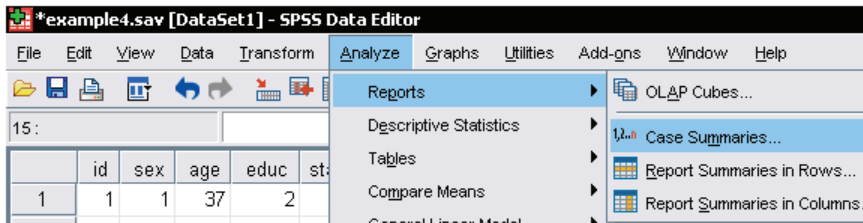
4.2 การคำนวณค่าสถิติเบื้องต้นด้วยคำสั่ง Analyze / Reports / Case Summaries

ผลการวิเคราะห์ข้อมูลด้วยคำสั่ง Analyze / Reports / Case Summaries จะได้รายละเอียดของข้อมูลจำแนกตามกลุ่ม พร้อมค่าสถิติเบื้องต้นต่างๆ เช่น ค่าเฉลี่ยเลขคณิต ค่าฐานนิยม ค่าส่วนเบี่ยงเบนมาตรฐาน ค่าสูงสุด ค่าต่ำสุด ที่คำนวณแบบจำแนกตามกลุ่มและแบบคิดรวมทั้งกลุ่ม

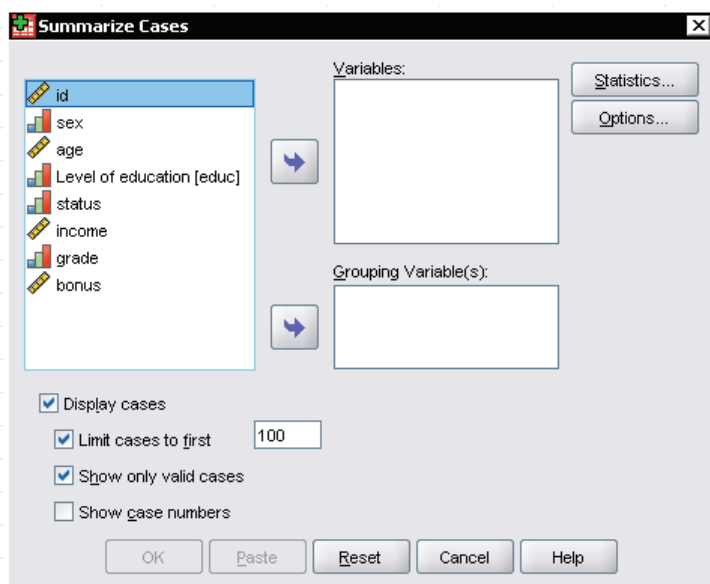
ขั้นที่ 1. เปิดแฟ้มข้อมูล example4.sav เข้ามาใน SPSS Data Editor



ขั้นที่ 2. เลือกคำสั่ง Analyze / Reports / Case Summaries



คลิกที่ Case Summaries จะได้เมนูย่อย Summarize Cases เป็น



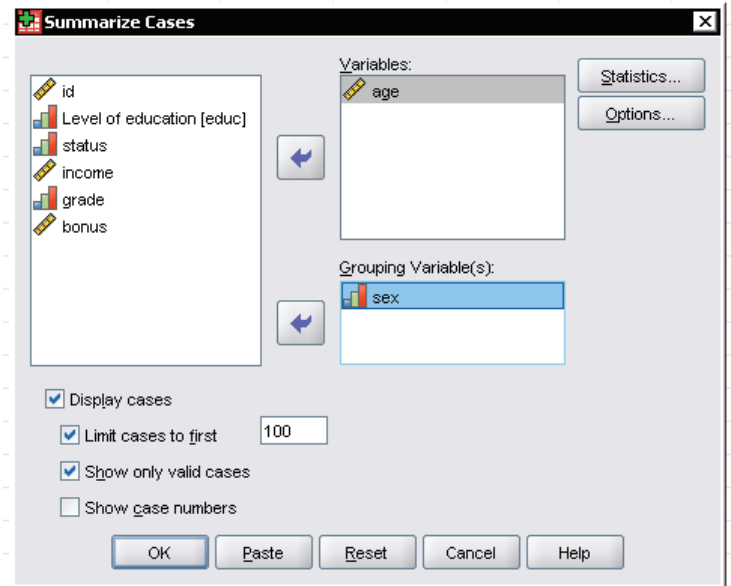
ขั้นที่ 3.

เลือกตัวแปร age ไปไว้ที่ช่อง Variables

เลือกตัวแปร sex ไปไว้ที่ช่อง

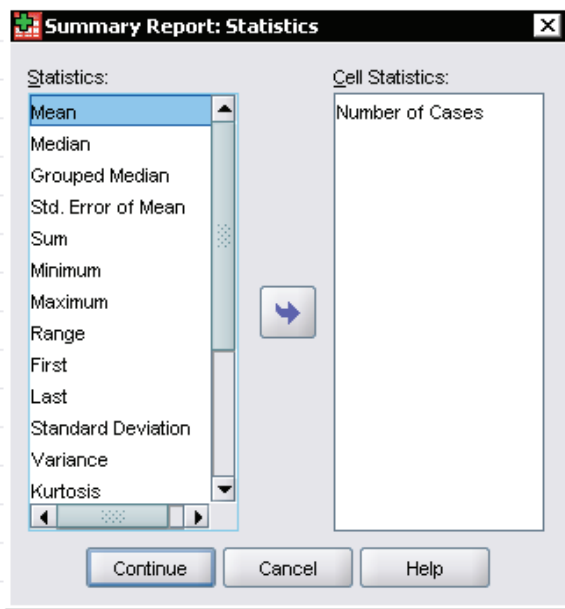
Grouping Variable(s)

บนจอภาพจะเป็น

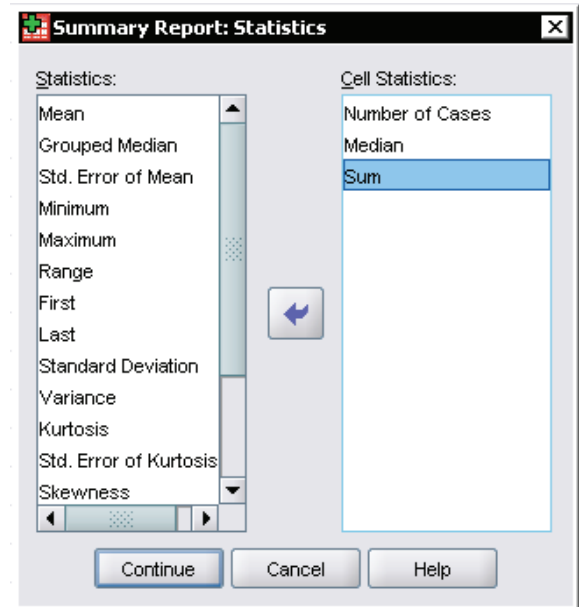


ขั้นที่ 4. คลิกปุ่ม Statistics

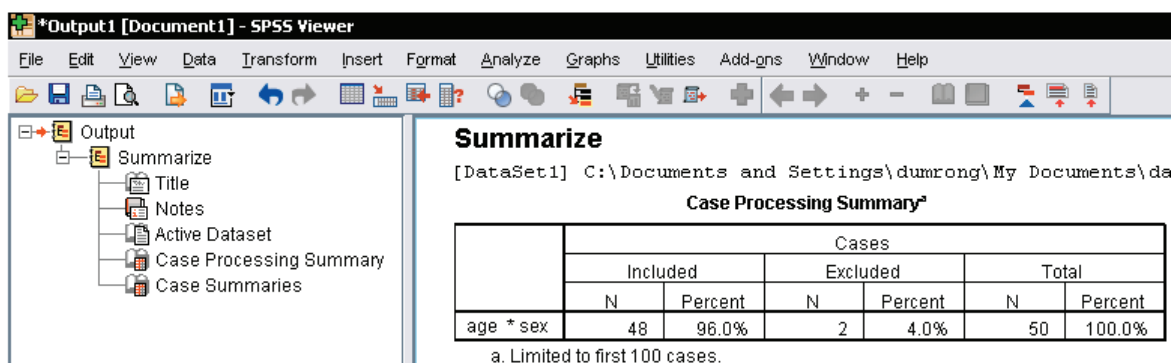
เพื่อให้ SPSS คำนวณค่าสถิติอื่นๆ เพิ่มเติม



เลือก Median, Sum มาไว้ทางขวามือ



เสร็จแล้วคลิก Continue และ OK ตามลำดับ
จะได้ผลการคำนวณดังนี้



ผลการคำนวณทั้งหมดคือ

Summarize

[DataSet1] C:\Documents and Settings\dumrong\My Documents\dataspss16\example4.sav

Case Processing Summary^a

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
age * sex	48	96.0%	2	4.0%	50	100.0%

a. Limited to first 100 cases.

Case Summaries

sex	age
Male	37
	34
	50
	24
	46
	32
	42
	38
	54
	43
	40
	37
	28
	44
	56
	35
	21
	39
	45
	31
	51
	23
	40
	47
	53
	29
	40
	31
	45
	22
	54
	Total
	N
	31
	Median
	40.00
	Sum
	1211

	Female	1		29
		2		48
		3		33
		4		45
		5		38
		6		23
		7		43
		8		37
		9		41
		10		32
		11		22
		12		42
		13		27
		14		30
		15		53
		16		34
		17		33
		Total	N	17
			Median	34.00
			Sum	610
	Total	N		48
		Median		38.00
		Sum		1821

a Limited to first 100 cases.

จากตัวอย่างนี้จะเห็นว่าผลการวิเคราะห์ข้อมูลด้วยคำสั่ง Analyze / Reports / Case Summaries จะได้รายละเอียดของข้อมูลทุกตัวในแฟ้มข้อมูล จำแนกตามกลุ่ม พร้อมค่าสถิติเบื้องต้นต่างๆ ที่คำนวณแบบจำแนกตามกลุ่ม และ แบบคิดรวมทั้งกลุ่ม

4.3 การคำนวณค่าสถิติเบื้องต้นด้วยคำสั่ง

Analyze / Reports / Report Summaries in Rows

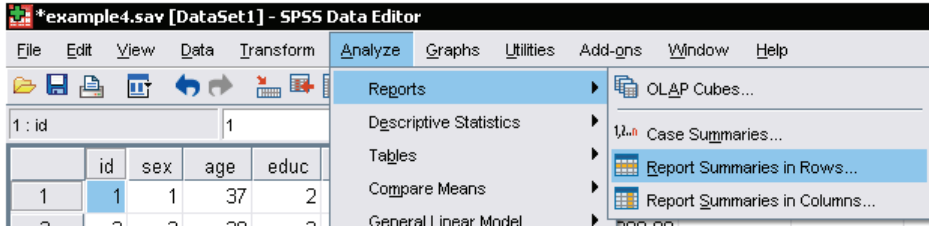
Analyze / Reports / Report Summaries in Columns

ผลการวิเคราะห์ข้อมูลด้วยคำสั่ง Analyze / Reports / Report Summaries in Rows จะได้รายละเอียดของข้อมูลจำแนกตามกลุ่ม พร้อมค่าสถิติเบื้องต้นต่างๆ เช่น ค่าเฉลี่ยเลขคณิต ค่าฐานนิยม ค่ามัธยฐาน ค่าส่วนเบี่ยงเบนมาตรฐาน ค่าสูงสุด ค่าต่ำสุด ที่คำนวณแบบจำแนกตามกลุ่มและแบบคิดรวมทั้งกลุ่ม

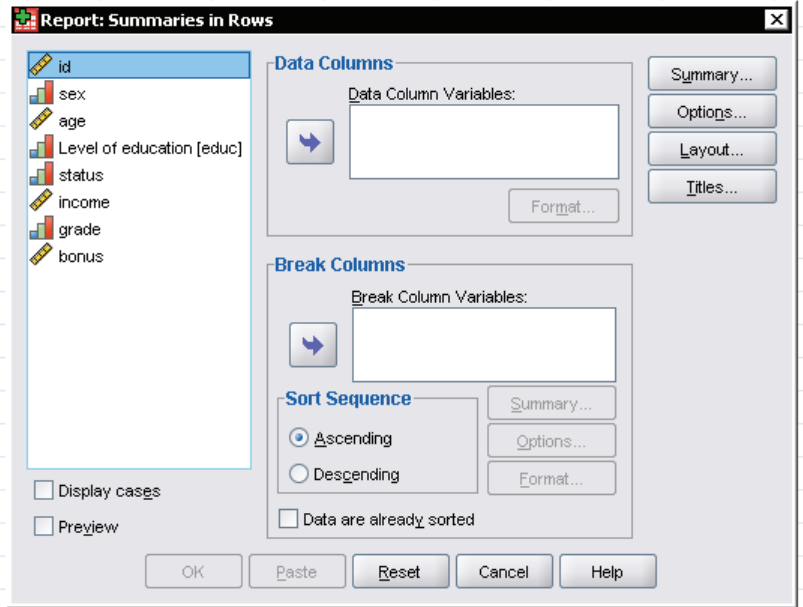
ขั้นที่ 1. เปิดแฟ้มข้อมูล example4.sav เข้ามาใน SPSS Data Editor

id	sex	age	educ	status	income	grade	bonus	var
1	1	37	2	4	5500	3.78	11000.00	

ขั้นที่ 2. เลือกคำสั่ง Analyze / Reports / Report Summaries in Rows



คลิกที่ Report Summaries in Rows
จะได้เมนูย่อยเป็น



ขั้นที่ 3.

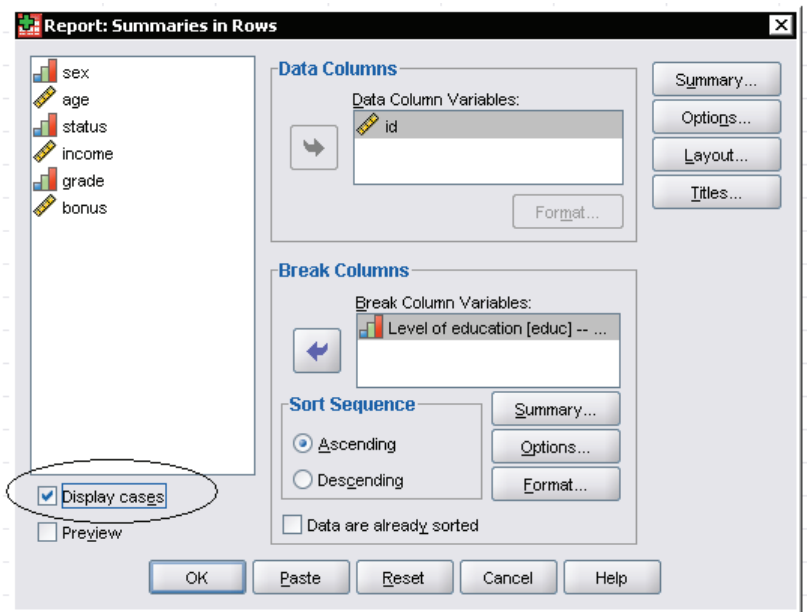
เลือกตัวแปร id

ไปไว้ที่ช่อง Data Column Variables

เลือกตัวแปร educ

ไปไว้ที่ช่อง Break Column Variables

และคลิกเลือก Display cases



คลิก OK จะได้ผลการคำนวณเป็นรายงาน ข้อมูลของ เลขที่ (id) และ ระดับการศึกษา (educ) ดังนี้

Level of education	ID
Under graduate	3
	4
	7
	15
	19
	22
	23

	26
	30
	33
	34
	43
	45
	48
Graduate	1
	5
	9
	10
	11
	13
	14
	16
	18
	20
	24
	28
	31
	32
	35
	38
	36
	40
	44
	46
	47
	50
Post graduate	2
	6
	8
	12
	21
	27
	29
	36
	37
	42
	49
Doctorate	17
	25
	41

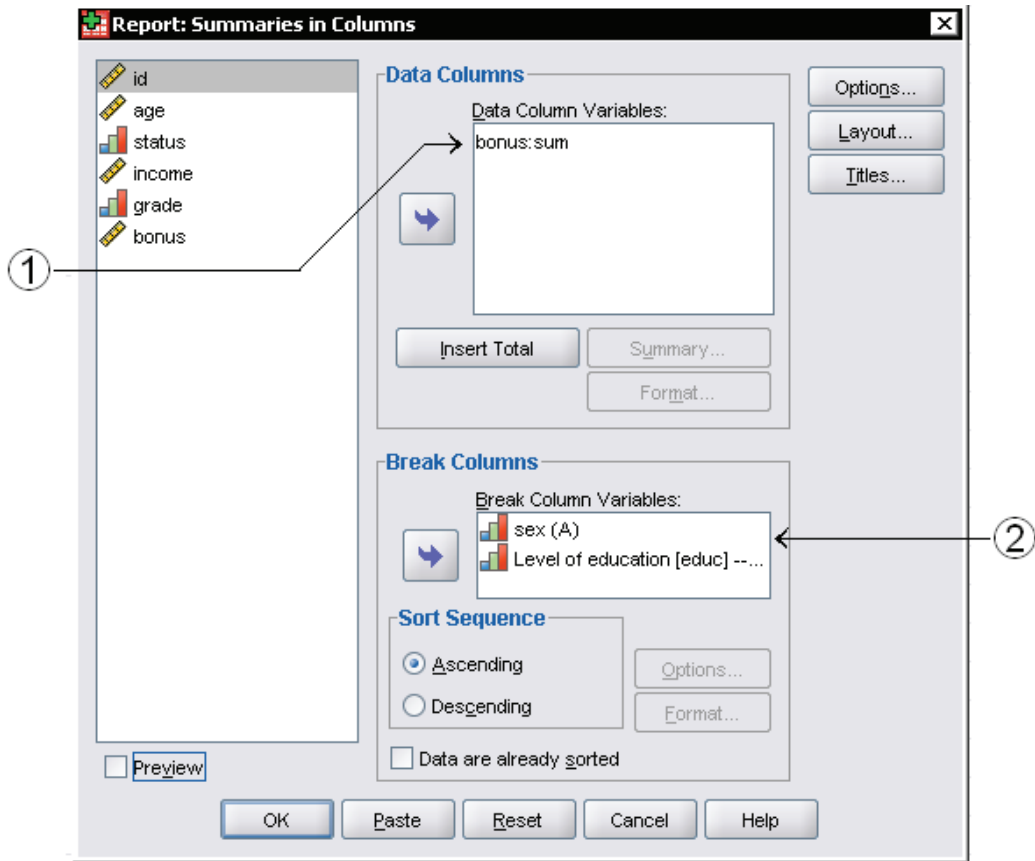
ในการทำงานเดียวกันคำสั่ง Analyze / Reports / Report Summaries in Columns

สามารถทำรายงานข้อมูลได้แบบเดียวกับ Analyze / Reports / Report Summaries in Rows

ตัวอย่างเช่น การจำแนกผลบวกของตัวแปร Bonus..... (1)

จำแนกตาม เพศ (sex) และ ระดับการศึกษา (educ)..... (2)

โดยการเลือกในเมนูย่อย Report Summaries in Columns ดังนี้



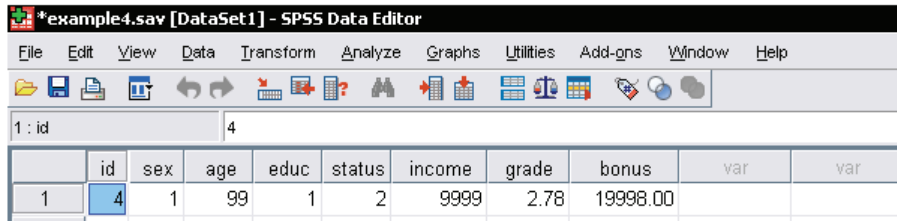
ผลการวิเคราะห์ข้อมูลที่ได้คือ

	Level of education	bonus Sum
Male	Under graduate	75898.00
	Graduate	248500.00
	Post graduate	71300.00
	Doctorate	69500.00
Female	Under graduate	66200.00
	Graduate	84297.00
	Post graduate	89400.00

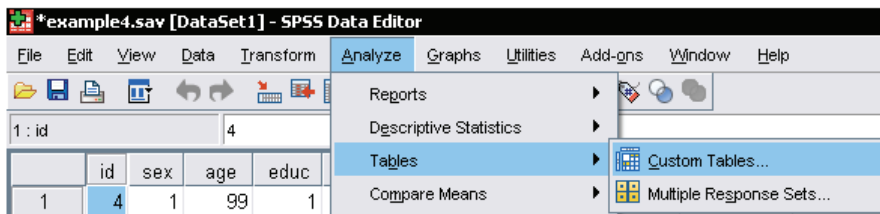
4.4 การนำเสนอข้อมูลด้วยคำสั่ง Analyze / Tables / Custom Tables

โปรแกรม SPSS for Windows มีคำสั่ง Analyze / Tables ช่วยในการนำเสนอข้อมูลในรูปแบบตารางที่จำแนกข้อมูลออกเป็น กลุ่มๆ และ หมวดหมู่ต่างๆ ได้อย่างดีและสวยงาม ตัวอย่างเช่น

ขั้นที่ 1. นำแฟ้มข้อมูล example4.sav เข้ามาใน SPSS Data Editor

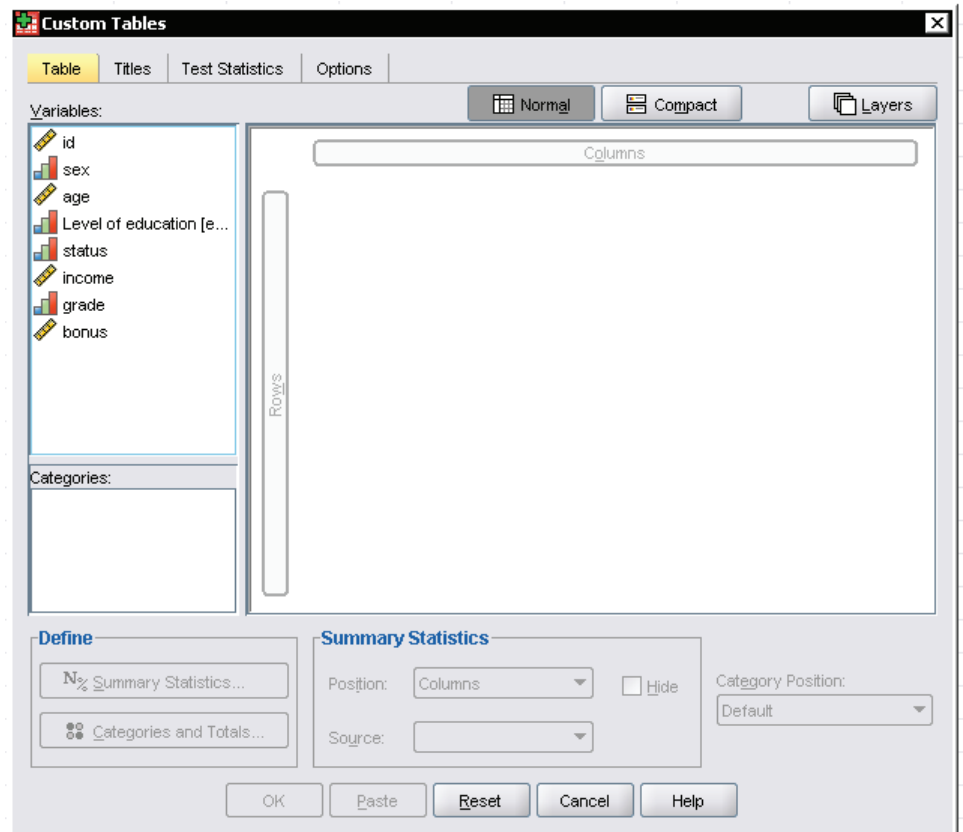


ขั้นที่ 2. เลือกคำสั่ง Analyze / Tables / Custom Tables



จะได้เมนูย่อย

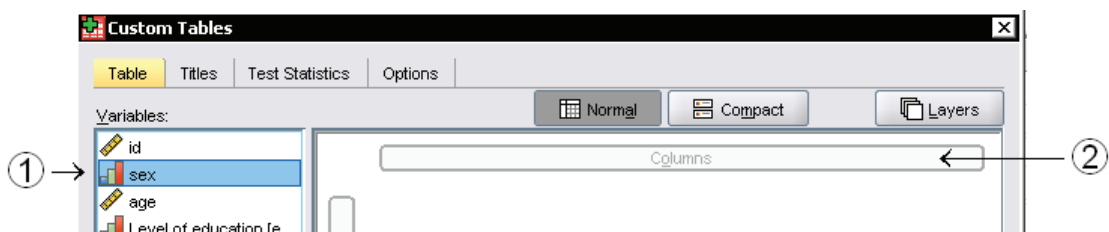
Custom Tables เป็นดังนี้



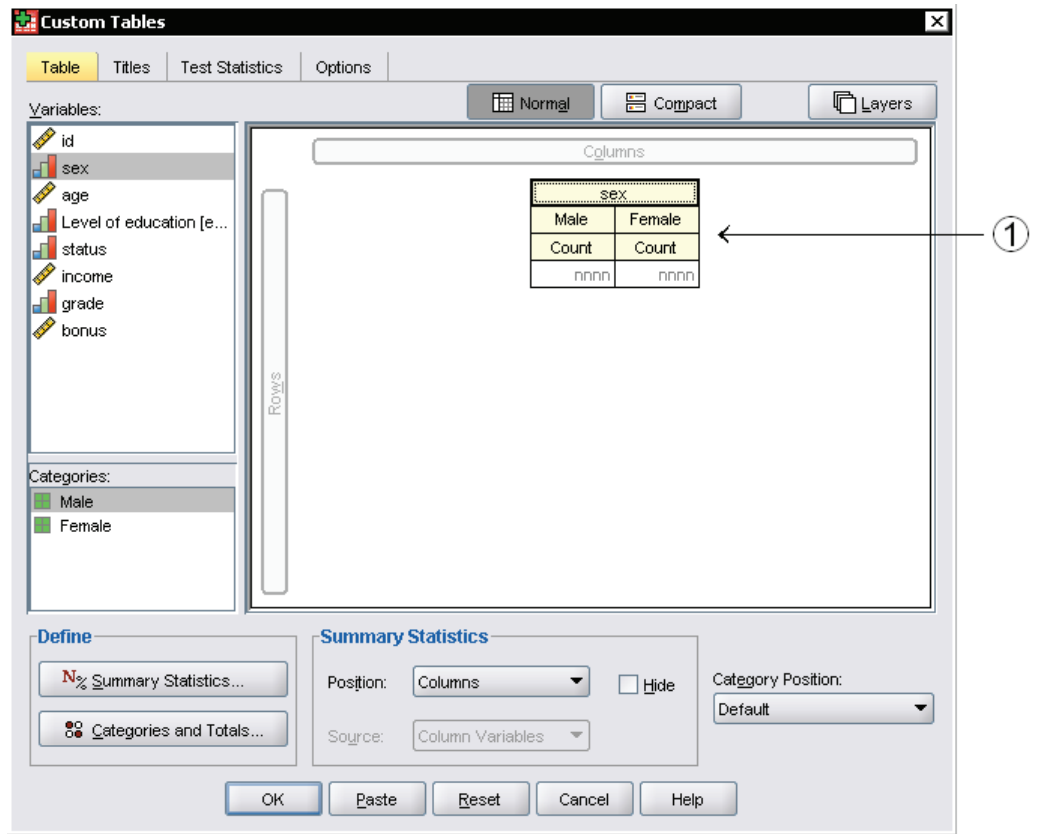
ขั้นที่ 3. เลือกตัวแปร

ที่ต้องการวิเคราะห์

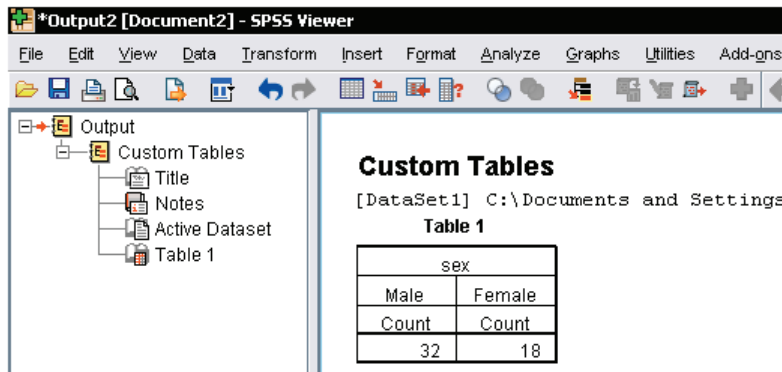
1. คลิกเมาส์ที่ตัวแปร sex
2. คลิกเมาส์ค้างไว้ แล้วลากตัวแปร sex ที่เลือกไว้ไปที่ช่อง columns



จะได้ผลที่เปลี่ยนแปลง
ไปดังนี้



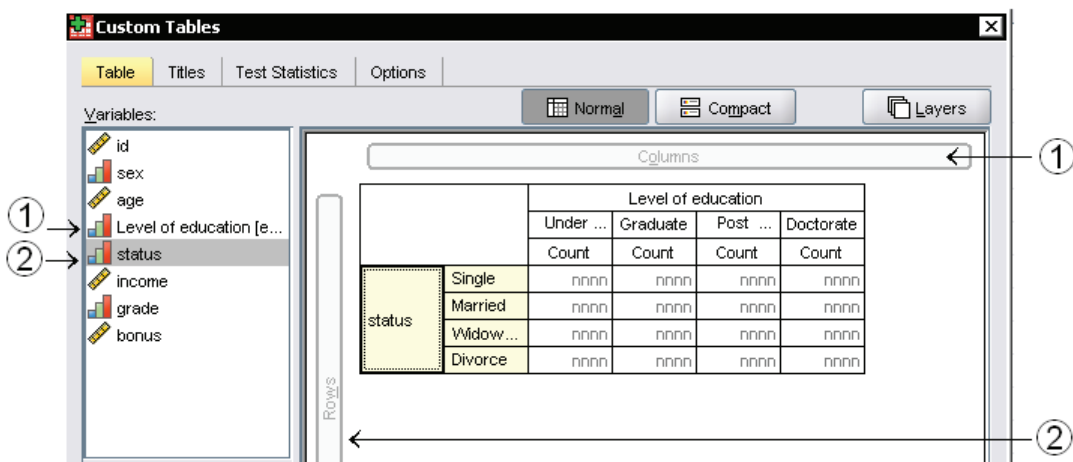
ขั้นที่ 4. คลิก OK
ได้ผลการคำนวณเป็นดังนี้



การใช้งานในรูปแบบต่าง ๆ ของตารางเช่น

ตัวอย่างที่ 4.4.1 การแจกแจงความถี่ของตัวแปร educ และ status ให้เลือกตัวแปรดังนี้

1. คลิกเมาส์ที่ตัวแปร educ คลิกเมาส์ค้างไว้ แล้วลากตัวแปร educ ที่เลือกไว้ไปไว้ที่ช่อง columns
2. คลิกเมาส์ที่ตัวแปร status คลิกเมาส์ค้างไว้ แล้วลากตัวแปร status ที่เลือกไว้ไปไว้ที่ช่อง rows จะได้ผลดังนี้



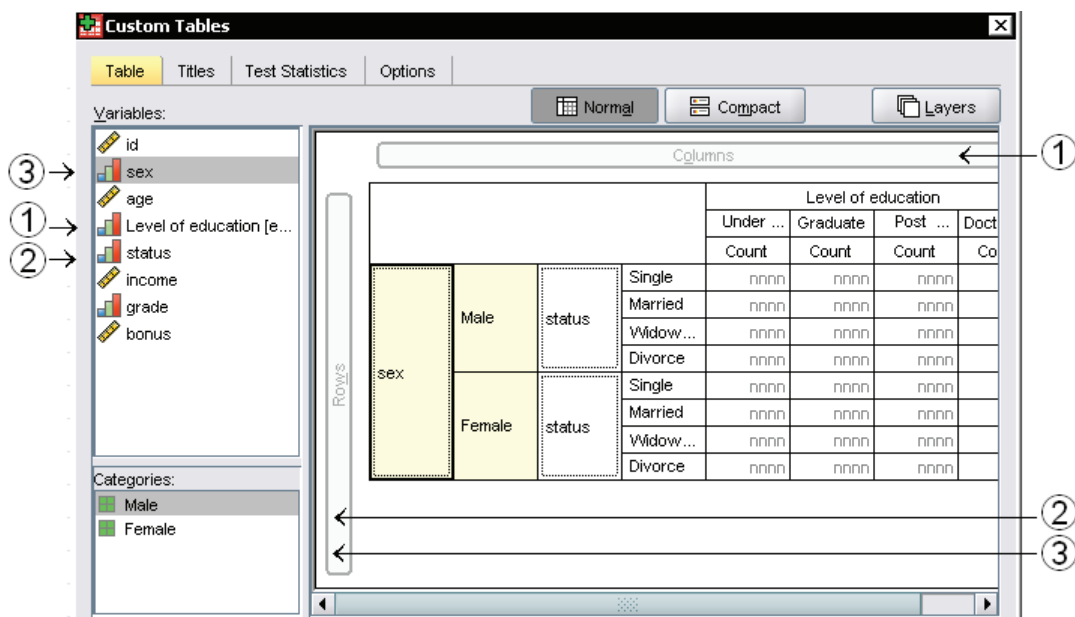
คลิก OK ได้ผลการคำนวณเป็นดังนี้

Table 1

		Level of education			
		Under graduate	Graduate	Post graduate	Doctorate
		Count	Count	Count	Count
status	Single	6	3	4	1
	Married	6	11	5	1
	Widowhood	0	4	0	0
	Divorce	1	3	1	0

ตัวอย่างที่ 4.4.2 การแจกแจงความถี่ของตัวแปร educ, status และ จำแนกกลุ่มตามเพศ ชาย หญิง ให้เลือกตัวแปรดังนี้

1. คลิกเมาส์ที่ตัวแปร educ คลิกเมาส์ค้างไว้ แล้วลากตัวแปร educ ที่เลือกไว้ไปไว้ที่ช่อง columns
2. คลิกเมาส์ที่ตัวแปร status คลิกเมาส์ค้างไว้ แล้วลากตัวแปร status ที่เลือกไว้ไปไว้ที่ช่อง rows
3. คลิกเมาส์ที่ตัวแปร sex คลิกเมาส์ค้างไว้ แล้วลากตัวแปร sex ที่เลือกไว้ไปไว้ที่ช่อง rows จะได้ผลดังนี้



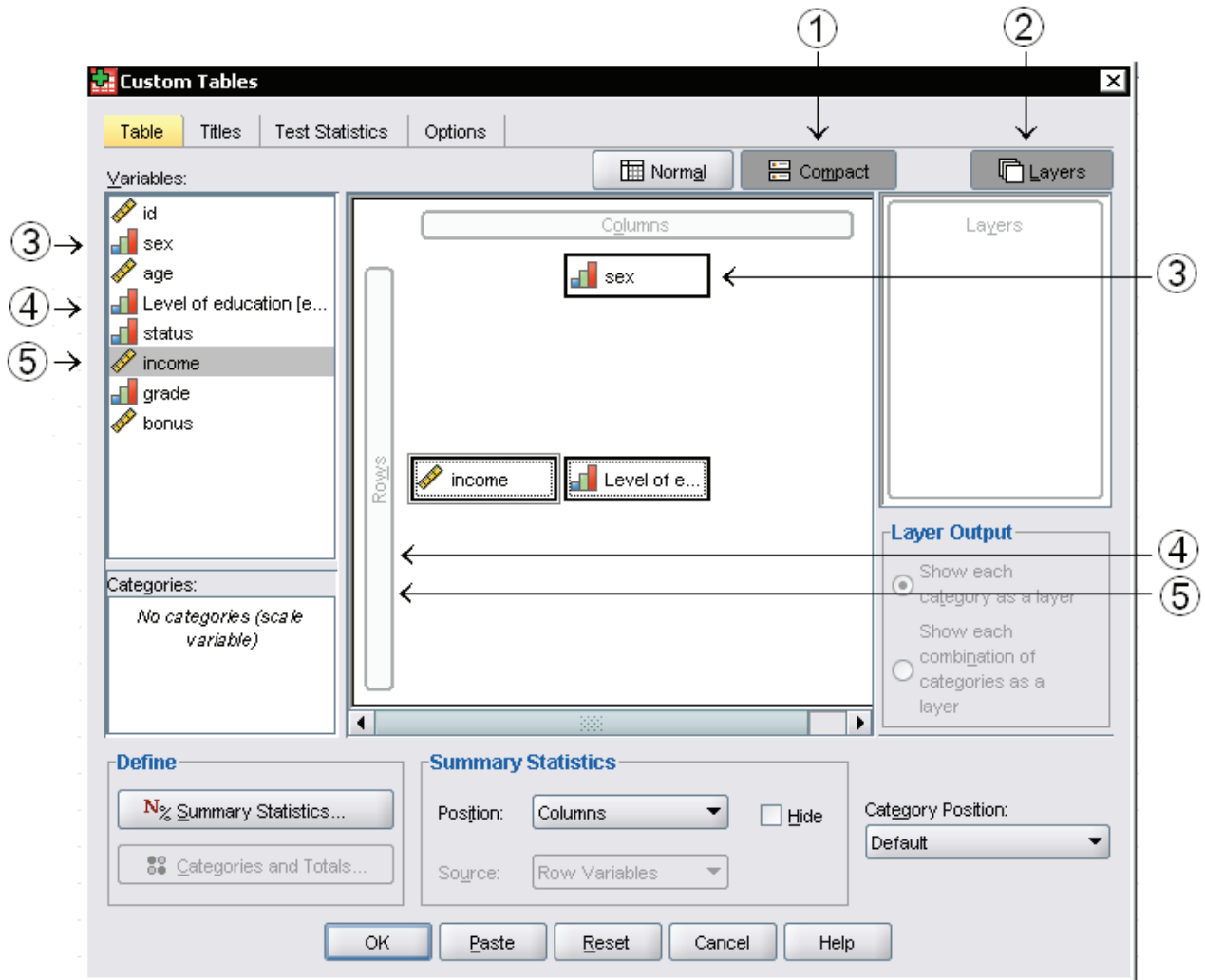
คลิก OK ได้ผลการคำนวณเป็นดังนี้

Table 1

				Level of education			
				Under graduate	Graduate	Post graduate	Doctorate
				Count	Count	Count	Count
sex	Male	status	Single	3	2	2	1
			Married	4	9	3	1
			Widowhood	0	3	0	0
			Divorce	0	3	0	0
	Female	status	Single	3	1	2	0
			Married	2	2	2	0
			Widowhood	0	1	0	0
			Divorce	1	0	1	0

ตัวอย่างที่ 4.4.3 การหารายได้เฉลี่ย (income) จำแนกตามกลุ่ม (ชาย หญิง) และ ระดับการศึกษา

1. เลือก Option Compact
2. เลือก Option Layers
3. คลิกเมาส์ที่ตัวแปร sex คลิกเมาส์ค้างไว้ แล้วลากตัวแปร sex ไปไว้ที่ช่อง columns
4. คลิกเมาส์ที่ตัวแปร educ คลิกเมาส์ค้างไว้ แล้วลากตัวแปร educ ไปไว้ที่ช่อง rows
5. คลิกเมาส์ที่ตัวแปร income คลิกเมาส์ค้างไว้ แล้วลากตัวแปร income ไปไว้ที่ช่อง rows จะได้ผลดังนี้



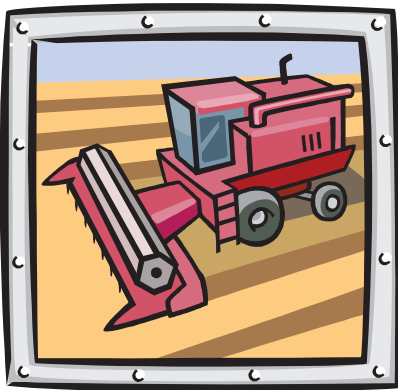
คลิก OK ได้ผลการคำนวณเป็นดังนี้

Table 1

			sex	
			Male	Female
			Mean	Mean
income	Level of education	Under graduate	3767	4186
		Graduate	5635	4525
		Post graduate	6400	5517
		Doctorate	6833	.

บทที่ 5

การปรับปรุงแฟ้มข้อมูลด้วยคำสั่ง Edit, Data และ Transform



การทำงานกับแฟ้มข้อมูลของ SPSS for Windows เราสามารถทำการแก้ไข และ ปรับปรุงแฟ้มข้อมูล เช่น เพิ่มตัวแปร ลดตัวแปร เพิ่มค่าสังเกต ลดค่าสังเกต สร้างตัวแปรใหม่จากตัวแปรเก่า เปลี่ยนแปลงค่าของตัวแปร ฯลฯ การปรับปรุงแฟ้มข้อมูลแบบต่างๆ เหล่านี้เราใช้คำสั่ง Edit, Data และ Transform ตัวอย่างเช่น คำสั่ง Edit มีคำสั่งย่อยของคำสั่ง Edit ที่ใช้งานกันมากคือ

Edit / Insert Variable แทรกตัวแปรใหม่

Edit / Insert Cases แทรกค่าสังเกต

คำสั่ง Data มีคำสั่งย่อยของคำสั่ง Data ที่ใช้งานกันมากคือ

Data / Sort Cases เรียงลำดับข้อมูล

Data / Merge Files รวมแฟ้ม 2 เข้าด้วยกัน แบบ Add Variables หรือ Add Cases

Data / Weight Cases กำหนดตัวแปรน้ำหนักของข้อมูล

คำสั่ง Transform มีคำสั่งย่อยของคำสั่ง Transform ที่ใช้งานกันมากคือ

Transform / Compute Variable นำค่าจากตัวแปรเก่าไปสร้างเป็นค่าของตัวแปรใหม่

Transform / Recode into Same Variables กำหนดค่าใหม่ให้กับตัวแปรเดิมตามเงื่อนไขที่กำหนด

Transform / Recode into Different Variables กำหนดค่าใหม่ให้กับตัวแปรใหม่เดิมตามเงื่อนไขที่กำหนด

เพื่อความสะดวกในการเรียนคำสั่ง Data และ Transform ขอให้สร้างข้อมูลตัวอย่างและ Save ไว้ในดิสก์ (หรือ download จาก <http://www.math.sc.chula.ac.th/~tdumrong/2301286data>)

file_xy1.sav มีตัวแปร x, y และค่าสังเกต 5 ตัว และ file_xy2.sav มีตัวแปร x, y และค่าสังเกต 3 ตัว file_zw.sav มีตัวแปร z, w และค่าสังเกต 5 ตัว โดยมีข้อมูลดังต่อไปนี้

file_xy1.sav

x	y
2.00	15.00
3.00	17.00
7.00	23.00
9.00	45.00
12.00	58.00

file_xy2.sav

x	y
8.00	32.00
9.00	48.00
15.00	67.00

file_zw.sav

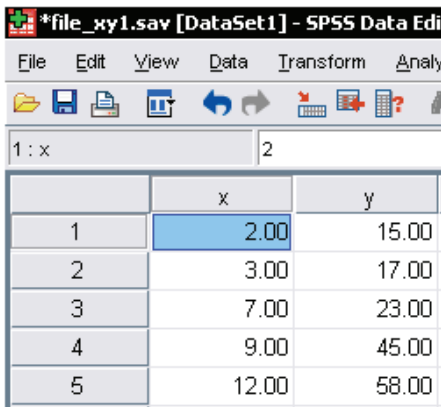
z	w
100.00	17.00
250.00	35.00
370.00	64.00
420.00	72.00
550.00	89.00

5.1 การเพิ่มตัวแปร การลดตัวแปร การแทรกตัวแปร

การทำงานกับเพิ่มข้อมูลงานที่เราจะต้องทำเพิ่มเติมกับเพิ่มข้อมูลคือ การเพิ่มตัวแปร การลดตัวแปร และการแทรกตัวแปร

5.1.1 การแทรกตัวแปร และ การเพิ่มตัวแปร

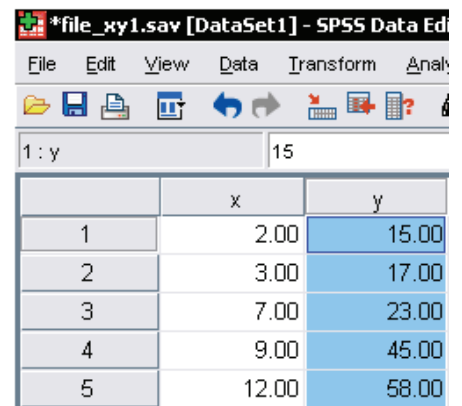
ขั้นที่ 1. เปิดเพิ่มข้อมูล file_xy1.sav เข้ามาใน SPSS Data Editor



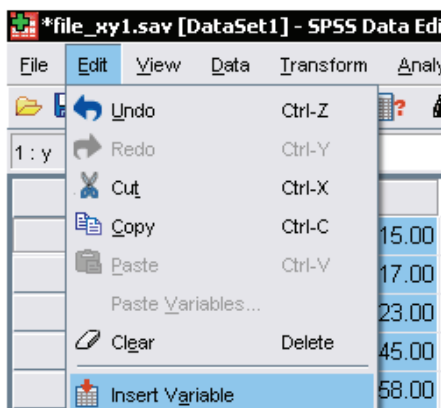
สมมติว่าเราต้องการแทรกตัวแปร t ระหว่างตัวแปร x และ y เพื่อให้เพิ่มข้อมูลใหม่ มีข้อมูลดังนี้

x	t	y
2.00	12.00	15.00
3.00	18.00	17.00
7.00	19.00	23.00
9.00	21.00	45.00
12.00	23.00	58.00

ขั้นที่ 2. ให้เลื่อนเมาส์ไปคลิกที่หัว column ของตัวแปร y



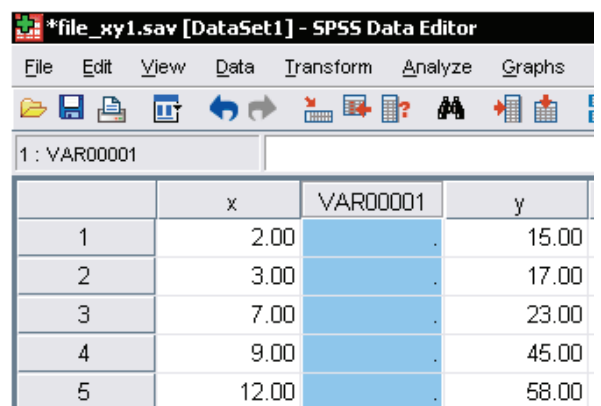
ขั้นที่ 3. เลือกคำสั่ง Edit / Insert Variable



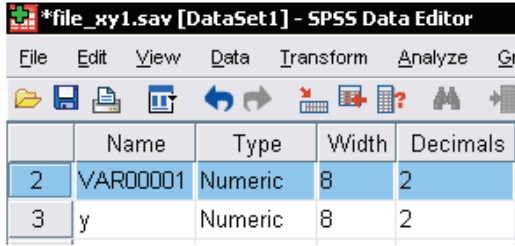
คลิกที่ Insert Variable

บนจอภาพจะแทรกช่องตัวแปร VAR00001

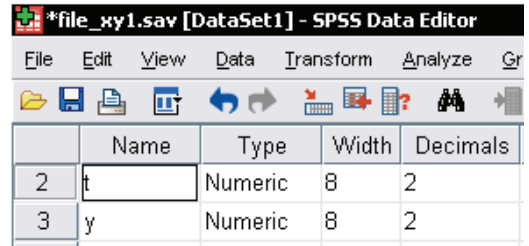
ระหว่างแปร x และ y



ขั้นที่ 4. กดดับเบิลคลิกที่ตัวแปร VAR00001 จอภาพจะเปลี่ยนเป็น Window ของ Variable View

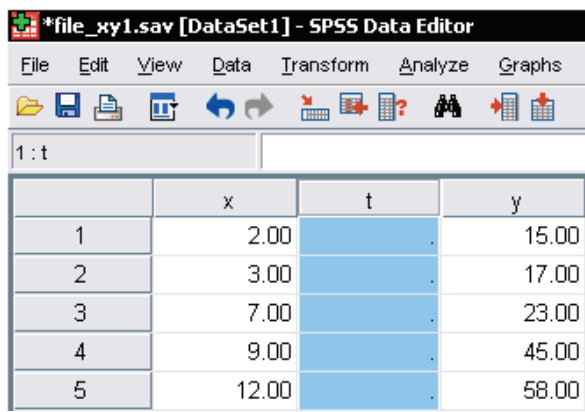
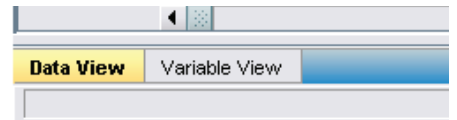


ขั้นที่ 5. พิมพ์ชื่อตัวแปรใหม่ t แทนที่ VAR00001

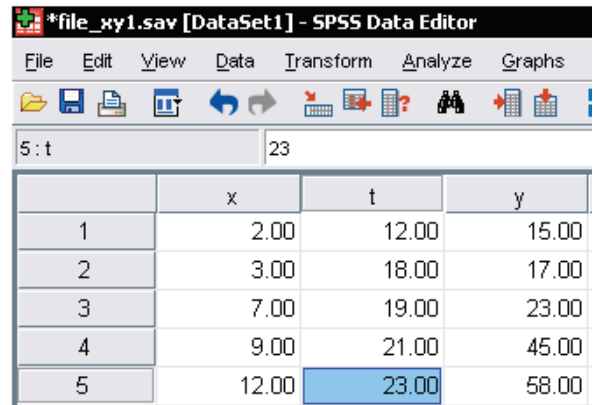


หมายเหตุ ในขั้นตอนนี้เราสามารถเปลี่ยนค่า Type, Width, Decimals และค่าอื่น ๆ ของตัวแปรได้

ขั้นที่ 6. คลิกที่เมนู Data View จะได้ผลบนจอภาพเป็น



พิมพ์ข้อมูลใหม่เข้าไปผลสุดท้ายบนจอภาพคือ



เสร็จแล้วขอให้ Save ไว้ที่ชื่อ file_xty.sav

5.1.2 การลดตัวแปร

สมมติว่าเราต้องการลดตัวแปร y ออกจากแฟ้ม file_xty.sav และ save ใหม่เป็นแฟ้มชื่อ file_xt.sav

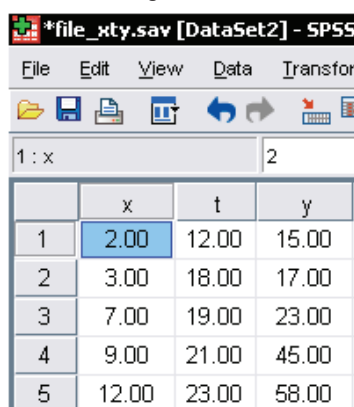
file_xty.sav

x	t	y
2.00	12.00	15.00
3.00	18.00	17.00
7.00	19.00	23.00
9.00	21.00	45.00
12.00	23.00	58.00

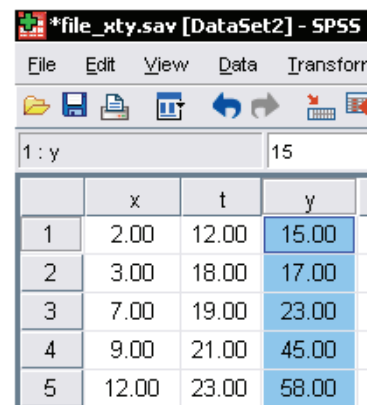
file_xt.sav

x	t
2.00	12.00
3.00	18.00
7.00	19.00
9.00	21.00
12.00	23.00

ขั้นที่ 1. เปิดแฟ้มข้อมูล file_xty.sav เข้ามาใน SPSS Data Editor



ขั้นที่ 2. คลิกที่ช่องตัวแปร y



กดปุ่ม Delete จะเห็นว่าตัวแปร y หายไป เสร็จแล้วขอให้ Save ไว้ที่ชื่อ file_xt.sav

	x	t
1	2.00	12.00
2	3.00	18.00
3	7.00	19.00
4	9.00	21.00
5	12.00	23.00

	x	y
1	2.00	15.00
2	3.00	17.00
3	7.00	23.00
4	9.00	45.00
5	12.00	58.00

5.1.3 การแทรกคำสั่ง

ขั้นที่ 1. เปิดเพิ่มข้อมูล file_xy1.sav

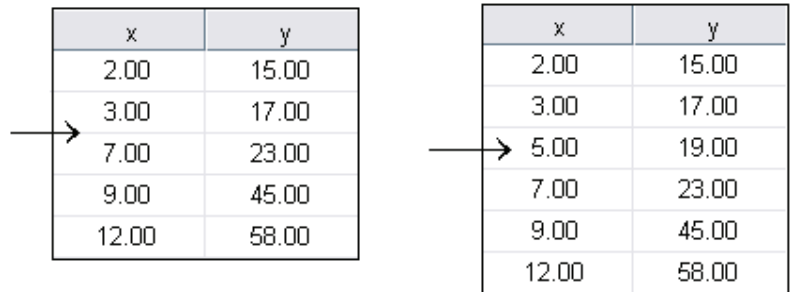
เข้ามาใน SPSS Data Editor

สมมติว่าเราต้องการแทรกคำสั่ง

$x = 5, y = 19$

ระหว่างคำสั่งแถวที่ 2 และ 3

ขั้นที่ 2. คลิกที่ตำแหน่งแถวที่ 3



	x	y
1	2.00	15.00
2	3.00	17.00
3	7.00	23.00
4	9.00	45.00
5	12.00	58.00

ขั้นที่ 3. เลือกคำสั่ง Edit / Insert Cases

File	Edit	View	Data	Transform	Analy
Undo	Ctrl-Z				
Redo	Ctrl-Y				
Cut	Ctrl-X				
Copy	Ctrl-C				
Paste	Ctrl-V				
Paste Variables...					
Clear	Delete				
Insert Variable					
Insert Cases					

ขั้นที่ 4. คลิกที่ Insert Cases จอภาพจะเป็น

	x	y
1	2.00	15.00
2	3.00	17.00
3		
4	7.00	23.00
5	9.00	45.00
6	12.00	58.00

ขั้นที่ 5. พิมพ์ค่า $x = 5$ และ $y = 19$

	x	y
1	2.00	15.00
2	3.00	17.00
3	5.00	19.00
4	7.00	23.00
5	9.00	45.00
6	12.00	58.00

เสร็จแล้ว Save ไว้ที่ชื่อ file_xy1 6obs.sav

5.2 การลบคำสั่งเกิด

การลบคำสั่งเกิดออกจากแฟ้มข้อมูล

ขั้นที่ 1. นำเมาส์ไปคลิกที่หมายเลขบรรทัดของคำสั่งเกิด

เช่นขณะนี้เราต้องการลบคำสั่งเกิดตัวที่ 5 ทิ้งไป

ขั้นที่ 2. คลิกเมาส์ตรงบรรทัดที่ 5

ขั้นที่ 3. กดปุ่ม Delete
จะได้ผลดังนี้

	x	y
1	2.00	15.00
2	3.00	17.00
3	5.00	19.00
4	7.00	23.00
5	12.00	58.00

	x	y
1	2.00	15.00
2	3.00	17.00
3	5.00	19.00
4	7.00	23.00
5	9.00	45.00
6	12.00	58.00

	x	y
1	2.00	15.00
2	3.00	17.00
3	5.00	19.00
4	7.00	23.00
5	9.00	45.00
6	12.00	58.00

คำสั่งเกิดที่ 5 ของเก่าหายไป
และ เลื่อนคำสั่งเกิดตัวที่ 6 ขึ้นมาเป็นคำสั่งเกิดตัวที่ 5

5.3 การรวมแฟ้มข้อมูลแบบเพิ่มตัวแปร

แฟ้มข้อมูล 2 แฟ้มที่มีตัวแปรต่างกันสามารถนำมารวมเป็นแฟ้มเดียวกันได้

โดยใช้คำสั่ง Data / Merge files / Add Variables ตัวอย่างเช่น

file_xy1.sav

x	y
2.00	15.00
3.00	17.00
7.00	23.00
9.00	45.00
12.00	58.00

file_zw.sav

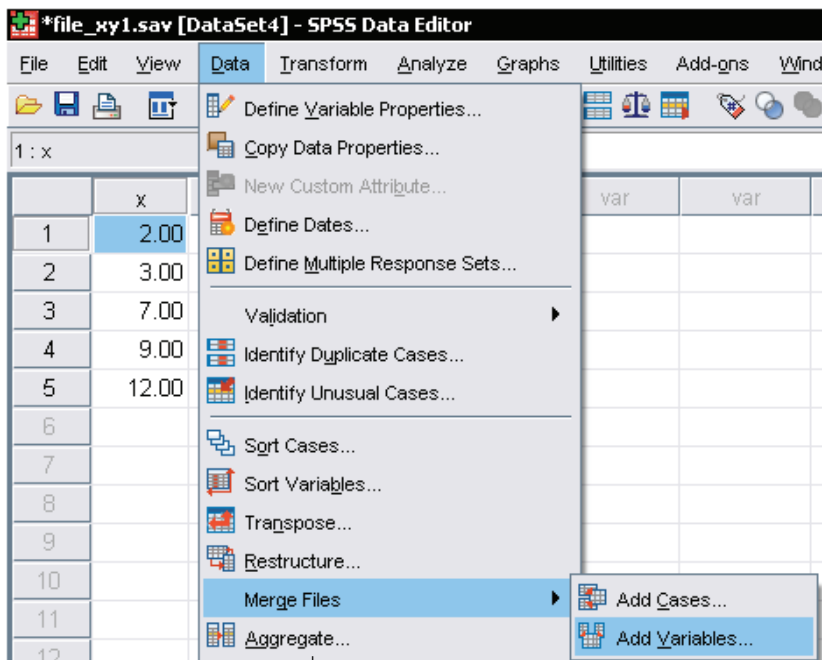
z	w
100.00	17.00
250.00	35.00
370.00	64.00
420.00	72.00
550.00	89.00

file_xyzw.sav

x	y	z	w
2.00	15.00	100.00	17.00
3.00	17.00	250.00	35.00
7.00	23.00	370.00	64.00
9.00	45.00	420.00	72.00
12.00	58.00	550.00	89.00

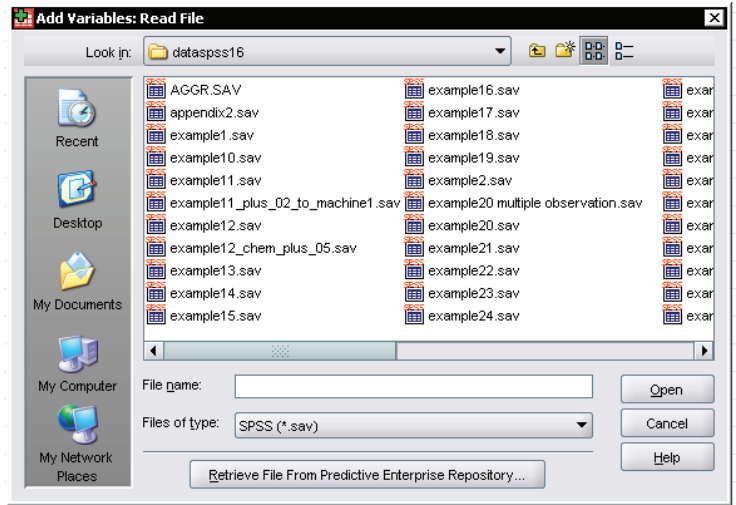
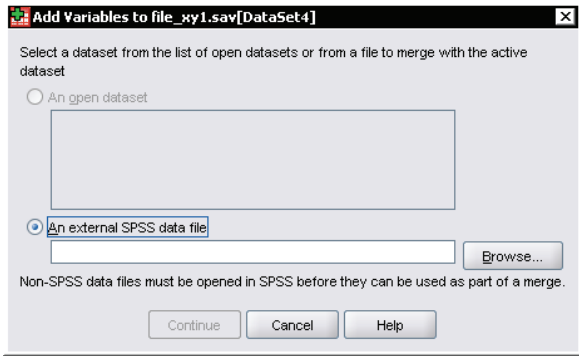
ขั้นที่ 1. เปิดแฟ้มข้อมูล file_xy1.sav เข้ามาใน SPSS Data Editor

ขั้นที่ 2. เลือกคำสั่ง Data / Merge Files / Add Variables



	x	y
1	2.00	15.00
2	3.00	17.00
3	7.00	23.00
4	9.00	45.00
5	12.00	58.00

ขั้นที่ 3. คลิกที่ Add Variables จะได้เมนูย่อย Add Variables to file_xy1.sav
 ต่อไปคลิกที่ Browse จะได้เมนู Add Variables : Read File



หมายเหตุ

ในดิสก์ต้องมีแฟ้ม file_zw.sav อยู่ก่อน

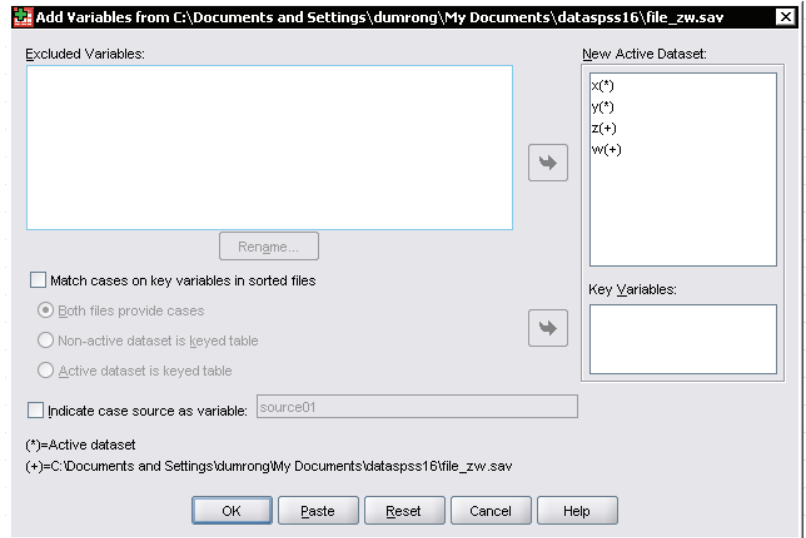
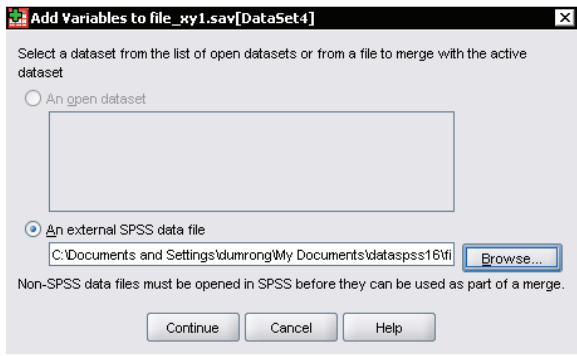
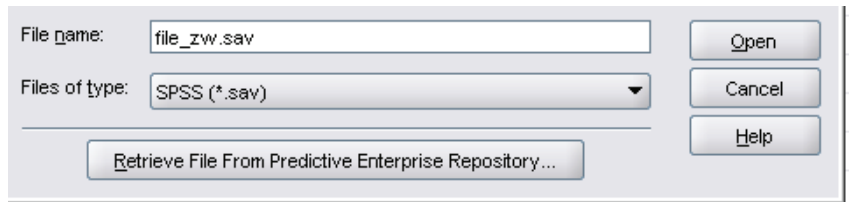
ขั้นที่ 4. พิมพ์ชื่อแฟ้ม file_zw.sav

แล้วคลิก Open

จะได้เมนูย่อย

Add Variables to file_xy1.sav

ต่อไปให้คลิก Continue จะได้



ข้อสังเกต คำอธิบายของ SPSS บอกว่า
 แฟ้มใหม่ที่ได้ประกอบด้วยตัวแปร 4 ตัว
 คือ x, y, z, w

ขั้นที่ 5. คลิก OK จะได้แฟ้มใหม่เป็นดังนี้

หมายเหตุ 1. ข้อมูลเดิมมีตัวแปร 2 ตัวคือ x และ y

ขณะนี้ได้มีตัวแปรเพิ่มมาอีก 2 ตัวคือ z และ w

2. ชื่อแฟ้มข้อมูลยังเป็นชื่อเดิมคือ file_xy1.sav

3. ถ้าต้องการเปลี่ยนชื่อแฟ้มข้อมูลใหม่
 ควรใช้ Save as โดยใช้ชื่อใหม่ว่า file_xyzw.sav

↓

	x	y	z	w
1	2.00	15.00	100.00	17.00
2	3.00	17.00	250.00	35.00
3	7.00	23.00	370.00	64.00
4	9.00	45.00	420.00	72.00
5	12.00	58.00	550.00	89.00

5.4 การรวมแฟ้มข้อมูลแบบเพิ่มคำสั่งเกิด

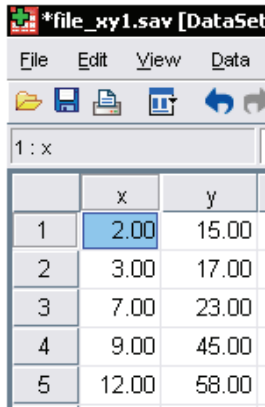
แฟ้มข้อมูล 2 แฟ้มที่มีโครงสร้างตัวแปรเหมือนกันเราสามารถรวมแฟ้มเข้าด้วยกันเพื่อให้จำนวนคำสั่งเกิดเพิ่มขึ้นได้โดยใช้คำสั่ง Data / Merge Files / Add Cases ตัวอย่างเช่น

file_xy1.sav	x	y
1	2.00	15.00
2	3.00	17.00
3	7.00	23.00
4	9.00	45.00
5	12.00	58.00

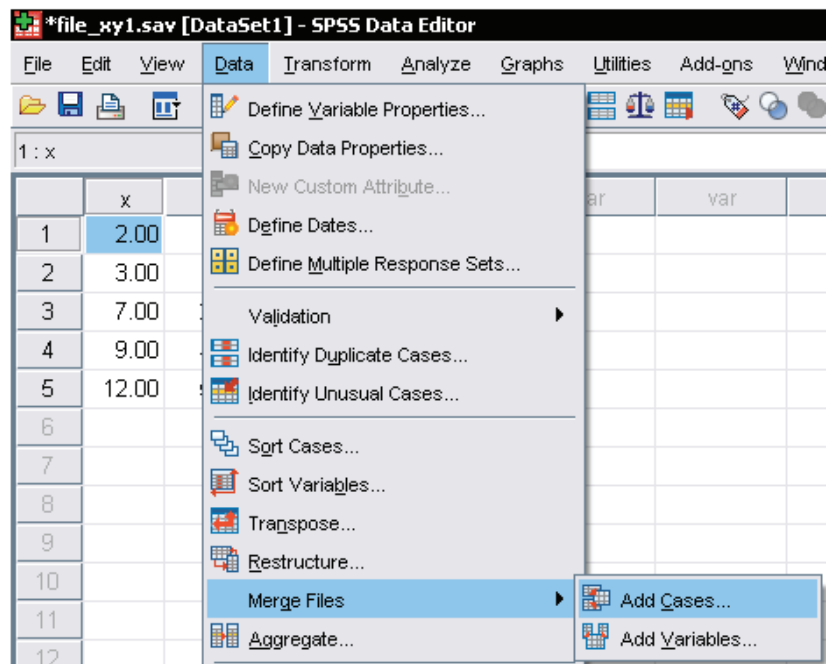
file_xy2.sav	x	y
1	8.00	32.00
2	9.00	48.00
3	15.00	67.00

	x	y
1	2.00	15.00
2	3.00	17.00
3	7.00	23.00
4	9.00	45.00
5	12.00	58.00
6	8.00	32.00
7	9.00	48.00
8	15.00	67.00

ขั้นที่ 1. เปิดแฟ้มข้อมูล file_xy1.sav เข้ามาใน SPSS Data Editor



ขั้นที่ 2. เลือกคำสั่ง Data / Merge Files / Add Cases

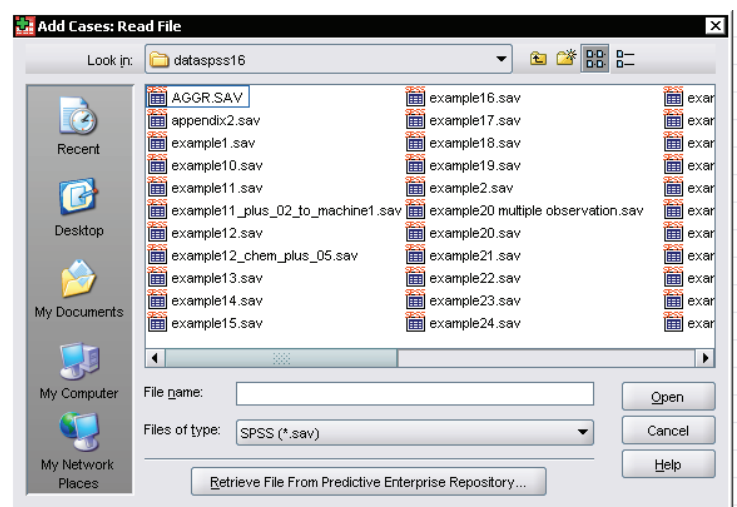
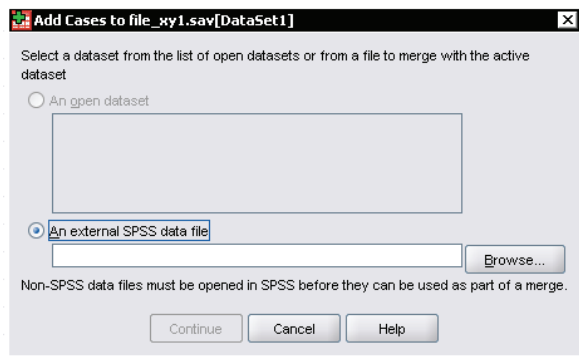


คลิกที่ Add Cases จะได้เมนูย่อย

Add Cases to file_xy1.sav

ต่อไปให้คลิก Browse

จะได้เมนูย่อย Add Cases : Read File



หมายเหตุ ในดิสก์ต้องมีแฟ้ม

file_xy2.sav อยู่ก่อน

ขั้นที่ 3. พิมพ์ชื่อแฟ้ม

file_xy2.sav แล้วคลิก Open

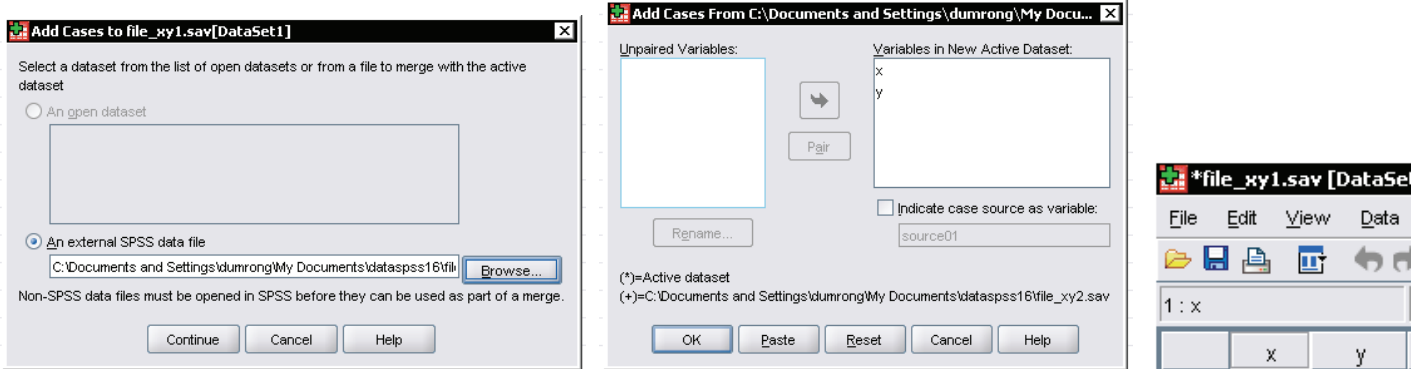
จะกลับเมนูย่อย

Add Cases to file_xy1.sav



ต่อไปให้คลิก Continue จะได้เมนูย่อย Add Cases

ข้อสังเกต คำอธิบายของ SPSS บอกว่าแฟ้มใหม่ที่จะได้ประกอบด้วยตัวแปร 2 ตัว คือ x, y



ขั้นที่ 4. คลิก OK จะได้แฟ้มใหม่เป็น

หมายเหตุ 1. ข้อมูลเดิมมีตัวแปร 2 ตัวคือ x และ y

ข้อมูลใหม่ยังมี 2 ตัวแปรเดิม แต่มี 8 คำสังเกต

2. ชื่อแฟ้มข้อมูลยังเป็นชื่อเดิมคือ file_xy1.sav

3. ควรเปลี่ยนชื่อแฟ้มข้อมูลใหม่โดยใช้ Save as และใช้ชื่อใหม่ว่า file_xy 8obs.sav

5.5 การเรียงลำดับข้อมูล

คำสั่ง Data / Sort Cases เป็นคำสั่งที่ช่วยในการเรียงลำดับข้อมูล

เราต้องการเรียงลำดับข้อมูลในตัวแปร x จากมากไปน้อย

ขั้นที่ 1. เปิดแฟ้มข้อมูล file_xy1.sav

ขั้นที่ 2. เลือกคำสั่ง Data / Sort Cases

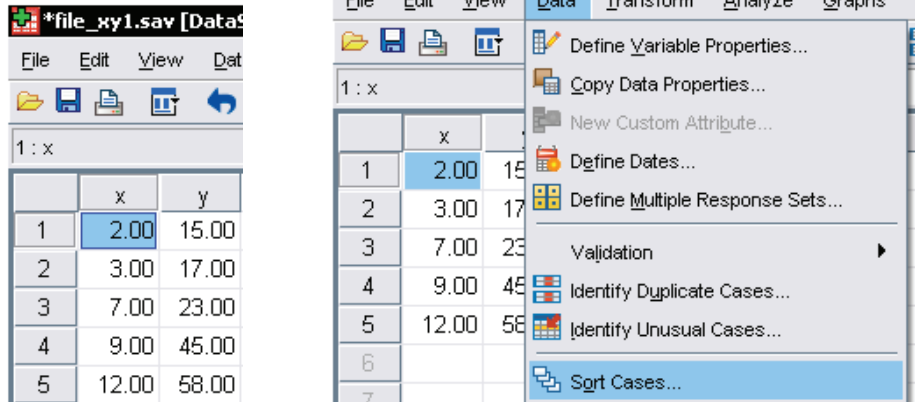
เข้ามาใน SPSS Data Editor

file_xy1.sav

x	y
2.00	15.00
3.00	17.00
7.00	23.00
9.00	45.00
12.00	58.00

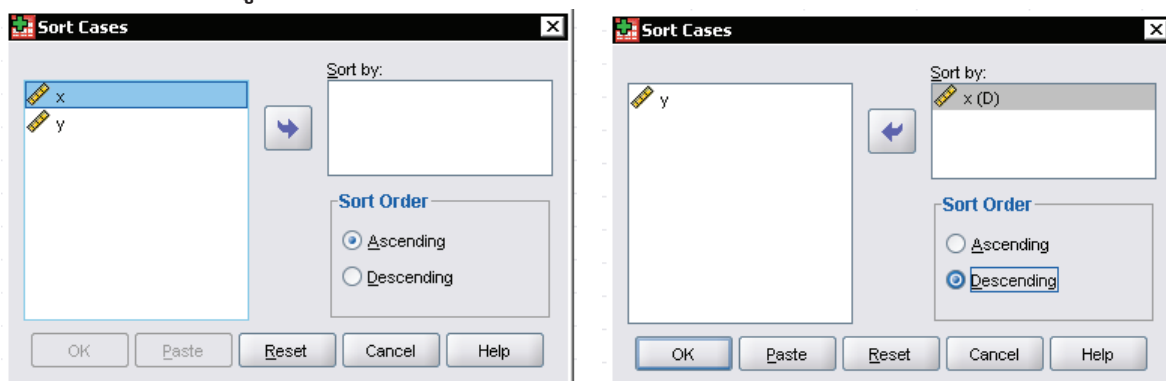
→

x	y
12.00	58.00
9.00	45.00
7.00	23.00
3.00	17.00
2.00	15.00



ขั้นที่ 3. คลิกที่ Sort Cases
จะได้เมนูย่อยเป็น

ขั้นที่ 4. เลือกตัวแปร x มาไว้ช่อง Sort by
และเลือก Sort Order เป็น Descending



หมายเหตุ Descending เรียงจากมากไปน้อย และ Ascending เรียงจากน้อยไปมาก

ขั้นที่ 5. คลิก OK จะได้ผลดังนี้

หมายเหตุ 1. ชื่อแฟ้มข้อมูลยังเป็นชื่อเดิมคือ file_xy1.sav

2. การเรียงลำดับของตัวแปร x มีผลทำให้ตัวแปร y มีการเรียงลำดับตามไปด้วย

	x	y
1	12.00	58.00
2	9.00	45.00
3	7.00	23.00
4	3.00	17.00
5	2.00	15.00

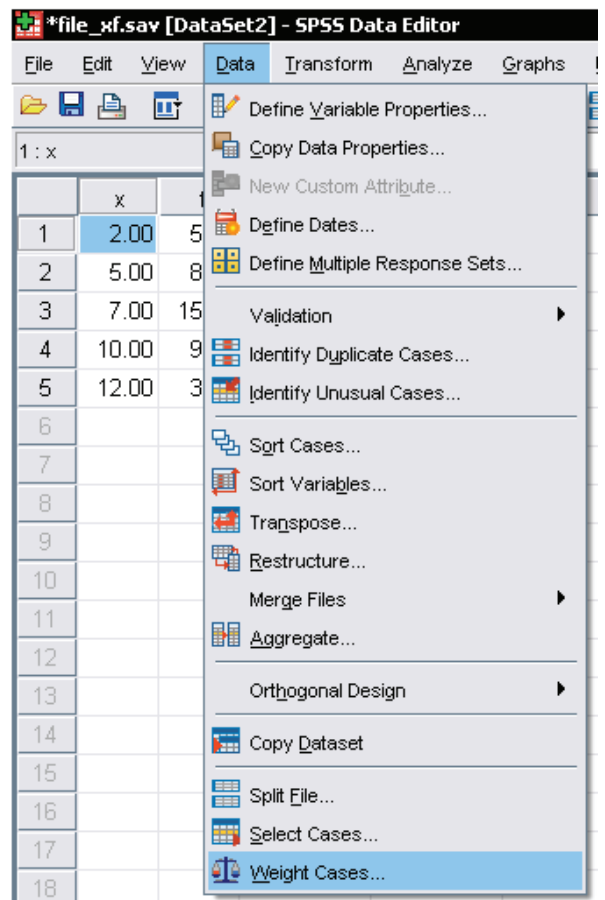
5.6 การกำหนดตัวแปรน้ำหนัก

ข้อมูลในรูปแบบของคะแนนและความถี่ ตัวอย่างเช่น การหาค่าเฉลี่ยของคะแนน (x) ที่มีความถี่ตามที่กำหนด ต้องกำหนดให้ตัวแปรความถี่ (f) เป็นค่าน้ำหนัก

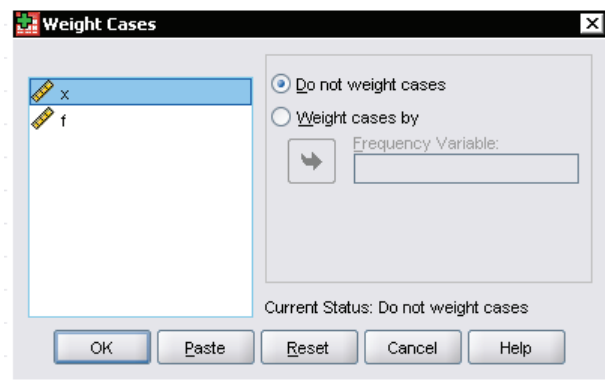
การกำหนดค่าตัวแปร f เป็นค่าน้ำหนักเราใช้คำสั่ง Data / Weight Cases

ขั้นที่ 1. สร้างแฟ้มข้อมูล file_xf.sav ขั้นที่ 2. เลือกคำสั่ง Data / Weight Cases

	x	f
1	2.00	5.00
2	5.00	8.00
3	7.00	15.00
4	10.00	9.00
5	12.00	3.00

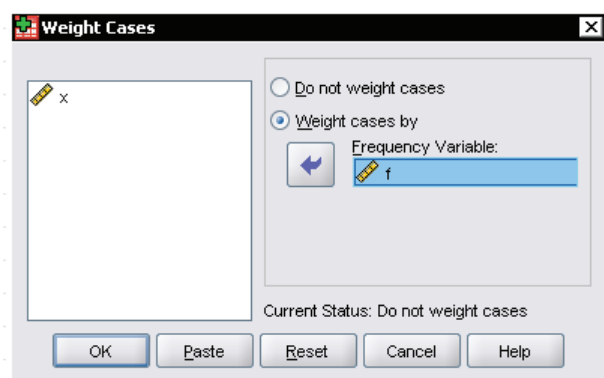


ขั้นที่ 3. คลิกที่ Weight Cases จะได้เมนูย่อยเป็น



ขั้นที่ 4. คลิกที่ Weight cases by

เสร็จแล้วเลือกตัวแปร f มาไว้ที่ช่อง frequency Variable

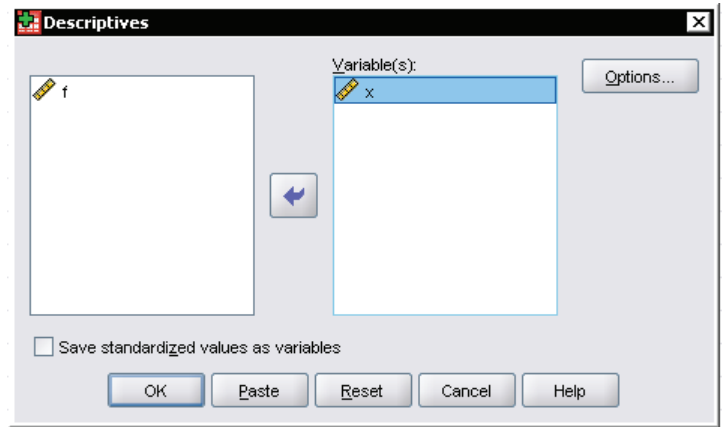
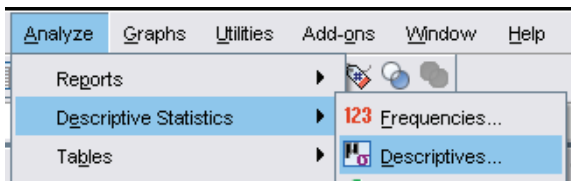


ขั้นที่ 5. คลิก OK จะได้ผลเป็นดังนี้

	x	f
1	2.00	5.00
2	5.00	8.00
3	7.00	15.00
4	10.00	9.00
5	12.00	3.00

ขณะนี้เรากำหนดตัวแปรน้ำหนักเสร็จแล้วต่อไปลองคำนวณค่าเฉลี่ยของตัวแปร x โดยใช้คำสั่ง

Analyze / Descriptive Statistics / Descriptives



เลือกตัวแปร x มาที่ช่อง Variable(s)
แล้วคลิก OK จะได้ผลการคำนวณเป็นดังนี้

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
x	40	2.00	12.00	7.0250	2.83284
Valid N (listwise)	40				

หมายเหตุ ค่าเฉลี่ยมาจากสูตร $\frac{(2)(5) + (5)(8) + (7)(15) + (10)(9) + (12)(3)}{5 + 8 + 15 + 9 + 3} = \frac{281}{40} = 7.0250$

ในกรณีที่ไมกำหนดค่าน้ำหนักให้กับตัวแปร x จะได้ผลลัพธ์ของ Descriptive Statistics เป็นดังนี้

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
x	5	2.00	12.00	7.2000	3.96232
Valid N (listwise)	5				

5.7 การนำค่าจากตัวแปรเก่าไปสร้างเป็นค่าของตัวแปรใหม่

คำสั่งที่ใช้ในการนำค่าจากตัวแปรเก่าไปสร้างเป็นตัวแปรใหม่คือ คำสั่ง Transform / Compute Variable

ตัวอย่างเช่น ในแฟ้ม file_xy1.sav มีตัวแปร x, y

เราต้องการสร้างตัวแปรใหม่เพิ่มอีกตัวคือ xplusy ที่มีสูตรเป็น $x + y$

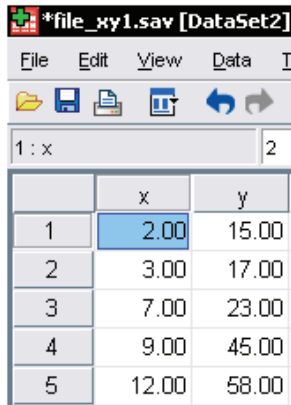
file_xy1.sav

x	y
2.00	15.00
3.00	17.00
7.00	23.00
9.00	45.00
12.00	58.00

เพิ่มตัวแปร xplusy เป็น

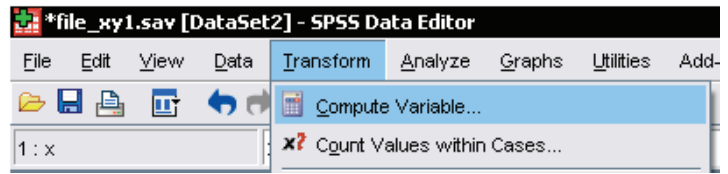
x	y	xplusy
2.00	15.00	17.00
3.00	17.00	20.00
7.00	23.00	30.00
9.00	45.00	54.00
12.00	58.00	70.00

ขั้นที่ 1. เปิดแฟ้มข้อมูล file_xy1.sav
เข้ามาใน SPSS Data Editor

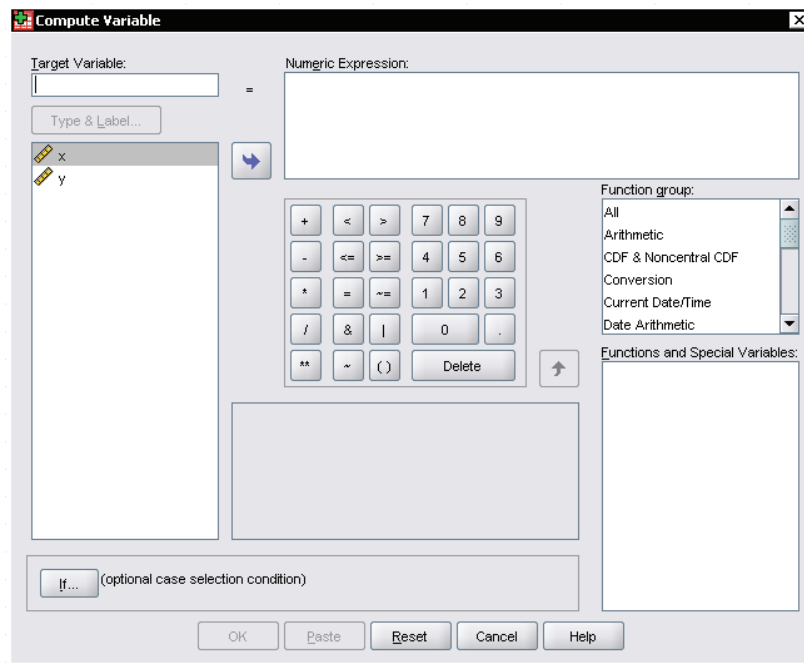


	x	y
1	2.00	15.00
2	3.00	17.00
3	7.00	23.00
4	9.00	45.00
5	12.00	58.00

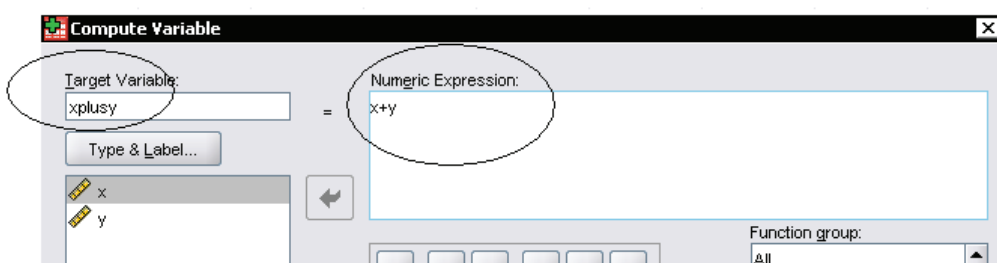
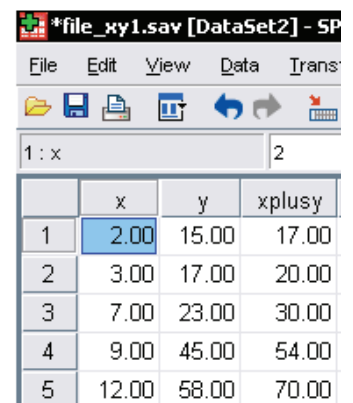
ขั้นที่ 2. เลือกคำสั่ง Transform / Compute Variable



ขั้นที่ 3. คลิกที่คำสั่ง Compute จะได้เมนูย่อยเป็น



ขั้นที่ 4. พิมพ์ชื่อตัวแปรใหม่ xplusy ที่ช่อง Target Variable พิมพ์สูตร $x + y$ ที่ช่อง Numeric Expression

	x	y	xplusy
1	2.00	15.00	17.00
2	3.00	17.00	20.00
3	7.00	23.00	30.00
4	9.00	45.00	54.00
5	12.00	58.00	70.00

ขั้นที่ 5. คลิก OK จะได้ผลดังนี้

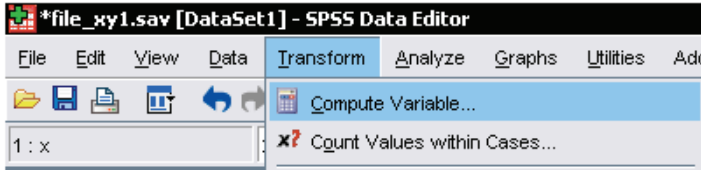
ขอให้ Save as ข้อมูลใหม่เป็นชื่อ file_xplusy.sav

การสร้างตัวแปรใหม่โดยใช้ค่าจากตัวแปรเก่าและเลือกใช้ฟังก์ชันสำเร็จรูปของ SPSS ในการแปลงค่า

ตัวอย่างเช่นเปลี่ยน x เป็น $\ln x$

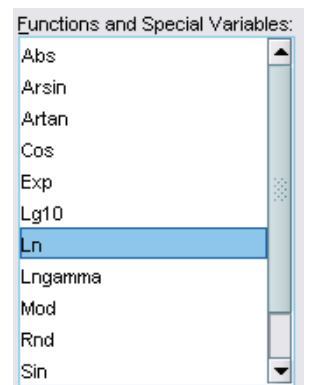
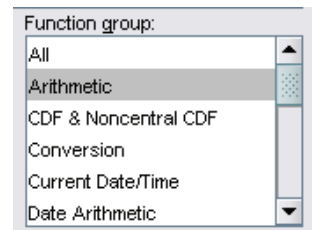
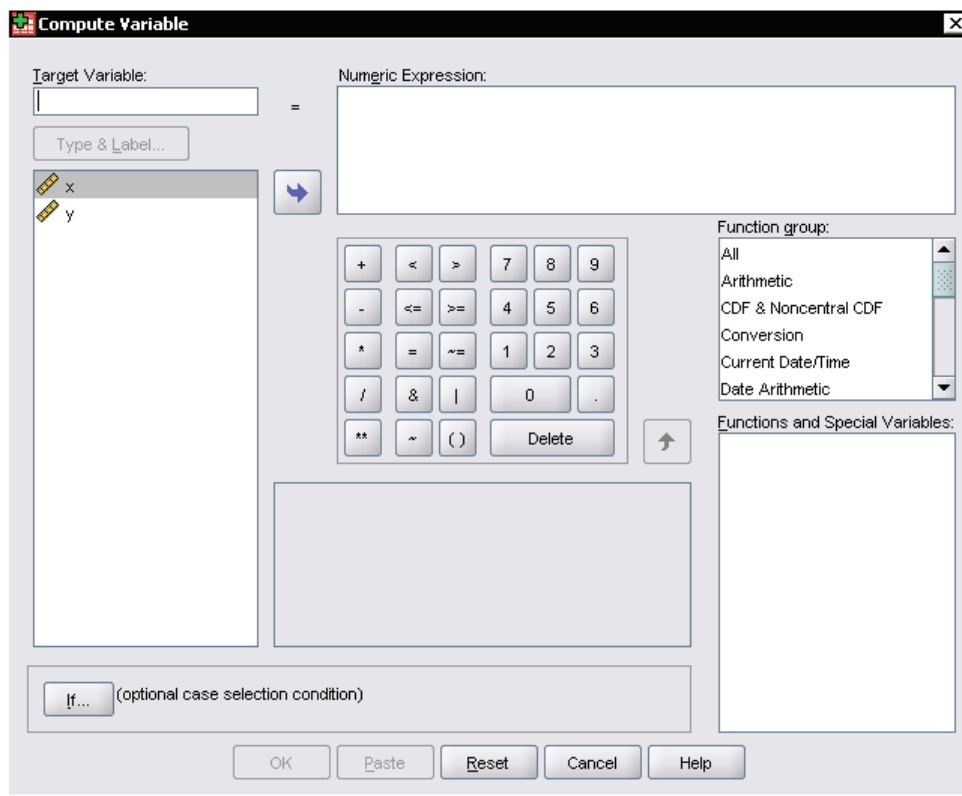
ขั้นที่ 1. เปิดแฟ้มข้อมูล file_xy1.sav เข้ามาใน SPSS Data Editor

ขั้นที่ 2. เลือกคำสั่ง Transform / Compute Variable

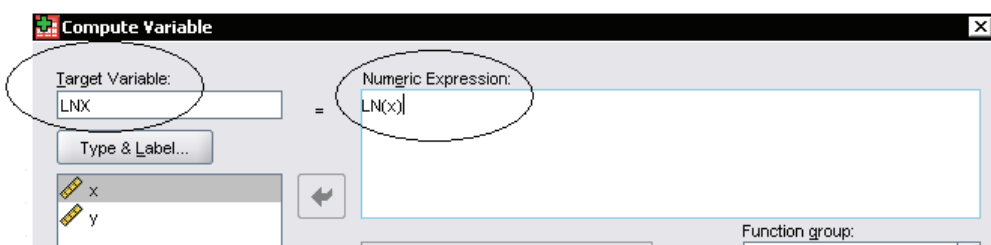


	x	y
1	2.00	15.00
2	3.00	17.00
3	7.00	23.00
4	9.00	45.00
5	12.00	58.00

ขั้นที่ 3. คลิกที่คำสั่ง Compute จะได้เมนูย่อยเป็น



ขั้นที่ 4. พิมพ์ชื่อตัวแปรใหม่ LN_X ที่ช่อง Target Variable พิมพ์สูตร LN(x) ที่ช่อง Numeric Expression



หมายเหตุ การเลือกฟังก์ชันที่มีใน SPSS ดูได้จากแถบเมนูย่อย

Function group และ Function and Spacial Variables

ขั้นที่ 5. คลิก OK จะได้ผลดังนี้

	x	y	LN _X
1	2.00	15.00	0.69
2	3.00	17.00	1.10
3	7.00	23.00	1.95
4	9.00	45.00	2.20
5	12.00	58.00	2.48

5.8 การปรับเปลี่ยนค่าของตัวแปรด้วยคำสั่ง Transform

คำสั่งที่ใช้ในการเปลี่ยนค่าเก่าของตัวแปรไปเป็นค่าใหม่ คือ Transform / Recode into Same Variables ตัวอย่างเช่น ในแฟ้ม file_xy1.sav มีตัวแปร x, y

เราต้องการ เปลี่ยนค่า x เป็น 1 ถ้า $x < 5$ และ เปลี่ยนค่า x เป็น 2 ถ้า $x \geq 5$

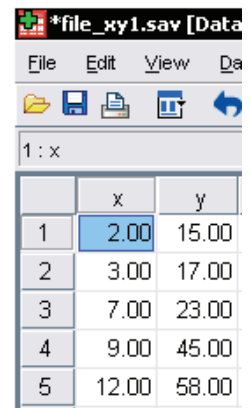
file_xy1.sav

x	y
2.00	15.00
3.00	17.00
7.00	23.00
9.00	45.00
12.00	58.00

เปลี่ยนค่า x แล้วจะได้เป็น

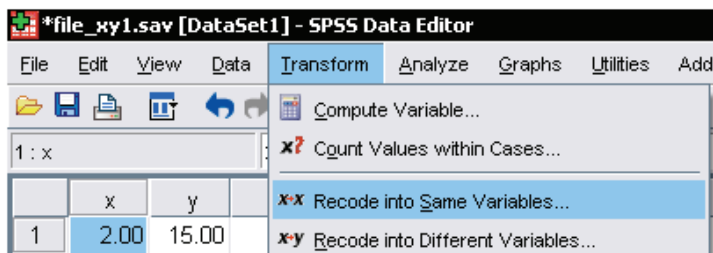
file_xy1_recode.sav

x	y
1.00	15.00
1.00	17.00
2.00	23.00
2.00	45.00
2.00	58.00

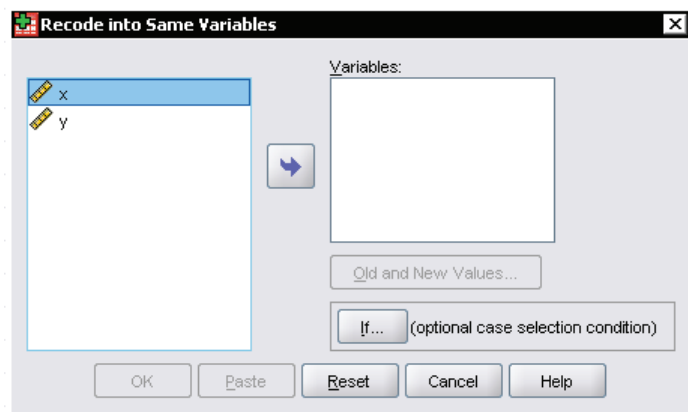


ขั้นที่ 1. เปิดแฟ้มข้อมูล file_xy1.sav เข้ามาใน SPSS Data Editor

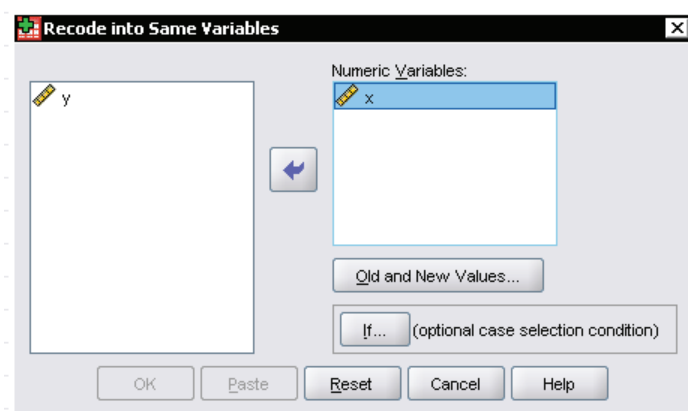
ขั้นที่ 2. เลือกคำสั่ง Transform / Recode into Same Variables



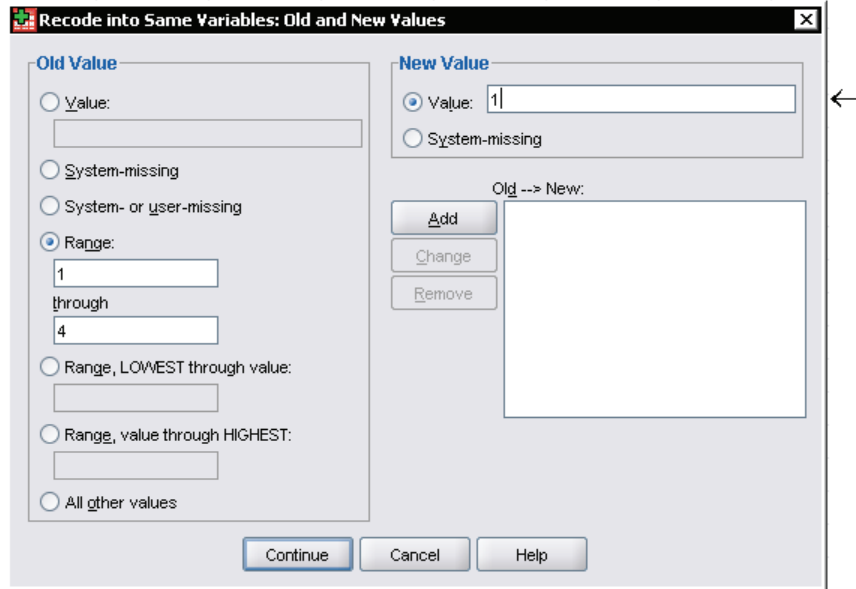
ขั้นที่ 3. คลิกที่คำสั่ง Recode into Same Variables จะได้เมนูย่อยเป็น



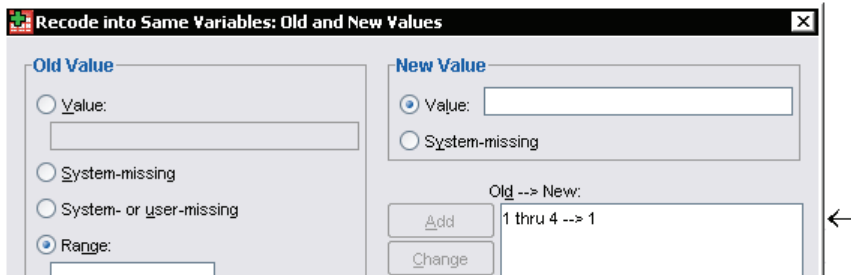
ขั้นที่ 4. เลือกตัวแปร x มาไว้ที่ช่อง Numeric Variables



ขั้นที่ 7. ในช่อง New Value ให้พิมพ์ค่าเป็น 1 จะสังเกตเห็นว่าปุ่ม Add จะมีสีเข้มขึ้นมา



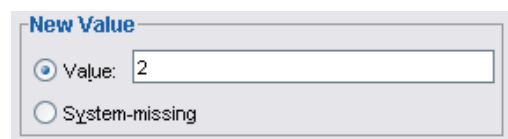
ขั้นที่ 8. คลิกที่ปุ่ม Add จะได้ผลดังนี้



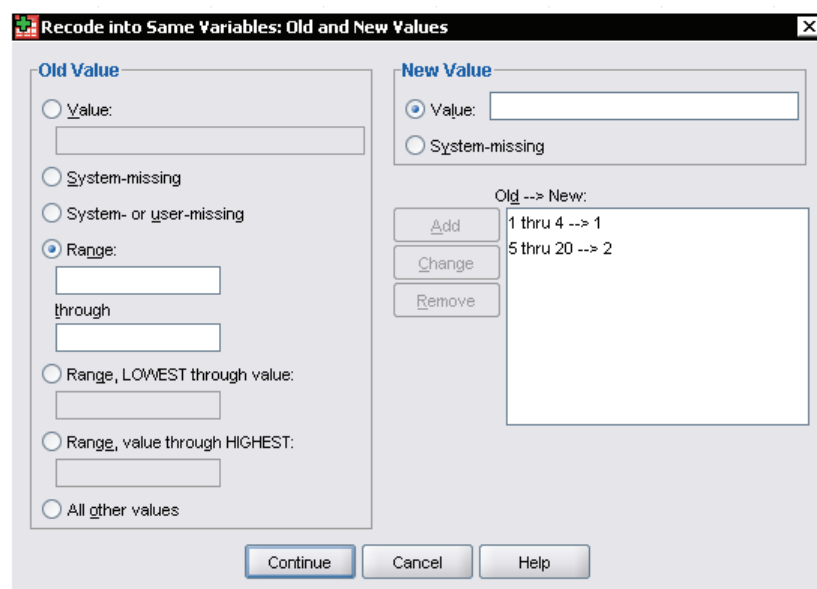
ในทำนองเดียวกัน ให้พิมพ์ค่า 5, 20 ในช่อง Range



ในช่อง New Value พิมพ์ค่า 2



เสร็จแล้วคลิก Add จะได้ผลบนจอภาพเป็นดังนี้



คลิก Continue และ OK ตามลำดับ จะได้ผลบนจอภาพเป็น

การเปลี่ยนแปลงค่าตัวแปรเก่าและเก็บค่านั้นไว้ที่ตัวแปรใหม่

ตัวอย่างเช่น ในแฟ้ม file_xy1.sav มีตัวแปร x, y

เราต้องการ เปลี่ยนค่า x เป็น 1 ถ้า $x < 5$

และ เปลี่ยนค่า x เป็น 2 ถ้า $x \geq 5$

โดยค่าที่เปลี่ยนแปลงแล้วเก็บไว้ที่ตัวแปรใหม่ชื่อ newx

	x	y
1	1.00	15.00
2	1.00	17.00
3	2.00	23.00
4	2.00	45.00
5	2.00	58.00

file_xy1.sav

x	y
2.00	15.00
3.00	17.00
7.00	23.00
9.00	45.00
12.00	58.00

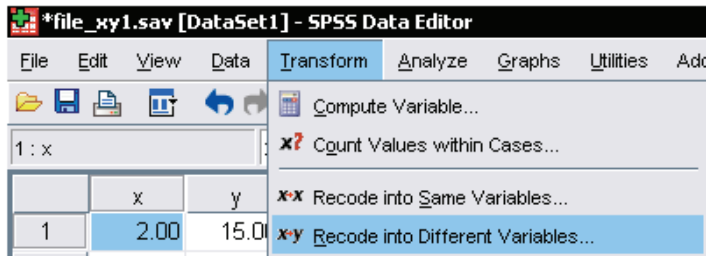
เปลี่ยนค่า x
ไว้ที่ตัวแปร newx แล้วจะได้เป็น

file_xy1_recode_newx.sav

x	y	newx
2.00	15.00	1.00
3.00	17.00	1.00
7.00	23.00	2.00
9.00	45.00	2.00
12.00	58.00	2.00

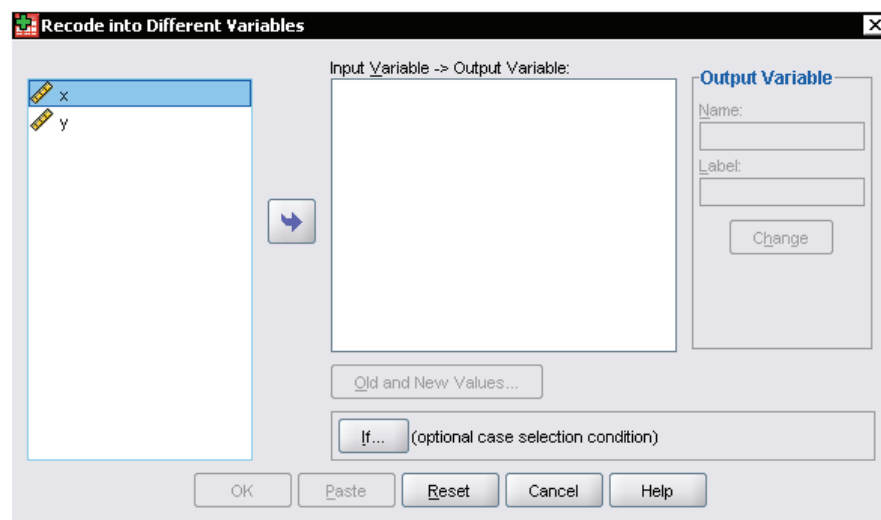
ขั้นที่ 1. เปิดแฟ้มข้อมูล file_xy1.sav เข้ามาใน SPSS Data Editor

ขั้นที่ 2. เลือกคำสั่ง Transform / Recode into Different Variables

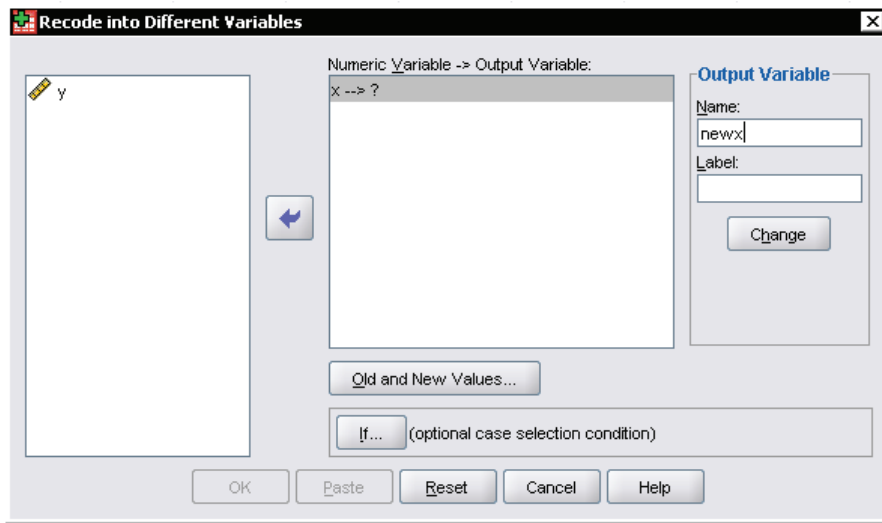


	x	y
1	2.00	15.00
2	3.00	17.00
3	7.00	23.00
4	9.00	45.00
5	12.00	58.00

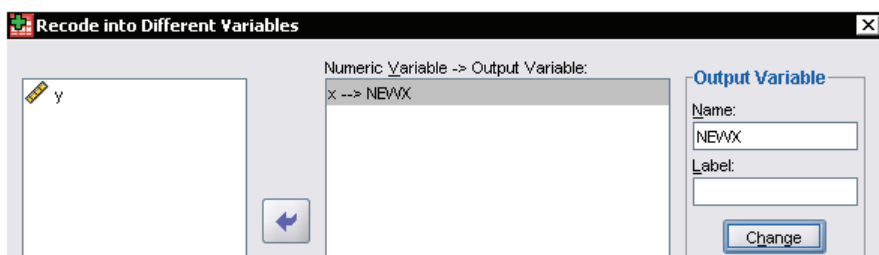
ขั้นที่ 3. คลิกที่คำสั่ง Recode into Different Variables จะได้เมนูย่อยเป็น



ขั้นที่ 4. เลือกตัวแปร x มาไว้ที่ช่อง Numeric Variable → Output Variable
พิมพ์ชื่อตัวแปรใหม่ newx ในช่อง Output Variable

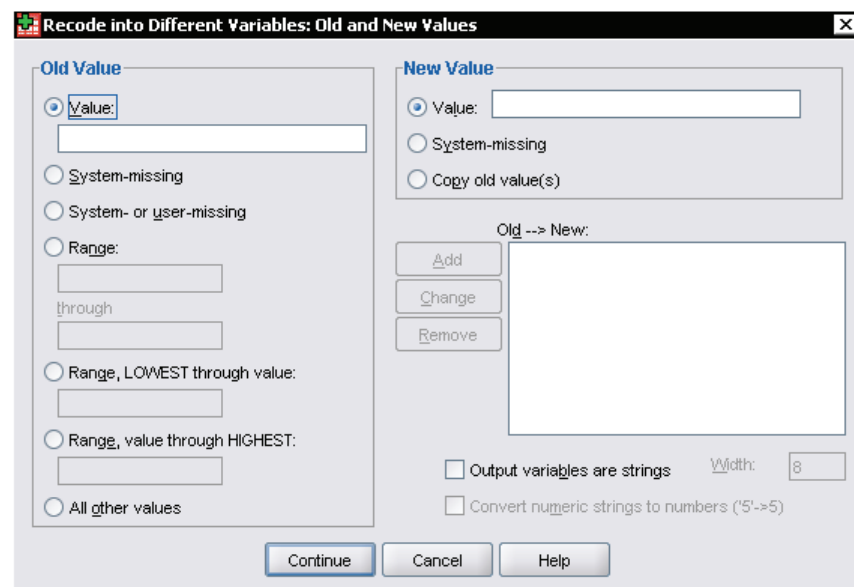


เสร็จแล้วคลิกที่ Change จะได้ผลบนจอภาพเป็น



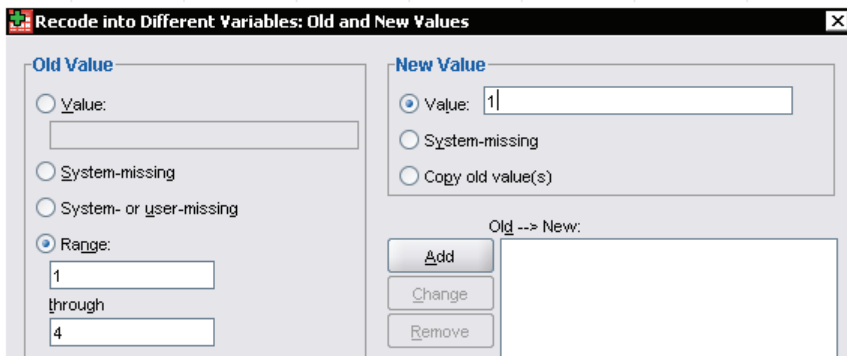
หมายเหตุ x → newx หมายความว่าค่าจากตัวแปร x เมื่อเปลี่ยนแปลงแล้วจะเก็บไว้ที่ตัวแปร newx

ขั้นที่ 5. คลิกที่ Old and New Values จะได้เมนูย่อยเป็น



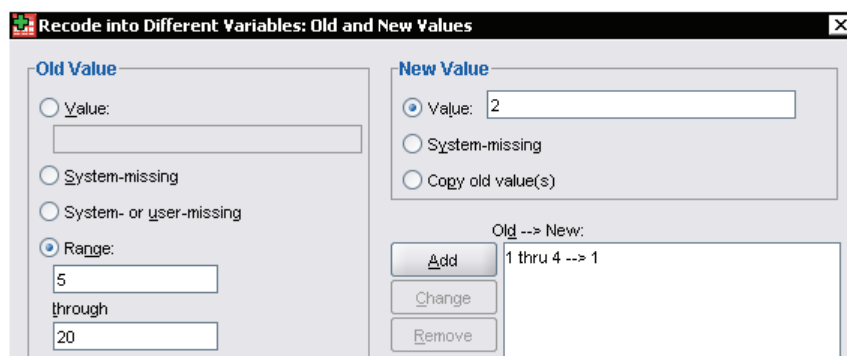
ขั้นที่ 6. เลือก Option Range:

ให้พิมพ์ค่า 1, 4 ในช่อง Range ของ Old Value และ ในช่อง New Value พิมพ์ค่า 1

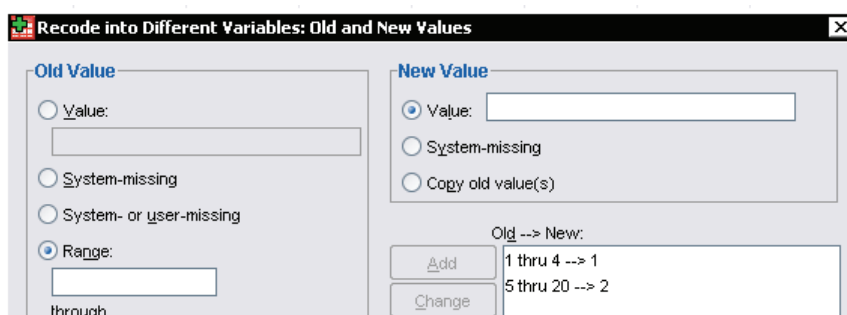


เสร็จแล้วคลิก Add

ขั้นที่ 7. ให้พิมพ์ค่า 5, 20 ในช่อง Range ของ Old Value และ ในช่อง New Value พิมพ์ค่า 2



เสร็จแล้วคลิก Add จะได้ผลบนจอภาพเป็นดังนี้



ขั้นที่ 8. คลิก Continue และ OK ตามลำดับ จะได้ผลบนจอภาพดังนี้

คำแนะนำ ในการทำงานควรจะใช้

คำสั่ง Transform / Recode into Different Variables

ดีกว่า คำสั่ง Transform / Recode into Same Variables

เพราะว่าหากมีข้อผิดพลาดจากการเปลี่ยนแปลงค่า

เรายังมีตัวแปรเก๋ออ้างอิงและใช้งานต่อไปได้

*file_xy1.sav [DataSet1] - SPSS Da

1 : x		2	
	x	y	newx
1	2.00	15.00	1.00
2	3.00	17.00	1.00
3	7.00	23.00	2.00
4	9.00	45.00	2.00
5	12.00	58.00	2.00

บทที่ 6

การหาช่วงความเชื่อมั่น (1- α)100% ของค่าพารามิเตอร์



การทำงานทางด้านสถิติวิเคราะห์ห้มีงานเกี่ยวกับการประมาณค่าพารามิเตอร์เช่น ค่าเฉลี่ยประชากร μ ผลต่างของค่าเฉลี่ยประชากร $\mu_1 - \mu_2$ ฯลฯ การประมาณค่าเราสามารถทำได้โดยการหาช่วงความเชื่อมั่น (1- α)100% ของค่าเฉลี่ย μ ช่วงความเชื่อมั่น (1- α)100% ของผลต่างค่าเฉลี่ย $\mu_1 - \mu_2$ ในบทนี้จะเป็นการใช้โปรแกรม SPSS ช่วยในการหาช่วงความเชื่อมั่นของค่าพารามิเตอร์ หากค่าสถิติเบื้องต้นแบบจำแนกตามกลุ่ม และแบบรวมกลุ่ม และทำการวิเคราะห์ความแปรปรวนทดสอบว่าค่าเฉลี่ยของประชากรทุกกลุ่มเท่ากันหรือไม่ และ ทดสอบสมมติฐาน $H_0 : \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \dots = \sigma_k^2$

คำสั่งสำคัญที่ใช้คือ **Analyze / Compare Means**

โดยมีคำสั่งย่อยและความสามารถของคำสั่งย่อย เช่น

Analyze / Compare Means / Means

คำนวณค่าสถิติเบื้องต้นจำแนกตามกลุ่ม

ทำตาราง ANOVA เพื่อทดสอบสมมติฐาน $H_0 : \mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$

Analyze / Compare Means / One-Sample T Test

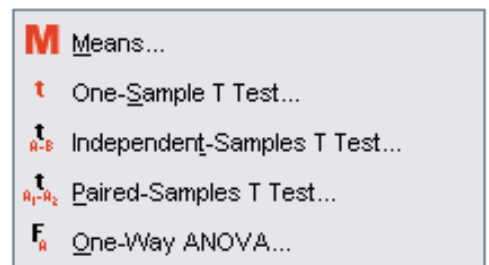
คำนวณค่าสถิติเบื้องต้น

หาช่วงความเชื่อมั่น (1- α)100% ของค่าเฉลี่ย μ

Analyze / Compare Means / Independent-Samples T Test

คำนวณค่าสถิติเบื้องต้นจำแนกตามกลุ่ม

หาช่วงความเชื่อมั่น (1- α)100% ของผลต่างค่าเฉลี่ย $\mu_1 - \mu_2$ กรณีข้อมูลอิสระต่อกัน



Analyze / Compare Means / Paired-Samples T Test

คำนวณค่าสถิติเบื้องต้นจำแนกตามกลุ่ม

หาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของผลต่างค่าเฉลี่ย μ_D กรณีข้อมูลไม่อิสระต่อกัน

หาค่าสหสัมพันธ์(Correlation)

Analyze / Compare Means / One-Way ANOVA

คำนวณค่าสถิติเบื้องต้นจำแนกตามกลุ่ม

หาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่าเฉลี่ย μ จำแนกตามกลุ่มและรวมกลุ่ม

ทำตาราง ANOVA เพื่อทดสอบสมมติฐาน $H_0 : \mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$

ทดสอบสมมติฐาน $H_0 : \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \dots = \sigma_k^2$

6.1 การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่าเฉลี่ย μ

หลักการทางทฤษฎีในเนื้อหาวิชาของความน่าจะเป็นและสถิติของการหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่าเฉลี่ย μ จำแนกเป็นกรณีต่างๆ ดังนี้

1. กรณีประชากรมีการแจกแจงปกติ และ รู้ค่าความแปรปรวน σ^2

สุ่มตัวอย่างขนาด n คำนวณค่าเฉลี่ยของตัวอย่าง \bar{x}

ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่าเฉลี่ย μ คือ

$$\bar{x} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

2. กรณีประชากรมีการแจกแจงปกติ และ ไม่รู้ค่าความแปรปรวน σ^2

สุ่มตัวอย่างขนาด n คำนวณค่าเฉลี่ยของตัวอย่าง \bar{x} และ ค่าความแปรปรวนของตัวอย่าง s^2

2.1 ขนาดตัวอย่าง $n \geq 30$ ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่าเฉลี่ย μ คือ

$$\bar{x} - z_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} < \mu < \bar{x} + z_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$$

2.2 ขนาดตัวอย่าง $n < 30$ ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่าเฉลี่ย μ คือ

$$\bar{x} - t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} ; df = n - 1$$

3. กรณีไม่ได้กำหนดว่าข้อมูลมีการแจกแจงปกติ และ รู้ค่าความแปรปรวน σ^2

สุ่มตัวอย่างขนาด $n \geq 30$ คำนวณค่าเฉลี่ยตัวอย่าง \bar{x}

ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่าเฉลี่ย μ คือ

$$\bar{x} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

4. กรณีไม่ได้กำหนดว่าข้อมูลมีการแจกแจงปกติ และ ไม่รู้ค่าความแปรปรวน σ^2

สุ่มตัวอย่างขนาด $n \geq 30$ คำนวณค่าเฉลี่ยตัวอย่าง \bar{x} และ ค่าความแปรปรวนตัวอย่าง s^2

และประมาณค่า σ^2 ด้วย s^2

ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่าเฉลี่ย μ คือ

$$\bar{x} - z_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} < \mu < \bar{x} + z_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$$

การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่าเฉลี่ย μ ด้วย SPSS for Windows

1. ข้อมูลที่นำมาทำการวิเคราะห์จะมีการแจกแจงแบบปกติหรือไม่ก็ได้
2. สุ่มตัวอย่างขนาด n คำนวณค่าเฉลี่ยของตัวอย่าง \bar{x} และค่าความแปรปรวนของตัวอย่าง s^2
3. ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่าเฉลี่ย μ คือ $\bar{x} - t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$ เมื่อ $df = n - 1$

หมายเหตุ โปรแกรม SPSS จะคิดว่าข้อมูลที่นำมาคำนวณเป็นข้อมูลตัวอย่าง และมีคำสั่งใช้หลายแบบเช่น

- โดยการใช้คำสั่ง Analyze / Descriptive Statistics / Explore
- โดยการใช้คำสั่ง Analyze / Compare Means / One-Sample T Test

ตัวอย่าง 6.1.1 อายุหลอดไฟฟ้ามีการแจกแจงปกติ ค่าความแปรปรวนของประชากร $\sigma^2 = 1600$

สุ่มตัวอย่างหลอดไฟฟ้าจำนวน 30 หลอด หาค่าเฉลี่ยของตัวอย่างได้เท่ากับ 780 ชั่วโมง

จงหาช่วงความเชื่อมั่น 95% ของค่าเฉลี่ย μ

วิธีทำ โดยใช้หลักการทางทฤษฎีของความน่าจะเป็นและสถิติ

กรณีประชากรมีการแจกแจงปกติ และ รู้ค่าความแปรปรวน σ^2

สุ่มตัวอย่างขนาด $n = 30$ ค่าเฉลี่ยของตัวอย่าง $\bar{x} = 780$ ช่วงความเชื่อมั่น 95% ของค่าเฉลี่ย μ คือ

$$\bar{x} - z_{0.025} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{0.025} \frac{\sigma}{\sqrt{n}}$$

$$780 - 1.96\left(\frac{40}{\sqrt{30}}\right) < \mu < 780 + 1.96\left(\frac{40}{\sqrt{30}}\right)$$

$$765.68 < \mu < 794.31$$

ตัวอย่าง 6.1.2 ในการประมาณค่าเฉลี่ยของประชากรหลอดไฟฟ้า

ผู้ทดลองได้ทำการสุ่มตัวอย่างหลอดไฟฟ้าจำนวน 30 หลอด ได้ข้อมูลดังนี้

826.30	793.70	829.90	780.00	750.70	810.20	717.80	786.30	835.80	739.00
770.10	722.80	804.40	786.90	732.50	823.70	726.60	725.60	799.80	801.40
765.50	724.10	811.00	829.20	818.30	730.40	785.70	822.30	731.60	818.40

จงหาช่วงความเชื่อมั่น 95% ของค่าเฉลี่ย μ

วิธีทำ การวิเคราะห์ข้อมูลโดย SPSS for Windows

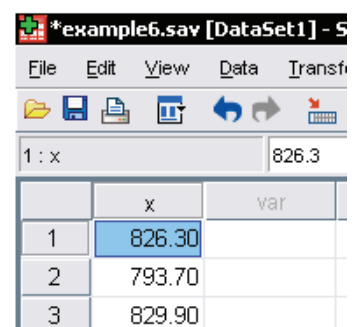
โดยการใช้คำสั่ง Analyze / Descriptive Statistics / Explore

ขั้นที่ 1. สร้างแฟ้มข้อมูล

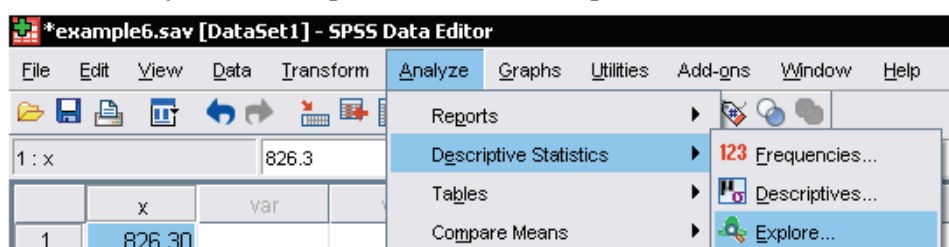
ประกอบด้วย 1 ตัวแปรคือตัวแปร x มี 30 คำสั่งเกิด

และบันทึกลงแฟ้มข้อมูลชื่อ example6.sav

ขั้นที่ 2. ใช้คำสั่ง Analyze / Descriptive Statistics / Explore

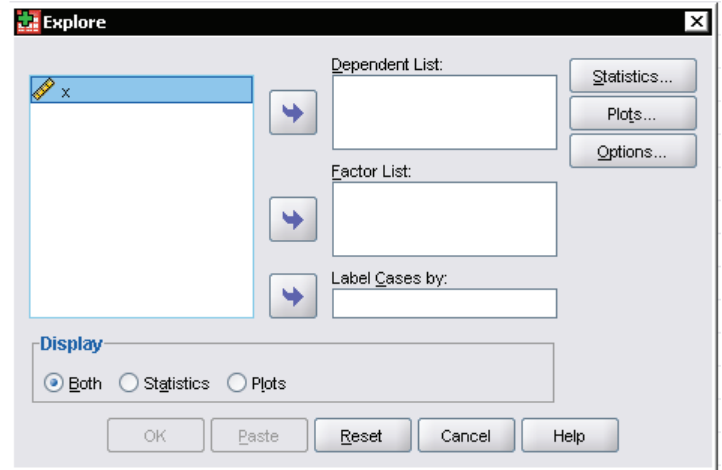


	x	var
1	826.30	
2	793.70	
3	829.90	



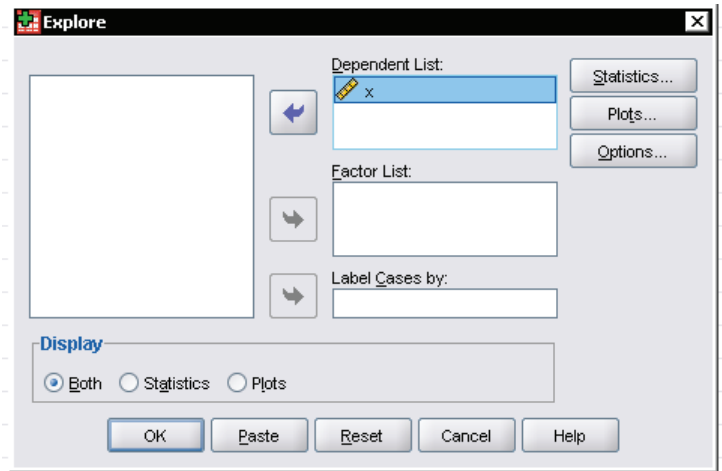
ขั้นที่ 3.

คลิกที่ Explore บนจอภาพ
จะได้เมนูย่อยเป็นดังนี้



ขั้นที่ 4.

เลือกตัวแปร x
ไปไว้ที่ช่อง Dependent List



ขั้นที่ 5. คลิก OK จะได้ผลการคำนวณที่ SPSS Viewer เป็นดังนี้

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
x	30	100.0%	0	.0%	30	100.0%

ผลการคำนวณโดยละเอียดคือ

Explore

[DataSet1] C:\Documents and Settings\dumrong\My Documents\dataspss16\example6.sav

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
x	30	100.0%	0	.0%	30	100.0%

Descriptives

		Statistic	Std. Error
x	Mean	780.0000	7.30434
	95% Confidence Interval for Mean	Lower Bound 765.0610	
		Upper Bound 794.9390	
	5% Trimmed Mean	780.3722	
	Median	786.6000	
	Variance	1600.601	
	Std. Deviation	40.00751	
	Minimum	717.80	
	Maximum	835.80	
	Range	118.00	
	Interquartile Range	86.05	
	Skewness	-.274	.427
	Kurtosis	-1.473	.833

สรุป ช่วงความเชื่อมั่น 95% ของค่าเฉลี่ย μ คือ $(765.0610, 794.9390)$

หมายเหตุ ที่มาของช่วง $(765.0610, 794.9390)$ ได้มาจากสูตร

$$\bar{x} - t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} ; df = n - 1$$

$\alpha = 0.05, \frac{\alpha}{2} = 0.025, df = 30 - 1 = 29, t_{0.025, 29} = 2.045, \bar{x} = 780.00, s = 40.0075$

ช่วงความเชื่อมั่น 95% ของค่าเฉลี่ย μ คือ $\bar{x} - t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$

$$780.0000 - 2.045 \left(\frac{40.0075}{\sqrt{30}} \right) < \mu < 780.0000 + 2.045 \left(\frac{40.0075}{\sqrt{30}} \right)$$

$$780.0000 - 14.937 < \mu < 780.0000 + 14.937$$

$$765.063 < \mu < 794.937$$

โดยใช้คำสั่ง Analyze / Compare Means / One-Sample T Test

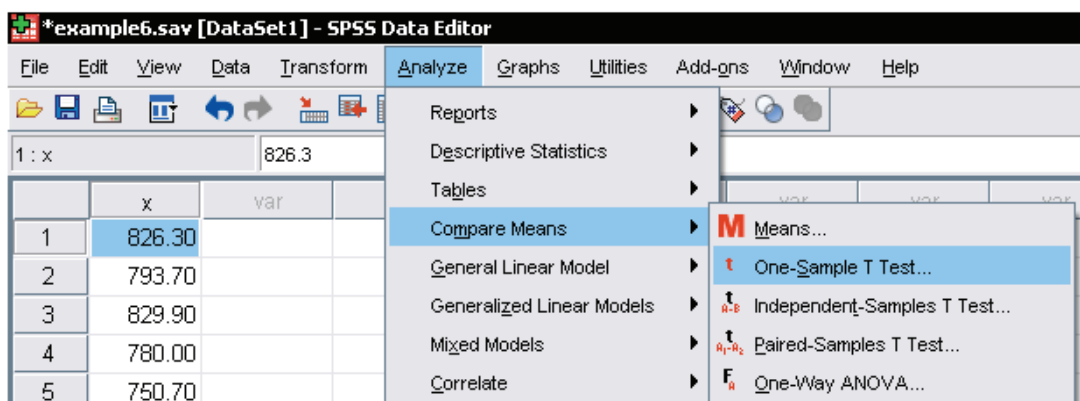
ขั้นที่ 1. สร้างแฟ้มข้อมูล

ประกอบด้วย 1 ตัวแปรคือตัวแปร x มี 30 ค่าสังเกต

และบันทึกลงแฟ้มข้อมูลชื่อ example6.sav

ขั้นที่ 2. เลือกคำสั่ง Analyze / Compare Means / One-Sample T Test

	x	var
1	826.30	
2	793.70	
3	829.90	

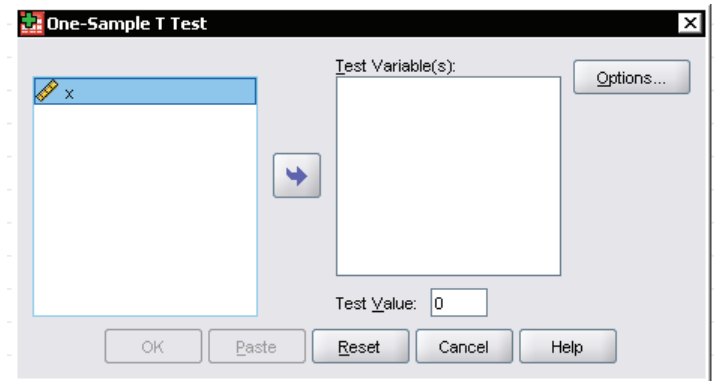


ขั้นที่ 3.

คลิกที่ One-Sample T Test

จอภาพจะขึ้นเมนูย่อยของ

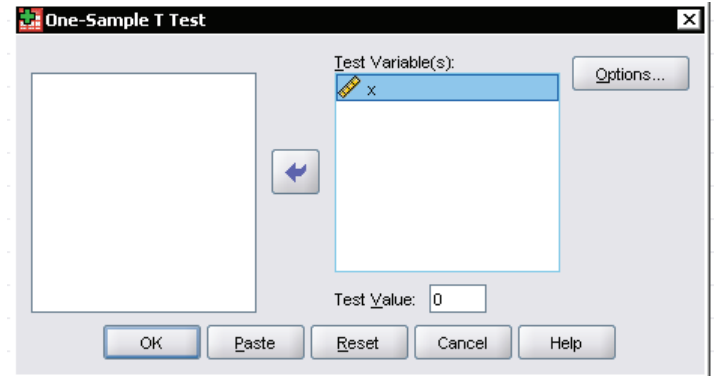
คำสั่ง One-Sample T Test



ขั้นที่ 4.

เลือกตัวแปร x

มาไว้ที่ช่อง Test Variable(s)



ขั้นที่ 5. คลิก OK จะได้ผลที่จอ

SPSS Viewer ดังนี้

ผลการคำนวณ

ทั้งหมดคือ

T-Test
[DataSet1] C:\Documents and Settings\dumrong\My Documents\dataspss16\example6.sav

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
x	30	780.0000	40.00751	7.30434

T-Test

[DataSet1] C:\Documents and Settings\dumrong\My Documents\dataspss16\example6.sav

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
x	30	780.0000	40.00751	7.30434

One-Sample Test

	Test Value = 0					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
x	106.786	29	.000	780.0000	765.0610	794.9390

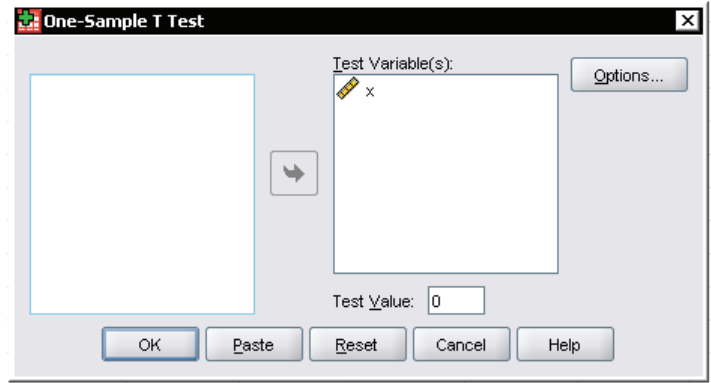
สรุป ช่วงความเชื่อมั่น 95% ของค่าเฉลี่ย μ คือ (765.0610, 794.9390)

หมายเหตุ คำสั่ง Analyze / Compare Means / One-Sample T Test

สามารถเปลี่ยนเปอร์เซ็นต์ของช่วงความเชื่อมั่นได้จากขั้นที่ 4.

เมื่อเลือกตัวแปร x มาไว้ที่

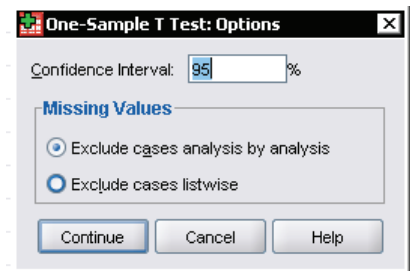
ช่อง Test Variable(s) เสร็จแล้ว



ขั้นที่ 4.1 การเปลี่ยนเปอร์เซ็นต์

ของช่วงความเชื่อมั่นให้

คลิกที่ปุ่ม Options บนจอภาพจะขึ้นเมนูย่อย



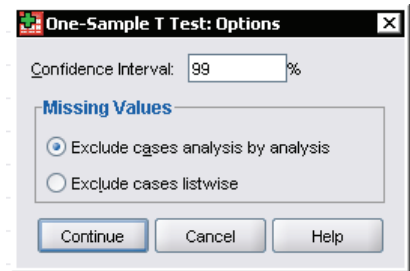
ขั้นที่ 4.2 ในช่อง Confidence Interval

ให้เปลี่ยนจาก 95% เป็น 99%

ขั้นที่ 4.3 คลิกที่ Continue และ OK ตามลำดับ

จะได้ผลที่จอ SPSS Viewer และที่ตาราง

One-Sample Test ได้ช่วงความเชื่อมั่นใหม่ดังนี้



One-Sample Test

	Test Value = 0					
	t	df	Sig. (2-tailed)	Mean Difference	99% Confidence Interval of the Difference	
					Lower	Upper
x	106.786	29	.000	780.00000	759.8664	800.1336

สรุป ช่วงความเชื่อมั่น 99% ของค่าเฉลี่ย μ คือ (759.8664, 800.1366)

หมายเหตุ ที่มาของช่วงความเชื่อมั่น 99% ของค่าเฉลี่ย (759.8664, 800.1366) ได้จากสูตร

$$\bar{x} - t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} ; df = n - 1$$

$\alpha = 0.01, \frac{\alpha}{2} = 0.005, df = 30 - 1 = 29, t_{0.005, 29} = 2.756, \bar{x} = 780.00, s = 40.0075$

ช่วงความเชื่อมั่น 99% ของค่าเฉลี่ย μ คือ $\bar{x} - t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$

$$780.0000 - 2.765 \left(\frac{40.0075}{\sqrt{30}} \right) < \mu < 780.0000 + 2.765 \left(\frac{40.0075}{\sqrt{30}} \right)$$

$$780.0000 - 20.131 < \mu < 780.0000 + 20.131$$

$$759.869 < \mu < 800.131$$

6.2 การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของผลต่างค่าเฉลี่ย $\mu_1 - \mu_2$

กรณีประชากร 2 ชุดเป็นอิสระต่อกัน

หลักการทางทฤษฎีของความน่าจะเป็นและสถิติในการหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของผลต่างค่าเฉลี่ย $\mu_1 - \mu_2$ จะจำแนกออกเป็น 2 กรณีคือ

- กรณีที่ประชากร 2 ชุดเป็นอิสระต่อกัน
- กรณีที่ประชากร 2 ชุดไม่เป็นอิสระต่อกัน

สุ่มตัวอย่างขนาด n_1 จากประชากรชุดที่ 1 และ สุ่มตัวอย่างขนาด n_2 จากประชากรชุดที่ 2 หาค่าเฉลี่ยของตัวอย่าง \bar{x}_1 และ \bar{x}_2

1. กรณี $n_1 \geq 30$ และ $n_2 \geq 30$

1.1 กรณีประชากร 2 ชุดมีการแจกแจงปกติ และ รู้ค่าความแปรปรวน σ_1^2 และ σ_2^2 ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของผลต่างค่าเฉลี่ย $\mu_1 - \mu_2$ คือ

$$(\bar{x}_1 - \bar{x}_2) - z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

1.2. กรณีประชากร 2 ชุดมีการแจกแจงปกติ และ ไม่รู้ค่าความแปรปรวน σ_1^2 และ σ_2^2

หาค่าความแปรปรวนของตัวอย่าง s_1^2 และ s_2^2 และประมาณ σ_1^2 และ σ_2^2 ด้วย s_1^2 และ s_2^2 ตามลำดับ ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของผลต่างค่าเฉลี่ย $\mu_1 - \mu_2$ คือ

$$(\bar{x}_1 - \bar{x}_2) - z_{\frac{\alpha}{2}} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + z_{\frac{\alpha}{2}} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

2. กรณี $n_1 < 30$ หรือ $n_2 < 30$

2.1 กรณีประชากร 2 ชุดมีการแจกแจงปกติ และ รู้ค่าความแปรปรวน σ_1^2 และ σ_2^2 ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของผลต่างค่าเฉลี่ย $\mu_1 - \mu_2$ คือ

$$(\bar{x}_1 - \bar{x}_2) - z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

2.2 กรณีประชากร 2 ชุดมีการแจกแจงปกติ และ ไม่รู้ค่าความแปรปรวน σ_1^2 และ σ_2^2

หาค่าความแปรปรวนของตัวอย่าง s_1^2 และ s_2^2

2.2.1 ภายใต้ข้อกำหนด $\sigma_1^2 = \sigma_2^2$ ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของผลต่างค่าเฉลี่ย $\mu_1 - \mu_2$ คือ

$$(\bar{x}_1 - \bar{x}_2) - t_{\frac{\alpha}{2}} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + t_{\frac{\alpha}{2}} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

เมื่อ $s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2}$ และ $df = n_1 + n_2 - 2$

2.2.2 ภายใต้ข้อกำหนด $\sigma_1^2 \neq \sigma_2^2$ ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของผลต่างค่าเฉลี่ย $\mu_1 - \mu_2$ คือ

$$(\bar{x}_1 - \bar{x}_2) - t_{\frac{\alpha}{2}} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + t_{\frac{\alpha}{2}} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$\text{เมื่อ } df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\left(\frac{s_1^2}{n_1}\right)^2 \frac{1}{(n_1-1)} + \left(\frac{s_2^2}{n_2}\right)^2 \frac{1}{(n_2-1)}}$$

การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของผลต่างค่าเฉลี่ย $\mu_1 - \mu_2$ ด้วย SPSS for Windows

1. ข้อมูลที่นำมาทำการวิเคราะห์จะมีการแจกแจงแบบปกติหรือไม่ก็ได้
2. สุ่มตัวอย่างขนาด n_1 จากประชากรชุดที่ 1 และ สุ่มตัวอย่างขนาด n_2 จากประชากรชุดที่ 2
3. หาค่าเฉลี่ยของตัวอย่าง \bar{x}_1 และ \bar{x}_2 หาค่าความแปรปรวนของตัวอย่าง s_1^2 และ s_2^2

ภายใต้ข้อกำหนด $\sigma_1^2 = \sigma_2^2$ ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของผลต่างค่าเฉลี่ย $\mu_1 - \mu_2$ คือ

$$(\bar{x}_1 - \bar{x}_2) - t_{\frac{\alpha}{2}} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + t_{\frac{\alpha}{2}} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$\text{เมื่อ } s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2} \text{ และ } df = n_1 + n_2 - 2$$

ภายใต้ข้อกำหนด $\sigma_1^2 \neq \sigma_2^2$ ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของผลต่างค่าเฉลี่ย $\mu_1 - \mu_2$ คือ

$$(\bar{x}_1 - \bar{x}_2) - t_{\frac{\alpha}{2}} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + t_{\frac{\alpha}{2}} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$\text{เมื่อ } df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\left(\frac{s_1^2}{n_1}\right)^2 \frac{1}{(n_1-1)} + \left(\frac{s_2^2}{n_2}\right)^2 \frac{1}{(n_2-1)}}$$

การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของผลต่างค่าเฉลี่ย $\mu_1 - \mu_2$

โดยใช้คำสั่ง Analyze / Compare Means / Independent-Samples T Test

ตัวอย่าง 6.2.1 ทำการทดลองสุ่มตัวอย่างข้อมูล 2 ชุด

ตัวอย่างขนาด $n_1 = 9$ จากประชากรชุดที่ 1 มีข้อมูลเป็นดังนี้

61.36	57.76	71.94	61.77	58.66	71.61	71.52	58.67	62.77
-------	-------	-------	-------	-------	-------	-------	-------	-------

ตัวอย่างขนาด $n_2 = 16$ จากประชากรชุดที่ 2 มีข้อมูลเป็นดังนี้

56.92	58.30	67.48	53.96	62.00	59.61	52.02	61.60
64.83	58.55	52.53	64.74	55.51	66.18	55.51	54.18

จงหาช่วงความเชื่อมั่น 95% ของผลต่างค่าเฉลี่ย $\mu_1 - \mu_2$

วิธีทำ การคำนวณด้วย SPSS for Windows

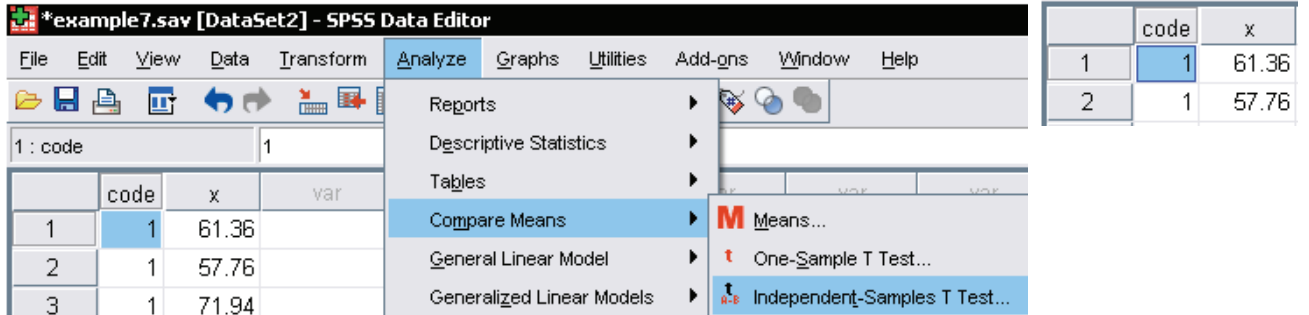
โดยใช้คำสั่ง Analyze / Compare Means / Independent-Samples T Test

ขั้นที่ 1. สร้างแฟ้มข้อมูล โดยกำหนดให้มีตัวแปร 2 ตัวคือ ตัวแปรจำแนกกลุ่มตัวอย่าง (code)

และ ตัวแปรข้อมูล (x) แล้วบันทึกไว้ที่แฟ้มชื่อ example7.sav

ขั้นที่ 2. เลือกคำสั่ง

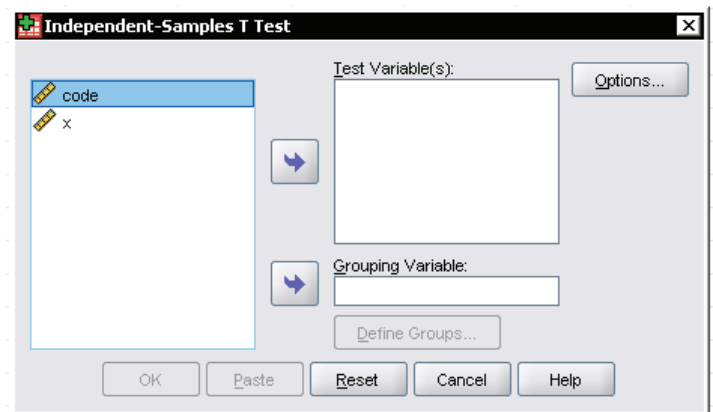
Analyze / Compare Means / Independent-Samples T Test



ขั้นที่ 3.

คลิกที่ Independent-Samples T Test
จะได้เมนูย่อยของคำสั่ง

Independent-Samples T Test ดังนี้



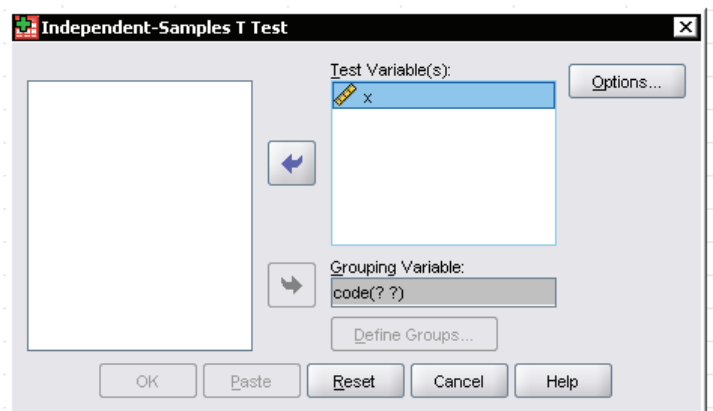
ขั้นที่ 4. คลิกที่ code(? ?)

เลือกตัวแปร code

มาไว้ที่ช่อง Grouping Variable

เลือกตัวแปร x

มาไว้ที่ช่อง Test Variable(s)

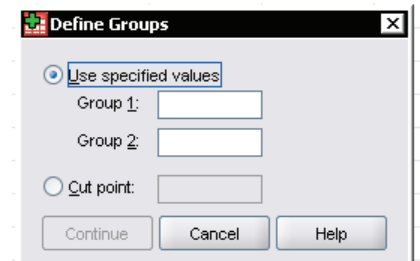


ขั้นที่ 5.

การเลือกหมายเลขของกลุ่มในตัวแปร code
ที่ต้องการวิเคราะห์ข้อมูล

ให้คลิกที่ Defined Groups

จะได้เมนูย่อยของการเลือกหมายเลขกลุ่มเป็นดังนี้

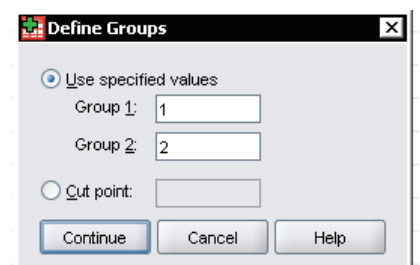


ขั้นที่ 6. ให้นำเมาส์มาคลิกที่ช่อง Group 1

และ พิมพ์หมายเลข 1 ในช่อง Group 1

นำเมาส์มาคลิกที่ช่อง Group 2

และ พิมพ์หมายเลข 2 ในช่อง Group 2



ขั้นที่ 7. คลิก Continue

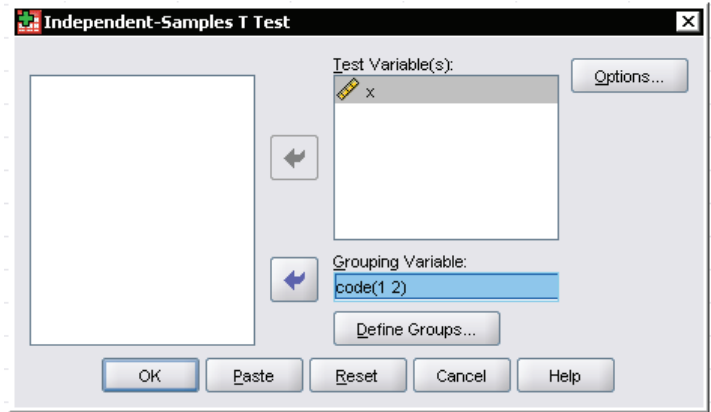
จะกลับมาที่เมนูย่อย

Independent-Samples T Test

ที่ช่อง Grouping Variable

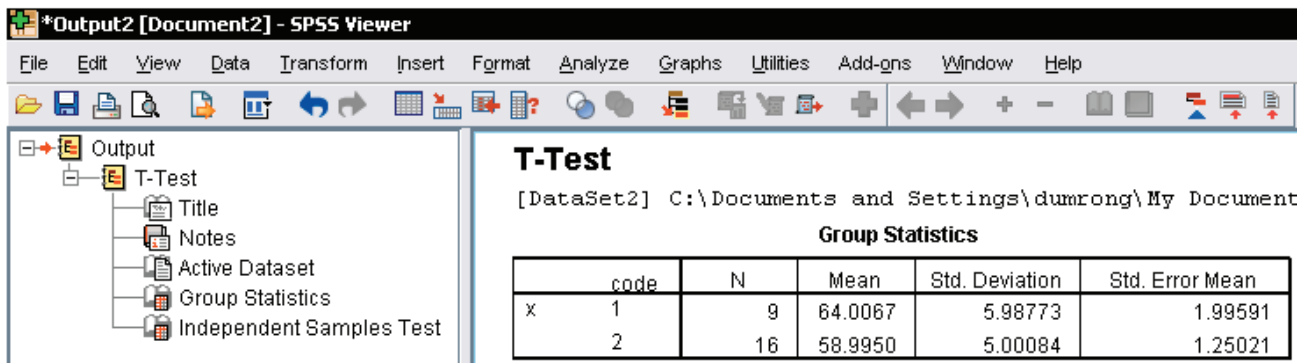
ที่ตัวแปร Code จะเปลี่ยนเป็น

Code(1 2)



ขั้นที่ 8. คลิก OK

จะได้ผลการคำนวณเป็นดังนี้



ผลการคำนวณทั้งหมดคือ

T-Test

[DataSet2] C:\Documents and Settings\dumrong\My Documents\dataspss16\example7.sav

Group Statistics

	code	N	Mean	Std. Deviation	Std. Error Mean
x	1	9	64.0067	5.98773	1.99591
	2	16	58.9950	5.00084	1.25021

Independent Samples Test

		x		
		Equal variances assumed	Equal variances not assumed	
Levene's Test for Equality of Variances	F	.800		
	Sig.	.380		
t-test for Equality of Means	t	2.242	2.128	
	df	23	14.333	
	Sig. (2-tailed)	.035	.051	
	Mean Difference	5.01167	5.01167	
	Std. Error Difference	2.23531	2.35514	
	95% Confidence Interval of the Difference	Lower	.38758	-.02863
		Upper	9.63575	10.05196

การนำผลการคำนวณของ SPSS ไปใช้งานต้องเลือกให้เหมาะสมกับข้อกำหนดของประชากร

กรณีที่ 1. ภายใต้เงื่อนไขว่าประชากรมีการแจกแจงปกติและมีความแปรปรวนเท่ากัน

ต้องใช้ผลสรุปใน Equal variances assumed

ช่วงความเชื่อมั่น 95% ของผลต่างค่าเฉลี่ย $\mu_1 - \mu_2$ คือ $0.38758 < \mu_1 - \mu_2 < 9.63575$

กรณีที่ 2. ภายใต้เงื่อนไขว่าประชากรมีการแจกแจงปกติและมีความแปรปรวนไม่เท่ากัน

ต้องใช้ผลสรุปใน Equal variances not assumed

ช่วงความเชื่อมั่น 95% ของผลต่างค่าเฉลี่ย $\mu_1 - \mu_2$ คือ $-0.02863 < \mu_1 - \mu_2 < 10.05196$

หมายเหตุ ถ้าไม่มีการกำหนดว่าความแปรปรวนเท่ากัน หรือไม่เท่ากัน การสรุปผลทางด้านสถิติเกี่ยวกับความแปรปรวนของประชากรทั้ง 2 ประชากรให้ดูจากค่าสถิติ Levene ในตาราง Independent Samples Test

Levene's Test for Equality of Variances	F	.800
	Sig.	.380

ถ้า Sig. ของค่า Leven's Test for Equality of Variances มีค่าน้อยกว่า $\alpha = 0.05$

แล้วเราสรุปได้ว่าประชากร 2 ชุดมีค่าความแปรปรวนไม่เท่ากัน โดยมีระดับนัยสำคัญ $\alpha = 0.05$

ที่มาของค่าสถิติตรวจสอบได้ด้วย MATHCAD ดังนี้

ตาราง Group Statistics

sample2 :=	(56.92)			
	58.30			
	67.48			
	53.96			
	62.00	sample1 :=	(61.36)	
	59.61		57.76	
	52.02		71.94	
	61.60		61.77	
	64.83		58.66	xbar1 := mean(sample1)
	58.55		71.61	xbar1 = 64.0067
	52.53		71.52	xbar2 := mean(sample2)
	64.74		58.67	xbar2 = 58.995
	55.51		(62.77)	n1 := length(sample1)
	66.18			n1 = 9
	55.51			n2 := length(sample2)
	(54.18)			n2 = 16
				s1 := Stdev(sample1)
				s1 = 5.9877
				s2 := Stdev(sample2)
				s2 = 5.0008
				Std_Error_Mean1 := $\frac{s1}{\sqrt{n1}}$
				Std_Error_Mean1 = 1.9959
				Std_Error_Mean2 := $\frac{s2}{\sqrt{n2}}$
				Std_Error_Mean2 = 1.2502

ตาราง Independent Samples Test กรณีที่ความแปรปรวนประชากรทั้งสองชุดเท่ากัน

$$sp := \sqrt{\frac{(n1 - 1) \cdot s1^2 + (n2 - 1) \cdot s2^2}{n1 + n2 - 2}} \quad sp = 5.3647 \quad t := \frac{\bar{x}1 - \bar{x}2}{sp \cdot \sqrt{\frac{1}{n1} + \frac{1}{n2}}} \quad t = 2.242$$

$$df := n1 + n2 - 2 \quad df = 23$$

$$Sig := 1 - pt(2.242, 23) \quad Sig = 0.0175 \quad Sig_2_tailed := 2 \cdot Sig \quad Sig_2_tailed = 0.0349$$

$$Mean_Difference := \bar{x}1 - \bar{x}2 \quad Mean_Difference = 5.0117$$

$$Std_Error_Difference := sp \cdot \sqrt{\frac{1}{n1} + \frac{1}{n2}} \quad Std_Error_Difference = 2.2353$$

$$\alpha := 0.05 \quad t_alpha_divide2 := qt\left(1 - \frac{\alpha}{2}, 23\right) \quad t_alpha_divide2 = 2.0687$$

$$Lower := (\bar{x}1 - \bar{x}2) - t_alpha_divide2 \cdot Std_Error_Difference \quad Lower = 0.3876$$

$$Upper := (\bar{x}1 - \bar{x}2) + t_alpha_divide2 \cdot Std_Error_Difference \quad Upper = 9.6357$$

ตาราง Independent Samples Test กรณีความแปรปรวนประชากรทั้งสองชุดไม่เท่ากัน

$$t := \frac{\bar{x}1 - \bar{x}2}{\sqrt{\frac{s1^2}{n1} + \frac{s2^2}{n2}}} \quad t = 2.128 \quad df := \frac{\left(\frac{s1^2}{n1} + \frac{s2^2}{n2}\right)^2}{\left(\frac{s1^2}{n1}\right)^2 \cdot \left(\frac{1}{n1 - 1}\right) + \left(\frac{s2^2}{n2}\right)^2 \cdot \left(\frac{1}{n2 - 1}\right)} \quad df = 14.3325$$

$$Sig := 1 - pt(2.128, 14.3325) \quad Sig = 0.0256 \quad Sig_2_tailed := 2 \cdot Sig \quad Sig_2_tailed = 0.0511$$

$$Mean_Difference := \bar{x}1 - \bar{x}2 \quad Mean_Difference = 5.0117$$

$$Std_Error_Difference := \sqrt{\frac{s1^2}{n1} + \frac{s2^2}{n2}} \quad Std_Error_Difference = 2.3551$$

$$\alpha := 0.05 \quad t_alpha_divide2 := qt\left(1 - \frac{\alpha}{2}, df\right) \quad t_alpha_divide2 = 2.1401$$

$$Lower := (\bar{x}1 - \bar{x}2) - t_alpha_divide2 \cdot Std_Error_Difference \quad Lower = -0.0286$$

$$Upper := (\bar{x}1 - \bar{x}2) + t_alpha_divide2 \cdot Std_Error_Difference \quad Upper = 10.0520$$

ตัวอย่าง 6.2.2 ข้อมูลปริมาณน้ำฝน (หน่วยเป็นนิ้ว) ของตำบลที่ 1 ในช่วง 15 ปีที่ผ่านมาเป็นดังนี้

2.40	2.42	1.87	2.50	2.29	1.68	2.57	1.60	1.65	1.41
1.66	1.32	2.43	1.83	1.41					

ข้อมูลปริมาณน้ำฝน (หน่วยเป็นนิ้ว) ของตำบลที่ 2 ในช่วง 10 ปีที่ผ่านมาเป็นดังนี้

0.79	1.25	0.72	0.84	1.32	1.35	1.29	0.72	0.96	1.13
------	------	------	------	------	------	------	------	------	------

สมมติว่าข้อมูลมีการแจกแจงปกติและมีค่าความแปรปรวนประชากรแตกต่างกัน จงหาช่วงความเชื่อมั่น 95 % ของความแตกต่างของค่าเฉลี่ยของปริมาณน้ำฝน

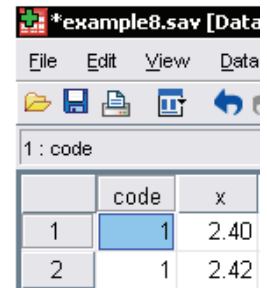
วิธีทำ การวิเคราะห์ข้อมูลด้วย SPSS for Windows

ขั้นที่ 1. สร้างแฟ้มข้อมูลที่ประกอบด้วย 2 ตัวแปร

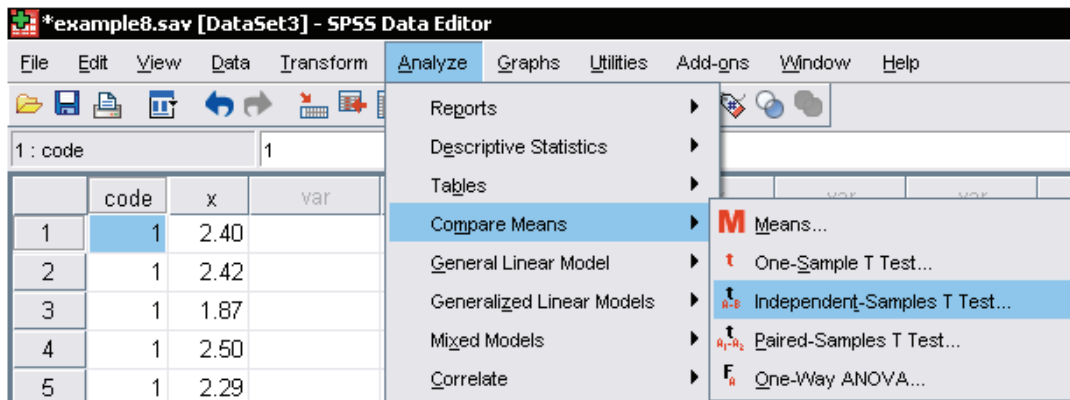
คือ code เป็นตัวแปรจำแนกกลุ่ม และ x เป็นตัวแปรปริมาณน้ำฝน

และ Save เป็นแฟ้มข้อมูลชื่อ example8.sav

ขั้นที่ 2. เลือกคำสั่ง Analyze / Compare Means / Independent-Samples T Test



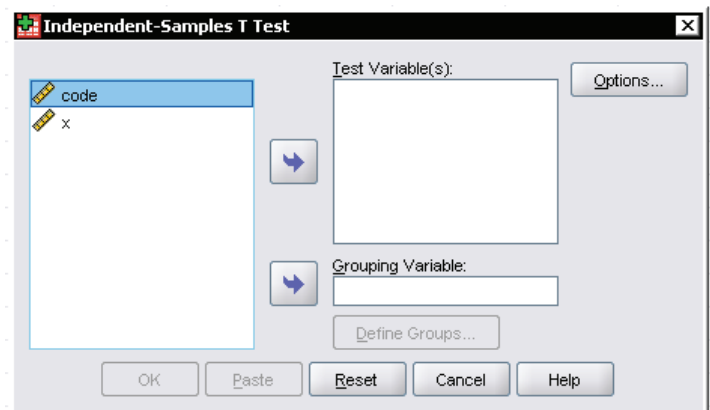
	code	x
1	1	2.40
2	1	2.42



ขั้นที่ 3.

คลิกที่ Independent-Samples T Test

จะได้เมนูย่อยดังนี้

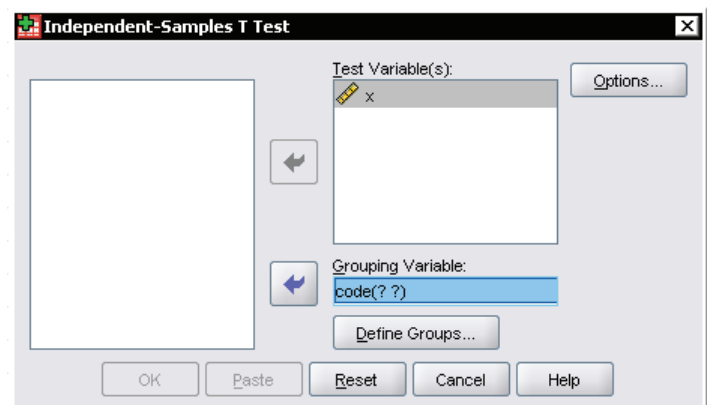


ขั้นที่ 4. เลือกตัวแปร x

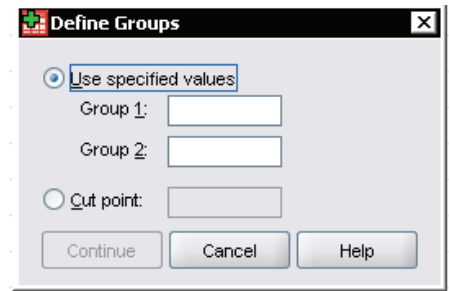
มาไว้ที่ช่อง Test Variable(s)

และเลือกตัวแปร code

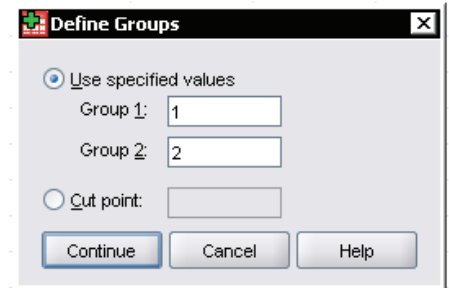
มาไว้ที่ช่อง Grouping Variable



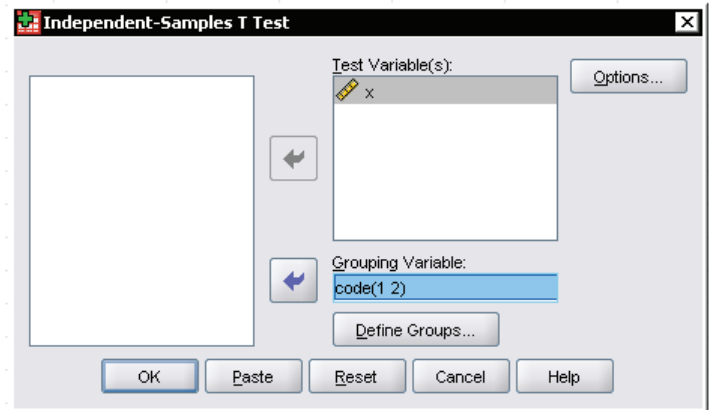
ขั้นที่ 5. ต่อไปเลือกหมายเลขของกลุ่มในตัวแปร code ที่เราต้องการวิเคราะห์ข้อมูล คลิกที่ code(? ?) และคลิกที่ Define groups จะได้เมนูย่อยของการเลือกหมายเลขกลุ่มเป็นดังนี้



ขั้นที่ 6. นำเมาส์มาคลิกที่ช่อง Group 1 และพิมพ์หมายเลข 1 ในช่อง Group 1 นำเมาส์มาคลิกที่ช่อง Group 2 และพิมพ์หมายเลข 2 ในช่อง Group 2



ขั้นที่ 7. คลิก Continue จะกลับมาที่เมนูย่อยที่ช่อง Grouping Variable จะเป็น code(1 2) แสดงว่าเป็นการวิเคราะห์ข้อมูลของกลุ่มที่ 1 และ 2



ขั้นที่ 8. คลิก OK จะได้ผลการคำนวณดังนี้

*Output3 [Document3] - SPSS Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Add-ons Window Help

Output

- T-Test
 - Title
 - Notes
 - Active Dataset
 - Group Statistics
 - Independent Samples Test

T-Test

[DataSet3] C:\Documents and Settings\dumrong\My Documents\

Group Statistics

	code	N	Mean	Std. Deviation	Std. Error Mean
x	1	15	1.9360	.44911	.11596
	2	10	1.0370	.25880	.08184

ผลการคำนวณทั้งหมดคือ

T-Test

[DataSet3] C:\Documents and Settings\dumrong\My Documents\datasps16\example8.sav

Group Statistics

	code	N	Mean	Std. Deviation	Std. Error Mean
x	1	15	1.9360	.44911	.11596
	2	10	1.0370	.25880	.08184

Independent Samples Test

		x		
		Equal variances assumed	Equal variances not assumed	
Levene's Test for Equality of Variances	F	7.769		
	Sig.	.010		
t-test for Equality of Means	t	5.705	6.334	
	df	23	22.671	
	Sig. (2-tailed)	.000	.000	
	Mean Difference	.89900	.89900	
	Std. Error Difference	.15758	.14193	
	95% Confidence Interval of the Difference	Lower	.57303	.60516
		Upper	1.22497	1.19284

สรุปช่วงความเชื่อมั่น 95 % ของความแตกต่างของค่าเฉลี่ยของปริมาณน้ำฝน คือ

$$0.60516 < \mu_1 - \mu_2 < 1.19284$$

การเปลี่ยนเปอร์เซ็นต์ของช่วงความเชื่อมั่นในการใช้งาน

ของคำสั่ง Analyze / Compare Means / Independent-Samples T Test

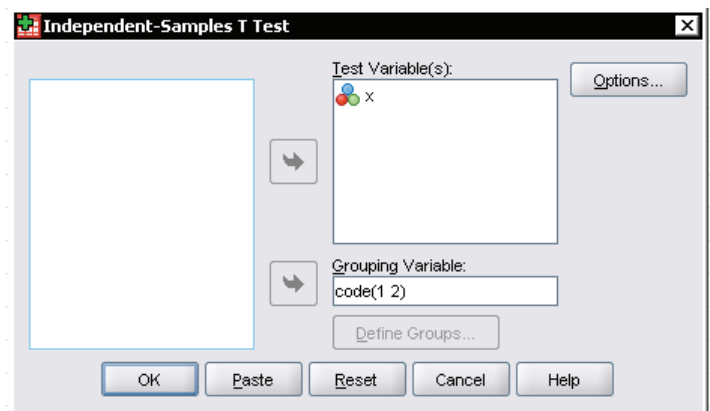
ตัวอย่างเช่นต้องการหาช่วงความเชื่อมั่น 99% ของผลต่างค่าเฉลี่ย $\mu_1 - \mu_2$

จาก ขั้นที่ 7.

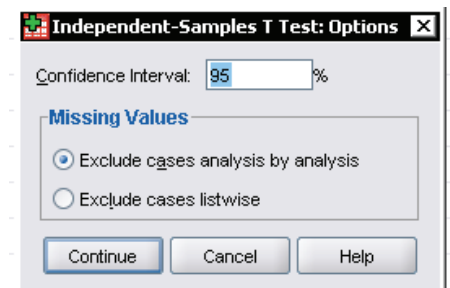
เมื่อเลือกตัวแปร

และกำหนดหมายเลขกลุ่ม

เสร็จแล้ว



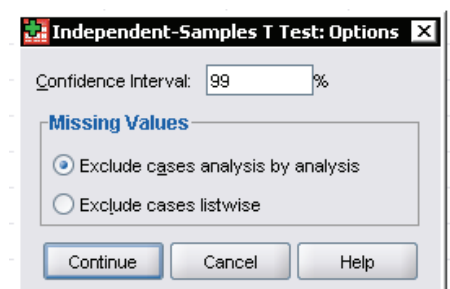
ขั้นที่ 7.1 ให้คลิกที่ Options จะได้เมนูย่อยเป็น



ขั้นที่ 7.2 ให้เปลี่ยนเปอร์เซ็นต์จาก 95% เป็น 99%

ขั้นที่ 7.3 คลิก Continue

จะกลับเมนูย่อย Independent-Samples T Test



ขั้นที่ 8. คลิก OK จะได้ผลการคำนวณใหม่ในส่วนของตาราง Independent-Samples Test ดังนี้

Independent Samples Test

		x		
		Equal variances assumed	Equal variances not assumed	
Levene's Test for Equality of Variances	F	7.769		
	Sig.	.010		
t-test for Equality of Means	t	5.705	6.334	
	df	23	22.671	
	Sig. (2-tailed)	.000	.000	
	Mean Difference	.89900	.89900	
	Std. Error Difference	.15758	.14193	
	99% Confidence Interval of the Difference	Lower	.45663	.50004
		Upper	1.34137	1.29796

สรุปช่วงความเชื่อมั่น 99% ของผลต่างค่าเฉลี่ย $\mu_1 - \mu_2$ คือ $(0.50004, 1.29796)$

6.3 การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของผลต่างค่าเฉลี่ย μ_D

กรณีประชากร 2 ชุดไม่เป็นอิสระต่อกัน

หลักการทางทฤษฎีของความน่าจะเป็นและสถิติในการหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของผลต่างค่าเฉลี่ย μ_D เมื่อประชากร 2 ชุดไม่เป็นอิสระต่อกัน มีขั้นตอนการทำงานดังนี้
 สุ่มตัวอย่างขนาด n จากประชากรชุดที่ 1 และ ประชากรชุดที่ 2 ได้ข้อมูลเป็น

ตัวอย่างจากประชากรชุดที่ 1.	ตัวอย่างจากประชากรชุดที่ 2.
x_1	y_1
x_2	y_2
x_3	y_3
:	:
x_n	y_n

- ขั้นตอนการคำนวณ
1. คำนวณค่าผลต่างของตัวอย่าง $d_i = x_i - y_i, i = 1, 2, \dots, n$
 2. คำนวณค่าเฉลี่ยของผลต่างของตัวอย่าง \bar{d}
 3. คำนวณค่าส่วนเบี่ยงเบนมาตรฐานของผลต่างของตัวอย่าง $s_{\bar{d}}$

กรณี $n \geq 30$ ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของผลต่างค่าเฉลี่ย μ_D คือ

$$\bar{d} - z_{\frac{\alpha}{2}} \frac{s_d}{\sqrt{n}} < \mu_D < \bar{d} + z_{\frac{\alpha}{2}} \frac{s_d}{\sqrt{n}}$$

กรณี $n < 30$ และภายใต้ข้อสมมติว่าประชากรมีการแจกแจงปกติ

ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของผลต่างค่าเฉลี่ย μ_D คือ

$$\bar{d} - t_{\frac{\alpha}{2}} \frac{s_d}{\sqrt{n}} < \mu_D < \bar{d} + t_{\frac{\alpha}{2}} \frac{s_d}{\sqrt{n}} \quad \text{เมื่อ } df = n - 1$$

การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของผลต่างค่าเฉลี่ย μ_D ด้วย SPSS for Windows

1. ข้อมูลที่นำมาทำการวิเคราะห์มีการแจกแจงแบบปกติ (ถ้าขนาดของตัวอย่างเล็กกว่า 30)
 2. สุ่มตัวอย่างขนาด n จากประชากรชุดที่ 1 และ ประชากรชุดที่ 2
 3. คำนวณค่าผลต่างของตัวอย่าง $d_i = x_i - y_i, i = 1, 2, \dots, n$
 4. คำนวณค่าเฉลี่ยของผลต่างของตัวอย่าง \bar{d}
 5. คำนวณค่าส่วนเบี่ยงเบนมาตรฐานของผลต่างของตัวอย่าง $s_{\bar{d}}$
- ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของผลต่างค่าเฉลี่ย μ_D คือ

$$\bar{d} - t_{\frac{\alpha}{2}} \frac{s_d}{\sqrt{n}} < \mu_D < \bar{d} + t_{\frac{\alpha}{2}} \frac{s_d}{\sqrt{n}} \quad \text{เมื่อ } df = n - 1$$

การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของ μ_D กรณีประชากร 2 ชุดไม่เป็นอิสระต่อกัน โดยใช้คำสั่ง Analyze / Compare Means / Paired – Samples T Test

ตัวอย่าง 6.3.1 จากตัวอย่างสุ่มของนิสิตที่เรียนสถิติ 10 คน

ได้ข้อมูลคะแนนการสอบย่อยครั้งที่ 1 และ ครั้งที่ 2 ของนิสิต 10 คน เป็นดังนี้

คนที่	คะแนนครั้งที่ 1.	คะแนนครั้งที่ 2.
1	76	81
2	60	52
3	85	87
4	58	70
5	91	86
6	75	77
7	82	90
8	64	63
9	79	85
10	88	83

จงหาช่วงความเชื่อมั่น 98% ของค่าผลต่างที่แท้จริงในการสอบย่อย

วิธีทำ กำหนด μ_1 เป็นค่าเฉลี่ยของประชากรชุดที่ 1 และ μ_2 เป็นค่าเฉลี่ยของประชากรชุดที่ 2

การคำนวณด้วย SPSS for Windows

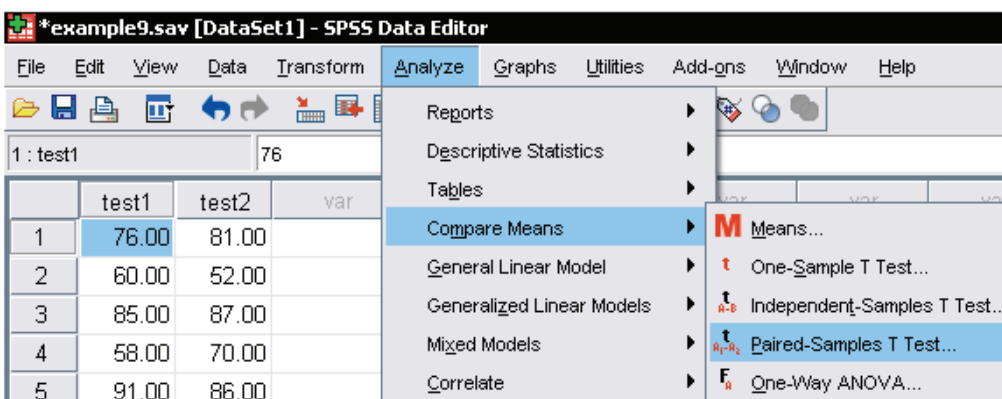
ขั้นที่ 1. สร้างแฟ้มข้อมูลโดยกำหนดให้มีตัวแปร 2 ตัวคือ

ตัวแปร test1 เป็นคะแนนสอบย่อยครั้งที่ 1

ตัวแปร test2 เป็นคะแนนสอบย่อยครั้งที่ 2 เสร็จแล้วบันทึกลงแฟ้ม example9.sav

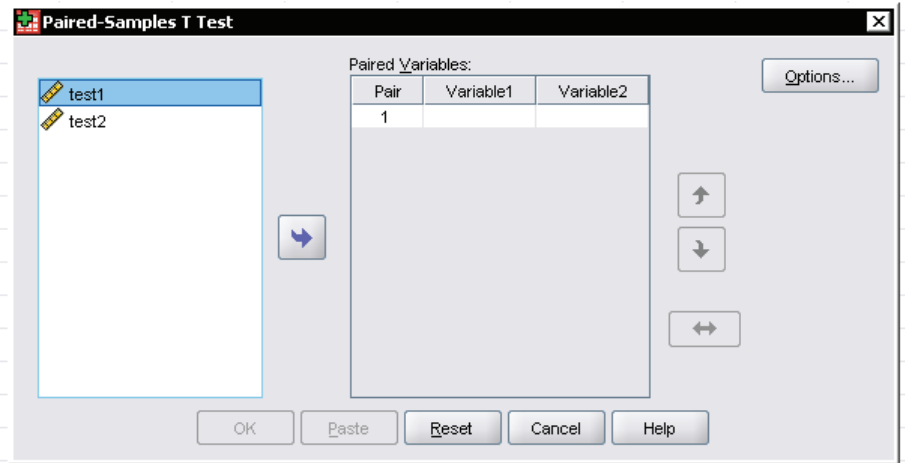
ขั้นที่ 2. เลือกคำสั่ง Analyze / Compare Means / Paired–Samples T Test

	test1	test2
1	76.00	81.00
2	60.00	52.00
3	85.00	87.00
4	58.00	70.00
5	91.00	86.00
6	75.00	77.00
7	82.00	90.00
8	64.00	63.00
9	79.00	85.00
10	88.00	83.00




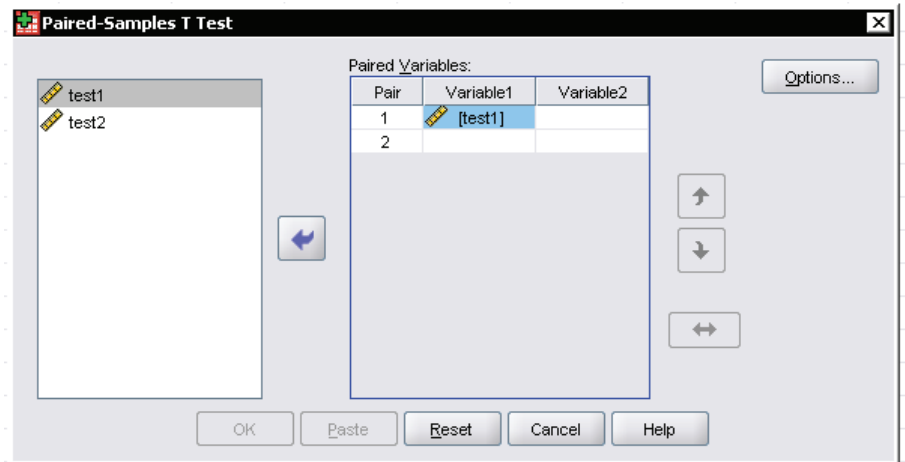
ขั้นที่ 3.

คลิกที่ Paired-Samples T Test
 จะได้เมนูย่อยของคำสั่ง
 Paired-Samples T Test
 ดังนี้



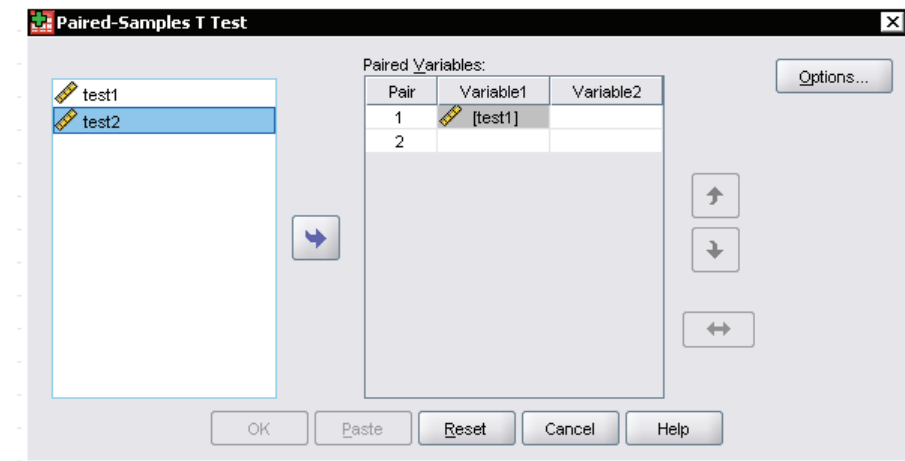
ขั้นที่ 4. การเลือกตัวแปร

คลิกที่ตัวแปร test1
 คลิกที่ 
 ตัวแปร test1 จะมาอยู่ที่
 ตำแหน่ง Variable 1 ก่อน
 เพื่อรอการเลือก Variable 2



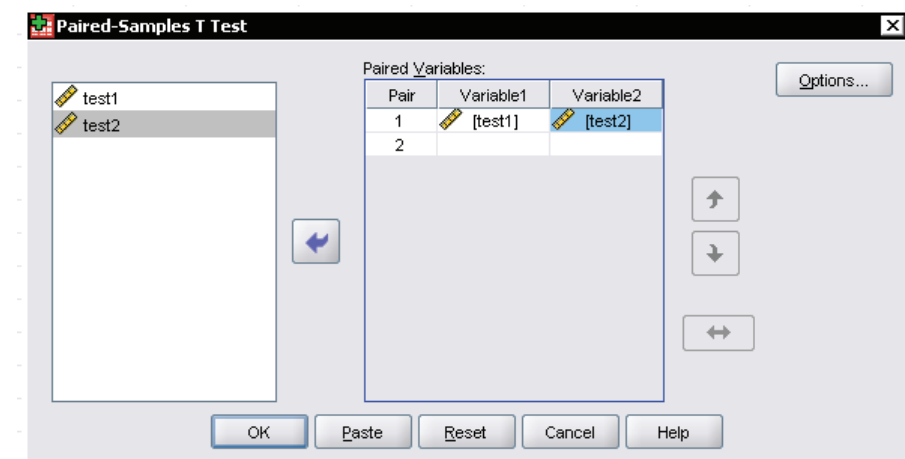
ขั้นที่ 5.

คลิกที่ตัวแปร test2



ขั้นที่ 6. คลิกที่ 

จะได้คู่ของตัวแปรที่ต้องการในช่อง
 Paired Variables: test1 – test2



ขั้นที่ 7. คลิก OK จะได้ผลการคำนวณดังนี้

The screenshot shows the SPSS Viewer interface. On the left, a tree view shows the 'Output' folder containing 'T-Test', which is expanded to show 'Paired Samples Statistics'. The main window displays the 'T-Test' results for the dataset 'C:\Documents and Settings\dumrong\My Documents\'. The 'Paired Samples Statistics' table is as follows:

	Mean	N	Std. Deviation	Std. Error Mean
Pair 1 test1	75.8000	10	11.64092	3.68118
test2	77.4000	10	12.17648	3.85054

ผลการคำนวณทั้งหมดคือ

T-Test

[DataSet1] C:\Documents and Settings\dumrong\My Documents\datasps16\example9.sav

Paired Samples Statistics

	Mean	N	Std. Deviation	Std. Error Mean
Pair 1 test1	75.8000	10	11.64092	3.68118
test2	77.4000	10	12.17648	3.85054

Paired Samples Correlations

	N	Correlation	Sig.
Pair 1 test1 & test2	10	.857	0.00151731

Paired Samples Test

		Pair 1	
		test1 - test2	
Paired Differences	Mean	-1.60000	
	Std. Deviation	6.38053	
	Std. Error Mean	2.01770	
	95% Confidence Interval of the Difference	Lower	-6.16435
		Upper	2.96435
t		-.793	
df		9	
Sig. (2-tailed)		0.44818721	

สรุป ช่วงความเชื่อมั่น 95% ของผลต่างค่าเฉลี่ย μ_D คือ $-6.16435 < \mu_D < 2.96435$

หมายเหตุ การแปลความหมายของตาราง Paired Sample Correlations

ตัวแปร test1 และ test2 มีค่าสหสัมพันธ์เชิงเส้นเท่ากับ 0.857

จากค่า Sig. = 0.001515 ซึ่งต่ำกว่า $\alpha = 0.05$ เราสามารถสรุปผลทางด้านสถิติได้ว่าตัวแปร test1 และ test2

มีความสัมพันธ์กันในรูปแบบเชิงเส้นตรง ที่ระดับนัยสำคัญ $\alpha = 0.05$

เพราะฉะนั้นคะแนนสอบครั้งที่ 1 และ 2 มีความสัมพันธ์กันในรูปแบบเชิงเส้นตรงที่ระดับนัยสำคัญ $\alpha = 0.05$

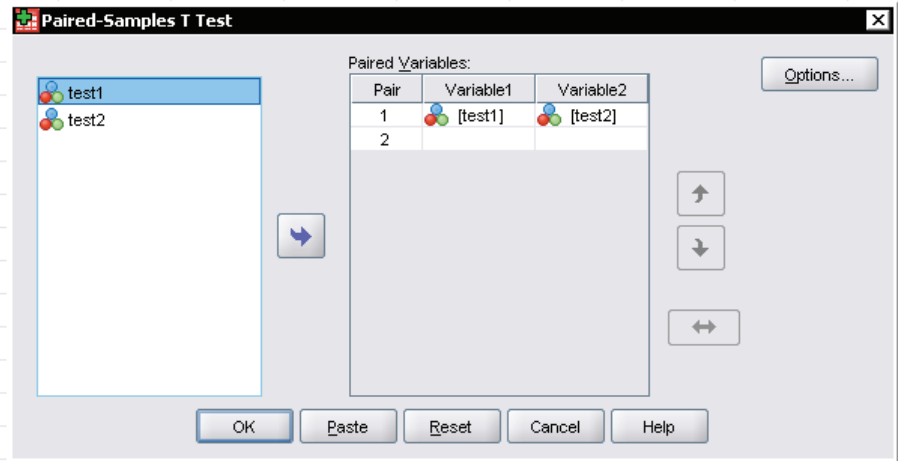
การเปลี่ยนเปอร์เซ็นต์ของช่วงความเชื่อมั่นในการใช้งานของคำสั่ง Analyze / Compare Means / Paired – Samples T Test

จาก ขั้นที่ 6.

เมื่อเลือกตัวแปร test1 – test2

ไว้ที่ช่อง Paired Variables

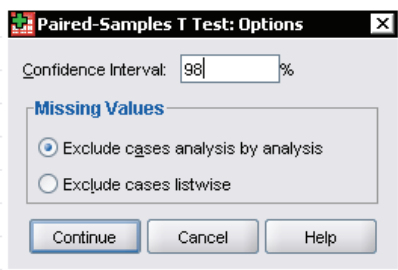
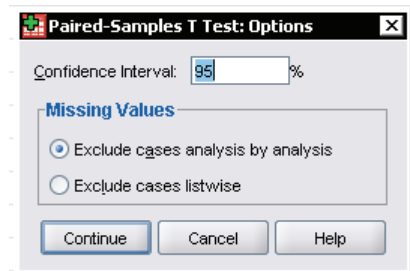
เสร็จแล้ว



ขั้นที่ 6.1 ให้คลิกที่ Options

จะได้เมนูย่อย Paired Samples T Test Options

ขั้นที่ 6.2 ให้เปลี่ยนเปอร์เซ็นต์จาก 95% เป็น 98%



ขั้นที่ 6.3 คลิก Continue

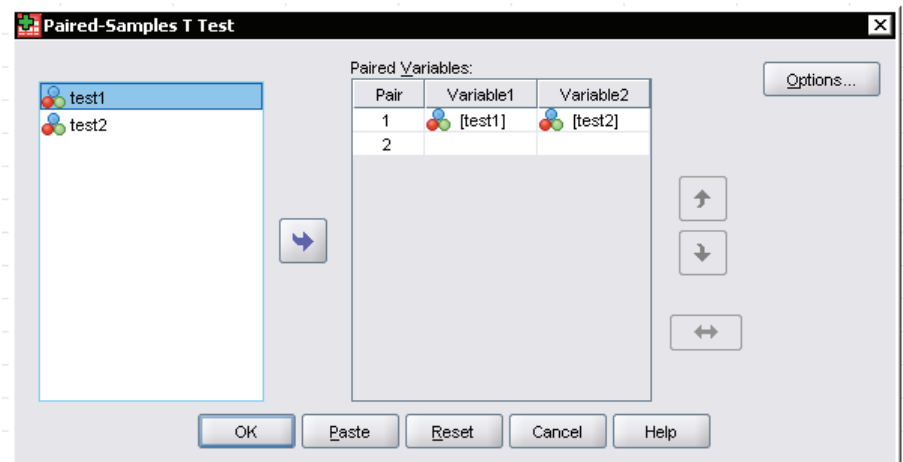
จอภาพจะกลับมาที่เมนูย่อย

Paired-Samples T Test

ขั้นที่ 6.4 คลิก OK จะได้

ผลการคำนวณในส่วนของตาราง

Paired Samples Test เป็นดังนี้



Paired Samples Test

		Pair 1	
		test1 - test2	
Paired Differences	Mean	-1.60000	
	Std. Deviation	6.38053	
	Std. Error Mean	2.01770	
	98% Confidence Interval of the Difference	Lower	-7.29281
		Upper	4.09281
t		-.793	
df		9	
Sig. (2-tailed)		0.44818721	

สรุปช่วงความเชื่อมั่น 98% ของผลต่างค่าเฉลี่ย μ_D คือ $-7.29281 < \mu_D < 4.09281$

6.4 การวิเคราะห์ข้อมูลด้วยคำสั่ง Analyze / Compare Means / Means

จากตัวอย่าง 6.2.1 ทำการทดลองสุ่มตัวอย่างข้อมูล 2 ชุด ตัวอย่างขนาด $n_1 = 9$

จากประชากรชุดที่ 1 มีข้อมูลเป็นดังนี้

61.36	57.76	71.94	61.77	58.66	71.61	71.52	58.67	62.77
-------	-------	-------	-------	-------	-------	-------	-------	-------

ตัวอย่างขนาด $n_2 = 16$ จากประชากรชุดที่ 2 มีข้อมูลเป็นดังนี้

56.92	58.30	67.48	53.96	62.00	59.61	52.02	61.60
64.83	58.55	52.53	64.74	55.51	66.18	55.51	54.18

โดยการใช้ คำสั่ง Analyze / Compare Means / Means

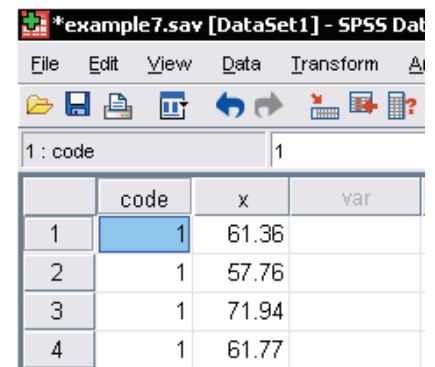
วิธีทำ ขั้นที่ 1. สร้างแฟ้มข้อมูล

โดยกำหนดให้มีตัวแปร 2 ตัวคือ

ตัวแปรจำแนกกลุ่มตัวอย่าง (code)

และ ตัวแปรข้อมูล (x)

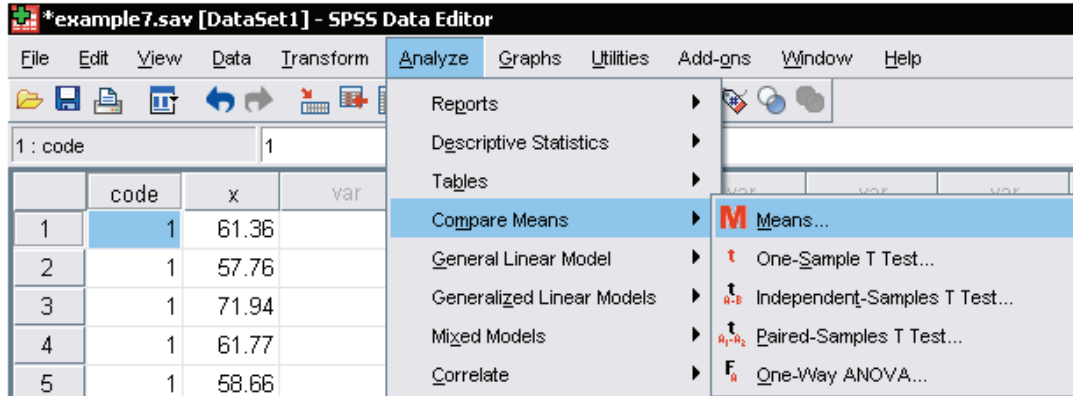
แล้วบันทึกไว้ที่แฟ้มชื่อ example7.sav



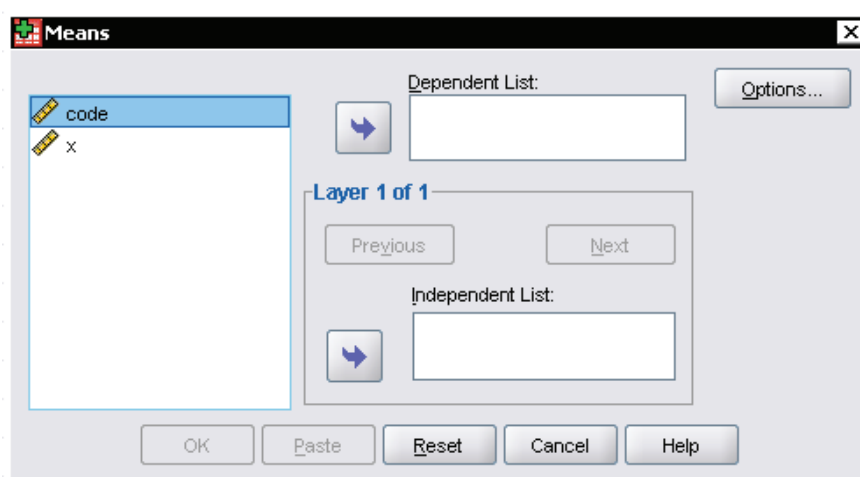
*example7.sav [DataSet1] - SPSS Data Editor

	code	x	var
1	1	61.36	
2	1	57.76	
3	1	71.94	
4	1	61.77	

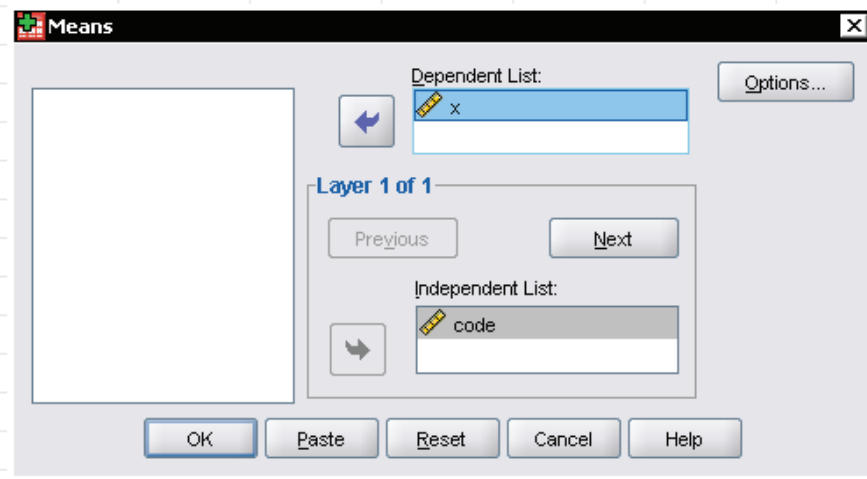
ขั้นที่ 2. เลือกคำสั่ง Analyze / Compare Means / Means



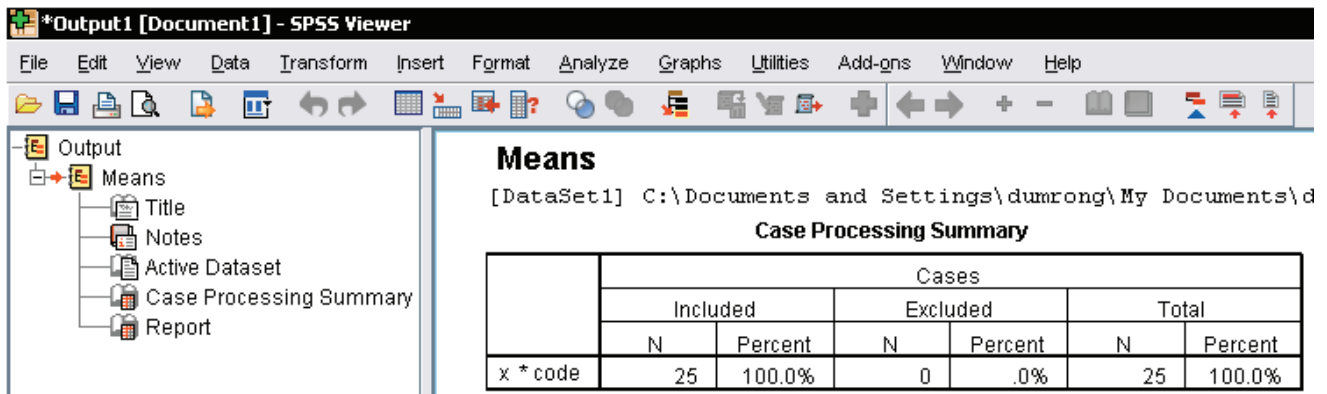
ขั้นที่ 3. คลิกที่ Means จะได้เมนูย่อยของคำสั่ง Analyze / Compare Means / Means



ขั้นที่ 4. เลือกตัวแปร x มาไว้ที่ช่อง Dependent List และ เลือกตัวแปร code มาไว้ที่ช่อง Independent List



ขั้นที่ 5. เสร็จแล้วคลิก OK จะได้ผลบนจอภาพเป็นดังนี้



ผลการคำนวณที่ได้คือ

Means

[DataSet1] C:\Documents and Settings\dumrong\My Documents\dataspss16\example7.sav

Case Processing Summary

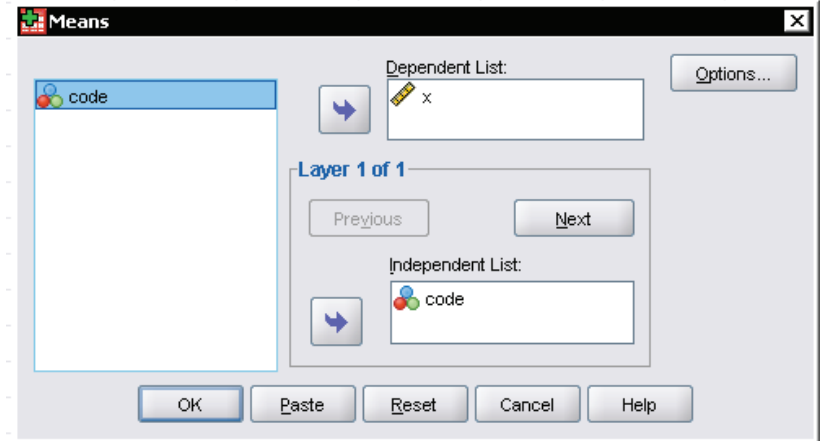
	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
x * code	25	100.0%	0	.0%	25	100.0%

Report

x			
code	Mean	N	Std. Deviation
1	64.0067	9	5.98773
2	58.9950	16	5.00084
Total	60.7992	25	5.79735

การเพิ่มเติมความสามารถในการคำนวณค่าสถิติต่างๆ ของคำสั่ง Analyze / Compare Means / Means จากขั้นที่ 4.

เมื่อเลือกตัวแปรเสร็จแล้ว



ขั้นที่ 4.1 ให้คลิก Options บนจอภาพ

จะขึ้นเมนูย่อยดังนี้

1. คลิกที่ค่าสถิติในช่อง Statistics

เช่นค่า Variance

คลิกที่  จะทำให้ค่าสถิติ Variance มาอยู่ที่ช่อง Cell Statistics

ในที่นี้จะขอเลือกค่าสถิติเพิ่มจากของเดิม คือ Variance, Minimum และ Maximum

2. คลิกที่ช่อง Anova table and eta

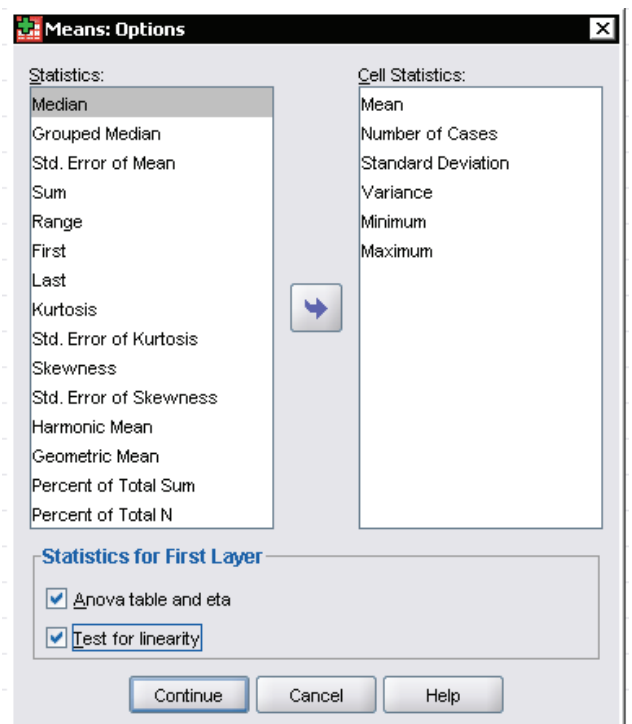
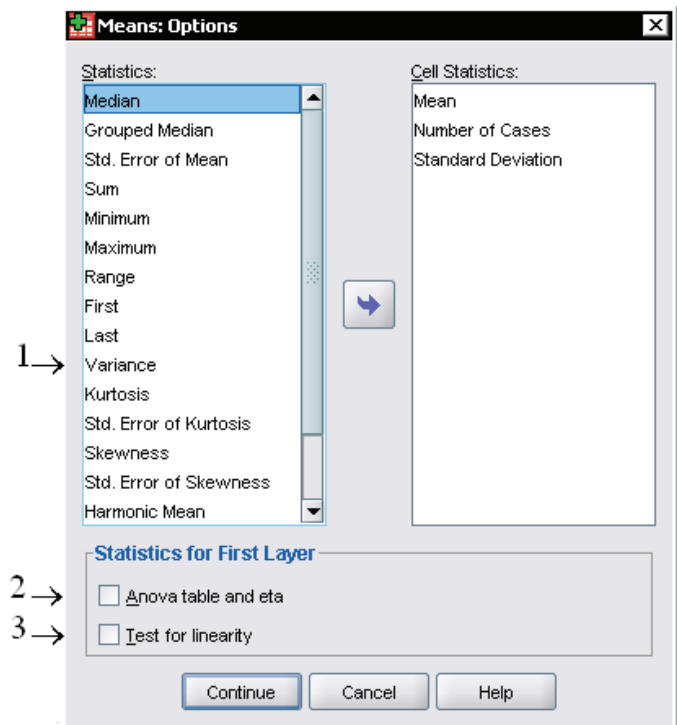
จะได้ตาราง Anova และ ค่าสถิติ eta เพื่อสรุปผล ค่าเฉลี่ย

และความสัมพันธ์ของข้อมูล

3. คลิกที่ Test for linearity เพื่อคำนวณค่าสถิติเกี่ยวกับความสัมพันธ์ของตัวแปร

ขั้นที่ 4.2 คลิก Continue

เพื่อกลับไปเมนู Means



ขั้นที่ 4.3 คลิก OK ซึ่งจะได้ผลการคำนวณทั้งหมดที่ได้คือ

Means

[DataSet1] C:\Documents and Settings\dumrong\My Documents\datspss16\example7.sav

Case Processing Summary

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
x * code	25	100.0%	0	.0%	25	100.0%

Report

code	Mean	N	Std. Deviation	Variance	Minimum	Maximum
1	64.0067	9	5.98773	35.853	57.76	71.94
2	58.9950	16	5.00084	25.008	52.02	67.48
Total	60.7992	25	5.79735	33.609	52.02	71.94

ANOVA Table^a

		Sum of Squares	df	Mean Square	F	Sig.
x * code	Between Groups (Combined)	144.673	1	144.673	5.027	0.03489679
	Within Groups	661.949	23	28.780		
	Total	806.621	24			

a. With fewer than three groups, linearity measures for x * code cannot be computed.

Measures of Association

	Eta	Eta Squared
x * code	.424	.179

การสรุปผลเกี่ยวกับสมมติฐานโดยใช้ข้อมูลในตาราง ANOVA

เกณฑ์ในการสรุปผล ถ้า $Sig. < \alpha$ แล้วค่าเฉลี่ยของประชากรทั้ง 2 ชุดไม่เท่ากัน ที่ระดับนัยสำคัญ α

เพราะว่า $Sig. = 0.03489679$ น้อยกว่า $\alpha = 0.05$

เพราะฉะนั้นค่าเฉลี่ยของประชากรทั้ง 2 ชุดไม่เท่ากัน ที่ระดับนัยสำคัญ $\alpha = 0.05$

หมายเหตุ สูตรที่มาของค่าต่างๆ ในตาราง ANOVA ขอให้ดูในบทที่ 9 เรื่องการวิเคราะห์ความแปรปรวน

การสรุปผลเกี่ยวกับสมมติฐานโดยใช้ข้อมูลในตาราง Measures of Association

Eta เป็นค่าสถิติที่ใช้อธิบายความสัมพันธ์ของประชากรในกลุ่มที่ 1 และ 2 ว่ามีความสัมพันธ์กันมากหรือน้อย

โดยมีสูตร $EtaSquare = \frac{\text{Sum of Squares Between Groups}}{\text{Sum of Square Within Groups}} = \frac{144.6728}{806.6214} = 0.179356$ และ ค่า Eta = 0.423505

เพราะฉะนั้นประชากรในกลุ่มที่ 1 และ 2 มีความสัมพันธ์

6.5 การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่าเฉลี่ย μ ด้วยคำสั่ง

Analyze / Compare Means / One-Way ANOVA

จากตัวอย่าง 6.2.1 ทำการทดลองสุ่มตัวอย่างข้อมูล 2 ชุด ตัวอย่างขนาด $n_1 = 9$

จากประชากรชุดที่ 1 มีข้อมูลเป็นดังนี้

61.36	57.76	71.94	61.77	58.66	71.61	71.52	58.67	62.77
-------	-------	-------	-------	-------	-------	-------	-------	-------

ตัวอย่างขนาด $n_2 = 16$ จากประชากรชุดที่ 2 มีข้อมูลเป็นดังนี้

56.92	58.30	67.48	53.96	62.00	59.61	52.02	61.60
64.83	58.55	52.53	64.74	55.51	66.18	55.51	54.18

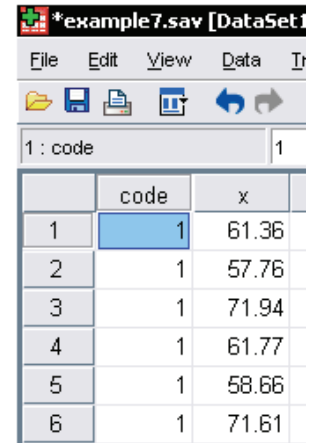
เราสามารถหาช่วงความเชื่อมั่น 95% ของค่าเฉลี่ย μ จำแนกตามกลุ่ม และทดสอบสมมติฐานว่าค่าเฉลี่ยเท่ากันหรือไม่เท่ากัน

โดยการใช้ คำสั่ง Analyze / Compare Means / One-Way ANOVA

วิธีทำ ขั้นที่ 1. สร้างแฟ้มข้อมูล โดยกำหนดให้มีตัวแปร 2 ตัวคือ

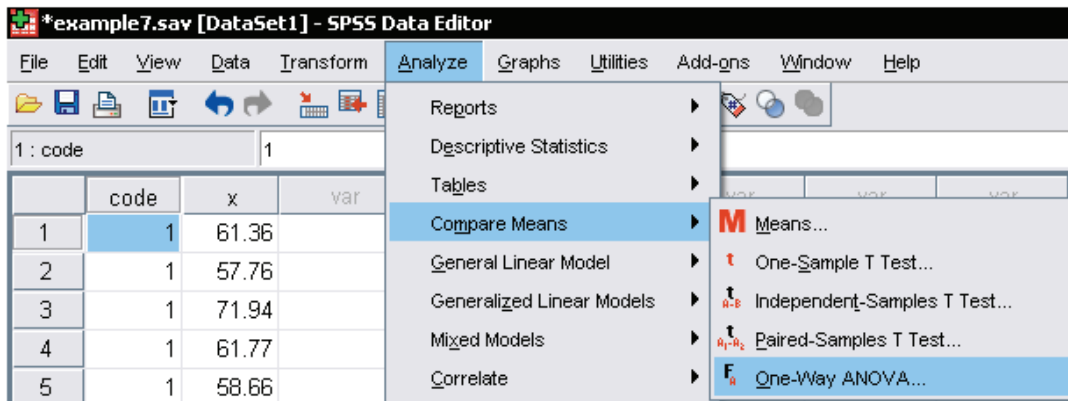
ตัวแปรจำแนกกลุ่มตัวอย่าง (code) และ ตัวแปรข้อมูล (x)

แล้วบันทึกไว้ที่แฟ้มชื่อ example7.sav หรือเปิดแฟ้มที่ Save ไว้แล้ว

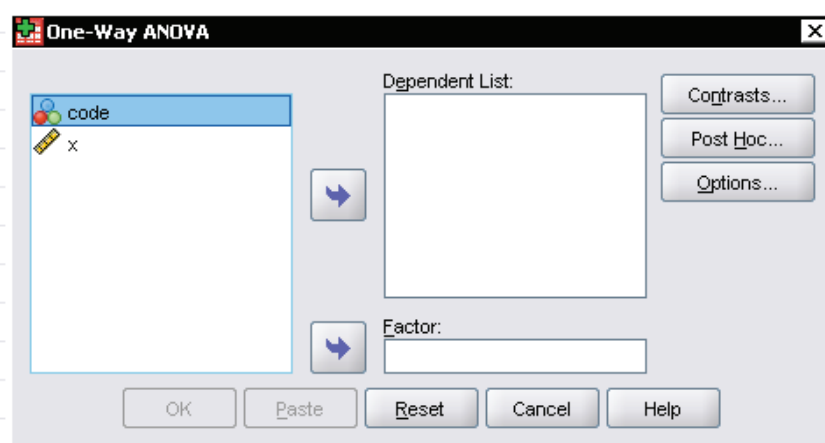


	code	x
1	1	61.36
2	1	57.76
3	1	71.94
4	1	61.77
5	1	58.66
6	1	71.61

ขั้นที่ 2. เลือกคำสั่ง Analyze / Compare Means / One-Way ANOVA

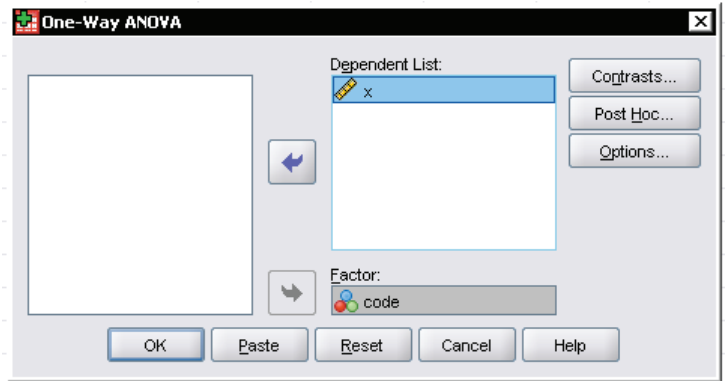


ขั้นที่ 3. คลิกที่ One-Way ANOVA จะได้เมนูย่อยของคำสั่ง One-Way ANOVA ดังนี้



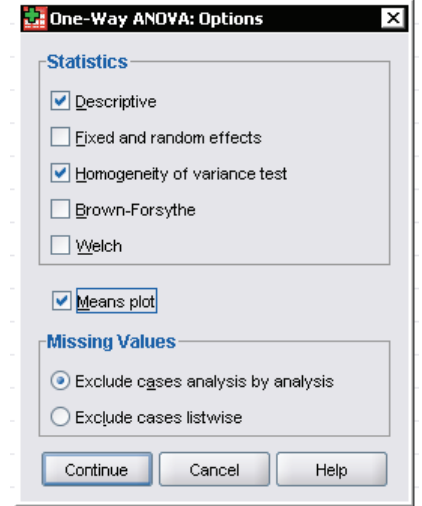
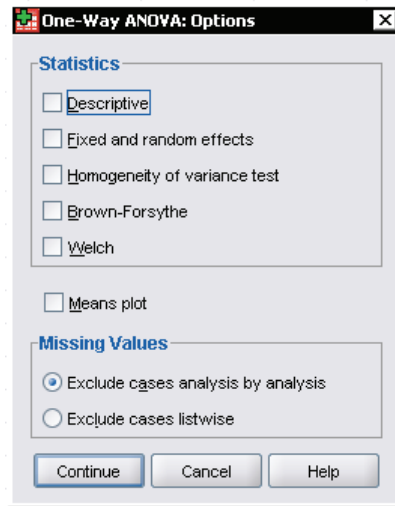
ขั้นที่ 4. เลือกตัวแปร code

มาไว้ที่ช่อง Factor
และ เลือกตัวแปร x
มาไว้ที่ช่อง Dependent List



ขั้นที่ 5. คลิก Options

เพื่อไปเมนูย่อยของการคำนวณค่าสถิติเพิ่มเติม



ขั้นที่ 6.

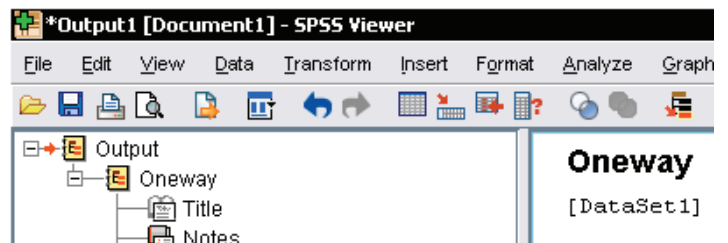
- เลือก Descriptive เพื่อคำนวณค่าสถิติเบื้องต้น และช่วงความเชื่อมั่น 95 % ของข้อมูลจำแนกตามกลุ่ม
- เลือก Homogeneity-of-variance test เพื่อคำนวณค่า

Levene Statistics สำหรับสรุปผลเกี่ยวกับความแปรปรวนของข้อมูลว่าเท่ากันหรือไม่

- เลือก Means plot เพื่อเขียนกราฟของค่าเฉลี่ยแต่ละกลุ่ม

ขั้นที่ 7. คลิก Continue และ OK ตามลำดับ

จะได้ผลการคำนวณเป็นดังนี้



ผลการคำนวณทั้งหมดคือ

Oneway

[DataSet1] C:\Documents and Settings\dumrong\My Documents\dataspss16\example7.sav

Descriptives

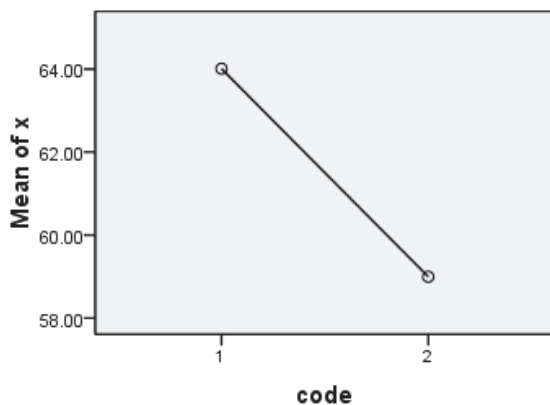
x		1	2	Total
N		9	16	25
Mean		64.0067	58.9950	60.7992
Std. Deviation		5.98773	5.00084	5.79735
Std. Error		1.99591	1.25021	1.15947
95% Confidence Interval for Mean	Lower Bound	59.4041	56.3302	58.4062
	Upper Bound	68.6092	61.6598	63.1922
Minimum		57.76	52.02	52.02
Maximum		71.94	67.48	71.94

Test of Homogeneity of Variances

Levene Statistic	df1	df2	Sig.
.800	1	23	0.38038311

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	144.673	1	144.673	5.027	0.03489679
Within Groups	661.949	23	28.780		
Total	806.621	24			



จากตาราง Descriptives จะได้ว่า

ช่วงความเชื่อมั่น 95% ของค่าเฉลี่ยของประชากรชุดที่ 1 คือ (59.4041, 68.6092)

ช่วงความเชื่อมั่น 95% ของค่าเฉลี่ยของประชากรชุดที่ 2 คือ (56.3302, 61.6598)

ช่วงความเชื่อมั่น 95% ของค่าเฉลี่ยของประชากรชุดที่ 1 และ 2 รวมกันคือ (58.4062, 63.1922)

หมายเหตุ

1. การสรุปผลเกี่ยวกับสมมติฐานโดยใช้ข้อมูลในตาราง Test of Homogeneity of variances

เกณฑ์ในการสรุปผล ถ้า $\text{Sig.} < \alpha$ แล้วความแปรปรวนของประชากรทั้ง 2 ชุดไม่เท่ากัน ที่ระดับนัยสำคัญ α เพราะว่า $\text{Sig.} = 0.3803831$ ไม่น้อยกว่า $\alpha = 0.05$

เพราะฉะนั้นยอมรับว่าความแปรปรวนของประชากรทั้ง 2 ชุดเท่ากัน ที่ระดับนัยสำคัญ $\alpha = 0.05$

2. การสรุปผลเกี่ยวกับสมมติฐานโดยใช้ข้อมูลในตาราง ANOVA

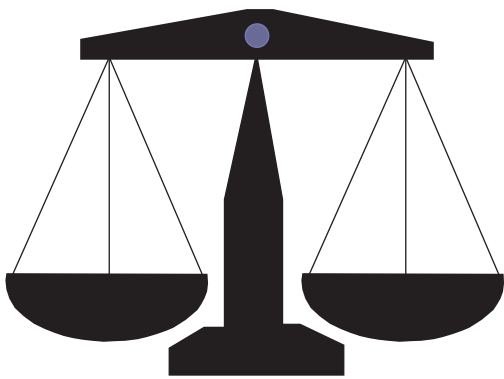
เกณฑ์ในการสรุปผล ถ้า $\text{Sig.} < \alpha$ แล้วค่าเฉลี่ยของประชากรทั้ง 2 ชุดไม่เท่ากัน ที่ระดับนัยสำคัญ α เพราะว่า $\text{Sig.} = 0.03489679$ น้อยกว่า $\alpha = 0.05$

เพราะฉะนั้นค่าเฉลี่ยของประชากรทั้ง 2 ชุดไม่เท่ากัน ที่ระดับนัยสำคัญ $\alpha = 0.05$

3. สูตรที่มาของค่าต่าง ๆ ในตาราง ANOVA และ การทดสอบค่าเฉลี่ยของประชากรหลายกลุ่มเท่ากันโดยใช้ผลการคำนวณจากตาราง ANOVA ขอให้อ่านในบทที่ 9 เรื่องการวิเคราะห์ความแปรปรวน

บทที่ 7

การทดสอบสมมติฐาน



การทดสอบสมมติฐานเป็นการทำงานทางสถิติที่สำคัญ เราจะทำการทดสอบสมมติฐานเมื่อเรามีข้อสงสัยเกี่ยวกับลักษณะต่างๆ ของประชากรเช่น ประชากรมีค่าเฉลี่ย $\mu = 50$ จริงหรือไม่ ค่าเฉลี่ยประชากรของประชากร 2 ชุดเท่ากันหรือไม่ การฉีดวัคซีนป้องกันอหิวาต์กับการเป็นโรคอหิวาต์เกี่ยวข้องกันหรือไม่ การนับถือศาสนาและถิ่นที่อยู่เกี่ยวข้องกันหรือไม่ ฯลฯ

การทดสอบสมมติฐานเป็นการนำข้อมูลตัวอย่างซึ่งอาจได้มาจากแบบสอบถาม การทดลอง เมื่อได้ข้อมูลตัวอย่างมาแล้วจึงทำการคำนวณค่าสถิติจากตัวอย่าง เพื่อนำมาช่วยในการตัดสินใจโดยใช้เหตุผลทางสถิติว่าจะยอมรับหรือปฏิเสธสมมติฐานที่เรากำหนดไว้ ตัวอย่างการทดสอบสมมติฐานแบบต่างๆ เช่น

การทดสอบสมมติฐานว่า $\mu = \mu_0$ จริงหรือไม่

การทดสอบสมมติฐานว่า $\mu_1 = \mu_2$ จริงหรือไม่

การทดสอบสมมติฐานว่า $\mu_D = 0$ จริงหรือไม่

การทดสอบสมมติฐานว่า $\sigma^2 = \sigma_0^2$ จริงหรือไม่

การทดสอบสมมติฐานว่า $\sigma_1^2 = \sigma_2^2$ จริงหรือไม่

การทดสอบสมมติฐานว่า $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$ จริงหรือไม่

การทดสอบสมมติฐานว่าข้อมูลเป็นอิสระต่อกัน จริงหรือไม่

การทดสอบสมมติฐานว่าข้อมูลมีการแจกแจงตามที่คาดไว้ (การทดสอบภาวะสารูปสนิทธิ) จริงหรือไม่

ในบทที่ 7 นี้เราจึงศึกษาเกี่ยวกับการทดสอบสมมติฐานแบบต่างๆ ทั้งหลักการขั้นตอนการทำงานทางทฤษฎี และ การนำ SPSS for Windows เข้ามาช่วยในการคำนวณ

7.1 การทดสอบสมมติฐาน $H_0 : \mu = \mu_0$

หลักการและขั้นตอนการทำงานทางทฤษฎีในหนังสือความน่าจะเป็นและสถิติ

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \mu = \mu_0$

กำหนดสมมติฐานอื่น $H_1 : \mu \neq \mu_0$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. ทำการสุ่มตัวอย่างขนาด n คำนวณค่าเฉลี่ย \bar{x} และ ค่าความแปรปรวนของตัวอย่าง s^2

ขั้นที่ 4. เลือกค่าสถิติที่เหมาะสม Z หรือ t ระดับชั้นความเสรี $df = n - 1$

ขั้นที่ 5. คำนวณค่าสถิติ $z_{\text{คำนวณ}}$ หรือ $t_{\text{คำนวณ}}$ ตามที่เลือกในขั้นที่ 4. จากข้อมูลตัวอย่าง

ขั้นที่ 6. เปิดตารางสถิติเพื่อหา ค่าวิกฤต และ บริเวณวิกฤต

6.1 กรณีใช้ค่า Z ค่าวิกฤตคือ $-z_{\frac{\alpha}{2}}$ และ $z_{\frac{\alpha}{2}}$ บริเวณวิกฤตคือ $Z < -z_{\frac{\alpha}{2}}$ หรือ $Z > z_{\frac{\alpha}{2}}$

6.2 กรณีใช้ค่า t ค่าวิกฤตคือ $-t_{\frac{\alpha}{2}}$ และ $t_{\frac{\alpha}{2}}$ บริเวณวิกฤตคือ $t < -t_{\frac{\alpha}{2}}$ หรือ $t > t_{\frac{\alpha}{2}}$

ขั้นที่ 7. สรุปผลโดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต โดยมีเกณฑ์การสรุปผล

7.1 กรณีใช้ค่า Z ถ้า $z_{\text{คำนวณ}} < -z_{\frac{\alpha}{2}}$ หรือ $z_{\text{คำนวณ}} > z_{\frac{\alpha}{2}}$ แล้วปฏิเสธ H_0

7.2 กรณีใช้ค่า t ถ้า $t_{\text{คำนวณ}} < -t_{\frac{\alpha}{2}}$ หรือ $t_{\text{คำนวณ}} > t_{\frac{\alpha}{2}}$ แล้วปฏิเสธ H_0

เกณฑ์การเลือกค่าสถิติที่เหมาะสมของขั้นที่ 4. (ในทางทฤษฎี)

1. กรณีประชากรมีการแจกแจงปกติ และ รู้ค่าความแปรปรวน σ^2 เลือก $Z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$

2. กรณีประชากรมีการแจกแจงปกติ และ ไม่รู้ค่าความแปรปรวน σ^2

2.1 ตัวอย่างขนาด $n \geq 30$ แทนค่า σ ด้วย s เลือก $Z = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$

2.2 ตัวอย่างขนาด $n < 30$ เลือก $t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$

3. กรณีไม่ได้กำหนดว่าข้อมูลมีการแจกแจงปกติ และ รู้ค่าความแปรปรวน σ^2

ตัวอย่างขนาด $n \geq 30$ เลือก $Z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$

4. กรณีไม่ได้กำหนดว่าข้อมูลมีการแจกแจงปกติ และ ไม่รู้ค่าความแปรปรวน σ^2

ต้องใช้ตัวอย่างขนาด $n \geq 30$ แทนค่า σ ด้วย s เลือก $Z = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$

หลักการและขั้นตอนของการทดสอบสมมติฐานด้วย SPSS for Windows

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \mu = \mu_0$

กำหนดสมมติฐานอื่น $H_1 : \mu \neq \mu_0$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

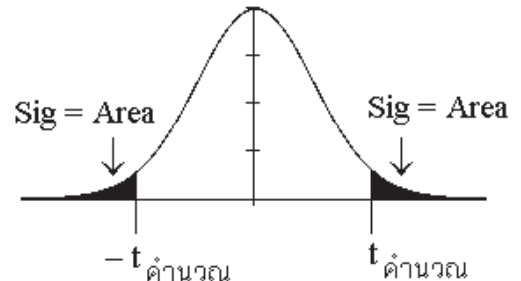
ขั้นที่ 3. ทำการสุ่มตัวอย่างขนาด n คำนวณค่าเฉลี่ย \bar{x} และ ค่าความแปรปรวนของตัวอย่าง s^2

ขั้นที่ 4. SPSS for Windows เลือกค่าสถิติ t เท่านั้น

ขั้นที่ 5. คำนวณค่าสถิติ $t = \frac{\bar{x} - \mu_0}{\left(\frac{s}{\sqrt{n}}\right)}$ ซึ่งเราเรียกว่า $t_{\text{คำนวณ}}$ และ ระดับขั้นความเสรี $df = n - 1$

ขั้นที่ 6. คำนวณค่า Sig. (2-tailed) ของค่าสถิติ $t_{\text{คำนวณ}}$

$$\begin{aligned} \text{Sig. (2-tailed)} &= 2 \text{ เท่าของพื้นที่ใต้โค้งที่} \\ &\text{ทางทางด้านขวาที่ระยะ } |t_{\text{คำนวณ}}| \\ &= 2P(t > |t_{\text{คำนวณ}}|) \end{aligned}$$



ขั้นที่ 7. การสรุปผลสามารถเลือกใช้เหตุผลได้ 2 วิธีคือ

1. โดยการเปรียบเทียบค่าสถิติ $t_{\text{คำนวณ}}$ จากตัวอย่าง กับ ค่าวิกฤตจากตารางสถิติ โดยมีเกณฑ์การสรุปผลว่า ถ้า $t_{\text{คำนวณ}} < -t_{\frac{\alpha}{2}}$ หรือ $t_{\text{คำนวณ}} > t_{\frac{\alpha}{2}}$ แล้วปฏิเสธ H_0

- หรือ 2. โดยการเปรียบเทียบค่า Sig. (2-tailed) กับ ระดับนัยสำคัญ α โดยมีเกณฑ์การสรุปผลว่า ถ้า Sig. (2-tailed) $< \alpha$ แล้วปฏิเสธ H_0

หมายเหตุ กรณีเป็นการทดสอบแบบ 1 ทาง

	$H_1 : \mu < \mu_0$	$H_1 : \mu > \mu_0$
ค่าวิกฤต Z	$-z_\alpha$	z_α
บริเวณวิกฤต	$Z < -z_\alpha$	$Z > z_\alpha$
ค่าวิกฤต t	$-t_\alpha$	t_α
บริเวณวิกฤต	$t < -t_\alpha$	$t > t_\alpha$
สรุปโดยใช้ Sig	ถ้า $t_{\text{คำนวณ}} < 0$, Sig. (2-tailed) $< 2\alpha$ แล้ว ปฏิเสธ H_0	ถ้า $t_{\text{คำนวณ}} > 0$, Sig. (2-tailed) $< 2\alpha$ แล้ว ปฏิเสธ H_0

ตัวอย่าง 7.1.1 เท่าที่ผ่านมานิสิตใช้เวลาลงทะเบียนโดยเฉลี่ย 50 นาที ขณะนี้มหาวิทยาลัยกำลังทดลองให้นิสิตลงทะเบียนเรียนโดยใช้ระบบคอมพิวเตอร์ เพื่อทดสอบว่าเวลาเฉลี่ยที่ใช้ในการลงทะเบียนโดยใช้ระบบคอมพิวเตอร์มีค่าเท่ากับ 50 นาทีหรือไม่ จึงทำการสุ่มตัวอย่างการลงทะเบียนของนิสิต 12 คนได้ข้อมูลดังนี้ 41, 42, 47, 41, 54, 26, 26, 65, 34, 49, 29, 50

กำหนดระดับนัยสำคัญ 0.05

วิธีทำ ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \mu = 50$
กำหนดสมมติฐานอื่น $H_1 : \mu \neq 50$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

ขั้นที่ 3. เข้าสู่ SPSS Data Editor

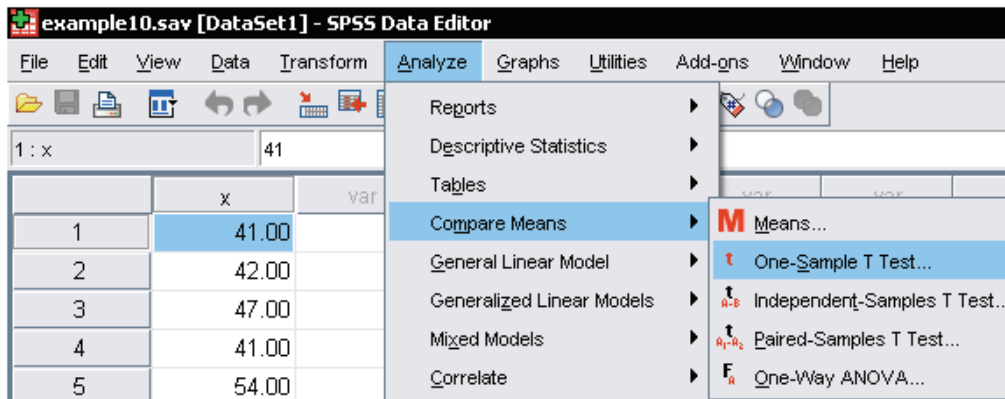
สร้างแฟ้มข้อมูลที่มีตัวแปร x เป็นเวลาที่ใช้ในการลงทะเบียน

และ Save ข้อมูลลงแฟ้มชื่อ example10.sav

ขั้นที่ 4. เลือกค่าสถิติ t

ขั้นที่ 4.1 วิเคราะห์ข้อมูลด้วยคำสั่ง

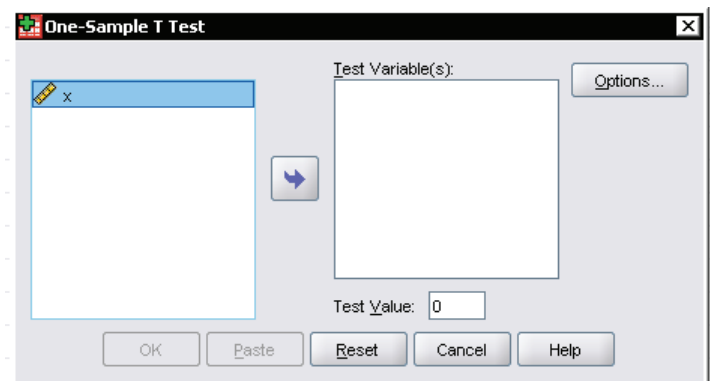
Analyze / Compare Means / One-Sample T Test



	x
1	41.00
2	42.00
3	47.00
4	41.00
5	54.00
6	26.00
7	26.00
8	65.00
9	34.00
10	49.00
11	29.00
12	50.00

ขั้นที่ 4.2 คลิก One-Sample T Test

จะได้ผลบนจอภาพเป็น



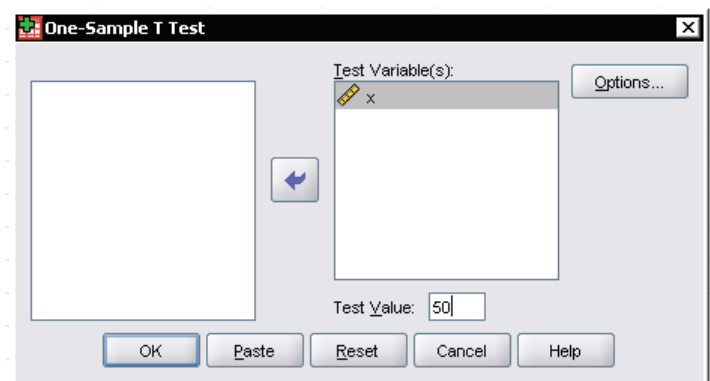
ขั้นที่ 4.3 เลือกตัวแปร x

มาไว้ที่ Test Variable(s)

เพราะว่ากำหนดสมมติฐานหลัก $\mu = 50$

เพราะฉะนั้นในช่อง Test Value

ให้พิมพ์ 50 ตามที่กำหนดสมมติฐานหลักไว้



ขั้นที่ 4.4 คลิก OK

จะได้ผลการคำนวณเป็น

One-Sample Statistics				
	N	Mean	Std. Deviation	Std. Error Mean
x	12	42.0000	11.90111	3.43555

ผลการคำนวณทั้งหมดคือ

T-Test

[DataSet1] C:\Documents and Settings\dumrong\My Documents\datasps16\example10.sav

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
x	12	42.0000	11.90111	3.43555

One-Sample Test

	Test Value = 50					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
x	-2.329	11	0.03996820	-8.00000	-15.5616	-.4384

ขั้นที่ 5. ผลการคำนวณของ SPSS ได้ว่า ค่าสถิติ $t_{\text{คำนวณ}} = -2.329$ ระดับชั้นความเสรี = 11

ขั้นที่ 6. คำนวณค่า Sig. (2-tailed) ของค่าสถิติ t

ผลการคำนวณของ SPSS ได้ว่า Sig. (2-tailed) = 0.03996820

หมายเหตุ ตรวจสอบการคำนวณด้วย Mathcad จะได้ว่าค่าของ

Sig. (2-tailed) = 0.03996820 มาจาก 2 เท่าของ Pvalue(-2.329)

ผลการคำนวณจากโปรแกรม Mathcad

T-distribution

TOL := 0.0000001 v := 11

Pvalue(T) := $0.5 - \int_0^{|T|} h(t) dt$

$$h(t) := \frac{\left(\frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right) \cdot \sqrt{\pi \cdot v}} \right) \cdot \left[1 + \left(\frac{t^2}{v} \right) \right]^{-\frac{v+1}{2}}}{1}$$

Sig_1_Tailed(T) := Pvalue(T)

Sig_1_Tailed(-2.329) = 0.020

Sig_2_Tailed(T) := 2 · Pvalue(T)

Sig_2_Tailed(-2.329) = 0.040

หรือโดยการคำนวณด้วยฟังก์ชันสำเร็จรูปของ Mathcad

One_tailed_Significant(T) := 1 - pt(|T|, 11)

One_tailed_Significant(-2.329) = 0.020

Two_Tailed_Significant(T) := 2 · One_tailed_Significant(T)

Two_Tailed_Significant(-2.329) = 0.040

ขั้นที่ 7. สรุปผล แบบที่ 1. โดยการเปรียบเทียบ $t_{\text{คำนวณ}} = -2.329$ กับ ค่าวิกฤต $t_{0.025, df=11} = 2.201$

เพราะว่า $H_1 : \mu \neq 50$ เพราะฉะนั้นบริเวณวิกฤตคือ $t < -2.201$ หรือ $t > 2.201$ สรุปปฏิเสธ H_0

แบบที่ 2. โดยการเปรียบเทียบค่า Sig(2 - tailed) กับ ระดับนัยสำคัญ α

เพราะว่า Sig(2 - tailed) = 0.02 < 0.05 = α เพราะฉะนั้นปฏิเสธ H_0

สรุปผลการทดสอบสมมติฐานจากข้อมูลที่เก็บมาได้ต้องปฏิเสธสมมติฐานที่กล่าวไว้ว่า

เวลาเฉลี่ยที่ใช้ในการลงทะเบียนโดยใช้ระบบคอมพิวเตอร์มีค่าเท่ากับ 50 นาที ที่ระดับนัยสำคัญ 0.05

7.2 การทดสอบสมมติฐาน $H_0 : \mu_1 = \mu_2$ กรณีที่ประชากร 2 ชุดเป็นอิสระต่อกัน

หลักการและขั้นตอนการทำงานทางทฤษฎีเกี่ยวกับการทดสอบสมมติฐาน

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \mu_1 = \mu_2$

กำหนดสมมติฐานอื่น $H_1 : \mu_1 \neq \mu_2$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. สุ่มตัวอย่างขนาด n_1 จากประชากรชุดที่ 1 หาค่าเฉลี่ยของตัวอย่าง \bar{x}_1

สุ่มตัวอย่างขนาด n_2 จากประชากรชุดที่ 2 หาค่าเฉลี่ยของตัวอย่าง \bar{x}_2

ขั้นที่ 4. เลือกค่าสถิติที่เหมาะสม Z หรือ t

ขั้นที่ 5. คำนวณค่าสถิติ $Z_{\text{คำนวณ}}$ หรือ $t_{\text{คำนวณ}}$ ตามที่เลือกในขั้นที่ 4. จากข้อมูลตัวอย่าง

ขั้นที่ 6. เปิดตารางสถิติเพื่อหา ค่าวิกฤต และ บริเวณวิกฤต

6.1 กรณีใช้ค่า Z ค่าวิกฤตคือ $-z_{\frac{\alpha}{2}}$ และ $z_{\frac{\alpha}{2}}$ บริเวณวิกฤตคือ $Z < -z_{\frac{\alpha}{2}}$ หรือ $Z > z_{\frac{\alpha}{2}}$

6.2 กรณีใช้ค่า t ค่าวิกฤตคือ $-t_{\frac{\alpha}{2}}$ และ $t_{\frac{\alpha}{2}}$ บริเวณวิกฤตคือ $t < -t_{\frac{\alpha}{2}}$ หรือ $t > t_{\frac{\alpha}{2}}$

ขั้นที่ 7. สรุปผลโดยการเปรียบเทียบค่าสถิติจากตัวอย่างกับค่าวิกฤต โดยมีเกณฑ์การสรุปผลว่า

7.1 กรณีใช้ค่า Z ถ้า $Z_{\text{คำนวณ}} < -z_{\frac{\alpha}{2}}$ หรือ $Z_{\text{คำนวณ}} > z_{\frac{\alpha}{2}}$ แล้วปฏิเสธ H_0

7.2 กรณีใช้ค่า t ถ้า $t_{\text{คำนวณ}} < -t_{\frac{\alpha}{2}}$ หรือ $t_{\text{คำนวณ}} > t_{\frac{\alpha}{2}}$ แล้วปฏิเสธ H_0

เกณฑ์การเลือกค่าสถิติที่เหมาะสมของขั้นที่ 4.

1. กรณี $n_1 \geq 30$ และ $n_2 \geq 30$

1.1 กรณีประชากร 2 ชุดมีการแจกแจงปกติ และ รู้ค่าความแปรปรวน σ_1^2 และ σ_2^2

$$\text{เลือกใช้ค่าสถิติ } Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

1.2. กรณีประชากร 2 ชุดมีการแจกแจงปกติ และ ไม่รู้ค่าความแปรปรวน σ_1^2 และ σ_2^2

หาค่าความแปรปรวนของตัวอย่าง s_1^2 และ ประมาณ σ_1^2 ด้วย s_1^2

หาค่าความแปรปรวนของตัวอย่าง s_2^2 และ ประมาณ σ_2^2 ด้วย s_2^2

$$\text{เลือกใช้ค่าสถิติ } Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

2. กรณี $n_1 < 30$ หรือ $n_2 < 30$

2.1 กรณีประชากร 2 ชุดมีการแจกแจงปกติ และ รู้ค่าความแปรปรวน σ_1^2 และ σ_2^2

$$\text{เลือกใช้ค่าสถิติ } Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

2.2 กรณีประชากร 2 ชุดมีการแจกแจงปกติ และ ไม่รู้ค่าความแปรปรวน σ_1^2 และ σ_2^2
 หาค่าความแปรปรวนของตัวอย่าง s_1^2 และ s_2^2

2.2.1 ภายใต้ข้อกำหนด $\sigma_1^2 = \sigma_2^2$

$$\text{เลือกใช้ค่าสถิติ } t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \text{ เมื่อ } s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \text{ และ } df = n_1 + n_2 - 2$$

2.2.2 ภายใต้ข้อกำหนด $\sigma_1^2 \neq \sigma_2^2$

$$\text{เลือกใช้ค่าสถิติ } t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \text{ เมื่อ } df = \frac{(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2})^2}{(\frac{s_1^2}{n_1})^2 \frac{1}{(n_1 - 1)} + (\frac{s_2^2}{n_2})^2 \frac{1}{(n_2 - 1)}}$$

หลักการและขั้นตอนการทดสอบสมมติฐานด้วย SPSS for Windows

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \mu_1 = \mu_2$

กำหนดสมมติฐานอื่น $H_1 : \mu_1 \neq \mu_2$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. สุ่มตัวอย่างขนาด n_1 จากประชากรชุดที่ 1 และ สุ่มตัวอย่างขนาด n_2 จากประชากรชุดที่ 2
 หาค่าเฉลี่ยของตัวอย่าง \bar{x}_1 และ \bar{x}_2 หาค่าความแปรปรวนของตัวอย่าง s_1^2 และ s_2^2

ขั้นที่ 4. เลือกค่าสถิติ t เท่านั้น

ภายใต้ข้อกำหนด $\sigma_1^2 = \sigma_2^2$

$$\text{เลือกใช้ค่าสถิติ } t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$\text{เมื่อ } s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \text{ และ } df = n_1 + n_2 - 2$$

ภายใต้ข้อกำหนด $\sigma_1^2 \neq \sigma_2^2$

$$\text{เลือกใช้ค่าสถิติ } t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \text{ เมื่อ } df = \frac{(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2})^2}{(\frac{s_1^2}{n_1})^2 \frac{1}{(n_1 - 1)} + (\frac{s_2^2}{n_2})^2 \frac{1}{(n_2 - 1)}}$$

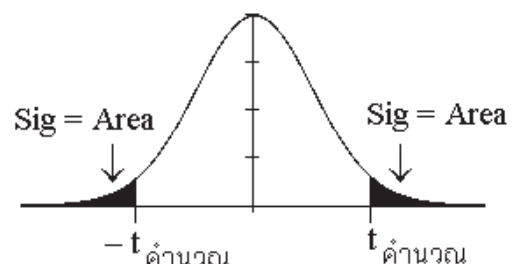
ขั้นที่ 5. คำนวณค่าสถิติ t ซึ่งเราเรียกว่า $t_{\text{คำนวณ}}$ และ ระดับชั้นความเสรี df

ขั้นที่ 6. คำนวณค่า Sig. (2-tailed) ของค่าสถิติ $t_{\text{คำนวณ}}$

Sig. (2-tailed)

$$= 2 \text{ เท่าของพื้นที่ใต้โค้ง ทางทางด้านขวาที่ระยะ } | t_{\text{คำนวณ}} |$$

$$= 2 P(t > | t_{\text{คำนวณ}} |)$$



ขั้นที่ 7. การสรุปผลสามารถเลือกใช้เหตุผลได้ 2 วิธีคือ

1. โดยการเปรียบเทียบค่าสถิติ t จากตัวอย่าง กับ ค่าวิกฤตจากตารางสถิติ โดยมีเกณฑ์การสรุปผลว่า ถ้า $t_{\text{คำนวณ}} < -t_{\frac{\alpha}{2}}$ หรือ $t_{\text{คำนวณ}} > t_{\frac{\alpha}{2}}$ แล้วปฏิเสธ H_0

หรือ 2. โดยการเปรียบเทียบค่า Sig. (2-tailed) กับ ระดับนัยสำคัญ α โดยมีเกณฑ์การสรุปผลว่า ถ้า Sig. (2-tailed) $< \alpha$ แล้วปฏิเสธ H_0

หมายเหตุ กรณีการทดสอบแบบ 1 ทาง

	$H_1 : \mu_1 - \mu_2 < d_0$	$H_1 : \mu_1 - \mu_2 > d_0$
ค่าวิกฤต Z	$-z_\alpha$	z_α
บริเวณวิกฤต	$Z < -z_\alpha$	$Z > z_\alpha$
ค่าวิกฤต t	$-t_\alpha$	t_α
บริเวณวิกฤต	$t < -t_\alpha$	$t > t_\alpha$
สรุปโดยใช้ Sig	ถ้า $t_{\text{คำนวณ}} < 0$, Sig. (2-tailed) $< 2\alpha$ แล้ว ปฏิเสธ H_0	ถ้า $t_{\text{คำนวณ}} > 0$, Sig. (2-tailed) $< 2\alpha$ แล้ว ปฏิเสธ H_0

ตัวอย่าง 7.2.1 โรงงานผลิตแป้งกระป๋องมีเครื่องจักร 2 เครื่อง

ให้ μ_1 เป็นค่าเฉลี่ยประชากรของน้ำหนักของแป้งที่บรรจุในกระป๋องที่ผลิตจากเครื่องจักรที่ 1

μ_2 เป็นค่าเฉลี่ยประชากรของน้ำหนักของแป้งที่บรรจุในกระป๋องที่ผลิตจากเครื่องจักรที่ 2

เพื่อทำการทดสอบว่าค่าเฉลี่ยของน้ำหนักแป้งในกระป๋องเท่ากันหรือไม่ จึงทำการสุ่มตัวอย่างแป้งกระป๋องจากเครื่องจักรเครื่องที่ 1 และ 2 มาอย่างละ 100 กระป๋อง ข้อมูลเก็บไว้ที่แฟ้มข้อมูลชื่อ example11.sav

จงทดสอบสมมติฐานว่าค่าเฉลี่ยของน้ำหนักแป้งในกระป๋องเท่ากันหรือไม่ ที่ระดับนัยสำคัญ 0.01

วิธีทำ ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \mu_1 = \mu_2$
กำหนดสมมติฐานอื่น $H_1 : \mu_1 \neq \mu_2$

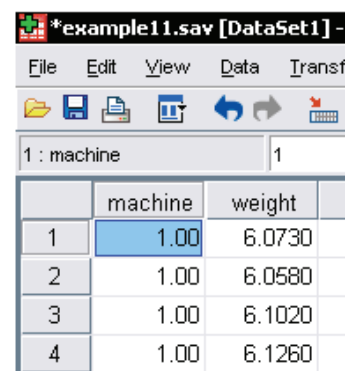
ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.01$

ขั้นที่ 3. นำข้อมูลเข้าสู่ SPSS Data Editor
machine เป็นตัวแปรจำแนกกลุ่มของโรงงาน

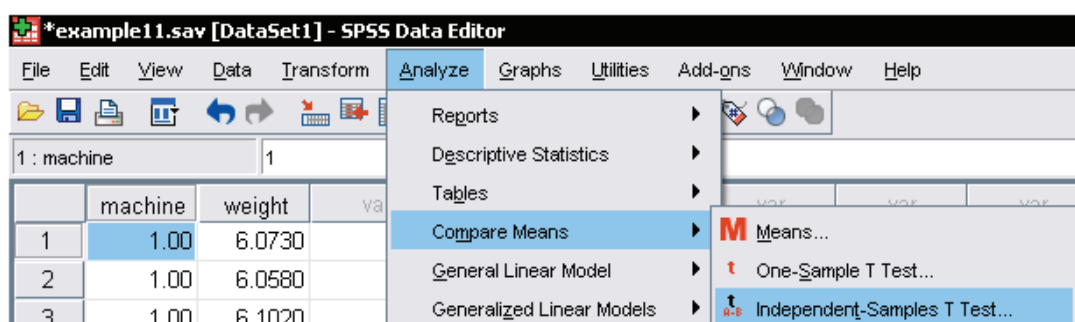
weight เป็นตัวแปรเก็บค่าน้ำหนักแป้ง

ขั้นที่ 4. เลือกค่าสถิติ t เท่านั้น

ขั้นที่ 4.1 เลือกคำสั่ง Analyze / Compare Means / Independent-Samples T Test

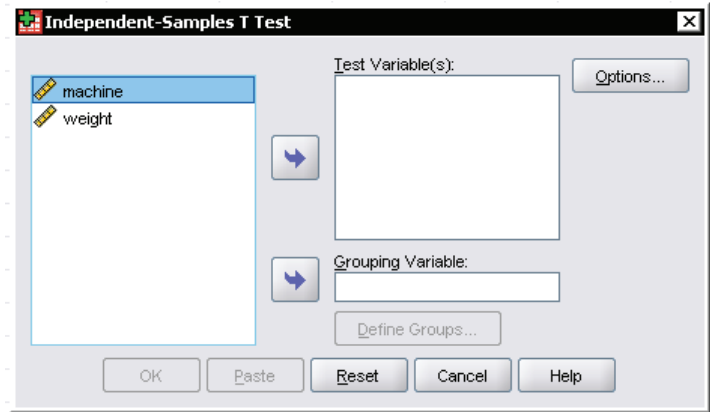


	machine	weight
1	1.00	6.0730
2	1.00	6.0580
3	1.00	6.1020
4	1.00	6.1260

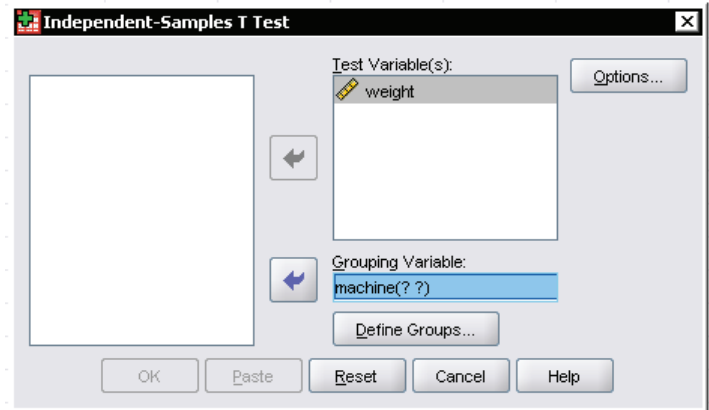


ขั้นที่ 4.2

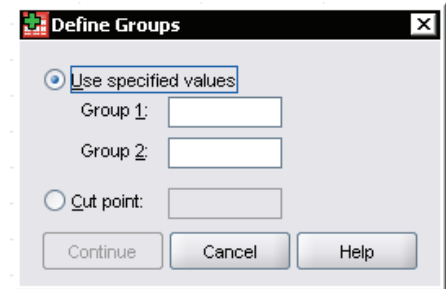
คลิก Independent-Samples T Test
จะได้ผลบนจอภาพเป็น



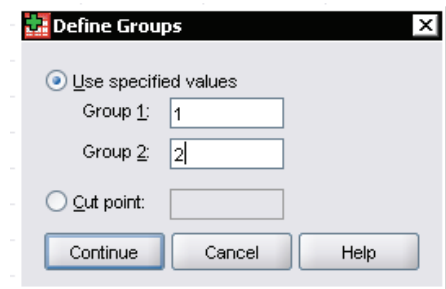
ขั้นที่ 4.3 เลือกตัวแปร weight
มาไว้ที่ Test Variable(s)
และ เลือกตัวแปร machine
มาไว้ที่ Grouping Variable



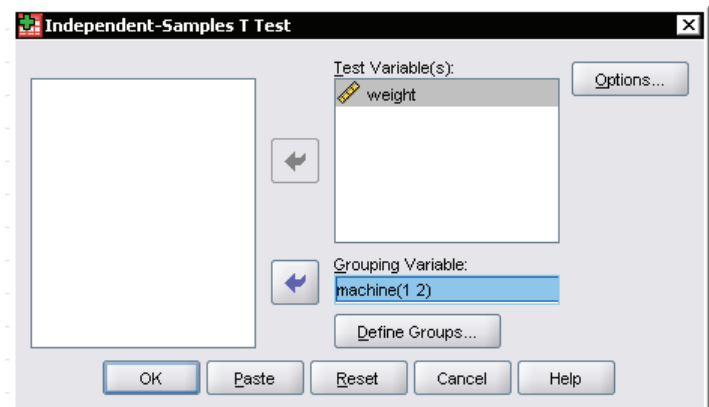
ขั้นที่ 4.4 คลิกที่ machine(? ?)
และ คลิกที่ Define Groups จอภาพจะมีเมนูย่อยเป็น



ขั้นที่ 4.5 การเลือกกลุ่มเพื่อทดสอบ
ให้พิมพ์ 1 ในช่อง Group 1.
และ พิมพ์ 2 ในช่อง Group 2.



ขั้นที่ 4.6 กด Continue จะเห็นได้ว่า
ตรงตัวแปร machine
จะกลายเป็น machine(1 2)



ขั้นที่ 4.7 คลิก OK

จะได้ผลการคำนวณที่

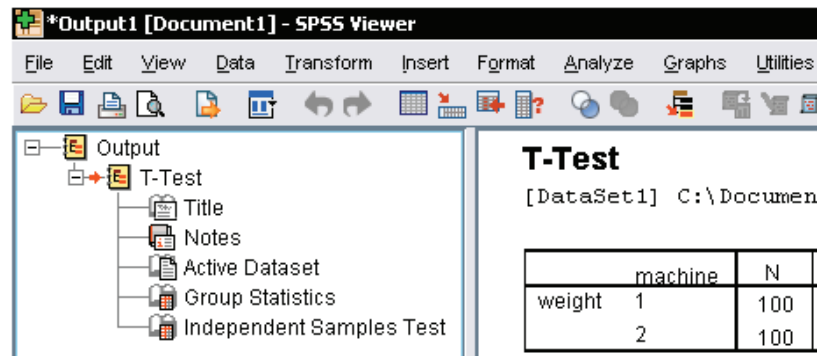
SPSS Viewer

เป็นดังนี้

ผลการคำนวณทั้งหมดคือ

T-Test

[DataSet1] C:\Documents and Settings\dumrong\My Documents\datspps16\example11.sav



Group Statistics

	machine	N	Mean	Std. Deviation	Std. Error Mean
weight	1	100	6.109930	.0399318	.0039932
	2	100	6.140250	.0500982	.0050098

Independent Samples Test

		weight	
		Equal variances assumed	Equal variances not assumed
Levene's Test for Equality of Variances	F	7.964764	
	Sig.	0.005256	
t-test for Equality of Means	t	-4.73266330	-4.73266330
	df	198	188.620
	Sig. (2-tailed)	0.00000421	0.00000434
	Mean Difference	-0.03032000	-0.03032000
	Std. Error Difference	0.00640654	0.00640654
	95% Confidence Interval of the Difference	Lower	-0.04295381
	Upper	-0.01768619	-0.01768233

การเลือกใช้ผลการวิเคราะห์ในกรณี Equal variances assumed หรือ Equal variances not assumed

เกณฑ์ในการสรุปผลคือ ถ้า Sig. ของ ค่า Levene's Test for Equality of Variances มีค่าน้อยกว่า α

แล้ว สรุปได้ว่าความแปรปรวนของประชากรทั้ง 2 ชุดมีค่าไม่เท่ากัน โดยมีระดับนัยสำคัญ α

การสรุปผล เกี่ยวกับความแปรปรวน

เพราะว่า Sig. ของ ค่า Levene's Test for Equality of Variances เท่ากับ 0.005256 มีค่าน้อยกว่า $\alpha = 0.01$

เพราะฉะนั้นแล้ว ความแปรปรวนของประชากรทั้ง 2 ชุดมีค่าไม่เท่ากัน โดยมีระดับนัยสำคัญ $\alpha = 0.01$

ภายใต้ข้อกำหนด $\sigma_1^2 \neq \sigma_2^2$

ขั้นที่ 5. เลือกใช้ค่าสถิติ $t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ เมื่อ $df = \frac{(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2})^2}{(\frac{s_1^2}{n_1})^2 \frac{1}{(n_1-1)} + (\frac{s_2^2}{n_2})^2 \frac{1}{(n_2-1)}}$

จากการคำนวณของ SPSS ได้ว่า $t_{\text{คำนวณ}} = -4.732663$, $df = 188.6201$

ขั้นที่ 6. คำนวณค่า Sig. (2-tailed) ของค่าสถิติ t

จากการคำนวณของ SPSS จะได้ค่า Sig. (2-tailed) = 0.0000043

ขั้นที่ 7. สรุปผล 1. โดยการเปรียบเทียบค่าสถิติ $t_{\text{คำนวณ}}$ จากตัวอย่าง กับ ค่าวิกฤตจากตารางสถิติ

จากตารางสถิติจะได้ค่า $t_{0.005, df=198} = 2.6008873$

เพราะว่ากำหนดสมมติฐานอื่น $H_1 : \mu_1 \neq \mu_2$

เพราะฉะนั้นบริเวณวิกฤตคือ $t < -2.6008873$ หรือ $t > 2.6008873$ สรุปปฏิเสธ H_0

หรือ 2. โดยการเปรียบเทียบค่า Sig. (2-tailed) กับ ระดับนัยสำคัญ α

เพราะว่ากำหนดสมมติฐานอื่น $H_1 : \mu_1 \neq \mu_2$

เพราะฉะนั้นต้องเปรียบเทียบค่า $\alpha = 0.01$ กับค่าของ Sig. (2-tailed)

เพราะว่า Sig. (2-tailed) = 0.0000043 < 0.01 = α สรุป ปฏิเสธ H_0

ผลสรุปค่าเฉลี่ยของน้ำหนักแป้งในกระป๋องทั้ง 2 ประชากรมีค่าไม่เท่ากันที่ระดับนัยสำคัญ 0.01

เสริมความรู้ของการคำนวณค่าทางสถิติด้วย Mathcad

$$x1 = \begin{array}{|c|c|} \hline & 1 \\ \hline 1 & 6.073 \\ \hline 2 & 6.058 \\ \hline 3 & 6.102 \\ \hline 4 & 6.126 \\ \hline 5 & 6.138 \\ \hline \end{array}$$

$$x1 = \begin{array}{|c|c|} \hline & 1 \\ \hline 96 & 6.125 \\ \hline 97 & 6.098 \\ \hline 98 & 6.126 \\ \hline 99 & 6.16 \\ \hline 100 & 6.123 \\ \hline \end{array}$$

$$x2 = \begin{array}{|c|c|} \hline & 1 \\ \hline 1 & 6.094 \\ \hline 2 & 6.075 \\ \hline 3 & 6.13 \\ \hline 4 & 6.16 \\ \hline 5 & 6.175 \\ \hline \end{array}$$

$$x2 = \begin{array}{|c|c|} \hline & 1 \\ \hline 96 & 6.159 \\ \hline 97 & 6.125 \\ \hline 98 & 6.16 \\ \hline 99 & 6.203 \\ \hline 100 & 6.156 \\ \hline \end{array}$$

$$xbar1 := \text{mean}(x1) \quad xbar1 = 6.10993$$

$$n1 := \text{length}(x1)$$

$$n1 = 100$$

$$xbar2 := \text{mean}(x2) \quad xbar2 = 6.14025$$

$$n2 := \text{length}(x2)$$

$$n2 = 100$$

$$s1 := \text{Stdev}(x1) \quad s1 = 0.0399318245$$

$$s2 := \text{Stdev}(x2) \quad s2 = 0.0500981612$$

$$\text{Std_Error_Mean1} := \frac{s1}{\sqrt{n1}}$$

$$\text{Std_Error_Mean1} = 0.003993182$$

$$\text{Std_Error_Mean2} := \frac{s2}{\sqrt{n2}}$$

$$\text{Std_Error_Mean2} = 0.005009816$$

Equal variances assumed

$$\text{Mean_Difference} := xbar1 - xbar2$$

$$\text{Mean_Difference} = -0.03032$$

$$sp := \sqrt{\frac{(n1 - 1) \cdot s1^2 + (n2 - 1) \cdot s2^2}{n1 + n2 - 2}}$$

$$sp = 0.0453010837$$

$$\text{Std_Error_Difference} := sp \cdot \sqrt{\frac{1}{n1} + \frac{1}{n2}}$$

$$\text{Std_Error_Difference} = 0.0064065407$$

$$t := \frac{xbar1 - xbar2}{sp \cdot \sqrt{\frac{1}{n1} + \frac{1}{n2}}}$$

$$t = -4.7326633$$

$$df := n1 + n2 - 2$$

$$df = 198$$

การหาค่า $t_{0.005, df=198}$ และ Sig. (2-tailed)

$$\alpha := 0.01$$

$$t_alpha_divide2 := qt\left(1 - \frac{\alpha}{2}, df\right) \quad t_alpha_divide2 = 2.600887278$$

เพราะฉะนั้น $t_{0.005, df=198} = 2.600887278$

$$\text{Sig} := 1 - pt(|-4.7326633|, 198) \quad \text{Sig} = 0.0000021055$$

$$\text{Sig}_{2_tailed} := 2 \cdot \text{Sig} \quad \text{Sig}_{2_tailed} = 0.000004211$$

เพราะฉะนั้น Sig. (2-tailed) = $2(0.0000021055) = 0.000004211$

การคำนวณ 95% Confidence Interval of the Difference of Mean

$$\alpha := 0.05 \quad t_alpha_divide2 := qt\left(1 - \frac{\alpha}{2}, df\right) \quad t_alpha_divide2 = 1.9720174778$$

$$\text{Lower} := (\bar{x}1 - \bar{x}2) - t_alpha_divide2 \cdot \text{Std_Error_Difference} \quad \text{Lower} = -0.04295381$$

$$\text{Upper} := (\bar{x}1 - \bar{x}2) + t_alpha_divide2 \cdot \text{Std_Error_Difference} \quad \text{Upper} = -0.01768619$$

Equal variances not assumed

$$t := \frac{\bar{x}1 - \bar{x}2}{\sqrt{\frac{s1^2}{n1} + \frac{s2^2}{n2}}} \quad t = -4.7326633$$

$$df := \frac{\left(\frac{s1^2}{n1} + \frac{s2^2}{n2}\right)^2}{\left(\frac{s1^2}{n1}\right)^2 \cdot \left(\frac{1}{n1-1}\right) + \left(\frac{s2^2}{n2}\right)^2 \cdot \left(\frac{1}{n2-1}\right)} \quad df = 188.6201257062$$

$$\text{Mean_Difference} := \bar{x}1 - \bar{x}2 \quad \text{Mean_Difference} = -0.03032$$

$$\text{Std_Error_Difference} := \sqrt{\frac{s1^2}{n1} + \frac{s2^2}{n2}} \quad \text{Std_Error_Difference} = 0.0064065407$$

การหาค่า $t_{0.005, df=188.62}$ และ Sig. (2-tailed)

$$\alpha := 0.01$$

$$t_alpha_divide2 := qt\left(1 - \frac{\alpha}{2}, df\right) \quad t_alpha_divide2 = 2.6021453165$$

เพราะฉะนั้น $t_{0.005, df=188.620} = 2.6021453165$

$$\text{Sig} := 1 - pt(|-4.7326633|, df) \quad \text{Sig} = 0.00000217$$

$$\text{Sig}_{2_tailed} := 2 \cdot \text{Sig} \quad \text{Sig}_{2_tailed} = 0.0000043378$$

เพราะฉะนั้น Sig. (2-tailed) = $2(0.00000217) = 0.0000043378$

การคำนวณ 95% Confidence Interval of the Difference of Mean

$$\alpha := 0.05 \quad t_{\alpha_divide2} := qt\left(1 - \frac{\alpha}{2}, df\right) \quad t_{\alpha_divide2} = 1.9726206787$$

$$\text{Lower} := (\bar{x}_1 - \bar{x}_2) - t_{\alpha_divide2} \cdot \text{Std_Error_Difference} \quad \text{Lower} = -0.04295767$$

$$\text{Upper} := (\bar{x}_1 - \bar{x}_2) + t_{\alpha_divide2} \cdot \text{Std_Error_Difference} \quad \text{Upper} = -0.01768233$$

หมายเหตุ ในกรณีที่เราต้องการทดสอบ สมมติฐานหลัก $H_0 : \mu_1 - \mu_2 = k$ แยกกับ $H_1 : \mu_1 - \mu_2 \neq k$ สามารถทำได้โดยนำเรื่องของการเปลี่ยนค่าตัวแปรมาช่วย ตัวอย่างเช่น จากตัวอย่าง 7.2.1 เราต้องการทดสอบ $H_0 : \mu_1 - \mu_2 = 0.02$ แยกกับ $H_1 : \mu_1 - \mu_2 \neq 0.02$ กำหนดระดับนัยสำคัญ 0.01

วิธีทำ ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \mu_1 - \mu_2 = 0.02$
กำหนดสมมติฐานอื่น $H_1 : \mu_1 - \mu_2 \neq 0.02$

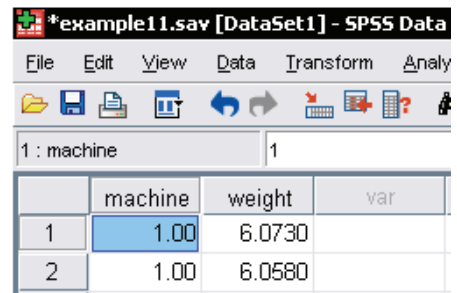
ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.01$

ขั้นที่ 3. นำแฟ้ม example11.sav ข้อมูลเข้าสู่ SPSS Data Editor

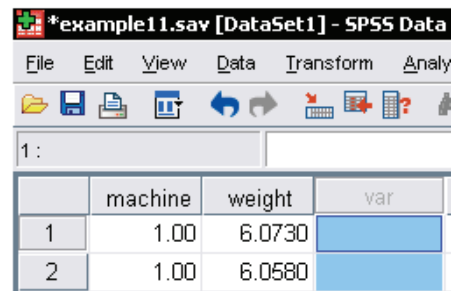
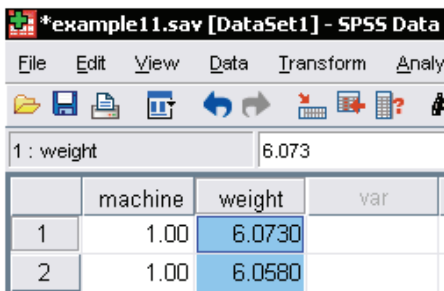
การ copy ตัวแปร weight ไปเป็นตัวแปรใหม่ชื่อ newweigh

ขั้นที่ 3.1 คลิกเมาส์ที่ชื่อตัวแปร weight

ขั้นที่ 3.2 เลือกคำสั่ง Edit และ Copy

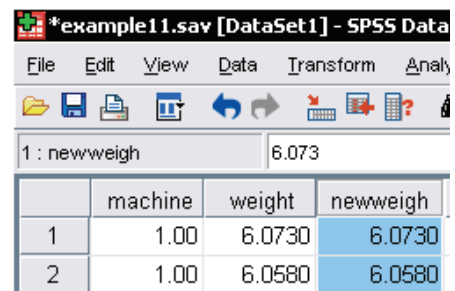
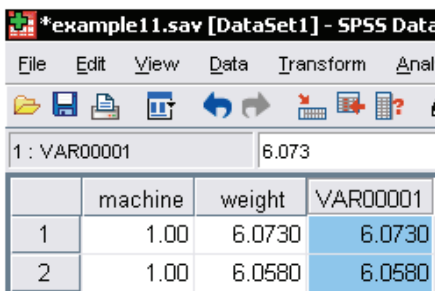


ขั้นที่ 3.3 คลิกเมาส์ที่ Column ของตัวแปรที่ยังว่างทางด้านขวา จะขึ้นแถบสีเข้มในช่องตัวแปรที่ว่าง



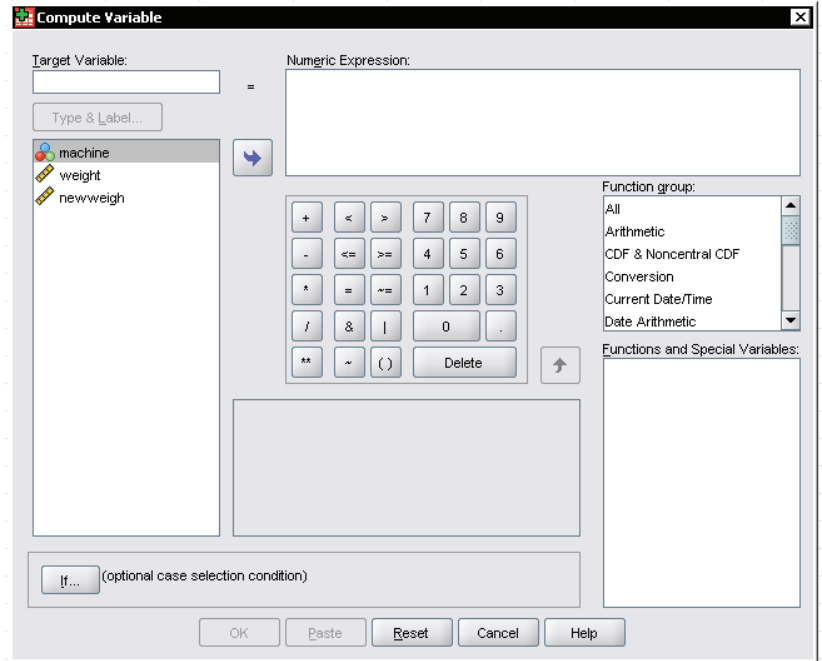
ขั้นที่ 3.4 เลือกคำสั่ง Edit / Paste จะได้ผลดังนี้

เปลี่ยนชื่อตัวแปร VAR00001 เป็น newweigh

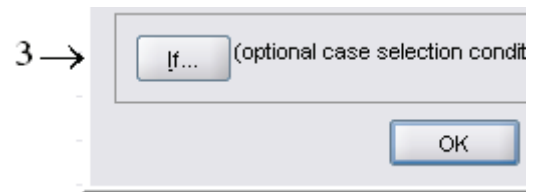


จะได้ข้อมูลของตัวแปร weight และ newweigh เหมือนกัน

การเพิ่มค่า newweigh อีก 0.02 เฉพาะกลุ่มเครื่องจักรเครื่องที่ 1
 ขั้นที่ 3.5 เลือกคำสั่ง Transform / Compute จะได้เมนูย่อยเป็น



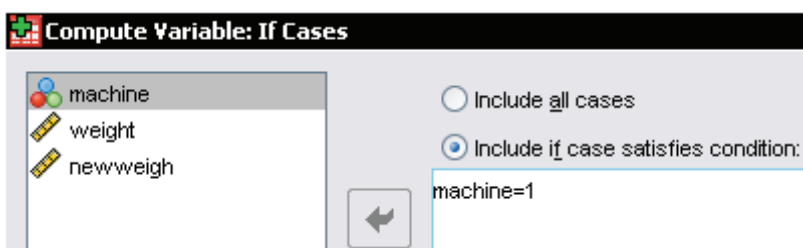
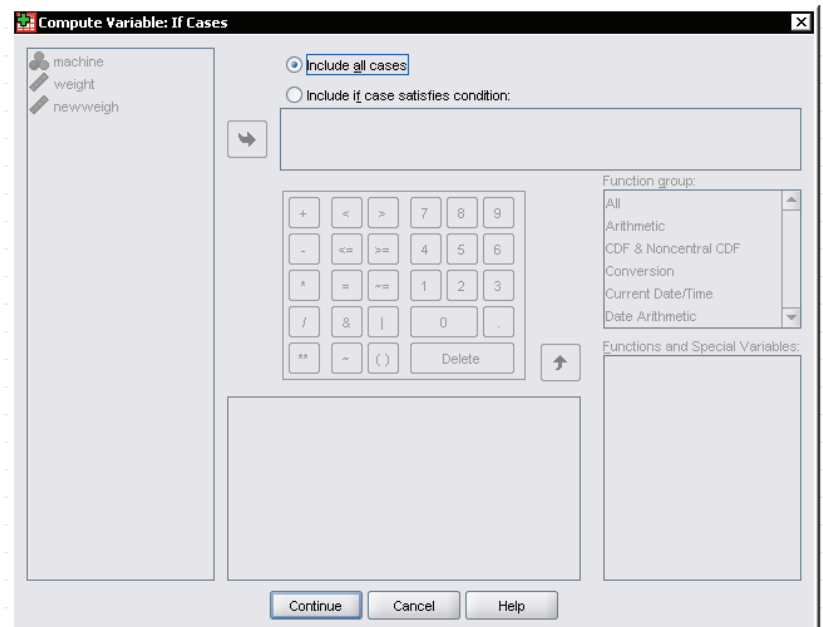
ขั้นที่ 3.6 1. ในช่อง Target variable :
 ให้พิมพ์ชื่อตัวแปร newweigh
 2. ในช่อง Numeric Expression :
 ให้พิมพ์ชื่อตัวแปร newweigh + 0.02



3. เราต้องการบวกเพิ่มเฉพาะน้ำหนัก
 ของกลุ่มที่ 1
 เพราะฉะนั้น ให้คลิกที่ If จะได้เมนูย่อย

ขั้นที่ 3.7

1. เลือก Option : Include if case
 satisfies condition
 2. พิมพ์ machine = 1
 ในช่องสี่เหลี่ยมด้านล่าง



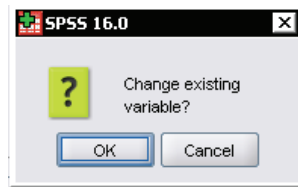
ขั้นที่ 3.8 คลิก Continue

และ OK ตามลำดับ

เมื่อมีคำถามว่า

ต้องการเปลี่ยนค่าตัวแปรจริงไม่

ให้คลิก OK



*example11.sav [DataSet1] - SPSS Data			
1 : machine			
	machine	weight	newweigh
1	1.00	6.0730	6.0930
2	1.00	6.0580	6.0780
99	1.00	6.1600	6.1800
100	1.00	6.1230	6.1430
101	2.00	6.0940	6.0940
102	2.00	6.0750	6.0750

จะได้ว่าข้อมูล newweigh กลุ่มที่ 1 ทุกตัวถูกบวกเพิ่ม 0.02

แต่ข้อมูล newweigh กลุ่มที่ 2 ยังเป็นค่าเดิม

เสร็จแล้ว save ไว้ที่ชื่อ example11_plus_02_to_machine1.sav

ขั้นที่ 4. เลือกค่าสถิติ t เท่านั้น

ขั้นที่ 4.1 เลือกคำสั่ง Analyze / Compare Means / Independent-Samples T Test

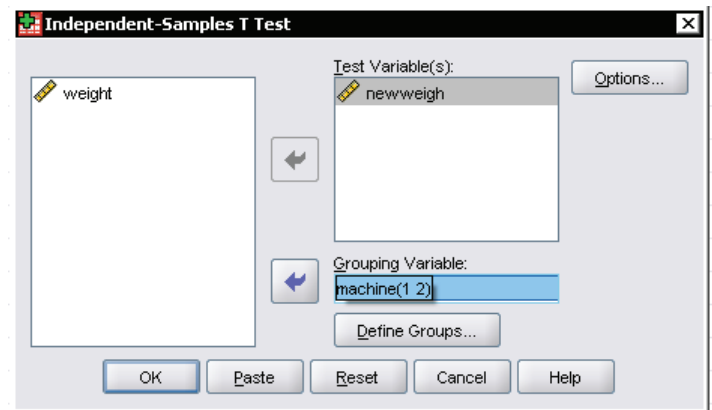
เลือกตัวแปร newweigh

ไปที่ช่อง Test Variable(s):

เลือกกลุ่มทดสอบเป็น

machine(1 2)

เสร็จแล้วคลิก OK



จะได้ผลการวิเคราะห์ข้อมูลเป็นดังนี้

T-Test

[DataSet1] C:\Documents and Settings\dumrong\My Documents\dataspss16\example11.sav

Group Statistics

machine		N	Mean	Std. Deviation	Std. Error Mean
newweigh	1	100	6.129930	.0399318	.0039932
	2	100	6.140250	.0500982	.0050098

Independent Samples Test

		newweigh		
		Equal variances assumed	Equal variances not assumed	
Levene's Test for Equality of Variances	F	7.96476419		
	Sig.	0.00525621		
t-test for Equality of Means	t	-1.61085373	-1.61085373	
	df	198	188.620	
	Sig. (2-tailed)	0.10880459	0.10888375	
	Mean Difference	-0.01032000	-0.01032000	
	Std. Error Difference	0.00640654	0.00640654	
	95% Confidence Interval of the Difference	Lower	-0.02295381	-0.02295767
		Upper	0.00231381	0.00231767

ภายใต้ข้อกำหนด $\sigma_1^2 \neq \sigma_2^2$

ขั้นที่ 6. จากตาราง Independent Samples Test

$$t_{\text{คำนวณ}} = -1.610854 \text{ และ Sig. (2-tailed)} = 0.108884$$

ขั้นที่ 7. เพราะว่ากำหนดสมมติฐานอื่น $H_1 : \mu_1 - \mu_2 \neq 0.02$

เพราะฉะนั้นต้องเปรียบเทียบค่า $\alpha = 0.01$ กับค่าของ Sig. (2-tailed)

เพราะว่า Sig. (2-tailed) = 0.108884 > 0.01 = α เพราะฉะนั้นไม่ปฏิเสธ H_0

สรุปค่าเฉลี่ยของน้ำหนักแบ่งในกระป๋องทั้ง 2 ประชากรแตกต่างกันไม่เกิน 0.02 ที่ระดับนัยสำคัญ 0.01

7.3 การทดสอบสมมติฐาน $H_0 : \mu_D = d_0$ กรณีประชากร 2 ชุดไม่เป็นอิสระต่อกัน

หลักการทางทฤษฎีของความน่าจะเป็นและสถิติ ของการทดสอบสมมติฐาน $H_0 : \mu_D = d_0$

กรณีที่ประชากร 2 ชุดไม่เป็นอิสระต่อกัน(ข้อมูลเป็นคู่) มีขั้นตอนการทำงานดังนี้

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \mu_D = d_0$

กำหนดสมมติฐานอื่น $H_1 : \mu_D \neq d_0$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. สุ่มตัวอย่างขนาด n จากประชากรชุดที่ 1 และ ประชากรชุดที่ 2 ได้ข้อมูลเป็น (ข้อมูลเป็นคู่)

ตัวอย่างจากประชากรชุดที่ 1.	ตัวอย่างจากประชากรชุดที่ 2.
x_1	y_1
x_2	y_2
x_3	y_3
:	:
x_n	y_n

ขั้นตอนการคำนวณ

1. คำนวณค่าผลต่างของตัวอย่าง $d_i = x_i - y_i, i = 1, 2, \dots, n$
2. คำนวณค่าเฉลี่ยของผลต่างของตัวอย่าง \bar{d}
3. คำนวณค่าส่วนเบี่ยงเบนมาตรฐานของผลต่างของตัวอย่าง s_d

ขั้นที่ 4. เลือกค่าสถิติที่เหมาะสม Z หรือ t

$$\text{กรณี } n \geq 30 \text{ เลือกใช้ค่าสถิติ } Z = \frac{\bar{d} - d_0}{\frac{s_d}{\sqrt{n}}}$$

กรณี $n < 30$ และภายใต้ข้อสมมติว่าผลต่างของข้อมูลมีการแจกแจงปกติ

$$\text{เลือก } t = \frac{\bar{d} - d_0}{\frac{s_d}{\sqrt{n}}} \text{ เมื่อ } df = n - 1$$

ขั้นที่ 5. คำนวณค่าสถิติ $t_{\text{คำนวณ}}$ หรือ $z_{\text{คำนวณ}}$ ที่เลือกจากข้อมูลตัวอย่าง

ขั้นที่ 6. เปิดตารางสถิติเพื่อหา ค่าวิกฤต และ บริเวณวิกฤต

6.1 กรณีใช้ค่า Z ค่าวิกฤตคือ $-z_{\frac{\alpha}{2}}$ และ $z_{\frac{\alpha}{2}}$ บริเวณวิกฤตคือ $Z < -z_{\frac{\alpha}{2}}$ หรือ $Z > z_{\frac{\alpha}{2}}$

6.2 กรณีใช้ค่า t ค่าวิกฤตคือ $-t_{\frac{\alpha}{2}}$ และ $t_{\frac{\alpha}{2}}$ บริเวณวิกฤตคือ $t < -t_{\frac{\alpha}{2}}$ หรือ $t > t_{\frac{\alpha}{2}}$

ขั้นที่ 7. สรุปผลโดยการเปรียบเทียบค่าสถิติจากตัวอย่างกับค่าวิกฤต โดยมีเกณฑ์การสรุปผลว่า

7.1 กรณีใช้ค่า Z ถ้า $z_{\text{คำนวณ}} < -z_{\frac{\alpha}{2}}$ หรือ $z_{\text{คำนวณ}} > z_{\frac{\alpha}{2}}$ แล้วปฏิเสธ H_0

7.2 กรณีใช้ค่า t ถ้า $t_{\text{คำนวณ}} < -t_{\frac{\alpha}{2}}$ หรือ $t_{\text{คำนวณ}} > t_{\frac{\alpha}{2}}$ แล้วปฏิเสธ H_0

หลักการและขั้นตอนการทำงานของโปรแกรมวิเคราะห์ข้อมูลด้วย SPSS for Windows

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \mu_D = d_0$

กำหนดสมมติฐานอื่น $H_1 : \mu_D \neq d_0$

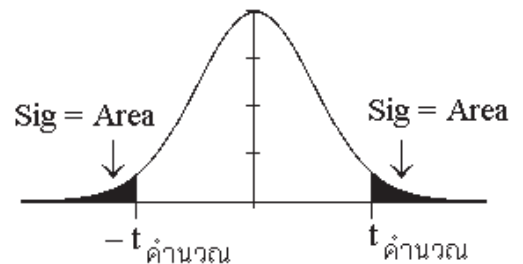
ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. สุ่มตัวอย่างขนาด n คำนวณค่า \bar{D} และ s_d

ขั้นที่ 4. เลือกค่าสถิติ t

ขั้นที่ 5. คำนวณค่าสถิติ t และ ระดับชั้นความเสรี df

$$\text{จากสูตร } t_{\text{คำนวณ}} = \frac{\bar{D} - d_0}{\frac{s_d}{\sqrt{n}}} \text{ และ } df = n - 1$$



ขั้นที่ 6. คำนวณค่า Sig. (2-tailed) ของค่าสถิติ $t_{\text{คำนวณ}}$

$$\text{Sig. (2-tailed)} = 2 \text{ เท่าของพื้นที่ใต้โค้งทางทางด้านขวาที่ระยะ } |t_{\text{คำนวณ}}| = 2 P(t > |t_{\text{คำนวณ}}|)$$

ขั้นที่ 7. การสรุปผลสามารถเลือกใช้เหตุผลได้ 2 วิธีคือ

1. โดยการเปรียบเทียบค่าสถิติ $t_{\text{คำนวณ}}$ จากตัวอย่าง กับ ค่าวิกฤต จากตารางสถิติ โดยมีเกณฑ์การสรุปผลว่า ถ้า $t_{\text{คำนวณ}} < -t_{\frac{\alpha}{2}}$ หรือ $t_{\text{คำนวณ}} > t_{\frac{\alpha}{2}}$ แล้วปฏิเสธ H_0

หรือ 2. โดยการเปรียบเทียบค่า Sig. (2-tailed) กับ ระดับนัยสำคัญ α

โดยมีเกณฑ์การสรุปผลว่า ถ้า Sig. (2-tailed) $< \alpha$ แล้วปฏิเสธ H_0

หมายเหตุ กรณีเป็นการทดสอบแบบ 1 ทาง

	$H_1 : \mu_D < d_0$	$H_1 : \mu_D > d_0$
ค่าวิกฤต Z	$-z_\alpha$	z_α
บริเวณวิกฤต	$Z < -z_\alpha$	$Z > z_\alpha$
ค่าวิกฤต t	$-t_\alpha$	t_α
บริเวณวิกฤต	$t < -t_\alpha$	$t > t_\alpha$
สรุปโดยใช้ Sig	ถ้า $t_{\text{คำนวณ}} < 0$, Sig. (2-tailed) $< 2\alpha$ แล้ว ปฏิเสธ H_0	ถ้า $t_{\text{คำนวณ}} > 0$, Sig. (2-tailed) $< 2\alpha$ แล้ว ปฏิเสธ H_0

ตัวอย่าง 7.3.1 จากตัวอย่างสารที่มีแร่เหล็ก 5 ตัวอย่างนำมาวิเคราะห์หาปริมาณเหล็ก 2 วิธี คือใช้รังสีเอ็กซ์ และ สารเคมี โดยแบ่งออกเป็น 2 กลุ่มตัวอย่างย่อยและใช้วิธีการวิเคราะห์หาปริมาณเหล็กได้ข้อมูลดังนี้

ตัวอย่างที่	1	2	3	4	5
วิธีวิเคราะห์ด้วยรังสีเอ็กซ์	2.0	2.0	2.3	2.1	2.4
วิธีวิเคราะห์ด้วยสารเคมี	2.2	1.9	2.5	2.3	2.4

ให้ μ_1 เป็นค่าเฉลี่ยประชากรของปริมาณเหล็กที่วิเคราะห์ด้วยรังสีเอ็กซ์

μ_2 เป็นค่าเฉลี่ยประชากรของปริมาณเหล็กที่วิเคราะห์ด้วยสารเคมี

สมมติว่าประชากรมีการแจกแจงปกติ จงทดสอบว่าการทดสอบสองวิธีให้ผลเหมือนกันที่ระดับนัยสำคัญ 0.05

วิธีทำ การวิเคราะห์ด้วย SPSS for Windows

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \mu_D = 0$

กำหนดสมมติฐานอื่น $H_1 : \mu_D \neq 0$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

ขั้นที่ 3. สุ่มตัวอย่างขนาด n คำนวณค่า \bar{D} และ s_d

ขั้นที่ 4. นำข้อมูลเข้ามาทำการวิเคราะห์เพื่อหาค่าสถิติ t

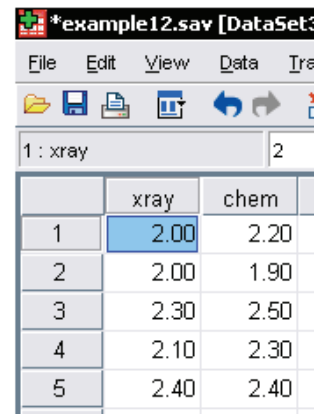
ขั้นที่ 4.1 สร้างแฟ้มข้อมูลใน SPSS Data Editor

กำหนดตัวแปร xray เป็นตัวแปรปริมาณแร่เหล็กโดยใช้วิธีวิเคราะห์ด้วยรังสีเอ็กซ์

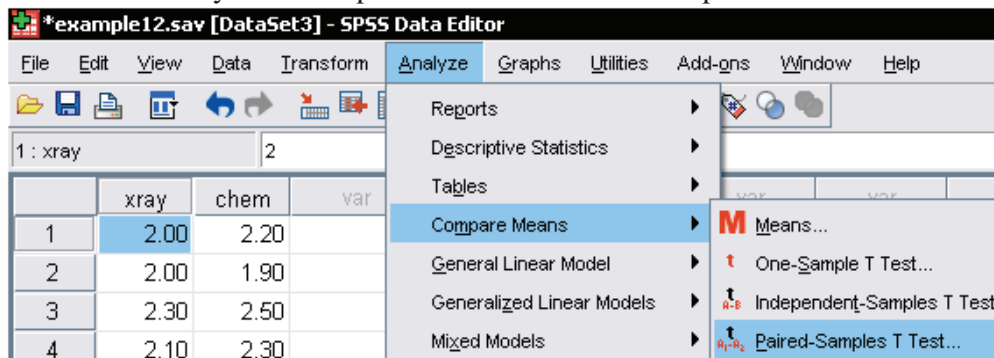
chem เป็นตัวแปรปริมาณแร่เหล็กโดยใช้วิธีวิเคราะห์ด้วยสารเคมี

และ Save แฟ้มข้อมูลชื่อ example12.sav

ขั้นที่ 4.2 เลือกคำสั่ง Analyze / Compare Means / Paired-Samples T Test

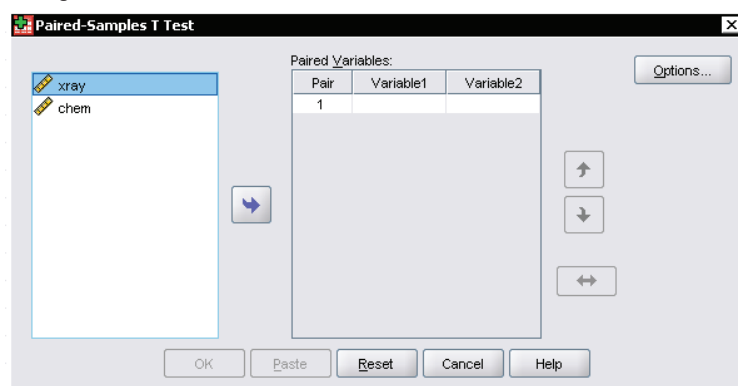


	xray	chem
1	2.00	2.20
2	2.00	1.90
3	2.30	2.50
4	2.10	2.30
5	2.40	2.40



ขั้นที่ 4.3 คลิกที่ Paired-Samples T Test

จอภาพจะขึ้นเมนูของคำสั่ง Analyze / Compare Means / Paired-Samples T Test

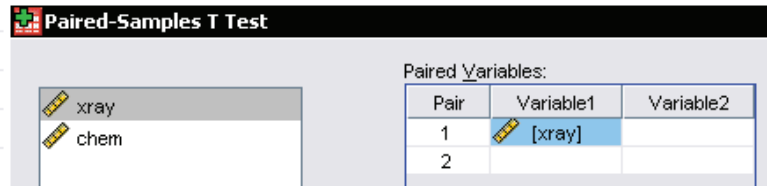


ขั้นที่ 4.4

เลือกตัวแปร xray

แล้วคลิกที่ 

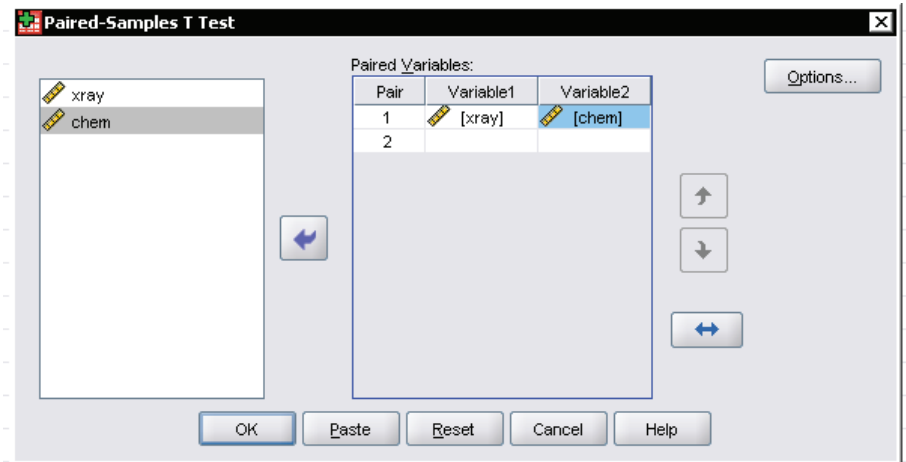
จะได้



ต่อไปเลือกตัวแปร chem

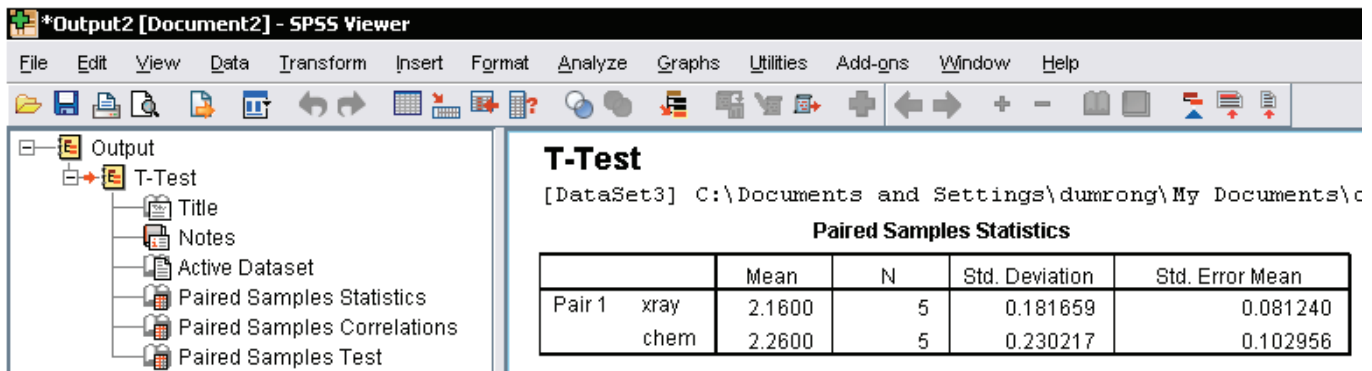
แล้วคลิกที่ 

จะได้



ขั้นที่ 4.5 คลิก OK จะได้

ผลการคำนวณเป็นดังนี้



ผลการคำนวณทั้งหมดคือ

T-Test

[DataSet3] C:\Documents and Settings\dumrong\My Documents\data spss16\example12.sav

Paired Samples Statistics

	Mean	N	Std. Deviation	Std. Error Mean
Pair 1 xray	2.1600	5	0.181659	0.081240
chem	2.2600	5	0.230217	0.102956

Paired Samples Correlations

	N	Correlation	Sig.
Pair 1 xray & chem	5	0.789076	0.112533

Paired Samples Test

		Pair 1	
		xray - chem	
Paired Differences	Mean	-0.100000	
	Std. Deviation	0.141421	
	Std. Error Mean	0.063246	
	95% Confidence Interval of the Difference	Lower	-0.275598
		Upper	0.075598
t		-1.581139	
df		4	
Sig. (2-tailed)		0.18900366	

ขั้นที่ 4. เลือกค่าสถิติ t

ขั้นที่ 5. จากตาราง Paired Samples Test จะได้ $t_{\text{คำนวณ}} = -1.581139$ และ ระดับชั้นความเสรี $df = 4$

ขั้นที่ 6. คำนวณค่า Sig. (2-tailed) ของค่าสถิติ t จากตาราง Sig. (2-tailed) = 0.18900366

ขั้นที่ 7. สรุปผล

1. จากการเปิดตาราง $t_{0.025, df=4} = 2.776$ บริเวณวิกฤตคือ $t < -2.776$ หรือ $t > 2.776$
 เพราะว่า $t_{\text{คำนวณ}} = 1.581$ ไม่อยู่ในบริเวณวิกฤต เพราะฉะนั้นไม่ปฏิเสธ H_0

หรือ 2. โดยการเปรียบเทียบค่า Sig. (2-tailed) กับ α
 เพราะว่า Sig. (2-tailed) = 0.18900366 $>$ 0.05 = α เพราะฉะนั้นไม่ปฏิเสธ H_0

หมายเหตุ การคำนวณด้วย Mathcad

ORIGIN:= 1

$$\text{xray} := \begin{pmatrix} 2.0 \\ 2.0 \\ 2.3 \\ 2.1 \\ 2.4 \end{pmatrix} \quad \text{chem} := \begin{pmatrix} 2.2 \\ 1.9 \\ 2.5 \\ 2.3 \\ 2.4 \end{pmatrix} \quad n := \text{length}(\text{xray}) \quad n = 5$$

$$\begin{aligned} \text{mean_xray} &:= \text{mean}(\text{xray}) & \text{mean_xray} &= 2.16 \\ \text{mean_chem} &:= \text{mean}(\text{chem}) & \text{mean_chem} &= 2.26 \\ \text{Std_Deviation_xray} &:= \text{Stdev}(\text{xray}) & \text{Std_Deviation_xray} &= 0.181659 \\ \text{Std_Deviation_chem} &:= \text{Stdev}(\text{chem}) & \text{Std_Deviation_chem} &= 0.230217 \\ \text{correlation} &:= \text{corr}(\text{xray}, \text{chem}) & \text{correlation} &= 0.789076 \end{aligned}$$

การคำนวณค่าผลต่าง d

$$d := \text{xray} - \text{chem}$$

$$d = \begin{pmatrix} -0.2 \\ 0.1 \\ -0.2 \\ -0.2 \\ 0 \end{pmatrix} \quad \begin{aligned} \text{dbar} &:= \text{mean}(d) & \text{dbar} &= -0.1 \\ \text{Std_Deviation} &:= \text{Stdev}(d) & \text{Std_Deviation} &= 0.141421 \\ \text{Std_Error_Mean} &:= \frac{\text{Std_Deviation}}{\sqrt{n}} & \text{Std_Error_Mean} &= 0.063246 \end{aligned}$$

การคำนวณค่าสถิติ t

$$t := \frac{\text{dbar} - 0}{\frac{\text{Std_Deviation}}{\sqrt{n}}} \quad t = -1.5811388301 \quad df := n - 1 \quad df = 4$$

การคำนวณค่า Sig. (2-tailed)

$$v := df \quad h(t) := \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right) \cdot \sqrt{\pi \cdot v}} \cdot \left(1 + \frac{t^2}{v}\right)^{-\frac{v+1}{2}}$$

$$\text{Significant} := 0.5 - \int_0^{|-1.5811388301|} h(t) dt \quad \text{Significant} = 0.0945018292$$

หรือโดยการใช้คำสั่งสำเร็จรูป pt ของ Mathcad

Sig := 1 - pt(| -1.5811388301 |, df) Sig = 0.0945018292

Sig_2_tailed := 2 · Sig Sig_2_tailed = 0.1890036585

การคำนวณ 95% Confidence Interval of the Difference

alpha := 0.05 t_alpha_divide2 := qt(1 - alpha/2, df) t_alpha_divide2 = 2.776445

Lower := dbar - t_alpha_divide2 · Std_Error_Mean Lower = -0.27559781

Upper := dbar + t_alpha_divide2 · Std_Error_Mean Upper = 0.07559781

หมายเหตุ การทดสอบ $H_0 : \mu_D = d_0$ แยกกับ $H_1 : \mu_D \neq d_0$

เมื่อ $d_0 \neq 0.5$ เช่น $d_0 = 0.5$ สามารถทำได้ตามขั้นตอนดังนี้

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \mu_D = 0.5$

กำหนดสมมติฐานอื่น $H_1 : \mu_D \neq 0.5$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

ขั้นที่ 3. สุ่มตัวอย่างขนาด n

ขั้นที่ 4. นำข้อมูลเข้ามาทำการวิเคราะห์ เพื่อหาค่าสถิติ t

ขั้นที่ 4.1

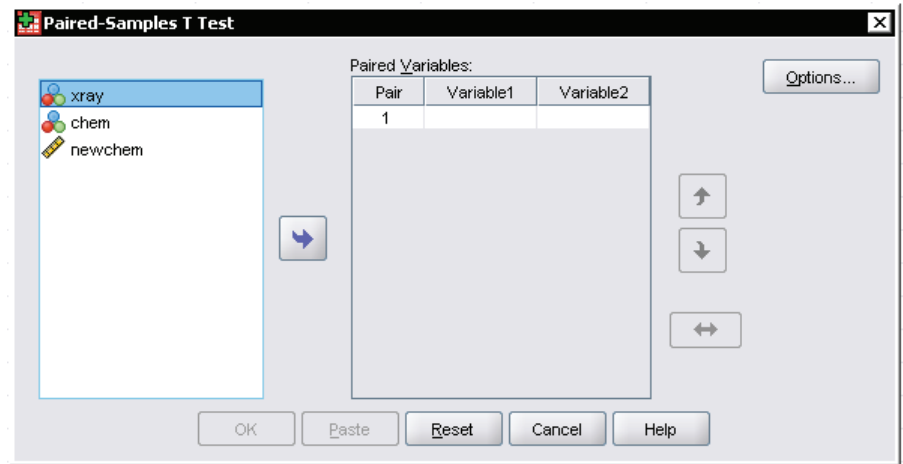
นำแฟ้มข้อมูล example12.sav เข้ามาใน SPSS Data Editor

สร้างตัวแปร newchem ให้มีค่าเท่ากับ chem + 0.05

	xray	chem	var
1	2.00	2.20	
2	2.00	1.90	
3	2.30	2.50	
4	2.10	2.30	
5	2.40	2.40	

	xray	chem	newchem
1	2.00	2.20	2.25
2	2.00	1.90	1.95
3	2.30	2.50	2.55
4	2.10	2.30	2.35
5	2.40	2.40	2.45

ขั้นที่ 4.2 เลือกคำสั่ง Analyze / Compare Means / Paired-Samples T Test

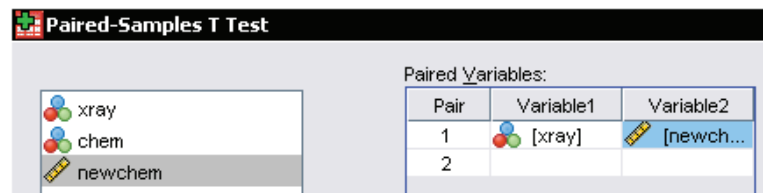


ขั้นที่ 4.3 เลือกคู่ของตัวแปรเพื่อทดสอบเป็น

xray - newchem

ขั้นที่ 4.4 คลิก OK

จะได้ผลการคำนวณทั้งหมดคือ



T-Test

[DataSet1] C:\Documents and Settings\dumrong\My Documents\datasps16\example12.sav

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	xray	2.1600	5	0.18165902	0.08124038
	newchem	2.3100	5	0.23021729	0.10295630

Paired Samples Correlations

	N	Correlation	Sig.
Pair 1 xray & newchem	5	0.78907636	0.11253311

Paired Samples Test

		Pair 1
		xray - newchem
Paired Differences	Mean	-0.15000000
	Std. Deviation	0.14142136
	Std. Error Mean	0.06324555
	95% Confidence Interval of the Difference	Lower Upper
t		-2.37170825
df		4
Sig. (2-tailed)		0.07667814

เพราะว่า Sig. (2-tailed) = 0.07667814 > 0.05 = α เพราะฉะนั้นไม่ปฏิเสธ $H_0 : \mu_D = 0.5$

7.4 การทดสอบสมมติฐาน $H_0 : \sigma^2 = \sigma_0^2$

หลักการและขั้นตอนการทำงานทางทฤษฎีของความน่าจะเป็นและสถิติ

- ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \sigma^2 = \sigma_0^2$
กำหนดสมมติฐานอื่น $H_1 : \sigma^2 \neq \sigma_0^2$
- ขั้นที่ 2. กำหนดระดับนัยสำคัญ α
- ขั้นที่ 3. ทำการสุ่มตัวอย่างขนาด n คำนวณค่าความแปรปรวนของตัวอย่าง s^2
- ขั้นที่ 4. เลือกค่าสถิติไคสแควร์
- ขั้นที่ 5. คำนวณค่าสถิติไคสแควร์ $\chi^2_{\text{คำนวณ}} = \frac{(n-1)s^2}{\sigma_0^2}$ ระดับชั้นความเสรี $df = n - 1$
- ขั้นที่ 6. เปิดตารางสถิติหาค่าวิกฤต $\chi^2_{1-\frac{\alpha}{2}}$ และ $\chi^2_{\frac{\alpha}{2}}$ บริเวณวิกฤตคือ $\chi^2 < \chi^2_{1-\frac{\alpha}{2}}$ หรือ $\chi^2 > \chi^2_{\frac{\alpha}{2}}$
- ขั้นที่ 7. การสรุปผล ถ้า $\chi^2_{\text{คำนวณ}} < \chi^2_{1-\frac{\alpha}{2}}$ หรือ $\chi^2_{\text{คำนวณ}} > \chi^2_{\frac{\alpha}{2}}$ แล้วปฏิเสธ H_0

ตัวอย่าง 7.4.1 ผู้ผลิตอ้างว่าอายุการใช้งานของแบตเตอรี่มีการแจกแจงปกติ และ มีส่วนเบี่ยงเบนมาตรฐานเป็น 0.9 ปี เพื่อทดสอบคำกล่าวอ้างของผู้ผลิตจึงทำการสุ่มตัวอย่างแบตเตอรี่ออกมา 10 ลูกได้อายุใช้งานดังนี้

5.25	3.76	5.36	3.67	6.05	3.89	3.39	6.12	6.49	6.03
------	------	------	------	------	------	------	------	------	------

จงทดสอบสมมติฐาน $\sigma^2 = 0.81$ ที่ระดับนัยสำคัญ 0.05

- วิธีทำ
- ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \sigma^2 = 0.81$
กำหนดสมมติฐานอื่น $H_1 : \sigma^2 \neq 0.81$
 - ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$
 - ขั้นที่ 3. สุ่มตัวอย่างขนาด 10 คำนวณค่าความแปรปรวนของตัวอย่าง s^2
 - ขั้นที่ 4. เลือกค่าสถิติไคสแควร์
 - ขั้นที่ 5. คำนวณค่าสถิติไคสแควร์ $\chi^2_{\text{คำนวณ}} = \frac{(n-1)s^2}{\sigma_0^2}$

ผลการคำนวณด้วย Mathcad

```

data := (
    5.25
    3.76
    5.36
    3.67
    6.05
    3.89
    3.39
    6.12
    6.49
    6.03
)
σ := 0.9
n := length(data)           n = 10
s := Stdev(data)           s = 1.2
chisquare :=  $\frac{(n-1) \cdot s^2}{\sigma^2}$    chisquare = 16.001
    
```

เพราะฉะนั้น $\chi^2_{\text{คำนวณ}} = 16.001$

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤต $\chi^2_{\text{คำนวณ}} = 19.023$, $\chi^2_{0.975} = 2.7$ เมื่อ $df = 10 - 1 = 9$
 บริเวณวิกฤตคือ $\chi^2 < 2.7$ หรือ $\chi^2 > 19.023$

ขั้นที่ 7. โดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต
 เพราะว่า $\chi^2_{\text{คำนวณ}}$ ไม่อยู่ในบริเวณวิกฤต เพราะฉะนั้น ไม่ปฏิเสธ H_0

7.5 การทดสอบสมมติฐาน $H_0 : \sigma_1^2 = \sigma_2^2$

หลักการและขั้นตอนการทำงานทางทฤษฎีของความน่าจะเป็นและสถิติ

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \sigma_1^2 = \sigma_2^2$

กำหนดสมมติฐานอื่น $H_1 : \sigma_1^2 \neq \sigma_2^2$

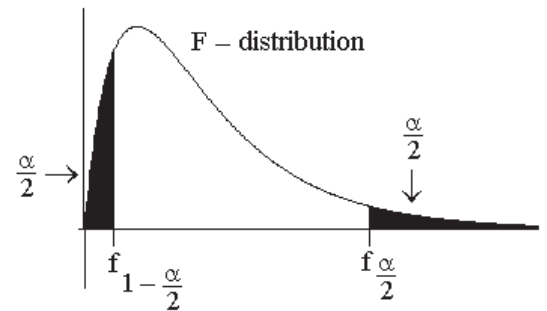
ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. ทำการสุ่มตัวอย่างขนาด n_1 และ n_2 คำนวณค่าความแปรปรวนของตัวอย่าง s_1^2 , s_2^2

ขั้นที่ 4. เลือกค่าสถิติเอฟ $F = \frac{s_1^2}{s_2^2}$

ขั้นที่ 5. คำนวณค่าสถิติเอฟ $f_{\text{คำนวณ}} = \frac{s_1^2}{s_2^2}$

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤตเอฟ
 ระดับชั้นความเสรี $v_1 = n_1 - 1$ และ $v_2 = n_2 - 1$
 ค่าวิกฤตคือ $f_{1-\frac{\alpha}{2}}$ และ $f_{\frac{\alpha}{2}}$



บริเวณวิกฤตคือบริเวณ $F < f_{1-\frac{\alpha}{2}}$ หรือ $F > f_{\frac{\alpha}{2}}$

ขั้นที่ 7. สรุปผลโดยการเปรียบเทียบค่าสถิติ $f_{\text{คำนวณ}}$ จากตัวอย่าง กับ ค่าวิกฤต

ถ้า $f_{\text{คำนวณ}} < f_{1-\frac{\alpha}{2}}$ หรือ $f_{\text{คำนวณ}} > f_{\frac{\alpha}{2}}$ แล้วปฏิเสธ H_0

ตัวอย่าง 7.5.1 วัตถุ 5 ชิ้นได้รับการปฏิบัติแบบที่ 1 ได้ผลการทดลองเป็นดังนี้

1.024 0.972 1.004 0.986 1.015

วัตถุ 6 ชิ้นได้รับการปฏิบัติแบบที่ 2 ได้ผลการทดลองเป็นดังนี้

1.017 0.991 1.018 1.018 0.983 0.975

จงทดสอบสมมติฐาน $H_0 : \sigma_1^2 = \sigma_2^2$ แยกกับ $H_1 : \sigma_1^2 \neq \sigma_2^2$ กำหนดระดับนัยสำคัญ 0.1

วิธีทำ

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \sigma_1^2 = \sigma_2^2$

กำหนดสมมติฐานอื่น $H_1 : \sigma_1^2 \neq \sigma_2^2$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.1$

ขั้นที่ 3. ทำการสุ่มตัวอย่างขนาด $n_1 = 5$ และ $n_2 = 6$ คำนวณค่าความแปรปรวนของตัวอย่าง s_1^2 , s_2^2

การคำนวณด้วย Mathcad

$$\begin{array}{l}
 x1 := \begin{pmatrix} 1.024 \\ 0.972 \\ 1.004 \\ 0.986 \\ 1.015 \end{pmatrix} \quad x2 := \begin{pmatrix} 1.017 \\ 0.991 \\ 1.018 \\ 1.018 \\ 0.983 \\ 0.975 \end{pmatrix} \\
 n1 := \text{length}(x1) \quad n1 = 5 \quad s1 := \text{Stdev}(x1) \quad s1 = 0.021194 \\
 n2 := \text{length}(x2) \quad n2 = 6 \quad s2 := \text{Stdev}(x2) \quad s2 = 0.019654 \\
 F_{\text{compute}} := \frac{s1^2}{s2^2} \quad F_{\text{compute}} = 1.162927
 \end{array}$$

ขั้นที่ 4. เลือกค่าสถิติเอฟ

ขั้นที่ 5. คำนวณค่าสถิติเอฟ $f_{\text{คำนวณ}} = \frac{s_1^2}{s_2^2} = 1.162927$

ขั้นที่ 6. เปิดตารางสถิติหาค่าวิกฤตเอฟ ระดับชั้นความเสรี $v_1 = n_1 - 1 = 4$ และ $v_2 = n_2 - 1 = 5$
 ค่าวิกฤตคือ $f_{0.05, (4, 5)} = 5.199$ และ $f_{0.95, (4, 5)} = \frac{1}{f_{0.05, (5, 4)}} = \frac{1}{6.26} = 0.1597444$

บริเวณวิกฤตคือ $F < 0.1597444$ หรือ $F > 5.199$

ขั้นที่ 7. สรุปผลโดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต
 เพราะว่า $f_{\text{คำนวณ}}$ ไม่อยู่ในบริเวณวิกฤต เพราะฉะนั้นไม่ปฏิเสธ H_0

หลักการและขั้นตอนของการวิเคราะห์ด้วย SPSS for Windows

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \sigma_1^2 = \sigma_2^2$
 กำหนดสมมติฐานอื่น $H_1 : \sigma_1^2 \neq \sigma_2^2$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. ทำการสุ่มตัวอย่างขนาด n_1 และ n_2 จากประชากรแต่ละชุด

ขั้นที่ 4. เลือกค่าสถิติ Levene

ขั้นที่ 5. คำนวณค่าสถิติ Levene และค่า Sig

ขั้นที่ 6. ทำการวิเคราะห์ข้อมูลโดย SPSS for Windows

ขั้นที่ 7. สรุปผลโดยการเปรียบเทียบค่า Sig กับ ค่านัยสำคัญ α
 ถ้า $\text{Sig} < \alpha$ แล้วปฏิเสธ H_0

จากตัวอย่าง 7.5.1 ต้องสร้างแฟ้มข้อมูลแบบ 2 ตัวแปรคือ ตัวแปร code จำแนกกลุ่มประชากร และตัวแปร x เก็บข้อมูลที่วัดได้จากการทดลอง

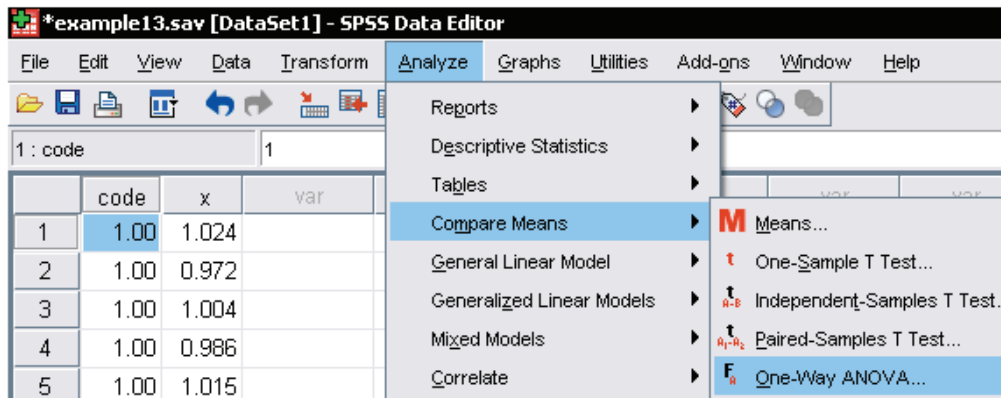
การทดสอบสมมติฐาน $H_0 : \sigma_1^2 = \sigma_2^2$ แยกกับ $H_1 : \sigma_1^2 \neq \sigma_2^2$ กำหนดระดับนัยสำคัญ 0.1

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \sigma_1^2 = \sigma_2^2$
 กำหนดสมมติฐานอื่น $H_1 : \sigma_1^2 \neq \sigma_2^2$

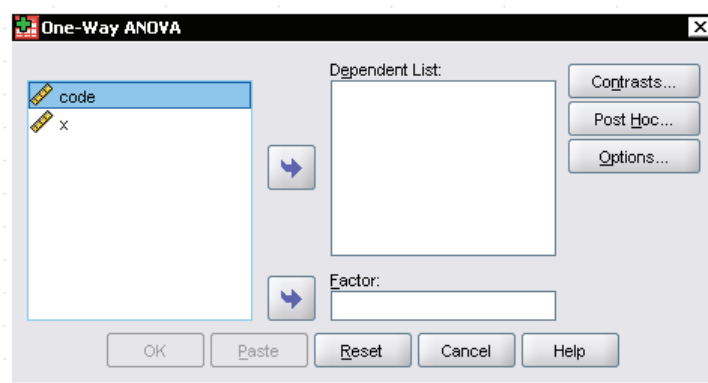
- ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.1$
- ขั้นที่ 3. ทำการสุ่มตัวอย่างขนาด $n_1 = 5$ และ $n_2 = 6$
- ขั้นที่ 4. เลือกค่าสถิติ Levene
- ขั้นที่ 5. คำนวณค่าสถิติ Levene และค่า Sig
- ขั้นที่ 6. สร้างแฟ้มข้อมูลโดยมี
ตัวแปร code เป็นตัวแปรจำแนกกลุ่ม
ตัวแปร x เป็นตัวแปรข้อมูลที่ต้องการวิเคราะห์

เสร็จแล้ว Save ข้อมูลไว้ที่แฟ้มข้อมูลชื่อ example13.sav

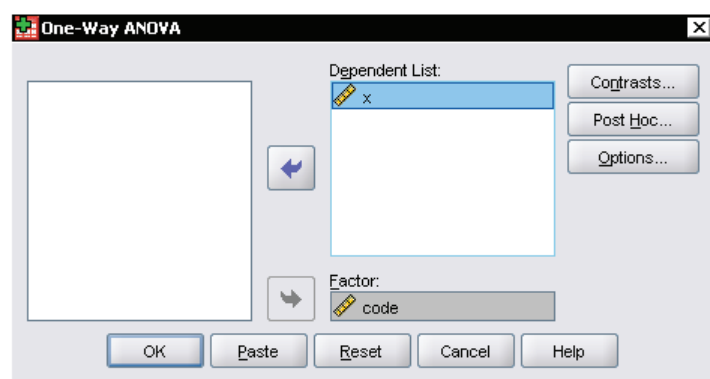
- ขั้นที่ 6.1 เลือกคำสั่ง Analyze / Compare Means / One-Way ANOVA



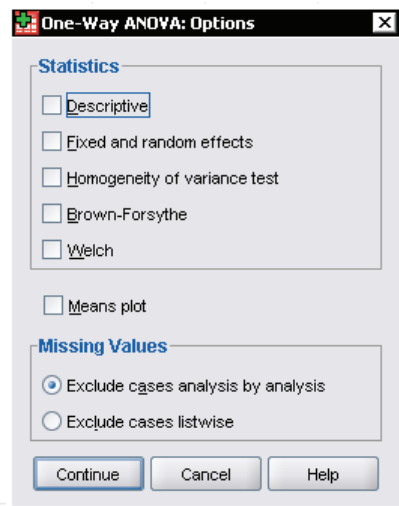
- ขั้นที่ 6.2 คลิกที่คำสั่ง One-Way ANOVA จะได้เมนูย่อยเป็น



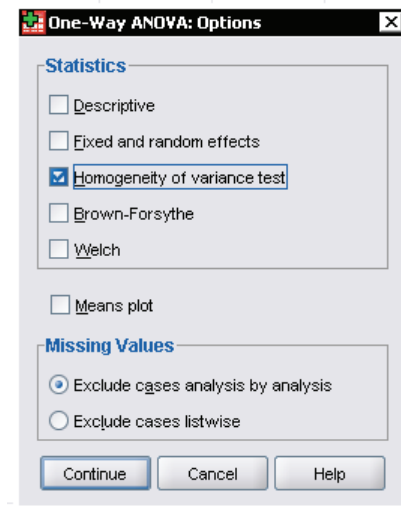
- ขั้นที่ 6.3 เลือกตัวแปร x ไปไว้ที่ Dependent List เลือกตัวแปร code ไปไว้ที่ Factor



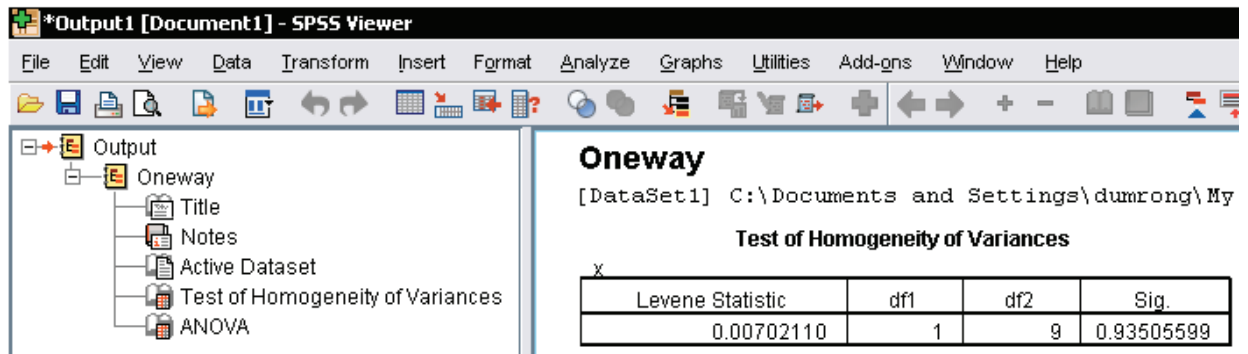
ขั้นที่ 6.4 คลิกปุ่ม Options ได้เมนูเป็น



ขั้นที่ 6.5 คลิกที่ Homogeneity of variance test



ขั้นที่ 6.6 คลิก Continue และ OK จะได้ผลการคำนวณเป็น



ผลการคำนวณทั้งหมดคือ

Oneway

[DataSet1] C:\Documents and Settings\dumrong\My Documents\dataspss16\example13.sav

Test of Homogeneity of Variances

X	Levene Statistic	df1	df2	Sig.
	0.00702110	1	9	0.93505599

ANOVA

X	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	0.00000005	1	0.00000005	0.00011705	0.99160405
Within Groups	0.00372813	9	0.00041424		
Total	0.00372818	10			

จากตารางผลการคำนวณค่า Levene Statistics = 0.00702110 และ Sig = 0.93505599

ขั้นที่ 7. สรุปผลโดยการเปรียบเทียบค่า Sig = 0.93505599 กับระดับนัยสำคัญ $\alpha = 0.1$

เพราะว่า Sig = 0.93505599 > $\alpha = 0.1$ เพราะฉะนั้น ไม่ปฏิเสธ $H_0 : \sigma_1^2 = \sigma_2^2$

7.6 การทดสอบภาวะสารูปสนิทธิ (Test of Goodness Of Fit)

การทดสอบภาวะสารูปสนิทธิเป็นการทดสอบสมมติฐานว่าข้อมูลมีการแจกแจงความน่าจะเป็นตามที่คาดไว้หรือไม่ เช่น ข้อมูลมีการแจกแจงทวินามจริงหรือไม่ ข้อมูลมีการแจกแจงปกติจริงหรือไม่ ข้อมูลมีการแจกแจงปัวส์ซองจริงหรือไม่ ข้อมูลมีการแจกแจงตามอัตราส่วนที่คาดไว้จริงหรือไม่

หลักการและขั้นตอนการทำงานทางทฤษฎีของความน่าจะเป็นและสถิติ

- ขั้นที่ 1. กำหนดสมมติฐานหลัก H_0 : ข้อมูลมีการแจกแจงความน่าจะเป็นตามที่คาดไว้
กำหนดสมมติฐานอื่น H_1 : ข้อมูลไม่มีการแจกแจงความน่าจะเป็นตามที่คาดไว้
- ขั้นที่ 2. กำหนดระดับนัยสำคัญ α
- ขั้นที่ 3. ทำการสุ่มตัวอย่างเพื่อหาค่าสังเกต o_i , $i = 1, 2, 3, \dots, k$
- ขั้นที่ 4. เลือกค่าสถิติไคสแควร์
- ขั้นที่ 5. คำนวณค่าความถี่ที่คาดว่าจะได้ e_i และค่าสถิติไคสแควร์ $\chi^2_{\text{คำนวณ}} = \sum_{i=1}^k \frac{(o_i - e_i)^2}{e_i}$
- ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤต χ^2_{α} เมื่อ $df = k -$ จำนวนพารามิเตอร์ที่ต้องประมาณค่า $- 1$
บริเวณวิกฤตคือ $\chi^2 > \chi^2_{\alpha}$
- ขั้นที่ 7. สรุปผลโดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต ถ้า $\chi^2_{\text{คำนวณ}} > \chi^2_{\alpha}$ แล้วปฏิเสธ H_0

ตัวอย่าง 7.6.1 การทดลองโยนเหรียญ 3 อัน 240 ครั้ง

ให้ x เป็นจำนวนหัวที่ได้ในการโยนเหรียญแต่ละครั้งผลการทดลองบันทึกไว้ที่แฟ้มข้อมูลชื่อ example14.sav
จงทดสอบสมมติฐานว่า เหรียญทั้งสามอันมีความเที่ยงตรง กำหนดระดับนัยสำคัญ 0.05

วิธีทำ การคำนวณโดย Mathcad

- ขั้นที่ 1. กำหนดสมมติฐานหลัก H_0 : เหรียญทั้งสามอันมีความเที่ยงตรง
กำหนดสมมติฐานอื่น H_1 : เหรียญทั้งสามอันไม่มีความเที่ยงตรง
- ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$
- ขั้นที่ 3. ทำการสุ่มตัวอย่างเพื่อหาค่าสังเกต o_i
- ขั้นที่ 4. เลือกค่าสถิติไคสแควร์
- ขั้นที่ 5. คำนวณค่าความถี่ที่คาดว่าจะได้ e_i ภายใต้เงื่อนไขที่กำหนดว่าเหรียญทั้งสามอันมีความเที่ยงตรง
เพราะฉะนั้น $x = 0, 1, 2, 3$ มีการแจกแจงแบบทวินาม ดังนั้นค่าความถี่ที่คาดว่าจะได้คือ

X	$P(X = x)$	e_i
0	$\frac{1}{8}$	$\frac{1}{8} (240) = 30$
1	$\frac{3}{8}$	$\frac{3}{8} (240) = 90$
2	$\frac{3}{8}$	$\frac{3}{8} (240) = 90$
3	$\frac{1}{8}$	$\frac{1}{8} (240) = 30$
		240

การคำนวณโดยใช้ Mathcad ทำได้ดังนี้

ORIGIN:= 1

$$o := \begin{pmatrix} 24 \\ 98 \\ 95 \\ 23 \end{pmatrix} \quad e := \begin{pmatrix} 30 \\ 90 \\ 90 \\ 30 \end{pmatrix} \quad k := 4 \quad \text{chisquare} := \sum_{i=1}^k \frac{(o_i - e_i)^2}{e_i} \quad \text{chisquare} = 3.822$$

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤต $\chi_{0.05}^2 = 7.815$, $df = 3$ บริเวณวิกฤตคือ $\chi^2 > 7.815$

ขั้นที่ 7. โดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต
 เพราะว่า $\chi_{\text{คำนวณ}}^2 > 7.815$ เพราะฉะนั้นไม่ปฏิเสธ H_0

หลักการและขั้นตอนการทำงานด้วย SPSS for Windows

ขั้นที่ 1. กำหนดสมมติฐานหลัก H_0 : ข้อมูลมีการแจกแจงความน่าจะเป็นตามที่คาดไว้
 กำหนดสมมติฐานอื่น H_1 : ข้อมูลไม่มีการแจกแจงความน่าจะเป็นตามที่คาดไว้

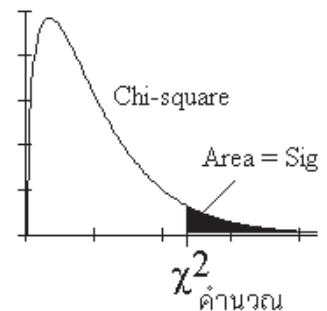
ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. ทำการสุ่มตัวอย่างเพื่อหาค่าสังเกต o_i

ขั้นที่ 4. เลือกค่าสถิติไคสแควร์

ขั้นที่ 5. คำนวณค่าความถี่ที่คาดว่าจะได้ e_i

ค่าสถิติ $\chi_{\text{คำนวณ}}^2 = \sum_{i=1}^k \frac{(o_i - e_i)^2}{e_i}$ และค่า Sig ของ $\chi_{\text{คำนวณ}}^2$



ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤต χ_{α}^2
 $df = k -$ จำนวนพารามิเตอร์ที่ต้องประมาณค่า $- 1$
 บริเวณวิกฤตคือ $\chi^2 > \chi_{\alpha}^2$

ขั้นที่ 7. สรุปผลทำได้ 2 แบบคือ
 1. ถ้า $\chi_{\text{คำนวณ}}^2 > \chi_{\alpha}^2$ แล้วปฏิเสธ H_0
 หรือ 2. ถ้า $\text{Sig} < \alpha$ แล้วปฏิเสธ H_0

จากตัวอย่าง 7.6.1. การทดลองโยนเหรียญ 3 อัน 240 ครั้ง

ให้ x เป็นจำนวนหัวที่ได้ในการโยนเหรียญแต่ละครั้ง ผลการทดลองบันทึกไว้ที่แฟ้มข้อมูลชื่อ example14.sav

จงทดสอบสมมติฐานว่าเหรียญทั้งสามอันมีความเที่ยงตรง กำหนดระดับนัยสำคัญ 0.05

วิธีทำ

การวิเคราะห์ข้อมูลด้วย SPSS for Windows

ขั้นที่ 1. กำหนดสมมติฐานหลัก H_0 : เหรียญทั้งสามอันมีความเที่ยงตรง
 กำหนดสมมติฐานอื่น H_1 : เหรียญทั้งสามอันไม่มีความเที่ยงตรง

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

ขั้นที่ 3. ทำการสุ่มตัวอย่างเพื่อหาค่าสังเกต o_i

หมายเหตุ การเก็บข้อมูล ครั้งที่ 1 ชั้นหัว 2 ครั้ง
 ครั้งที่ 2 ชั้นหัว 3 ครั้ง
 :
 ครั้งที่ 240 ชั้นหัว 3 ครั้ง

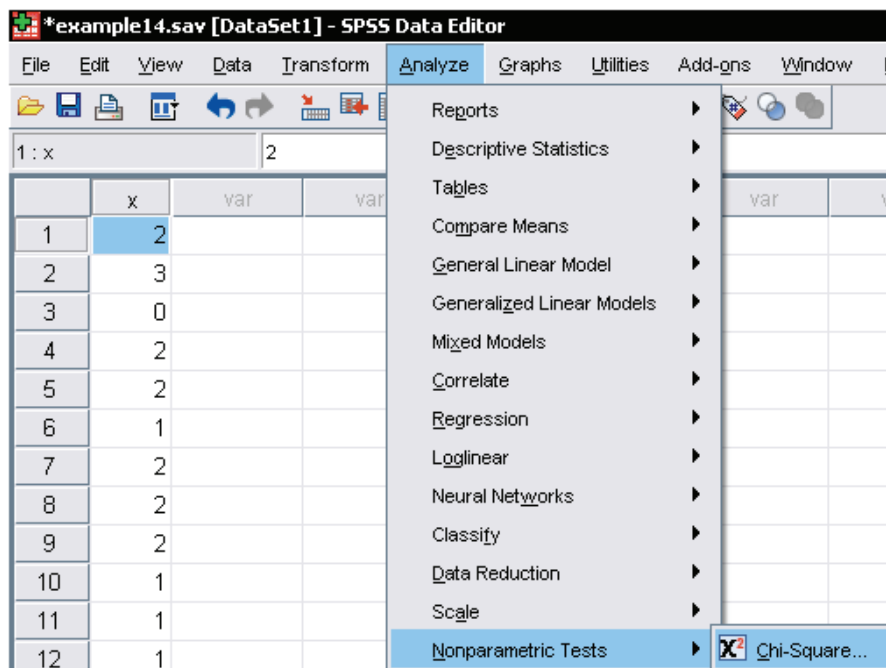
บันทึกแฟ้มชื่อ example14.sav

ขั้นที่ 4. เลือกค่าสถิติไคสแควร์

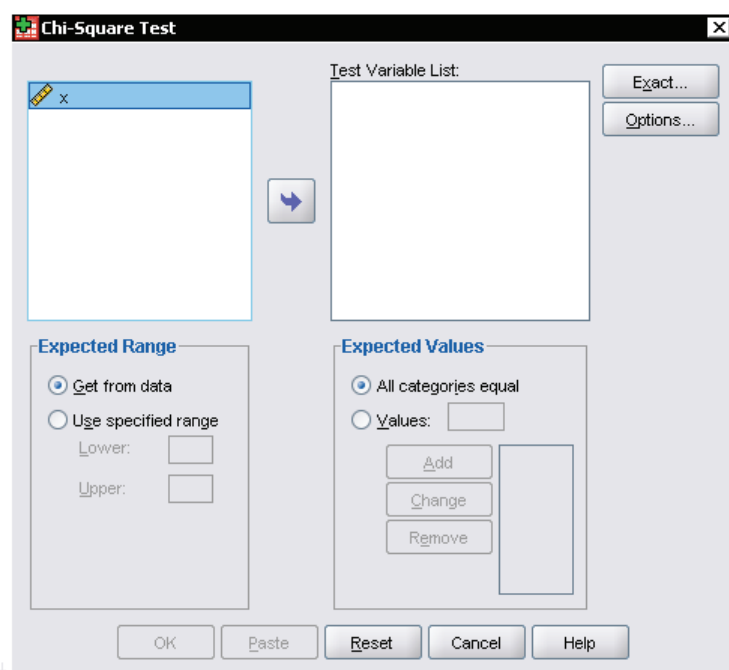
ขั้นที่ 5. คำนวณค่าความถี่ที่คาดว่าจะได้ e_i

ขั้นที่ 5.1 เลือกคำสั่ง Analyze / Nonparametric Tests / Chi-Square

	x	var
1	2	
2	3	
3	0	
238	1	
239	2	
240	3	

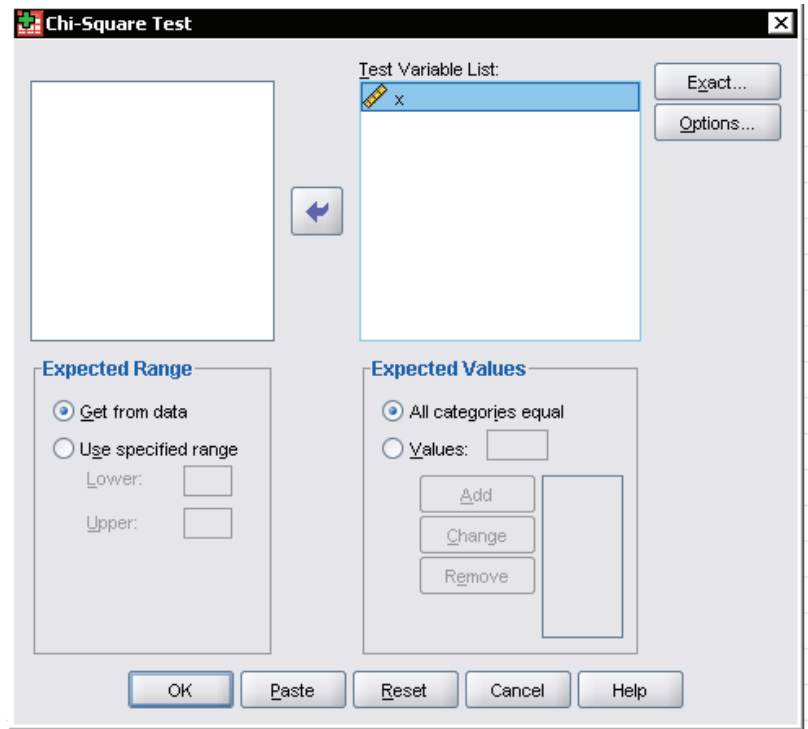


ขั้นที่ 5.2 คลิกที่ Chi-Square จะได้เมนูย่อย



ขั้นที่ 5.3 เลือกตัวแปร x มาไว้ที่ช่อง

Test Variable List



หมายเหตุ ความหมายของ Options ที่ต้องเลือกให้เหมาะสมกับการทดสอบ

Expected Range Get from data หมายความว่าให้เลือกกลุ่มของข้อมูลจากข้อมูลทั้งหมด

Use specified range ใช้กำหนดช่วงข้อมูลที่ต้องการโดยระบุค่าต่ำสุดและสูงสุด

Expected Value All categories equal หมายความว่าค่าคาดคะเน หรือสัดส่วน ของทุกกลุ่มเท่ากัน

Values ให้เรากำหนดค่าความถี่ที่คาดไว้หรือค่าสัดส่วน ด้วยการพิมพ์เข้าไปใหม่

ขั้นที่ 5.4 เพราะว่าเราต้องการใช้ข้อมูลทุกตัวในแฟ้ม

เพราะฉะนั้นเลือก Option Get from data

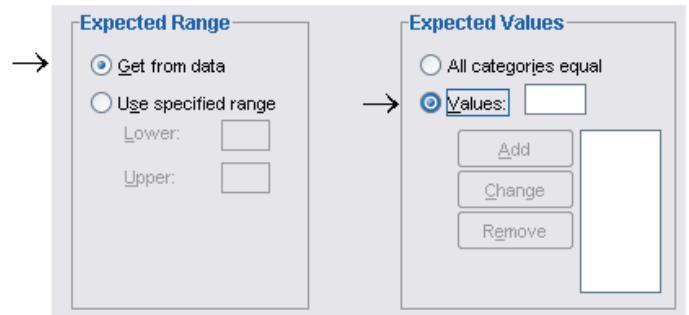
เพราะว่าความน่าจะเป็นที่จะขึ้นหัว 0, 1, 2, 3 เหรียญ

มีค่าเป็น $\frac{1}{8}, \frac{3}{8}, \frac{3}{8}, \frac{1}{8}$

เพราะฉะนั้นสัดส่วนของกลุ่มที่ 1, 2, 3, 4

คือ $\frac{1}{8} : \frac{3}{8} : \frac{3}{8} : \frac{1}{8} = 1 : 3 : 3 : 1$

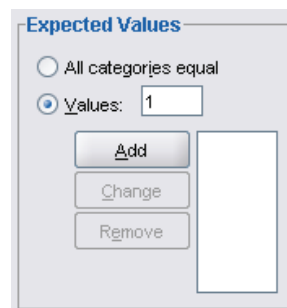
เพราะฉะนั้นต้องเลือก Option Values



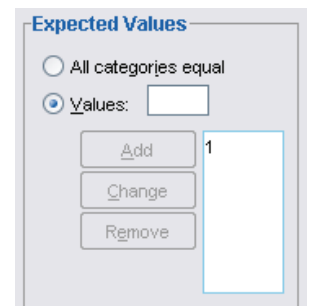
ขั้นที่ 5.5 พิมพ์ค่าสัดส่วนตัวแรกคือ 1 ลงในช่อง Values :

ผลบนจอภาพที่ช่อง Expected values จะเป็นดังนี้

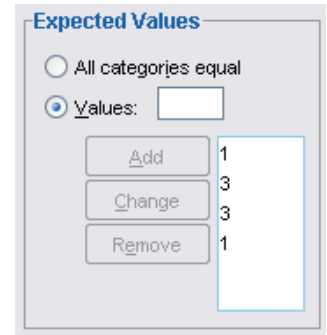
จะเห็นว่าปุ่ม Add จะเปลี่ยนจากสีเทาจางๆ เป็นสีเข้ม



ขั้นที่ 5.6 คลิก Add ตัวเลข 1 จะเข้าไปในช่องสี่เหลี่ยมด้านล่าง



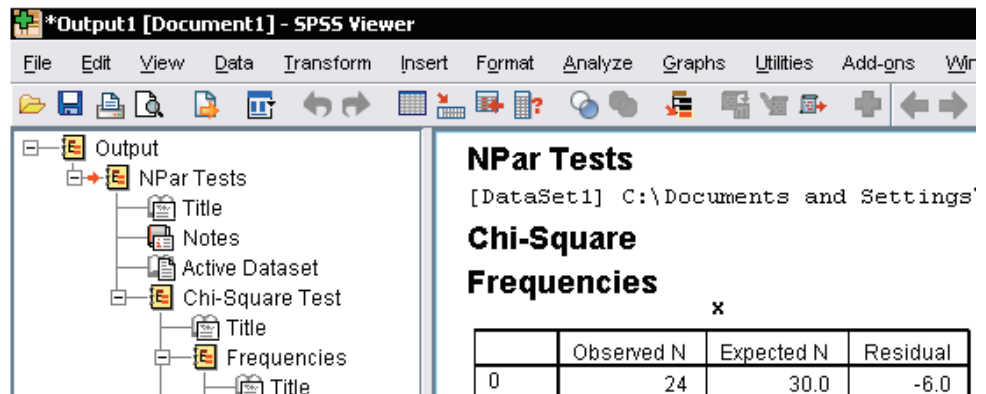
ในทำนองเดียวกัน พิมพ์ 3 ในช่อง Values แล้วคลิก Add
 พิมพ์ 3 ในช่อง Values แล้วคลิก Add
 พิมพ์ 1 ในช่อง Values แล้วคลิก Add



ผลบนจอภาพที่ช่อง Expected Values จะเป็นดังนี้

หมายเหตุ

1. ถ้าต้องการลบตัวเลขสัดส่วน ให้คลิกที่ตัวเลขนั้นแล้วกด Remove
 2. ถ้าต้องการแก้ไขตัวเลขสัดส่วน ให้คลิกที่ตัวเลข แล้วพิมพ์ค่าใหม่ในช่อง Values แล้วกด Change
- ขั้นที่ 6. คลิก OK จะได้ผลการคำนวณดังนี้



ผลการคำนวณทั้งหมดคือ

NPar Tests

[DataSet1] C:\Documents and Settings\dumrong\My Documents\dataspss16\example14.sav

Chi-Square Test Frequencies

	x		
	Observed N	Expected N	Residual
0	24	30.0	-6.0
1	98	90.0	8.0
2	95	90.0	5.0
3	23	30.0	-7.0
Total	240		

Test Statistics

	x
Chi-Square	3.8222222 ^a
df	3
Asymp. Sig.	0.28131187

a. 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 30.0.

จากตาราง Test Statistics จะได้ $\chi^2_{\text{คำนวณ}} = \sum_{i=1}^k \frac{(o_i - e_i)^2}{e_i} = 3.822222$, df = 3

และ Asymp. Sig. = 0.281312

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤต $\chi^2_{0.05} = 7.815$, $df = 3$ บริเวณวิกฤตคือ $\chi^2 > 7.815$

ขั้นที่ 7. เพราะว่า $\chi^2_{คำนวณ} < 7.815$ เพราะฉะนั้นไม่ปฏิเสธ H_0

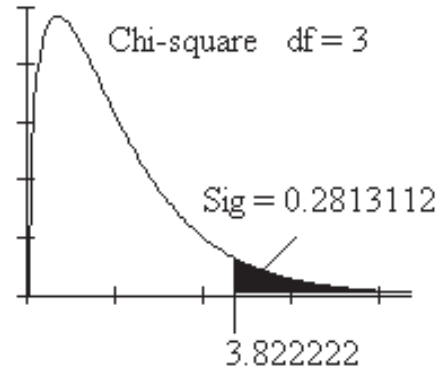
หรือ เพราะว่า Sig = 0.281312 มากกว่า 0.05 เพราะฉะนั้นไม่ปฏิเสธ H_0

หมายเหตุ ค่า Sig = พื้นที่ใต้โค้งทางทางด้านขวาของเส้นโค้ง Chi square, $df = 3$ ที่ระยะ 3.822222 จากการคำนวณด้วย Mathcad

$$v := 3 \quad f(x) := \left(\frac{1}{2^{\frac{v}{2}} \cdot \Gamma\left(\frac{v}{2}\right)} \right) \cdot x^{\left(\frac{v}{2}\right)-1} \cdot e^{-\frac{x}{2}}$$

$$\text{Sig} := 1 - \int_0^{3.822222} f(x) dx \quad \text{Sig} = 0.281312$$

$$\text{Sig} := 1 - \text{pchisq}(3.822222, 3) \quad \text{Sig} = 0.281312$$



7.7 การทดสอบสมมติฐานว่าข้อมูลเป็นอิสระต่อกันหรือไม่

ในกรณีที่เรารู้ต้องการทดสอบความสัมพันธ์ของตัวแปรตั้งแต่ 2 ตัวเกี่ยวข้องกันหรือไม่ ตัวอย่างเช่น

- การฉีดวัคซีนป้องกันอหิวาต์ กับ การเป็นโรคอหิวาต์ เกี่ยวข้องกันหรือไม่
- การนับถือศาสนา และ ถิ่นที่อยู่ เกี่ยวข้องกันหรือไม่

หลักการและขั้นตอนการทำงานทางทฤษฎีความน่าจะเป็นและสถิติในการทดสอบความเป็นอิสระ

ขั้นที่ 1. กำหนดสมมติฐาน H_0 : เหตุการณ์ A และเหตุการณ์ B เป็นอิสระต่อกัน

H_1 : เหตุการณ์ A และเหตุการณ์ B ไม่เป็นอิสระต่อกัน

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. ทำการสุ่มตัวอย่างเพื่อหาค่าสังเกต o_{ij}

ขั้นที่ 4. เลือกค่าสถิติไคสแควร์

ขั้นที่ 5. คำนวณค่าความถี่ที่คาดหวังจะได้ e_{ij} และค่าสถิติไคสแควร์

โดยทำการแจกแจงความถี่เพื่อหาค่า o_{ij} และ $e_{ij} = \frac{R_i C_j}{N}$ เมื่อ $i = 1, 2, 3, \dots, r$ และ $j = 1, 2, 3, \dots, c$

	A ₁	A ₂	A ₃	...	A _c	
B ₁	o_{11}	o_{12}	o_{13}	...	o_{1c}	R ₁
B ₂	o_{21}	o_{22}	o_{23}	...	o_{2c}	R ₂
:	:	:	:		:	:
:	:	:	:		:	:
B _r	o_{r1}	o_{r2}	o_{r3}	...	o_{rc}	R _r
	C ₁	C ₂	C ₃	...	C _c	N

หมายเหตุ e_{ij} ควรมีค่ามากกว่าหรือเท่ากับ 5 และ N ควรมีค่ามากกว่าหรือเท่ากับ 50

$$\chi^2_{\text{คำนวณ}} = \sum_{i=1}^r \sum_{j=1}^c \frac{(o_{ij} - e_{ij})^2}{e_{ij}} \text{ และ ระดับชั้นความเสรี } v = (r - 1)(c - 1)$$

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤต χ^2_{α} , $df = (r - 1)(c - 1)$ และบริเวณวิกฤตคือ $\chi^2 > \chi^2_{\alpha}$

ขั้นที่ 7. สรุปผลโดยการเปรียบเทียบค่าสถิติ $\chi^2_{\text{คำนวณ}}$ จากตัวอย่าง กับ ค่าวิกฤต

ถ้า $\chi^2_{\text{คำนวณ}} > \chi^2_{\alpha}$ แล้วปฏิเสธ H_0

หลักการและขั้นตอนการทดสอบสมมติฐานด้วย SPSS for Windows

ขั้นที่ 1. กำหนดสมมติฐานหลัก H_0 : ข้อมูลเป็นอิสระต่อกัน

กำหนดสมมติฐานอื่น H_1 : ข้อมูลไม่เป็นอิสระต่อกัน

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. ทำการสุ่มตัวอย่างเพื่อหาค่าสังเกต o_{ij}

ขั้นที่ 4. เลือกค่าสถิติไคสแควร์

ขั้นที่ 5. คำนวณค่าความถี่ที่คาดว่าจะได้ e_{ij}

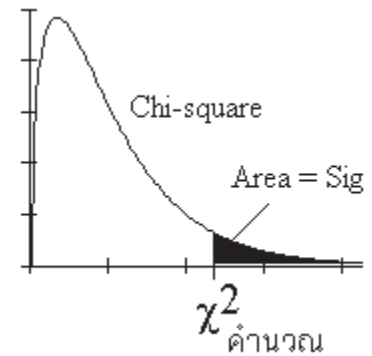
$$\text{และค่าสถิติไคสแควร์ } \chi^2_{\text{คำนวณ}} = \sum_{i=1}^r \sum_{j=1}^c \frac{(o_{ij} - e_{ij})^2}{e_{ij}}$$

และค่า Sig (หมายเหตุ Sig = $P(\chi^2 > \chi^2_{\text{คำนวณ}})$)

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤต χ^2_{α} , $df = (r - 1)(c - 1)$ และบริเวณวิกฤตคือ $\chi^2 > \chi^2_{\alpha}$

ขั้นที่ 7. สรุปผลทำได้ 2 แบบคือ 1. ถ้า $\chi^2_{\text{คำนวณ}} > \chi^2_{\alpha}$ แล้วปฏิเสธ H_0

2. ถ้า Sig < α แล้วปฏิเสธ H_0



ตัวอย่าง 7.7.1 แฟ้มข้อมูล example15.sav บันทึกข้อมูลเพื่อศึกษาความสัมพันธ์ระหว่างการนับถือศาสนาและถิ่นที่อยู่ (download จาก <http://www.math.sc.chula.ac.th/~tdumrong/2301286data>)

การนับถือศาสนาจำแนกเป็น 3 กลุ่ม คือ โปรเตสตันท์ คาทอลิก และ ยิว

ถิ่นที่อยู่อาศัยจำแนกเป็น 2 กลุ่ม คือ ฝั่งตะวันออก และ ฝั่งตะวันตก

จงทดสอบสมมติฐานว่า การนับถือศาสนาและถิ่นที่อยู่มีความสัมพันธ์กันหรือไม่ กำหนดนัยสำคัญ 0.05

วิธีทำ การคำนวณโดย SPSS for Windows

ขั้นที่ 1. กำหนดสมมติฐานหลัก H_0 : การนับถือศาสนาและถิ่นที่อยู่อาศัย ไม่มีความสัมพันธ์กัน

กำหนดสมมติฐานอื่น H_1 : การนับถือศาสนาและถิ่นที่อยู่อาศัย มีความสัมพันธ์กัน

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

ขั้นที่ 3. ทำการสุ่มตัวอย่างเพื่อหาค่าสังเกต o_{ij}

เมื่อสุ่มตัวอย่างมาแล้วต้องสร้าง แฟ้มข้อมูลประกอบด้วย 2 ตัวแปร (สร้างให้แล้วชื่อ example15.sav)

x เป็นตัวแปรจำแนก ถิ่นที่อยู่อาศัย โดยมีค่า Value Label : 1 = ฝั่งตะวันออก และ 2 = ฝั่งตะวันตก

y เป็นตัวแปรจำแนก ศาสนา โดยมีค่า Value Label : 1 = โปรเตสตันท์, 2 = คาทอลิก และ 3 = ยิว

นำแฟ้มชื่อ example15.sav เข้ามาทำงานใน SPSS Data Editor

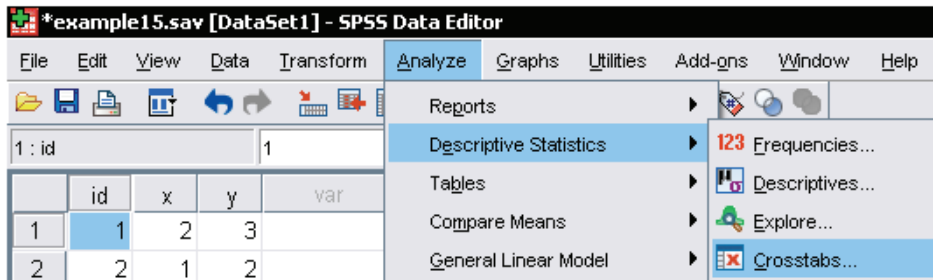
ขั้นที่ 4. เลือกค่าสถิติไคสแควร์

ขั้นที่ 5. คำนวณค่าความถี่ที่คาดว่าจะได้ e_{ij}

$$\text{และค่าสถิติไคสแควร์ } \chi^2_{\text{คำนวณ}} = \sum_{i=1}^r \sum_{j=1}^c \frac{(o_{ij} - e_{ij})^2}{e_{ij}}$$

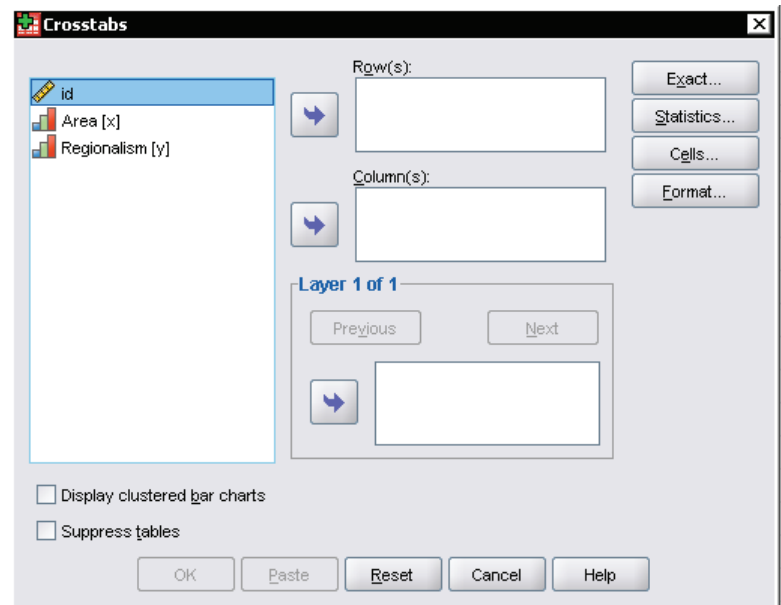
	id	x	y	var
1	1	2	3	
2	2	1	2	

ขั้นที่ 5.1 เลือก Analyze / Descriptive Statistics / Crosstabs



ขั้นที่ 5.2 คลิกที่ Crosstabs

จะได้เมนูดังนี้



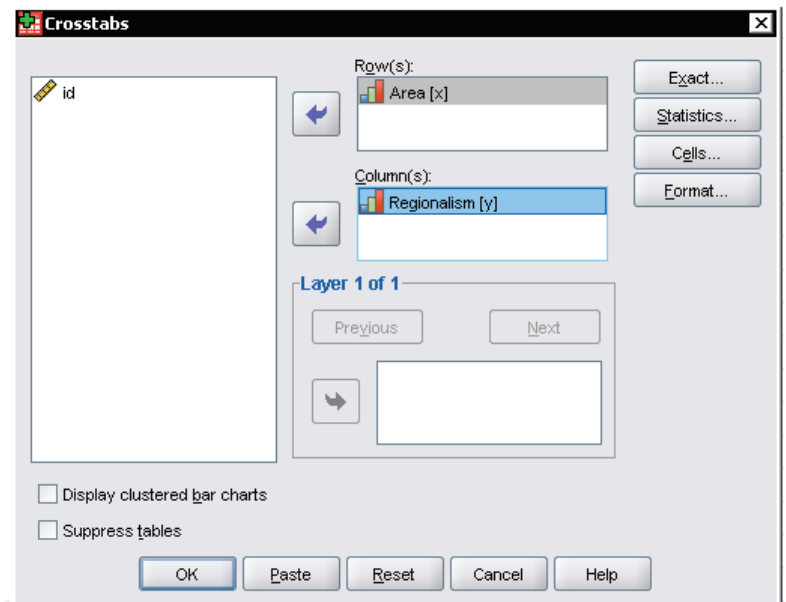
ขั้นที่ 5.3

เลือกตัวแปร x

ไปไว้ที่ช่อง Row(s)

เลือกตัวแปร y

ไปไว้ที่ช่อง Column(s)



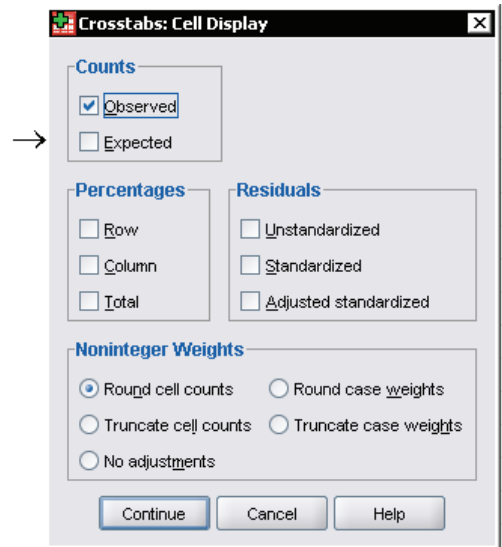
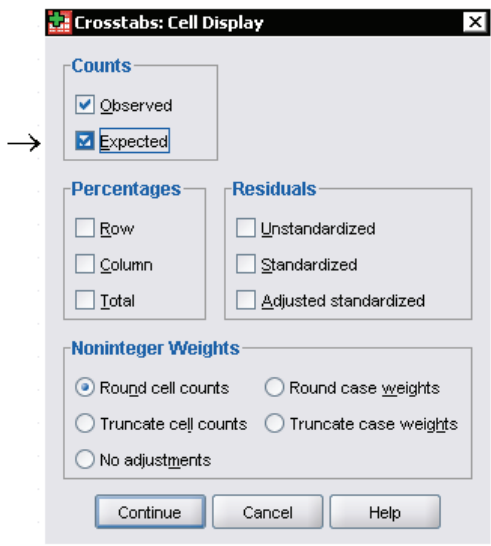
เพราะว่าเราต้องการให้คำนวณค่าคาดคะเนทางทฤษฎี

เพราะฉะนั้นต้องเลือก 

ขั้นที่ 5.4 คลิกที่ Cells

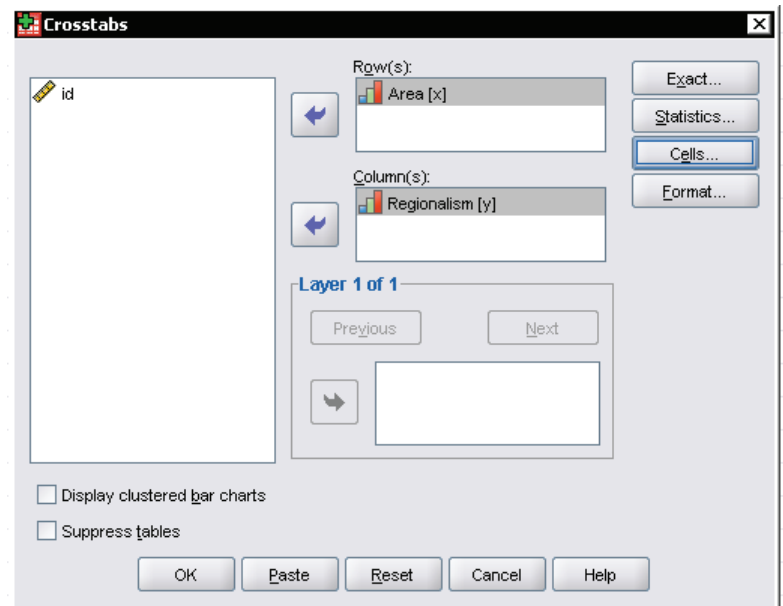
จะได้เมนูย่อย Crosstabs Cell Display เป็นดังนี้

ขั้นที่ 5.5 คลิกที่ Expected ให้เกิดเครื่องหมายถูก



ขั้นที่ 5.6 คลิก Continue

จะกลับมาที่เมนู Crosstabs

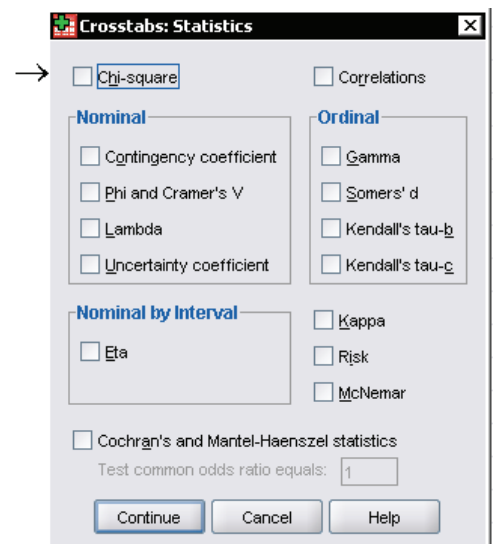


ขั้นที่ 5.7 เพราะเราต้องการ

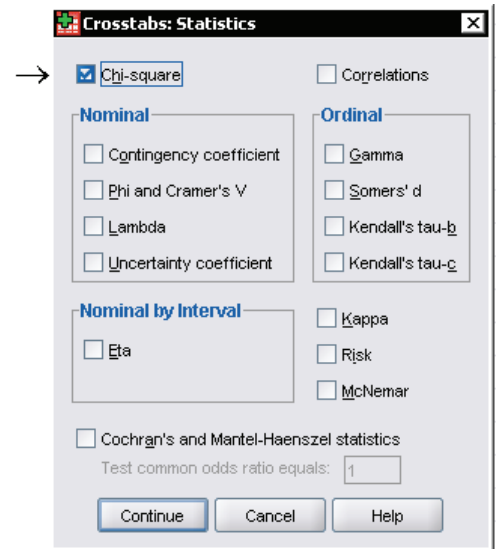
ค่าสถิติไคสแควร์

เพราะฉะนั้นคลิกปุ่ม 

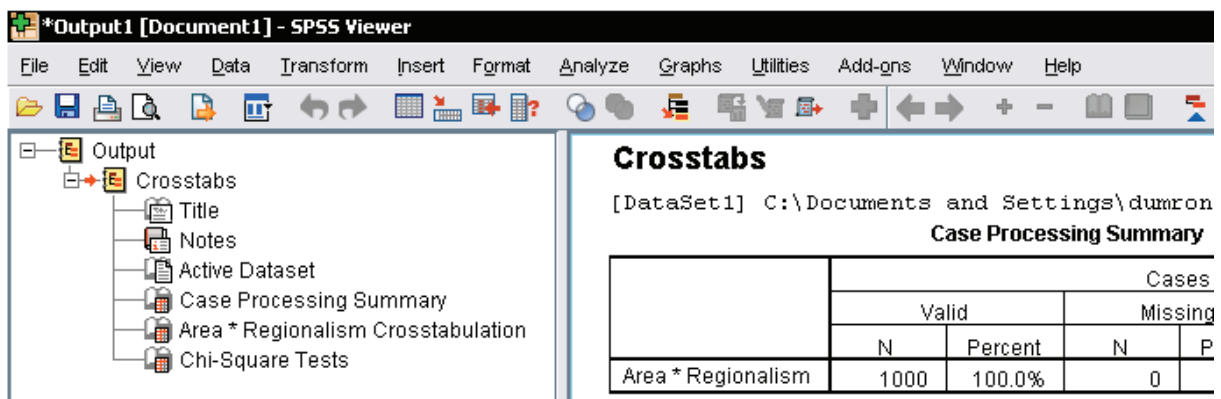
จะได้เมนูย่อย Crosstabs Statistics



ขั้นที่ 5.8 เลือก Chi-square



ขั้นที่ 5.9 คลิก Continue และ OK ตามลำดับ
จะได้ผลการคำนวณดังนี้



ผลการคำนวณทั้งหมดคือ

Crosstabs

[DataSet1] C:\Documents and Settings\dumrong\My Documents\dataspss16\example15.sav

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Area * Regionalism	1000	100.0%	0	.0%	1000	100.0%

Area * Regionalism Crosstabulation

			Regionalism			Total
			Protestant	Christ	Jew	
Area	East	Count	182	215	203	600
		Expected Count	201.6	210.6	187.8	600.0
	West	Count	154	136	110	400
		Expected Count	134.4	140.4	125.2	400.0
Total		Count	336	351	313	1000
		Expected Count	336.0	351.0	313.0	1000.0

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	8.069321 ^a	2	0.017692
Likelihood Ratio	8.053133	2	0.017835
Linear-by-Linear Association	7.773600	1	0.005302
N of Valid Cases	1000		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 125.20.

หมายเหตุ Remark ท้ายตาราง Chi-Square Tests เตือนว่าค่าคาดคะเนทุกเซลล์มีค่ามากกว่า 5 ทำให้ค่าไคสแควร์ที่คำนวณได้มีความน่าเชื่อถือสามารถนำไปสรุปผลการทดสอบได้

จากตาราง Chi-Square Tests $\chi^2_{\text{คำนวณ}} = \sum_{i=1}^2 \sum_{j=1}^3 \frac{(o_{ij} - e_{ij})^2}{e_{ij}} = 8.069321$, df = 2 และ Sig = 0.017692

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤต $\chi^2_{0.05} = 5.991$, df = 2 และบริเวณวิกฤตคือ $\chi^2 > 5.991$

ขั้นที่ 7. แบบที่ 1 เพราะว่า $\chi^2_{\text{คำนวณ}} > 5.991$ เพราะฉะนั้นปฏิเสธ H_0

แบบที่ 2 เพราะว่า Sig < 0.05 เพราะฉะนั้นปฏิเสธ H_0

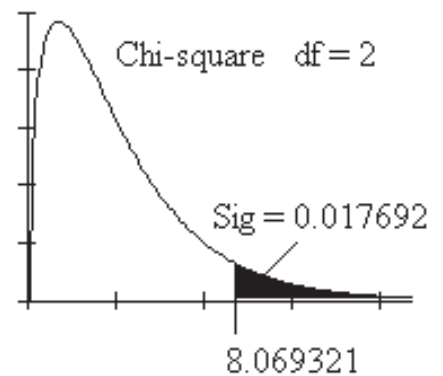
หมายเหตุ ความหมายและที่มาของค่า Asymp. Sig. (2-sided) = 0.017692

การคำนวณด้วย Mathcad

$$v := 2 \quad f(x) := \left(\frac{1}{2^{\frac{v}{2}} \cdot \Gamma\left(\frac{v}{2}\right)} \right) \cdot x^{\left(\frac{v}{2}\right)-1} \cdot e^{-\frac{x}{2}}$$

$$\text{Sig} := 1 - \int_0^{8.069321} f(x) dx \quad \text{Sig} = 0.017692$$

$$\text{Sig} := 1 - \text{pchisq}(8.069321, 2) \quad \text{Sig} = 0.017692$$



หมายเหตุ ในกรณีที่ข้อมูลแจกแจงความถี่แล้ว การคำนวณโดยใช้ Mathcad ทำได้ดังนี้

$$\text{ORIGIN} := 1 \quad r := 2 \quad i := 1..r \quad c := 3 \quad j := 1..c \quad o := \begin{pmatrix} 182 & 215 & 203 \\ 154 & 136 & 110 \end{pmatrix} \quad N := \sum_{i=1}^r \sum_{j=1}^c o_{(i,j)}$$

$$N = 1000 \quad R_i := \sum_{j=1}^c o_{(i,j)} \quad R = \begin{pmatrix} 600 \\ 400 \end{pmatrix} \quad C_j := \sum_{i=1}^r o_{(i,j)} \quad C = \begin{pmatrix} 336 \\ 351 \\ 313 \end{pmatrix} \quad e_{(i,j)} := \frac{R_i \cdot C_j}{N}$$

$$e = \begin{pmatrix} 201.6 & 210.6 & 187.8 \\ 134.4 & 140.4 & 125.2 \end{pmatrix} \quad \text{chisquare} := \sum_{i=1}^r \sum_{j=1}^c \frac{[o_{(i,j)} - e_{(i,j)}]^2}{e_{(i,j)}} \quad \text{chisquare} = 8.069321$$

7.8 การทดสอบสมมติฐานว่าข้อมูลมีการแจกแจงปกติจริงหรือไม่

การทำงานทางด้านสถิติส่วนใหญ่สมมติว่าข้อมูลมีการแจกแจงปกติ

แต่ถ้าเรามีข้อสงสัยว่าข้อมูลมีการแจกแจงปกติจริงหรือไม่ สามารถทำการทดสอบสมมติฐานได้ดังนี้

- ขั้นที่ 1. กำหนดสมมติฐาน H_0 : ข้อมูลมีการแจกแจงปกติ
 H_1 : ข้อมูลไม่ได้มีการแจกแจงปกติ
- ขั้นที่ 2. กำหนดระดับนัยสำคัญ α
- ขั้นที่ 3. ทำการสุ่มตัวอย่าง
- ขั้นที่ 4. เลือกค่าสถิติ Kolmogorov หรือ Shapiro-Wilk
ถ้าขนาดตัวอย่างมากกว่า 50 ใช้ค่าสถิติ Kolmogorov
ถ้าขนาดตัวอย่างน้อยกว่าหรือเท่ากับ 50 ใช้ค่าสถิติ Kolmogorov หรือ Shapiro-Wilk
- ขั้นที่ 5. คำนวณค่าสถิติที่ต้องการ
- ขั้นที่ 6. การสรุปผลใช้การเปรียบเทียบค่า Sig กับค่านัยสำคัญ α
- ขั้นที่ 7. การสรุปผล ถ้า Sig < α แล้วปฏิเสธ H_0

ตัวอย่าง 7.7.1 จากแฟ้มข้อมูล example4.sav

จงทดสอบว่าข้อมูลระดับคะแนน (grade) มีการแจกแจงปกติ กำหนดระดับนัยสำคัญ 0.05

วิธีทำ

- ขั้นที่ 1. กำหนดสมมติฐาน H_0 : ระดับคะแนน (grade) มีการแจกแจงปกติ
 H_1 : ระดับคะแนน (grade) ไม่ได้มีการแจกแจงปกติ
- ขั้นที่ 2. กำหนดระดับนัยสำคัญ α
- ขั้นที่ 3. ทำการสุ่มตัวอย่าง

การคำนวณโดย SPSS for Windows

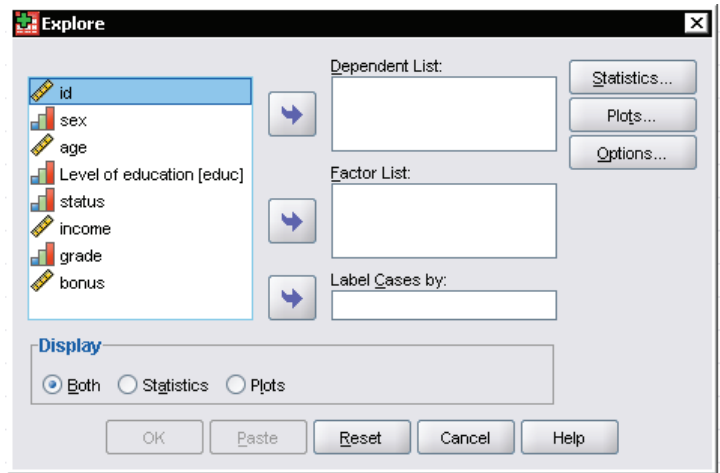
- ขั้นที่ 4. เพราะว่าขนาดตัวอย่าง $N = 50$ เพราะฉะนั้นเลือกค่าสถิติ Kolmogorov
- ขั้นที่ 5. เปิดแฟ้มข้อมูล example4.sav ใน SPSS Data Editor

	id	sex	age	educ	status	income	grade	bonus	var
1	1	1	37	2	4	5500	3.78	11000.00	
2	2	2	29	3	1	4100	3.89	12300.00	

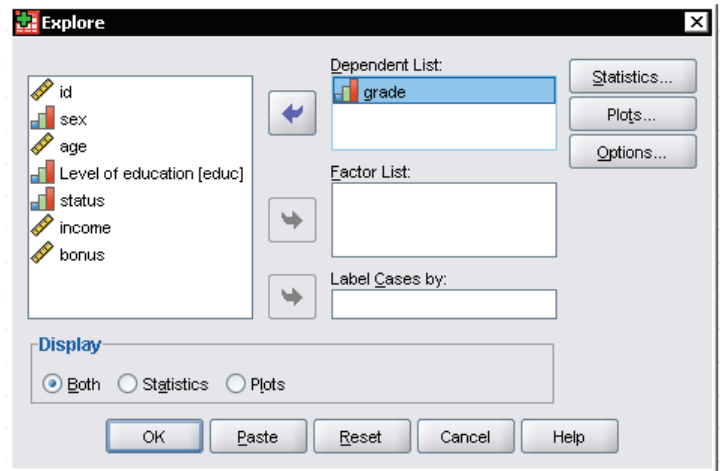
- ขั้นที่ 5.1 เลือกคำสั่ง Analyze / Descriptive Statistics / Explore

	id	sex	age	educ	s
1	1	1	37	2	

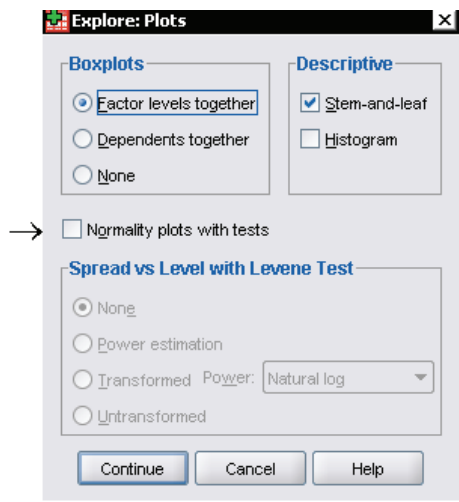
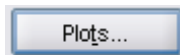
ขั้นที่ 5.2 คลิกที่ Explore จะได้เมนูย่อยเป็น



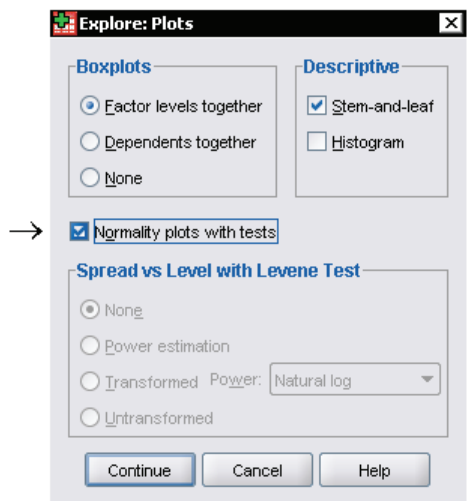
ขั้นที่ 5.3 เลือกตัวแปร grade มาไว้ที่ช่อง Dependent List



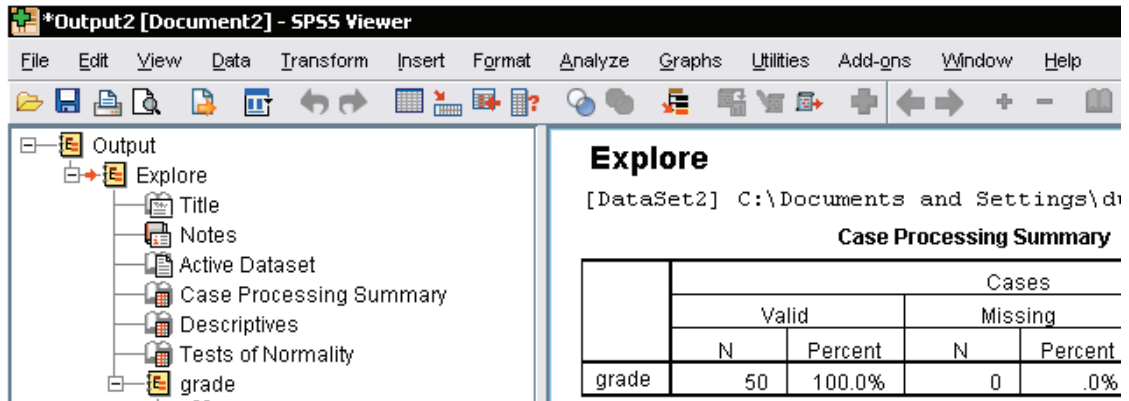
ขั้นที่ 5.4 คลิกปุ่ม จะได้เมนูย่อยเป็น



ขั้นที่ 5.5 เลือก Normality plots with tests



ขั้นที่ 5.6 คลิก Continue และ OK ตามลำดับจะได้ผลการวิเคราะห์บนจอภาพเป็นดังนี้



ผลการคำนวณที่ใช้ในการสรุปผลเราใช้ข้อมูลจากตารางสถิติ

Explore

[DataSet2] C:\Documents and Settings\dumrong\My Documents\dataspss16\example4.sav

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
grade	50	100.0%	0	.0%	50	100.0%

Descriptives

		Statistic	Std. Error
grade	Mean	2.9980	.06717
	95% Confidence Interval for Mean	Lower Bound 2.8630	Upper Bound 3.1330
	5% Trimmed Mean	2.9896	
	Median	2.8900	
	Variance	.226	
	Std. Deviation	.47496	
	Minimum	2.12	
	Maximum	3.89	
	Range	1.77	
	Interquartile Range	.82	
	Skewness	.478	.337
	Kurtosis	-.888	.662

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
grade	0.156879	50	0.003534	.929	50	0.004932

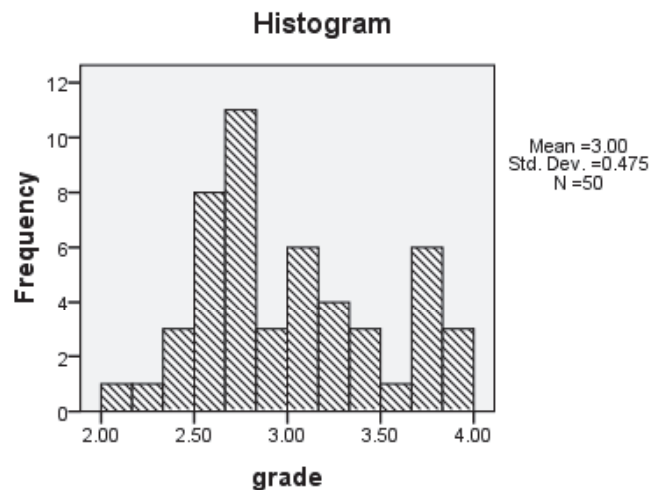
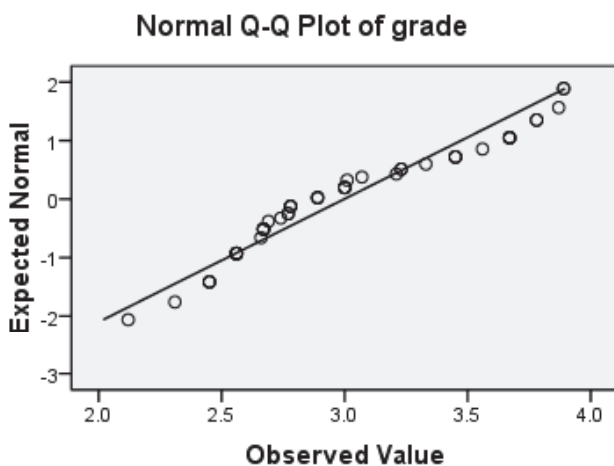
a. Lilliefors Significance Correction

grade

grade Stem-and-Leaf Plot

Frequency	Stem & Leaf
1.00	2 . 1
1.00	2 . 3
10.00	2 . 4445555555
12.00	2 . 666666777777
3.00	2 . 888
6.00	3 . 000000
4.00	3 . 2223
4.00	3 . 4445
6.00	3 . 666677
3.00	3 . 888

Stem width: 1.00
Each leaf: 1 case(s)



และตารางอื่น ๆ

หมายเหตุ ตารางที่สำคัญในการสรุปผลว่าข้อมูลมีการแจกแจงปกติจริงหรือไม่คือ ตาราง Test of Normality

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
grade	0.156879	50	0.003534	.929	50	0.004932

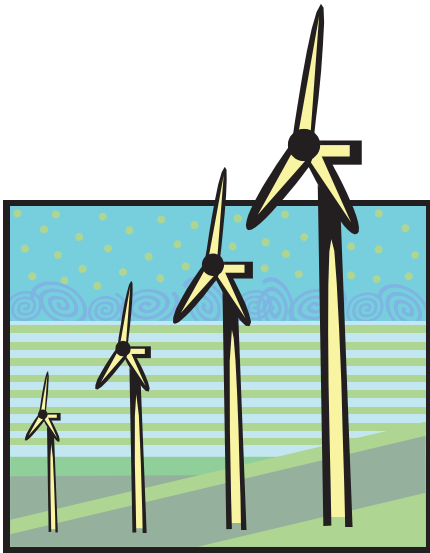
a. Lilliefors Significance Correction

ขั้นที่ 6. ค่าสถิติ Kolmogorov มีค่าเท่ากับ 0.156879 และมีค่า Sig = 0.003534

ขั้นที่ 7. เพราะว่า Sig = 0.003534 น้อยกว่าค่านัยสำคัญ $\alpha = 0.05$ ดังนั้นปฏิเสธ H_0 เพราะฉะนั้น ระดับคะแนน (grade) ไม่ได้มีการแจกแจงปกติ

บทที่ 8

สหสัมพันธ์และการถดถอยเชิงเส้น



ตัวแปรทางด้านสถิติที่เกิดจากการเก็บข้อมูลเช่น จากบริษัทประกันภัย ตัวแปรอาจเป็น รายได้ ผลกำไร ค่าโฆษณา เบี้ยประกันภัย ฯลฯ เมื่อพบว่าตัวแปรที่ได้จากหน่วยทดลองตั้งแต่ 2 ตัวขึ้นไปจึงเกิดความสนใจว่าตัวแปรเหล่านี้จะมีความสัมพันธ์กันหรือไม่ ถ้ามีความสัมพันธ์กันจะอยู่ในรูปใด การศึกษาเรื่องนี้จัดอยู่ในเรื่องของถดถอยซึ่งมีทั้งที่เป็นเส้นตรงและไม่เป็นเส้นตรง ในบทนี้จึงเป็นการศึกษาเกี่ยวกับ การถดถอยเชิงเส้นเชิงเดียว (simple linear regression) และ สหสัมพันธ์ (correlation)

การถดถอยเชิงเส้นเชิงเดียว (simple linear regression) ประกอบด้วยตัวแปรอิสระ 1 ตัว และตัวแปรตาม 1 ตัว รูปแบบของสมการความสัมพันธ์เชิงเส้นเชิงเดียวมีได้หลายรูปแบบเช่น

$$\bullet y = a + bx \quad \bullet \ln y = a + b \ln x \quad \bullet y = a + b \ln x \quad \bullet \ln y = a + bx$$

การถดถอยแบบพหุคูณ (Multiple regression) ประกอบด้วยตัวแปรอิสระมากกว่า 1 ตัว และตัวแปรตาม 1 ตัว รูปแบบของสมการความสัมพันธ์พหุคูณ เช่น

$$\bullet y = a + b_1 x_1 + b_2 x_2 + \dots + b_n x_n$$

การถดถอยแบบไม่เชิงเส้น (Non linear regression) ประกอบด้วยตัวแปรอิสระ 1 ตัว และตัวแปรตาม 1 ตัว รูปแบบของสมการความสัมพันธ์แบบไม่เชิงเส้น เช่น

$$\bullet y = a x^2 + bx + c$$

สหสัมพันธ์ (Simple correlation) เป็นการศึกษาระดับความสัมพันธ์และทิศทางความสัมพันธ์ของตัวแปรตั้งแต่ 2 ตัวแปรขึ้นไปโดยที่รูปแบบความสัมพันธ์อาจเป็น

$$\bullet y = a + bx \quad \bullet \ln y = a + b \ln x \quad \bullet y = a + b \ln x \quad \bullet \ln y = a + bx$$
$$\bullet y = a x^2 + bx + c \quad \bullet y = a + b_1 x_1 + b_2 x_2 + \dots + b_n x_n$$

8.1 การหาสมการเส้นถดถอยเชิงเส้นเชิงเดียว(Simple Linear Regression) และสหสัมพันธ์ (Correlation)

ให้ X เป็นตัวแปรอิสระ และ Y เป็นตัวแปรตาม โดยมีความสัมพันธ์ที่แท้จริงของ X, Y คือ $\mu_{Y|X} = \alpha + \beta x$ β เรียกว่า สัมประสิทธิ์การถดถอย (regression coefficients) และ α เรียกว่า ระยะเวลาตัดแกน (Intercept)

ในทางสถิติเราจะใช้ข้อมูลตัวอย่างประมาณความสัมพันธ์ $\mu_{Y|X} = \alpha + \beta x$ ด้วย $\hat{y} = a + bx$

สหสัมพันธ์ (correlation) เป็นตัวบอกระดับและทิศทางของความสัมพันธ์ระหว่างตัวแปรใช้สัญลักษณ์แทนด้วย ρ โดยใช้ข้อมูลตัวอย่างเราจะประมาณค่า ρ ด้วย r

หมายเหตุ 1. $-1 \leq \rho \leq 1$

2. $|\rho|$ มีค่ามาก แสดงว่า X และ Y มีความสัมพันธ์กันมาก (ความสัมพันธ์เชิงเส้นตรง $\mu_{Y|X} = \alpha + \beta x$)

3. $\rho = 0$ แสดงว่า X และ Y ไม่มีความสัมพันธ์ (ความสัมพันธ์เชิงเส้นตรง $\mu_{Y|X} = \alpha + \beta x$)

4. $\rho > 0$ แสดงว่าถ้า X มีค่าเพิ่มขึ้น แล้ว Y มีค่าเพิ่มขึ้น หรือ ถ้า X มีค่าลดลง แล้ว Y มีค่าลดลง

5. $\rho < 0$ แสดงว่าถ้า X มีค่าเพิ่มขึ้น แล้ว Y มีค่าลดลง หรือ ถ้า X มีค่าลดลง แล้ว Y มีค่าเพิ่มขึ้น

6. b และ r จะมีเครื่องหมายเหมือนกัน

7. b สามารถบอกอัตราการเพิ่มหรืออัตราการลดของตัวแปรตาม Y เทียบกับตัวแปรอิสระ X ได้

การหาสมการเส้นถดถอยเชิงเส้นเชิงเดียวและสหสัมพันธ์

หลักการทางทฤษฎีของความน่าจะเป็นและสถิติจากข้อมูลที่เก็บมาได้

ตัวอย่างจากประชากรชุดที่ 1.	ตัวอย่างจากประชากรชุดที่ 2.
x_1	y_1
x_2	y_2
:	:
x_n	y_n

ขั้นตอนการคำนวณ ค่า a และ b ที่ทำให้ $\hat{y} = a + bx$ และ สัมประสิทธิ์สหสัมพันธ์ r

ขั้นที่ 1. คำนวณค่า $\sum_{i=1}^n x_i$, $\sum_{i=1}^n y_i$, $\sum_{i=1}^n x_i y_i$, $\sum_{i=1}^n x_i^2$, $\sum_{i=1}^n y_i^2$

ขั้นที่ 2. คำนวณค่า $b = \frac{n \sum_{i=1}^n x_i y_i - \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right)}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}$ และ $a = \bar{y} - b \bar{x}$

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

ตัวอย่างเช่น

X	1.50	1.80	2.40	3.00	3.50	3.90	4.40	4.80	5.00
Y	4.80	5.70	7.00	8.30	10.90	12.40	13.10	13.60	15.30

เราสามารถหาสมการ $\hat{y} = a + bx$ และค่า r ตามขั้นตอนการคำนวณดังนี้

การคำนวณด้วย Mathcad

แบบที่ 1. คำนวณค่าตามสูตร

$$\text{ORIGIN} := 1 \quad x := \begin{pmatrix} 1.50 \\ 1.80 \\ 2.40 \\ 3.00 \\ 3.50 \\ 3.90 \\ 4.40 \\ 4.80 \\ 5.00 \end{pmatrix} \quad y := \begin{pmatrix} 4.80 \\ 5.70 \\ 7.00 \\ 8.30 \\ 10.90 \\ 12.40 \\ 13.10 \\ 13.60 \\ 15.30 \end{pmatrix}$$

$$n := \text{length}(x) \quad n = 9$$

$$b := \frac{n \cdot \sum_{i=1}^n x_i \cdot y_i - \sum_{i=1}^n x_i \cdot \sum_{i=1}^n y_i}{n \cdot \sum_{i=1}^n (x_i)^2 - \left(\sum_{i=1}^n x_i \right)^2} \quad b = 2.93028 \quad a := \text{mean}(y) - b \cdot \text{mean}(x) \quad a = 0.256947$$

$$r := \frac{n \cdot \sum_{i=1}^n x_i \cdot y_i - \sum_{i=1}^n x_i \cdot \sum_{i=1}^n y_i}{\sqrt{n \cdot \sum_{i=1}^n (x_i)^2 - \left(\sum_{i=1}^n x_i \right)^2} \cdot \sqrt{n \cdot \sum_{i=1}^n (y_i)^2 - \left(\sum_{i=1}^n y_i \right)^2}} \quad r = 0.991089$$

แบบที่ 2. ใช้ฟังก์ชัน slope(x,y), intercept(x,y) และ corr(x, y) ของ Mathcad

$$b := \text{slope}(x, y) \quad b = 2.93028$$

$$a := \text{intercept}(x, y) \quad a = 0.256947$$

$$r := \text{corr}(x, y) \quad r = 0.991089$$

แบบที่ 3. การคำนวณค่า b และ r

$$S_{xx} = \sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n x_i^2 - \frac{\left(\sum_{i=1}^n x_i\right)^2}{n} \qquad S_{yy} = \sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n y_i^2 - \frac{\left(\sum_{i=1}^n y_i\right)^2}{n}$$

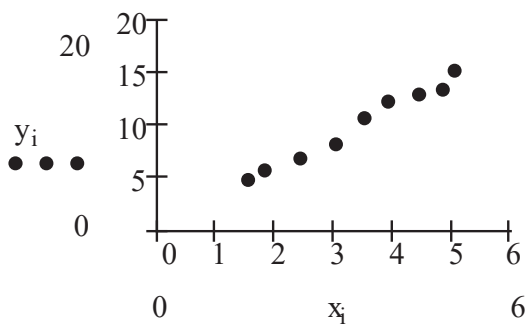
$$S_{xy} = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^n x_i y_i - \frac{\left(\sum_{i=1}^n x_i\right)\left(\sum_{i=1}^n y_i\right)}{n} \qquad b = \frac{S_{xy}}{S_{xx}} \quad \text{และ} \quad r = \frac{S_{xy}}{\sqrt{S_{xx}}\sqrt{S_{yy}}}$$

$$S_{xx} := \sum_{i=1}^n (x_i)^2 - \frac{\left(\sum_{i=1}^n x_i\right)^2}{n} \qquad S_{xx} = 13.1 \qquad S_{yy} := \sum_{i=1}^n (y_i)^2 - \frac{\left(\sum_{i=1}^n y_i\right)^2}{n} \qquad S_{yy} = 114.515556$$

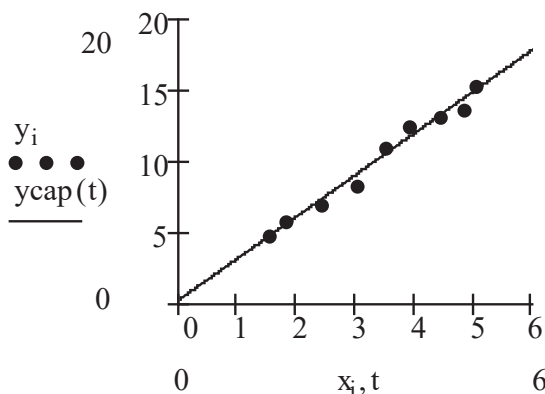
$$S_{xy} := \sum_{i=1}^n x_i \cdot y_i - \frac{\left(\sum_{i=1}^n x_i\right) \cdot \left(\sum_{i=1}^n y_i\right)}{n} \qquad S_{xy} = 38.386667$$

$$b := \frac{S_{xy}}{S_{xx}} \qquad b = 2.93028 \qquad r := \frac{S_{xy}}{\sqrt{S_{xx}}\sqrt{S_{yy}}} \qquad r = 0.991089$$

แผนภาพการกระจายของข้อมูลเป็นดังนี้



แผนภาพการกระจายของข้อมูลและสมการเส้นถดถอย $\hat{y} = a + bx$

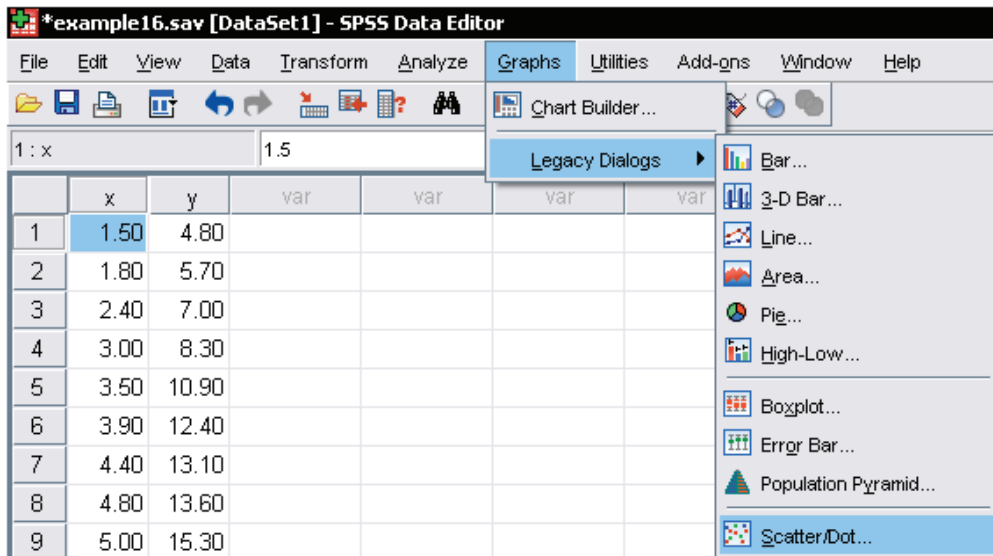


การเขียนกราฟของแผนภาพการกระจายด้วยโปรแกรม SPSS for Windows

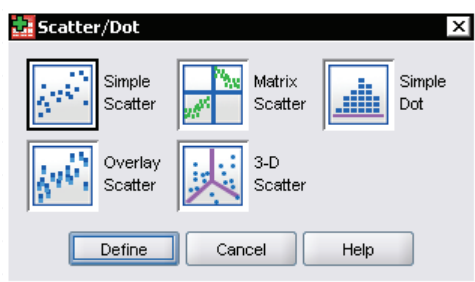
ขั้นที่ 1. สร้างแฟ้มข้อมูลประกอบด้วย 2 ตัวแปร x และ y ใน SPSS Data Editor

เสร็จแล้ว Save ไว้ที่ชื่อ example16.sav

ขั้นที่ 2. เลือกคำสั่ง Graphs \ Legacy Dialogs \ Scatter \ Dot



จะได้เมนูย่อยเป็นดังนี้

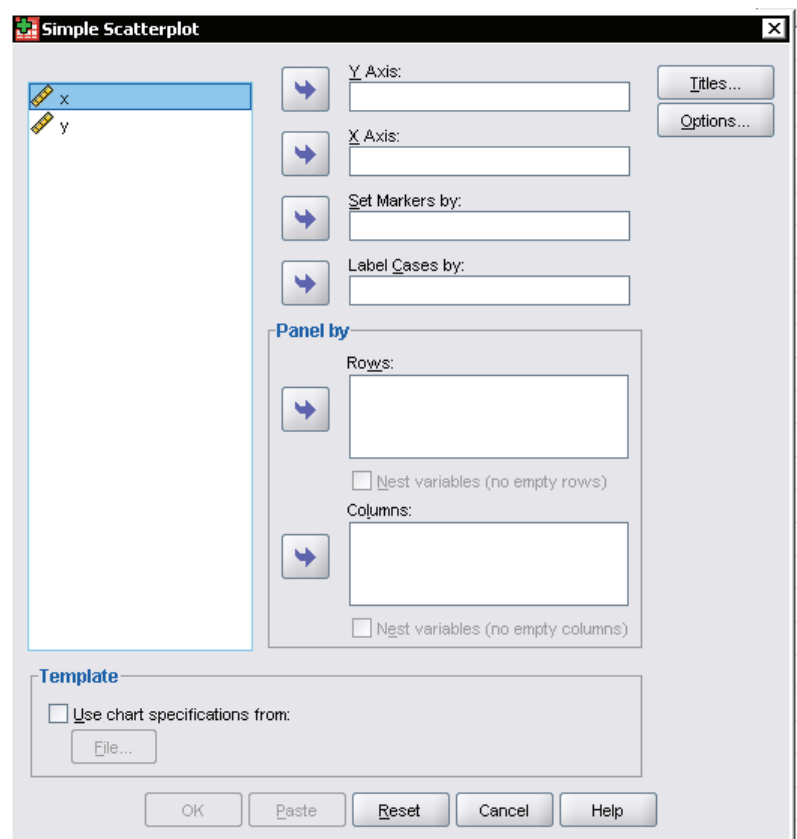


ขั้นที่ 3.

เลือกรูปแบบกราฟเป็น Simple Scatter

เสร็จแล้วคลิกปุ่ม Define

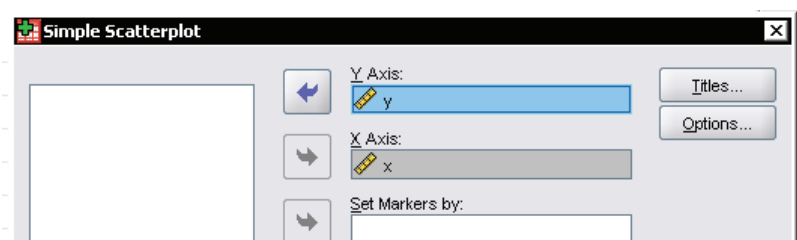
จะได้เมนูย่อยเป็นดังนี้



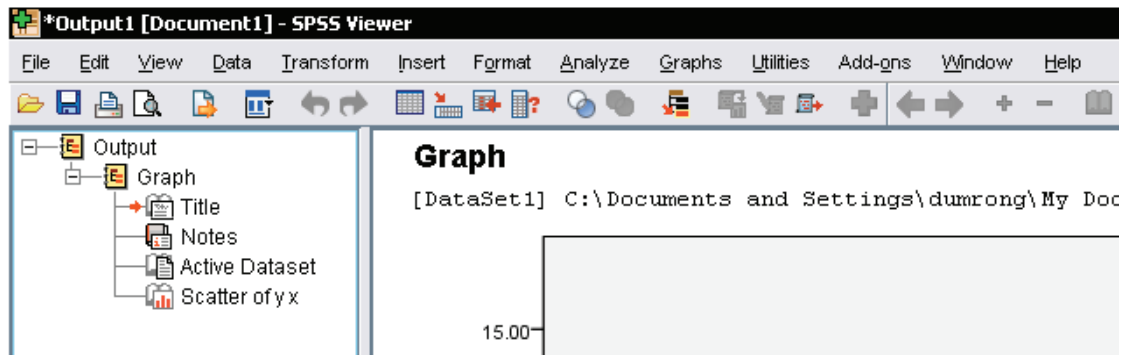
ขั้นที่ 4.

เลือกตัวแปร x ไว้ที่ X Axis

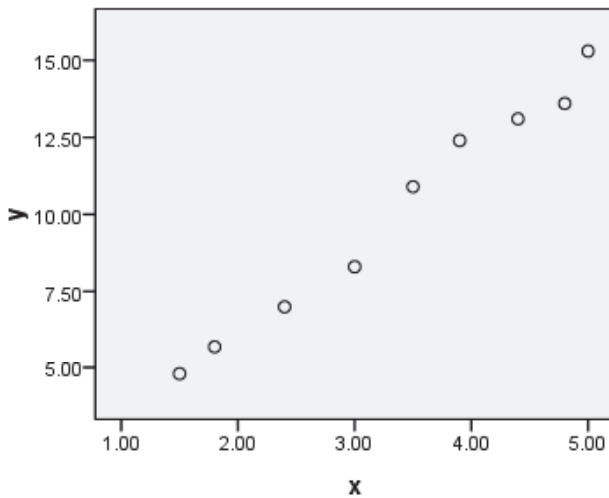
เลือกตัวแปร y ไว้ที่ Y Axis



ขั้นที่ 5. คลิก OK จะได้กราฟของแผนภาพการกระจายที่ SPSS Viewer



กราฟของแผนภาพการกระจายที่ได้คือ

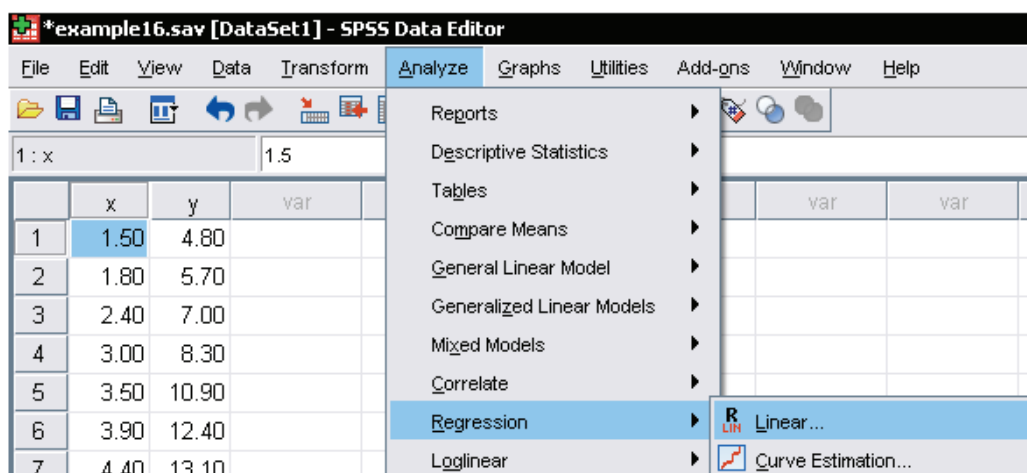


การคำนวณหาสมการถดถอยและสหสัมพันธ์ด้วย SPSS for Windows

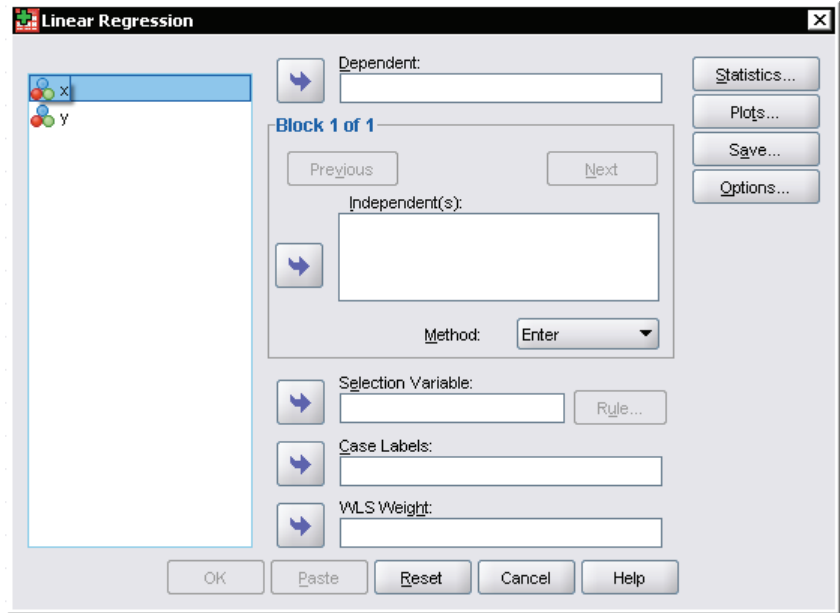
ขั้นที่ 1. สร้างแฟ้มข้อมูลประกอบด้วย 2 ตัวแปร x และ y ใน SPSS Data Editor เสร็จแล้ว Save ไว้ที่ชื่อ example16.sav

	x	y	
1	1.50	4.80	
2	1.80	5.70	
3	2.40	7.00	
4	3.00	8.30	
5	3.50	10.90	
6	3.90	12.40	
7	4.40	13.10	
8	4.80	13.60	
9	5.00	15.30	

ขั้นที่ 2. เลือกคำสั่ง Analyze / Regression / Linear

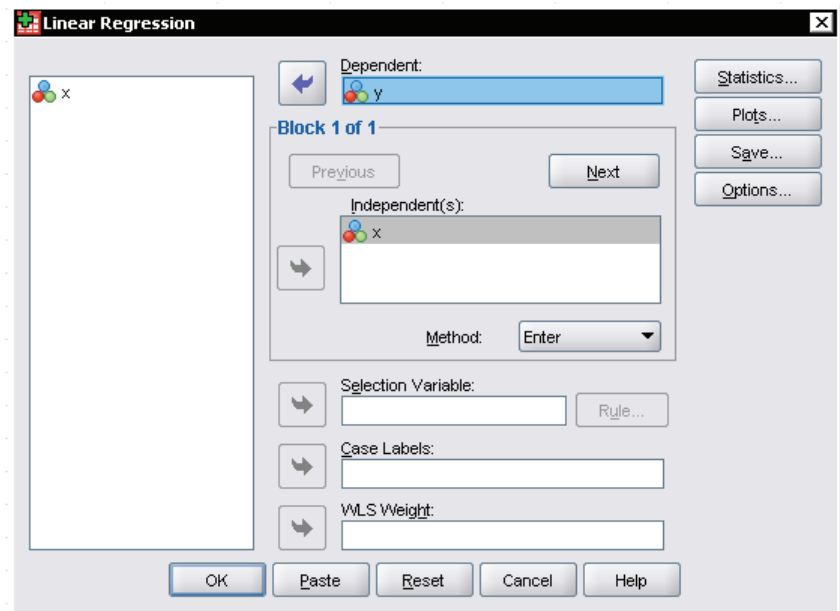


ขั้นที่ 3. คลิกที่ Linear
จะได้เมนูของคำสั่งดังนี้



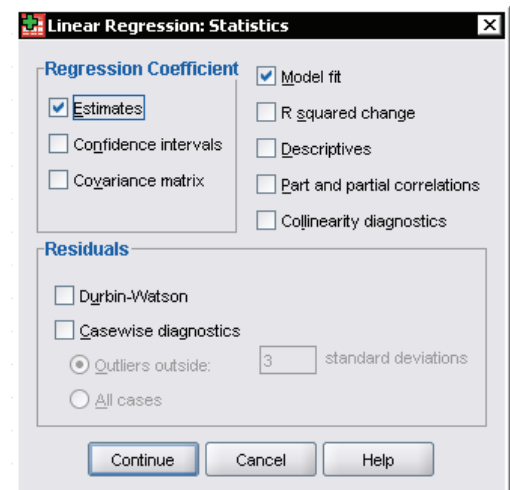
ขั้นที่ 4.

เลือกตัวแปร x เป็นตัวแปรอิสระ
นำไปไว้ที่ช่อง Independent(s)
เลือกตัวแปร y เป็นตัวแปรตาม
นำไปไว้ที่ช่อง Dependent

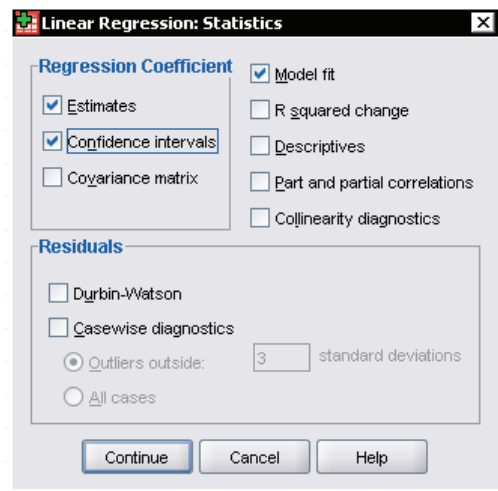


หมายเหตุ หากต้องการเฉพาะค่า a, b และ r ให้คลิก OK จะได้ผลการวิเคราะห์ทันที
แต่ถ้าต้องการให้มีการเขียนกราฟให้คลิก Plots
หรือ ต้องการหาช่วงความเชื่อมั่น
ของค่าพารามิเตอร์ α และ β
ให้คลิกที่ปุ่ม Statistics

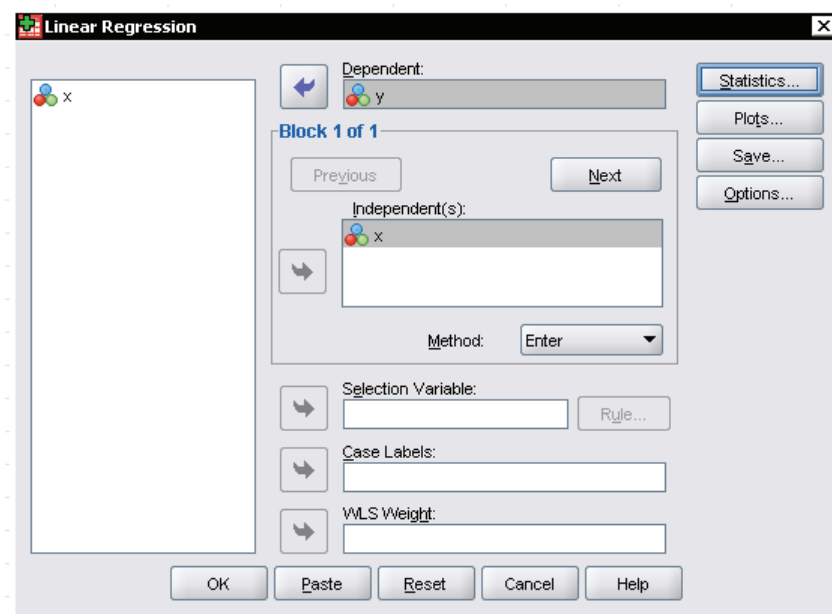
ขั้นที่ 5. คลิกปุ่ม Statistics
จะได้เมนูย่อยเป็นดังนี้



ขั้นที่ 5. เลือก Confidence Intervals เพื่อหาช่วงความเชื่อมั่นของค่าพารามิเตอร์ α และ β



ขั้นที่ 6. คลิก Continue เพื่อกลับไปเมนู Linear Regression



ขั้นที่ 7. คลิก OK จะได้ผลการคำนวณเป็นดังนี้

Regression

[DataSet1] C:\Documents and Settings\d\

Variables Entered/Removed^a

Model	Variables Entered	Variables Removed	Method
1	x ^a	.	Enter

a. All requested variables entered.
b. Dependent Variable: y

ผลการคำนวณทั้งหมดคือ

Regression

[DataSet1] C:\Documents and Settings\dumrong\My Documents\dataspss16\example16.sav

Variables Entered/Removed^b

Mode	Variables Entered	Variables Removed	Method
1	x ^a	.	Enter

a. All requested variables entered.

b. Dependent Variable: y

Model Summary

Mode	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	0.991089 ^a	0.982257	0.979722	0.538766

a. Predictors: (Constant), x

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	112.4837	1	112.483678	387.516277	0.00000022 ^a
	Residual	2.0319	7	0.290268		
	Total	114.5156	8			

a. Predictors: (Constant), x

b. Dependent Variable: y

Coefficients^a

		Model	
		1	
		(Constant)	x
Unstandardized Coefficients	B	0.256947	2.930280
	Std. Error	0.532353	0.148855
Standardized Coefficients	Beta		0.991089
t		0.482662	19.685433
Sig.		0.644064	0.000000
95% Confidence Interval for B	Lower Bound	-1.001867	2.578293
	Upper Bound	1.515761	3.282267

a. Dependent Variable: y

การวิเคราะห์ที่เกี่ยวกับสัมประสิทธิ์การถดถอยให้ดูในช่องตัวแปร x ของตาราง Coefficients

$b = 2.930280$ และ ช่วงความเชื่อมั่น 95 % ของค่า β คือ $2.578293 < \beta < 3.282267$

การวิเคราะห์ที่เกี่ยวกับสัมประสิทธิ์ระยะตัดแกนให้ดูในช่อง Constant ของตาราง Coefficients

$a = 0.256947$ และ ช่วงความเชื่อมั่น 95 % ของค่า α คือ $-1.0011867 < \alpha < 1.515761$

หมายเหตุ ค่าสหสัมพันธ์ r ให้ดูที่ตาราง Model Summary จะได้ค่าสหสัมพันธ์เฉพาะขนาดของตัวเลข

เพราะว่าเครื่องหมายของ r และ b เหมือนกัน เพราะฉะนั้นเครื่องหมายของ r ให้ดูจากเครื่องหมายของ b

เพราะฉะนั้นสหสัมพันธ์ $r = 0.991089$

ที่มาของค่าสถิติในตาราง Coefficients ค่าสถิติในช่องของตัวแปร X จากข้อมูล

$$\begin{array}{l} \text{ORIGIN} := 1 \\ x := \begin{pmatrix} 1.5 \\ 1.8 \\ 2.4 \\ 3.0 \\ 3.5 \\ 3.9 \\ 4.4 \\ 4.8 \\ 5.0 \end{pmatrix} \quad y := \begin{pmatrix} 4.8 \\ 5.7 \\ 7.0 \\ 8.3 \\ 10.9 \\ 12.4 \\ 13.1 \\ 13.6 \\ 15.3 \end{pmatrix} \end{array} \quad \begin{array}{l} n := \text{length}(x) \quad n = 9 \\ \text{mean}(x) = 3.36666667 \\ \text{mean}(y) = 10.12222222 \end{array}$$

Unstandardized Coefficients B คือค่าสัมประสิทธิ์ของการถดถอยเชิงเส้น b ที่คำนวณจากสูตร

$$b := \frac{n \cdot \sum_{i=1}^n x_i \cdot y_i - \sum_{i=1}^n x_i \cdot \sum_{i=1}^n y_i}{n \cdot \sum_{i=1}^n (x_i)^2 - \left(\sum_{i=1}^n x_i \right)^2} \quad b = 2.9302799 \quad a := \text{mean}(y) - b \cdot \text{mean}(x) \quad a = 0.25694656$$

Unstandardized Coefficients Std. Error คือส่วนเบี่ยงเบนมาตรฐานของค่าสถิติ b

ที่คำนวณจากสูตร $\sigma_b = \frac{S}{\sqrt{S_{xx}}}$ โดยมีขั้นตอนการคำนวณที่สำคัญดังนี้

$$\begin{array}{l} S_{xx} := \sum_{i=1}^n (x_i)^2 - \frac{\left(\sum_{i=1}^n x_i \right)^2}{n} \quad S_{xx} = 13.1 \quad S_{yy} := \sum_{i=1}^n (y_i)^2 - \frac{\left(\sum_{i=1}^n y_i \right)^2}{n} \quad S_{yy} = 114.5156 \\ S_{xy} := \sum_{i=1}^n x_i \cdot y_i - \frac{\left(\sum_{i=1}^n x_i \right) \cdot \left(\sum_{i=1}^n y_i \right)}{n} \quad S_{xy} = 38.3867 \\ SSE := \sum_{i=1}^n (y_i - a - b \cdot x_i)^2 \quad SSE = 2.03187786 \quad SSE := S_{yy} - b \cdot S_{xy} \quad SSE = 2.03187786 \end{array}$$

$$S_{\text{square}} := \frac{SSE}{n - 2} \quad S_{\text{square}} = 0.29026827 \quad S := \sqrt{\frac{SSE}{n - 2}} \quad S = 0.5387655$$

$$\text{Sigma} := \frac{S}{\sqrt{S_{xx}}} \quad \text{Sigma} = 0.14885524$$

สรุป $\sigma_b = 0.14885524$

Standardized Coefficients Beta ในกรณีของความสัมพันธ์เชิงเส้นตรงค่าของ Standardized Coefficients Beta (X) มีค่าเท่ากับค่าสัมประสิทธิ์สหสัมพันธ์

ค่า t ได้มาจากสูตร $t = \frac{b}{\frac{S}{\sqrt{S_{XX}}}}$, $t = 19.68543311$

ค่า Sig คือ 2 เท่าของพื้นที่ใต้โค้งทางทางด้านขวาของเส้นโค้ง t เมื่อ $df = n - 2 = 7$

$$v := 7 \quad h(t) := \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right) \cdot \sqrt{\pi \cdot v}} \cdot \left(1 + \frac{t^2}{v}\right)^{-\frac{v+1}{2}} \quad \text{Sig} := 0.5 - \int_0^{19.68543311} h(t) dt \quad \text{Sig} = 0.000000109$$

หรือคำนวณโดยใช้ฟังก์ชันสำเร็จรูป pt ของ Mathcad

Significant := 1 - pt(19.68543311, 7) Significant = 0.000000109

Significant_2_tailed := 2 · Significant Significant_2_tailed = 0.00000022

95% Confidence interval for B หมายถึงช่วงความเชื่อมั่น 95% ของค่าพารามิเตอร์ β มีสูตรเป็น

$$b - t_{\alpha/2} \frac{S}{\sqrt{S_{XX}}} < \beta < b + t_{\alpha/2} \frac{S}{\sqrt{S_{XX}}}$$

Std_Error_of_b := $\frac{S}{\sqrt{S_{XX}}}$ Std_Error_of_b = 0.14885524

alpha := 0.05 t_alpha_divide2 := qt $\left(1 - \frac{\text{alpha}}{2}, 7\right)$ t_alpha_divide2 = 2.36462425

Lower := b - t_alpha_divide2 · Std_Error_of_b Lower = 2.57829318

Upper := b + t_alpha_divide2 · Std_Error_of_b Upper = 3.28226661

ค่าสถิติในช่องของ Constant

Unstandardized Coefficients B คือค่าระยะตัดแกน Y จากสมการ $\hat{y} = a + bx$

a := mean(y) - b · mean(x) a = 0.25694656

Unstandardized Coefficients Std. Error คือส่วนเบี่ยงเบนมาตรฐานของค่าสถิติ a

ที่คำนวณจากสูตร $\sigma_a = s \sqrt{\frac{\sum_{i=1}^n x_i^2}{n S_{XX}}}$

Std_Error_of_a := $S \cdot \sqrt{\frac{\sum_{i=1}^n (x_i)^2}{n \cdot S_{XX}}}$ Std_Error_of_a = 0.53235263

ค่า t ได้มาจากสูตร $t = \frac{a}{S \sqrt{\frac{\sum_{i=1}^n x_i^2}{n S_{XX}}}}$, $t = 0.48266234$

ค่า Sig คือ 2 เท่าของพื้นที่ใต้โค้งทางทางด้านขวาของเส้นโค้งที่ ที่ระยะ $t = 0.48266234$ (จากที่คำนวณได้)
เมื่อ $df = n - 2 = 7$

$$v := 7 \quad h(t) := \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right) \cdot \sqrt{\pi \cdot v}} \cdot \left(1 + \frac{t^2}{v}\right)^{-\frac{v+1}{2}} \quad \text{Sig} := 0.5 - \int_0^{0.48266234} h(t) dt \quad \text{Sig} = 0.3220318163$$

หรือคำนวณโดยใช้ฟังก์ชันสำเร็จรูป pt ของ Mathcad

$$\text{Significant} := 1 - \text{pt}(0.48266234, 7)$$

$$\text{Significant} = 0.3220318163$$

$$\text{Significant}_2\text{-tailed} := 2 \cdot \text{Significant}$$

$$\text{Significant}_2\text{-tailed} = 0.64406363$$

95% Confidence interval for B หมายถึงช่วงความเชื่อมั่น 95% ของค่าพารามิเตอร์ α มีสูตรเป็น

$$a - t_{\frac{\alpha}{2}} s \sqrt{\frac{\sum_{i=1}^n x_i^2}{n S_{XX}}} < \alpha < a + t_{\frac{\alpha}{2}} s \sqrt{\frac{\sum_{i=1}^n x_i^2}{n S_{XX}}}$$

$$\text{Std_Error_of_a} := S \cdot \sqrt{\frac{\sum_{i=1}^n (x_i)^2}{n \cdot S_{XX}}}$$

$$\text{Std_Error_of_a} = 0.53235263$$

$$\alpha := 0.05 \quad t_{\alpha\text{-divide}2} := \text{qt}\left(1 - \frac{\alpha}{2}, 7\right) \quad t_{\alpha\text{-divide}2} = 2.36462425$$

$$\text{Lower} := a - t_{\alpha\text{-divide}2} \cdot \text{Std_Error_of_a}$$

$$\text{Lower} = -1.00186737$$

$$\text{Upper} := a + t_{\alpha\text{-divide}2} \cdot \text{Std_Error_of_a}$$

$$\text{Upper} = 1.5157605$$

ที่มาของค่าสถิติในตาราง **Model Summary**

R = ค่าสัมประสิทธิ์สหสัมพันธ์เป็นตัวเลขที่บอกระดับและทิศทางของความสัมพันธ์ระหว่างตัวแปร

หมายเหตุ เครื่องหมายของ r และ b ต้องเหมือนกัน

$$\text{สูตรของค่า } R \text{ คือ } r = \frac{n \sum_{i=1}^n x_i y_i - \left(\sum_{i=1}^n x_i\right) \left(\sum_{i=1}^n y_i\right)}{\sqrt{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i\right)^2} \sqrt{n \sum_{i=1}^n y_i^2 - \left(\sum_{i=1}^n y_i\right)^2}} = \frac{S_{xy}}{\sqrt{S_{xx}} \sqrt{S_{yy}}}, r = 0.9910887$$

หรือใช้ฟังก์ชัน $\text{corr}(x, y)$ ของ Mathcad จะได้ว่า $R = \text{corr}(x, y)$ เพราะฉะนั้น $R = 0.9910887$

R Square เป็นค่าสัมประสิทธิ์การตัดสินใจ ได้มาจากค่า R^2 เป็นตัวเลขที่ใช้ในการอธิบายว่า สมการเส้นถดถอย $\hat{y} = a + bx$ มีความเหมาะสมที่จะนำไปใช้ในการอธิบายความสัมพันธ์ได้ดีหรือไม่ กล่าวคือ

R^2 มีค่าเข้าใกล้ 1 แสดงว่าสมการเส้นถดถอย $\hat{y} = a + bx$ มีความเหมาะสมดีมาก

R^2 มีค่าเข้าใกล้ 0 แสดงว่าสมการเส้นถดถอย $\hat{y} = a + bx$ ไม่มีความเหมาะสม

ตัวอย่างการแปลความหมาย

$R^2 = 0.1$ สมการเส้นถดถอย $\hat{y} = a + bx$ ใช้อธิบายการเปลี่ยนแปลงของค่า y ได้ 10%

$R^2 = 0.98226$ สมการเส้นถดถอย $\hat{y} = a + bx$ ใช้อธิบายการเปลี่ยนแปลงของค่า y ได้ 98.226%

Adjusted R Squares เป็นค่าที่ใช้ในการปรับปรุงค่าของ R Squares ในกรณีที่ค่าของ n มีน้อยๆ

$$\text{สูตรของ Adjust R Square} = 1 - \frac{n-1}{n-2} \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

ขั้นตอนการคำนวณของ Mathcad

$$ycap(x) := a + b \cdot x$$

$$\text{Adjust_R_Square} := 1 - \frac{n-1}{n-2} \cdot \frac{\left[\sum_{i=1}^n (y_i - ycap(x_i))^2 \right]}{\left[\sum_{i=1}^n (y_i - \text{mean}(y))^2 \right]} \quad \text{Adjust_R_Square} = 0.979722$$

ที่มาของค่าสถิติในตาราง ANOVA

ANOVA TABLE

SOV	SS	DF	MS	F
REGRESSION	SSR	1	MSR = SSR	$\frac{MSR}{MSE}$
ERROR	SSE	$n - 2$	$MSE = \frac{SSE}{n - 2}$	
TOTAL	SST	$n - 1$		

$$SST := S_{yy} \quad SST = 114.515556 \quad SSR := b \cdot S_{xy} \quad SSR = 112.48367769$$

$$SSE := SST - SSR \quad SSE = 2.03187786 \quad MSR := SSR \quad MSR = 112.48367769$$

$$MSE := \frac{SSE}{n - 2} \quad MSE = 0.29026827 \quad F := \frac{MSR}{MSE} \quad F = 387.51627662$$

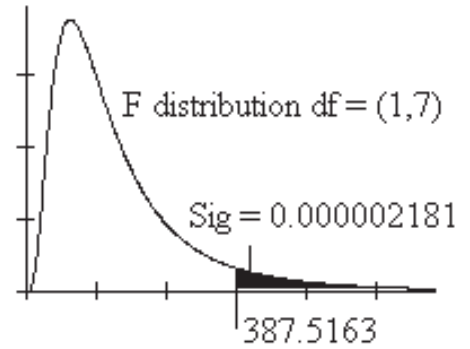
ค่า Sig เป็นค่าที่คำนวณมาจากพื้นที่ใต้โค้งทางทางด้านขวาของเส้นโค้งเอฟ

ที่มีระดับชั้นความเสรี $v_1 = 1$ และ $v_2 = 7$ ที่ระยะ $F = 387.5163$ จากค่าในตารางที่คำนวณได้

การคำนวณค่า Sig ของค่าสถิติ F ด้วย Mathcad

$v1 := 1 \quad v2 := 7 \quad TOL := 0.0000001$

$$h(f) := \frac{\Gamma\left(\frac{v1 + v2}{2}\right) \cdot \left(\frac{v1}{v2}\right)^{\frac{v1}{2}} \cdot f^{\left(\frac{v1}{2}\right) - 1}}{\Gamma\left(\frac{v1}{2}\right) \cdot \Gamma\left(\frac{v2}{2}\right) \cdot \left[1 + \left(\frac{v1}{v2}\right) \cdot f\right]^{\frac{v1+v2}{2}}}$$



Significant := $1 - \int_0^{387.5163} h(f) df$ Significant = 0.0000002181

Sig := $1 - pF(387.5163, 1, 7)$ Sig = 0.0000002181

หมายเหตุ ค่า F และ Sig ในตาราง ANOVA ใช้ในการทดสอบสมมติฐาน $H_0 : \beta = 0$ แยกกับ $H_1 : \beta \neq 0$ หรือสมมติฐาน $H_0 : \rho = 0$ แยกกับ $H_1 : \rho \neq 0$ โดยมีเกณฑ์ในการสรุปผลคือ ถ้า Sig จากตาราง ANOVA มีค่าน้อยกว่า α แล้วปฏิเสธ H_0 โดยมีนัยสำคัญการทดสอบ α

8.2 การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่าสัมประสิทธิ์การถดถอย β และระยะตัดแกน α

เราสามารถประมาณค่าของ β และ α โดยใช้ช่วงความเชื่อมั่นที่มีสูตรดังนี้

ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของสัมประสิทธิ์การถดถอย β คือ

$$b - t_{\frac{\alpha}{2}} \frac{S}{\sqrt{S_{XX}}} < \beta < b + t_{\frac{\alpha}{2}} \frac{S}{\sqrt{S_{XX}}} \quad (df = n - 2)$$

ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของระยะตัดแกน α คือ

$$a - t_{\frac{\alpha}{2}} s \sqrt{\frac{\sum_{i=1}^n x_i^2}{n S_{XX}}} < \alpha < a + t_{\frac{\alpha}{2}} s \sqrt{\frac{\sum_{i=1}^n x_i^2}{n S_{XX}}} \quad (df = n - 2)$$

ตัวอย่าง 8.2.1 ในการศึกษาความสัมพันธ์ระหว่างปริมาณอากาศเป็นพิษที่ถูกกำจัดออกไป(ไมโครกรัมต่อลูกบาศก์เมตร) กับปริมาณน้ำฝน (หน่วย 0.01 นิ้ว) ได้ข้อมูลดังนี้

ปริมาณน้ำฝน	ปริมาณอากาศเป็นพิษที่ถูกกำจัดออกไป
4.30	126.00
4.50	121.00
5.90	116.00
5.60	118.00
6.10	114.00
5.20	118.00
3.80	132.00
2.10	141.00
7.50	108.00

จงหา 1. สัมประสิทธิ์การถดถอยเชิงเส้น b และ a และ สมการของเส้นถดถอยเชิงเส้น $\hat{y} = a + bx$

2. สัมประสิทธิ์สหสัมพันธ์เชิงเส้น r

3. ช่วงความเชื่อมั่น 95% ของค่า β และ ช่วงความเชื่อมั่น 95% ของค่า α

วิธีทำ ขั้นที่ 1. สร้างแฟ้มข้อมูลประกอบด้วย 2 ตัวแปร

ตัวแปร rain แทนปริมาณน้ำฝน

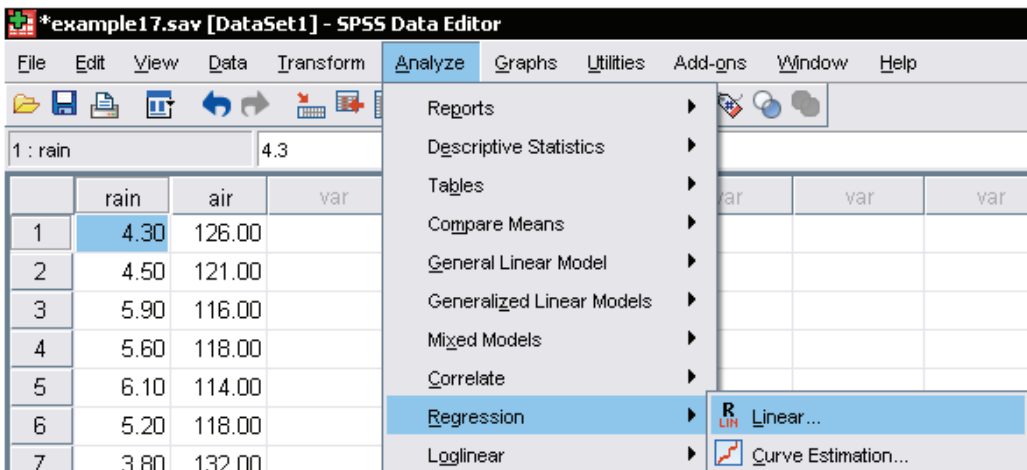
ตัวแปร air แทนปริมาณอากาศเป็นพิษที่ถูกกำจัดออกไป

เสร็จแล้ว Save ไว้ที่ชื่อ example17.sav

ขั้นที่ 2. เลือกคำสั่ง Analyze / Regression / Linear

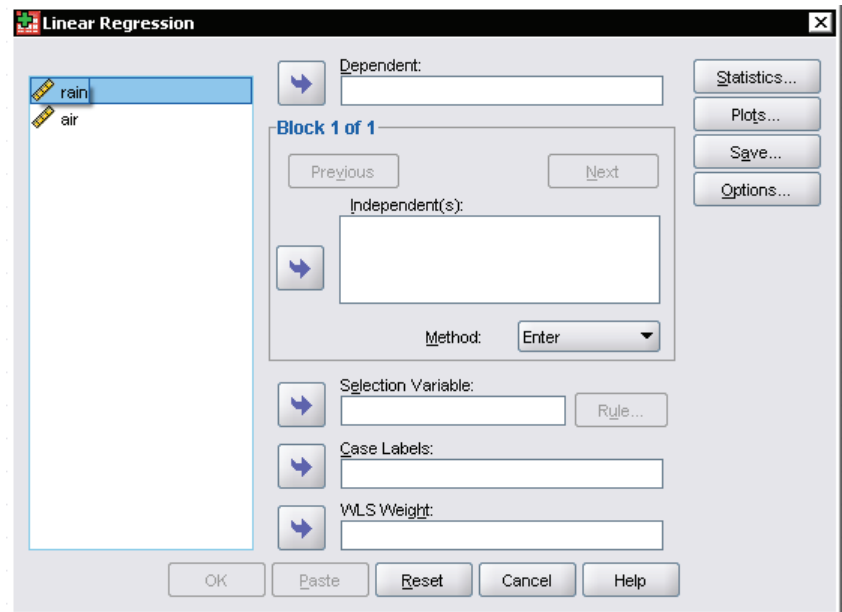
*example17.sav [DataSet1]

	rain	air
1	4.30	126.00
2	4.50	121.00
3	5.90	116.00
4	5.60	118.00
5	6.10	114.00
6	5.20	118.00
7	3.80	132.00
8	2.10	141.00
9	7.50	108.00



ขั้นที่ 3.

คลิกที่ Linear จะได้เมนู

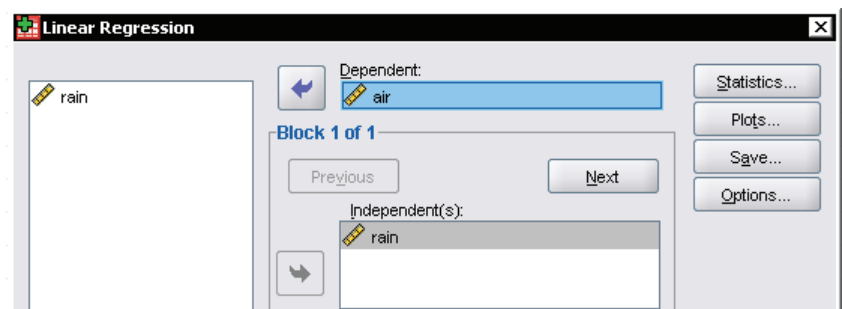


ขั้นที่ 4. เลือกตัวแปร air

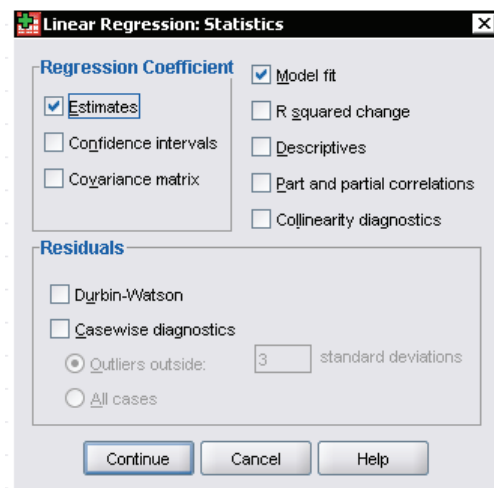
ไว้ที่ช่อง Dependent

เลือกตัวแปร rain

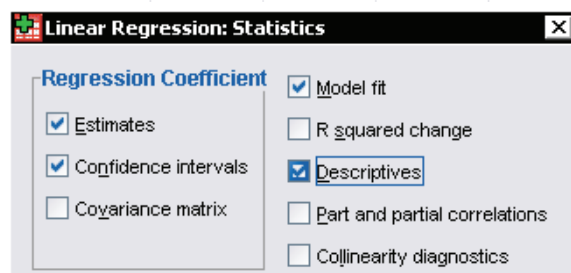
ไว้ที่ช่อง Independent[s]



ขั้นที่ 5. เพราะว่าเราต้องการช่วงความเชื่อมั่นของ β และ α เพราะฉะนั้นเราต้องเลือก Statistics เมื่อคลิกที่ปุ่ม Statistics จะได้เมนูย่อยดังนี้



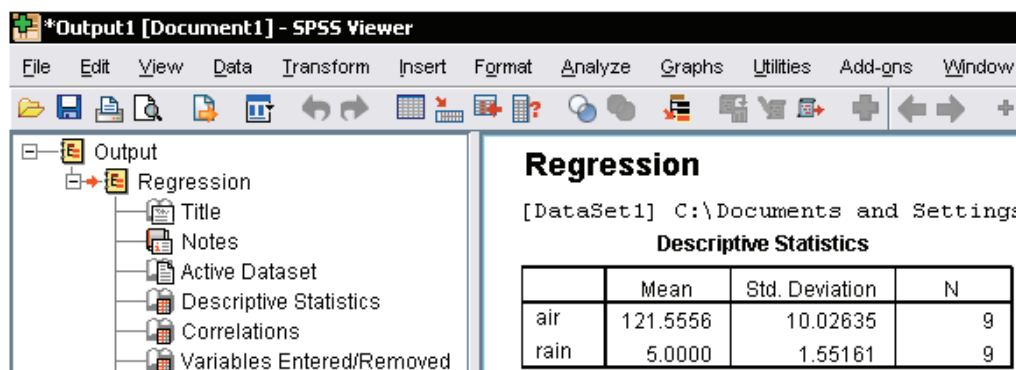
ขั้นที่ 6. คลิกในช่องสี่เหลี่ยมที่หน้า Confidence intervals
หมายเหตุ ถ้าต้องการค่าสถิติเบื้องต้นคลิกในช่องสี่เหลี่ยมที่หน้า Descriptives



ขั้นที่ 7. คลิก Continue จะกลับไปเมนูของคำสั่ง Analyze / Regression / Linear

ขั้นที่ 8. คลิก OK จะได้ผลการคำนวณดังนี้

ผลการคำนวณทั้งหมดคือ



Regression

[DataSet1] C:\Documents and Settings\dumrong\My Documents\datspss16\example17.sav

Descriptive Statistics

	Mean	Std. Deviation	N
air	121.5556	10.02635	9
rain	5.0000	1.55161	9

Correlations

		air	rain
Pearson Correlation	air	1.0000000000	-0.9786583643
	rain	-0.9786583643	1.0000000000
Sig. (1-tailed)	air	.	0.0000022896
	rain	0.0000022896	.
N	air	9	9
	rain	9	9

Variables Entered/Removed^b

Mode	Variables Entered	Variables Removed	Method
1	rain ^a	.	Enter

a. All requested variables entered.

b. Dependent Variable: air

Model Summary

Mode	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	0.97865836 ^a	0.95777219	0.95173965	2.20261338

a. Predictors: (Constant), rain

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	770.262	1	770.262	158.768	0.0000045792 ^a
	Residual	33.961	7	4.852		
	Total	804.222	8			

a. Predictors: (Constant), rain

b. Dependent Variable: air

หมายเหตุ จากตาราง ANOVA

เพราะว่า Sig = 0.0000045792 < 0.05 เพราะฉะนั้นสรุปได้ว่า $\rho \neq 0$ ที่ระดับนัยสำคัญ 0.05

เพราะฉะนั้นปริมาณอากาศเป็นพิษที่ถูกกำจัดออกไปกับปริมาณน้ำฝนมีความสัมพันธ์กันในรูปแบบเชิงเส้น

Coefficients^a

		Model	
		1	
		(Constant)	rain
Unstandardized Coefficients	B	153.175	-6.324
	Std. Error	2.615	.502
Standardized Coefficients	Beta		-.979
	t	58.583	-12.600
Sig.		0.0000000001	0.0000045792
95% Confidence Interval for B	Lower Bound	146.993	-7.511
	Upper Bound	159.358	-5.137

a. Dependent Variable: air

จากตาราง Coefficients จะได้ผลการวิเคราะห์ข้อมูลคือ

1. สัมประสิทธิ์การถดถอยเชิงเส้น $b = -6.324$ ระยะตัดแกน $a = 153.175$

สมการของเส้นถดถอยเชิงเส้น $\hat{y} = a + bx$ คือ $\hat{y} = 153.175 - 6.324x$

2. จากตาราง Model Summary สัมประสิทธิ์สหสัมพันธ์ $r = -0.97865836$

หมายเหตุ เครื่องหมายของ r ให้ดูจากเครื่องหมายของ b

3. ช่วงความเชื่อมั่น 95% ของค่าสัมประสิทธิ์การถดถอย β คือ $-7.511 < \beta < -5.137$

ช่วงความเชื่อมั่น 95% ของค่าระยะตัดแกน α คือ $146.993 < \alpha < 159.358$

8.3 การทดสอบสมมติฐาน $H_0 : \rho = 0$

การทดสอบสมมติฐานเกี่ยวกับความสัมพันธ์ของประชากร 2 ชุดว่ามีความสัมพันธ์กันในรูปแบบเชิงเส้นหรือไม่มีขั้นตอนในการทำงานดังนี้

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \rho = 0$

กำหนดสมมติฐานอื่น $H_1 : \rho \neq 0$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. ทำการสุ่มตัวอย่างขนาด n และ คำนวณค่า r

ขั้นที่ 4. เลือกค่าสถิติที่เหมาะสมคือ t

ขั้นที่ 5. คำนวณค่าสถิติที่เลือกจากข้อมูลตัวอย่าง $t_{\text{คำนวณ}} = r \sqrt{\frac{n-2}{1-r^2}}$, $df = n - 2$

คำนวณค่า Sig ของค่าสถิติ $t_{\text{คำนวณ}}$ เมื่อ $\text{Sig} = P(t > |t_{\text{คำนวณ}}|)$

ขั้นที่ 6. เปิดตารางสถิติหาค่าวิกฤตคือ $-t_{\frac{\alpha}{2}}$ และ $t_{\frac{\alpha}{2}}$ บริเวณวิกฤตคือ $t < -t_{\frac{\alpha}{2}}$ หรือ $t > t_{\frac{\alpha}{2}}$

ขั้นที่ 7. สรุปผล แบบที่ 1 โดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต

ถ้า $t_{\text{คำนวณ}} < -t_{\frac{\alpha}{2}}$ หรือ $t_{\text{คำนวณ}} > t_{\frac{\alpha}{2}}$ แล้วปฏิเสธ H_0

แบบที่ 2 โดยการเปรียบเทียบค่า Sig ของค่าสถิติ $t_{\text{คำนวณ}}$ กับค่านัยสำคัญ α

ถ้า $\text{Sig} < \frac{\alpha}{2}$ แล้วปฏิเสธ H_0

หมายเหตุ $t = r \sqrt{\frac{n-2}{1-r^2}}$ และ $t = \frac{b}{\frac{S}{\sqrt{S_{xx}}}}$ เป็นค่าเดียวกัน

กรณีของการทดสอบแบบ 1 ข้าง

	$H_1 : \rho < 0$	$H_1 : \rho > 0$
ค่าวิกฤต	$-t_{\alpha, n-2}$	$t_{\alpha, n-2}$
บริเวณวิกฤต	$t < -t_{\alpha, n-2}$	$t > t_{\alpha, n-2}$
ปฏิเสธ H_0	ถ้า $t_{\text{คำนวณ}} < 0$ และ $\text{Sig} < \alpha$	ถ้า $t_{\text{คำนวณ}} > 0$ และ $\text{Sig} < \alpha$

จากตัวอย่าง 8.2.1 ในการศึกษาความสัมพันธ์ระหว่างปริมาณอากาศเป็นพิษที่ถูกกำจัดออกไปกับปริมาณน้ำฝน จงทดสอบว่า ปริมาณอากาศเป็นพิษที่ถูกกำจัดออกไป กับ ปริมาณน้ำฝนไม่มีความสัมพันธ์กันที่ระดับนัยสำคัญ 0.05

วิธีทำ ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \rho = 0$

กำหนดสมมติฐานอื่น $H_1 : \rho \neq 0$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

ขั้นที่ 3. ทำการสุ่มตัวอย่างและนำข้อมูลเข้าสู่การคำนวณด้วย SPSS

ขั้นที่ 4. เลือกค่าสถิติที่เหมาะสมคือ t

ขั้นที่ 5. จากผลการคำนวณของ SPSS ข้างต้น

Coefficients^a

		Model	
		1	
		(Constant)	rain
Unstandardized Coefficients	B	153.175	-6.324
	Std. Error	2.615	.502
Standardized Coefficients	Beta		-.979
t		58.583	-12.600
Sig.		0.0000000001	0.0000045792
95% Confidence Interval for B	Lower Bound	146.993	-7.511
	Upper Bound	159.358	-5.137

← ①
← ②

a. Dependent Variable: air

จากตาราง Coefficient จะได้ 1. $t_{\text{คำนวณ}} = r \sqrt{\frac{n-2}{1-r^2}} = -12.600, df = 7$

และ 2. Sig = 0.0000045792

ขั้นที่ 6. ค่าวิกฤตคือ $-t_{0.025} = -2.365, t_{0.025} = 2.365$ บริเวณวิกฤตคือ $t < -2.365$ หรือ $t > 2.365$

ขั้นที่ 7. สรุปผล แบบที่ 1 เพราะว่า $t_{\text{คำนวณ}}$ อยู่ในบริเวณวิกฤต เพราะฉะนั้นปฏิเสธ H_0

แบบที่ 2 เพราะว่า Sig = 0.0000045792 < $\frac{\alpha}{2} = 0.025$ เพราะฉะนั้นปฏิเสธ H_0

เพราะฉะนั้นปริมาณอากาศเป็นพิษที่ถูกกำจัดออกไปกับปริมาณน้ำฝนมีความสัมพันธ์กัน

การทดสอบสมมติฐาน $H_0 : \rho = 0$ โดยใช้ค่าสถิติเอฟจากตาราง ANOVA

การทดสอบสมมติฐานเกี่ยวกับความสัมพันธ์ของประชากร 2 ชุดว่ามีความสัมพันธ์กันในรูปแบบเชิงเส้น หรือไม่มีขั้นตอนในการทำงานดังนี้

ขั้นที่ 1. กำหนดสมมติฐาน $H_0 : \rho = 0$

$H_1 : \rho \neq 0$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. สุ่มตัวอย่าง

ขั้นที่ 4. เลือกใช้ค่าสถิติเอฟ

ขั้นที่ 5. คำนวณค่าสถิติเอฟจากตาราง ANOVA (หมายเหตุ ผลบวกที่สำคัญในการทำตาราง ANOVA

คือ $S_{yy}, S_{xx}, S_{xy}, SST = \sum_{i=1}^n y_i^2 - \frac{\left(\sum_{i=1}^n y_i\right)^2}{n} = S_{yy}, SSR = b S_{xy} = b^2 S_{xx}, SSE = SST - SSR)$

ANOVA TABLE

SOV	SS	DF	MS	F	Sig
Regression	SSR	1	MSR	$f_{\text{คำนวณ}} = \frac{MSR}{MSE}$	$P(F > f_{\text{คำนวณ}})$
Residual	SSE	$n - 2$	MSE		
Total	SST	$n - 1$			

ขั้นที่ 6. ค่าวิกฤตคือ $f_{\alpha, (v_1, v_2)}$ บริเวณวิกฤตคือ $F > f_{\alpha, (v_1, v_2)}$ เมื่อ $v_1 = 1, v_2 = n - 2$

- ขั้นที่ 7. สรุปผล แบบที่ 1 ถ้า $f_{\text{คำนวณ}} > f_{\alpha, (v_1, v_2)}$ แล้วปฏิเสธ H_0
 แบบที่ 2 ถ้า $\text{Sig} < \alpha$ แล้วปฏิเสธ H_0

จากการคำนวณของ SPSS for Windows เราได้ตาราง ANOVA เป็น

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	770.262	1	770.262	158.768	0.0000045792 ^a
	Residual	33.961	7	4.852		
	Total	804.222	8			

a. Predictors: (Constant), rain

b. Dependent Variable: air

- ขั้นที่ 6. ค่าวิกฤตคือ $f_{0.05, (1, 7)} = 5.59$ บริเวณวิกฤตคือ $F > 5.59$

- ขั้นที่ 7. สรุปผล แบบที่ 1 เพราะว่า $f_{\text{คำนวณ}} = 158.768 > 5.59$ เพราะฉะนั้นปฏิเสธ H_0
 แบบที่ 2 เพราะว่า $\text{Sig} = 0.0000045792 < \alpha = 0.05$ เพราะฉะนั้นปฏิเสธ H_0

8.4 การทดสอบสมมติฐาน $H_0 : \beta = \beta_0$

เนื่องจากค่าสหสัมพันธ์เป็นการบอกระดับและทิศทางความสัมพันธ์ของตัวแปร แต่ไม่สามารถบอกอัตราการเปลี่ยนแปลงได้ ส่วนค่าสัมประสิทธิ์การถดถอย β สามารถบอกอัตราการเปลี่ยนแปลงของค่าตัวแปรได้ เช่น $y = 2 + 3x$ แสดงว่าถ้าค่า x เพิ่ม 1 หน่วยแล้วค่าของ y จะเพิ่มค่า 3 หน่วย ดังนั้นเราจึงศึกษาเกี่ยวกับการทดสอบสมมติฐาน $\beta = \beta_0$

หลักการและขั้นตอนการทดสอบสมมติฐานกรณี $\beta_0 = 0$

- ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \beta = 0$

กำหนดสมมติฐานอื่น $H_1 : \beta \neq 0$

- ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

- ขั้นที่ 3. ทำการสุ่มตัวอย่างขนาด n และ คำนวณค่า b, s, S_{xx}

- ขั้นที่ 4. เลือกค่าสถิติที่

- ขั้นที่ 5. คำนวณค่า $t_{\text{คำนวณ}} = \frac{b}{\frac{s}{\sqrt{S_{xx}}}}$, $df = n - 2$ และ $\text{Sig} = P(t > | t_{\text{คำนวณ}} |)$

- ขั้นที่ 6. เปิดตารางสถิติหาค่าวิกฤต $-t_{\frac{\alpha}{2}}$ และ $t_{\frac{\alpha}{2}}$ บริเวณวิกฤตคือ $t < -t_{\frac{\alpha}{2}}$ หรือ $t > t_{\frac{\alpha}{2}}$

- ขั้นที่ 7. สรุปผล

แบบที่ 1 โดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต

ถ้า $t_{\text{คำนวณ}} < -t_{\frac{\alpha}{2}}$ หรือ $t_{\text{คำนวณ}} > t_{\frac{\alpha}{2}}$ แล้วปฏิเสธ H_0

แบบที่ 2 ถ้า $\text{Sig} < \frac{\alpha}{2}$ แล้วปฏิเสธ H_0

หมายเหตุ ในกรณีที่ $\beta_0 = 0$ เราทำการทดสอบแบบเดียวกับการทดสอบ $\rho = 0$

ในทางทฤษฎีสามารถพิสูจน์ได้ว่า $t = r \sqrt{\frac{n-2}{1-r^2}}$ และ $t = \frac{b}{\frac{s}{\sqrt{S_{xx}}}}$ เป็นค่าเดียวกัน

กรณีของการทดสอบแบบ 1 ข้าง

	$H_1 : \beta < 0$	$H_1 : \beta > 0$
ค่าวิกฤต	$-t_{\alpha, n-2}$	$t_{\alpha, n-2}$
บริเวณวิกฤต	$t < -t_{\alpha, n-2}$	$t < t_{\alpha, n-2}$
ปฏิเสธ H_0	ถ้า $t_{\text{คำนวณ}} < 0$ และ $\text{Sig} < \alpha$	ถ้า $t_{\text{คำนวณ}} > 0$ และ $\text{Sig} < \alpha$

จากตัวอย่าง 8.2.1 ในการศึกษาความสัมพันธ์ระหว่างปริมาณอากาศเป็นพิษที่ถูกกำจัดออกไป (y) กับปริมาณน้ำฝน (x) จงทดสอบว่า สมการถดถอย $y = \alpha + \beta x$ มีค่าสัมประสิทธิ์การถดถอย $\beta = 0$ ที่ระดับนัยสำคัญ 0.05

วิธีทำ ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \beta = 0$
กำหนดสมมติฐานอื่น $H_1 : \beta \neq 0$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. ทำการสุ่มตัวอย่างขนาด n และ คำนวณค่า b, s, S_{xx}

ขั้นที่ 4. เลือกค่าสถิติที่

ขั้นที่ 5. คำนวณค่าสถิติที่เลือกจากข้อมูลตัวอย่าง $t_{\text{คำนวณ}} = \frac{b}{\frac{s}{\sqrt{S_{xx}}}}$, $df = n - 2$

Coefficients^a

		Model	
		1	
		(Constant)	rain
Unstandardized Coefficients	B	153.175	-6.324
	Std. Error	2.615	.502
Standardized Coefficients	Beta		-.979
t		58.583	-12.600 ← ①
Sig.		0.0000000001	0.0000045792 ← ②
95% Confidence Interval for B	Lower Bound	146.993	-7.511
	Upper Bound	159.358	-5.137

a. Dependent Variable: air

จากตาราง Coefficient จะได้ $t_{\text{คำนวณ}} = \frac{b}{\frac{s}{\sqrt{S_{xx}}}} = -12.600$, $df = 7$ และ $\text{Sig} = 0.0000045792$

ขั้นที่ 6. ค่าวิกฤตคือ $-t_{0.025} = -2.365$, $t_{0.025} = 2.365$ บริเวณวิกฤตคือ $t < -2.365$ หรือ $t > 2.365$

ขั้นที่ 7. สรุปผล แบบที่ 1 เพราะว่า $t_{\text{คำนวณ}}$ อยู่ในบริเวณวิกฤต เพราะฉะนั้นปฏิเสธ H_0

แบบที่ 2 เพราะว่า $\text{Sig} = 0.0000045792 < \frac{\alpha}{2} = 0.025$ เพราะฉะนั้นปฏิเสธ H_0

เพราะฉะนั้นสมการถดถอย $y = \alpha + \beta x$ มีค่าสัมประสิทธิ์การถดถอย $\beta \neq 0$ ที่ระดับนัยสำคัญ 0.05

หลักการและขั้นตอนการทดสอบสมมติฐาน $H_0 : \beta = \beta_0$ กรณี β_0 ไม่เท่ากับ 0

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \beta = \beta_0$

กำหนดสมมติฐานอื่น $H_1 : \beta \neq \beta_0$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. ทำการสุ่มตัวอย่างขนาด n และ คำนวณค่าโดยการนำข้อมูลเข้าสู่โปรแกรม SPSS

ขั้นที่ 4. เพราะว่าผลการคำนวณของ SPSS ไม่ให้ค่า $t = \frac{b - \beta_0}{\frac{s}{\sqrt{S_{xx}}}}$ ออกมาโดยตรง

เพราะฉะนั้นเราจึงใช้ช่วงความเชื่อมั่นของ β ช่วยในการสรุปสมมติฐาน

ขั้นที่ 5. ให้หาช่วงความเชื่อมั่น $(1 - \alpha)100\%$ ของค่า β

ขั้นที่ 6. ไม่มีการเปิดตารางสถิติเพื่อหาค่าวิกฤต

ขั้นที่ 7. สรุปผลโดยการดูว่า β_0 อยู่ในช่วงความเชื่อมั่น $(1 - \alpha)100\%$ ของค่า β ที่หาได้หรือไม่ ถ้า β_0 อยู่ในช่วงความเชื่อมั่น $(1 - \alpha)100\%$ ของค่า β ที่หาได้ แล้ว ไม่ปฏิเสธ H_0

จากตัวอย่างข้อมูล

X	1.50	1.80	2.40	3.00	3.50	3.90	4.40	4.80	5.00
Y	4.80	5.70	7.00	8.30	10.90	12.40	13.10	13.60	15.30

สมมติความสัมพันธ์ของตัวแปรในรูปแบบเชิงเส้นคือ $y = \alpha + \beta x$

งทดสอบสมมติฐาน $H_0 : \beta = 2.5$ แยกกับ $H_1 : \beta \neq 2.5$ ที่ระดับนัยสำคัญ 0.05

วิธีทำ ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \beta = 2.5$

กำหนดสมมติฐานอื่น $H_1 : \beta \neq 2.5$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

ขั้นที่ 3. ทำการสุ่มตัวอย่างขนาด n และ คำนวณค่าโดยการนำข้อมูลเข้าสู่โปรแกรม SPSS

ขั้นที่ 4. เพราะว่าผลการคำนวณของ SPSS ไม่ให้ค่า $t = \frac{b - \beta_0}{\frac{s}{\sqrt{S_{xx}}}}$ ออกมาโดยตรง

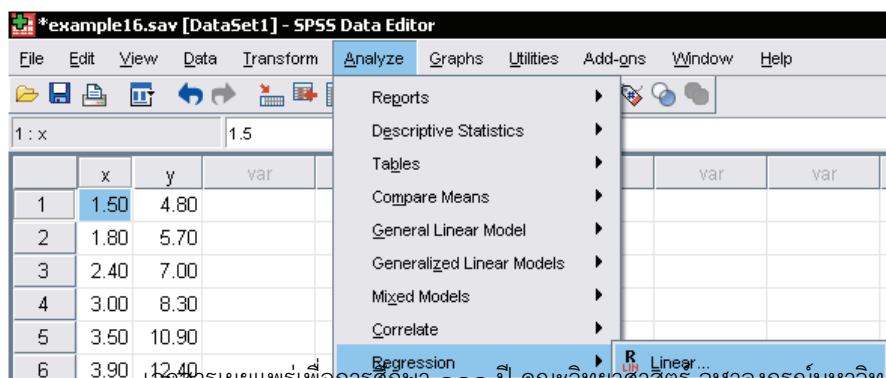
เพราะฉะนั้นเราจึงใช้ช่วงความเชื่อมั่นของ β ช่วยในการสรุปสมมติฐาน

ขั้นที่ 5. หาช่วงความเชื่อมั่น 95% ของค่า β (เพราะว่า $H_0 : \beta = 2.5$)

การหาช่วงความเชื่อมั่น 95% ของ β

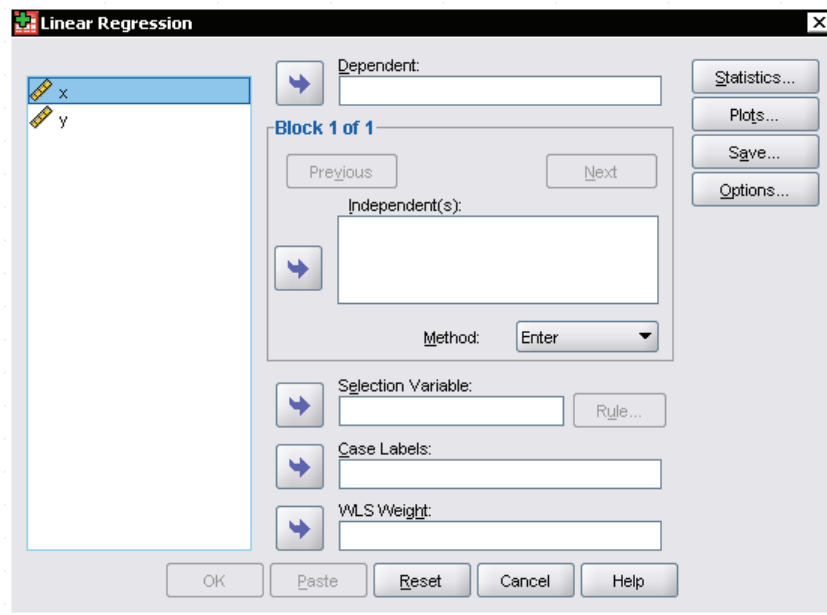
ขั้นที่ 5.1 นำข้อมูลเข้าสู่ example16.sav ที่สร้างไว้เข้าสู่ SPSS Data Editor

ขั้นที่ 5.2 เลือกคำสั่ง Analyze / Regression / Linear

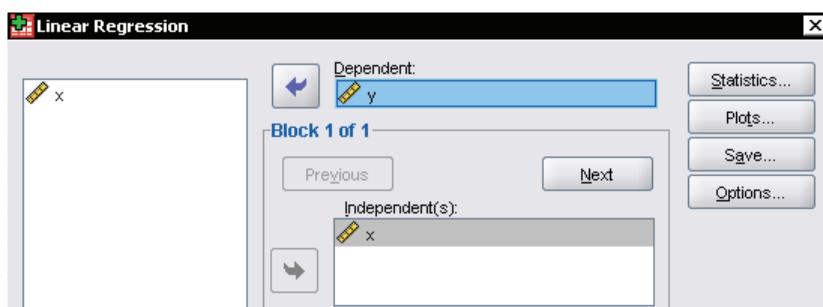


	x	y	var
1	1.50	4.80	
2	1.80	5.70	
3	2.40	7.00	
4	3.00	8.30	
5	3.50	10.90	
6	3.90	12.40	
7	4.40	13.10	
8	4.80	13.60	
9	5.00	15.30	

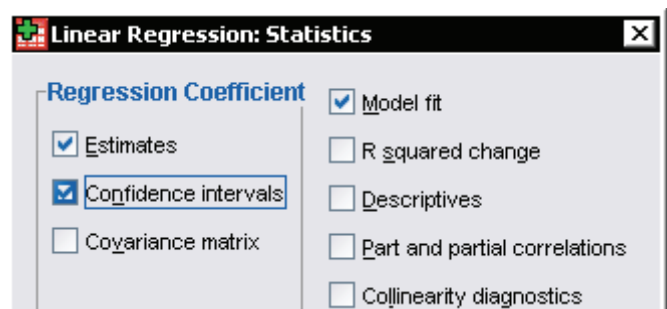
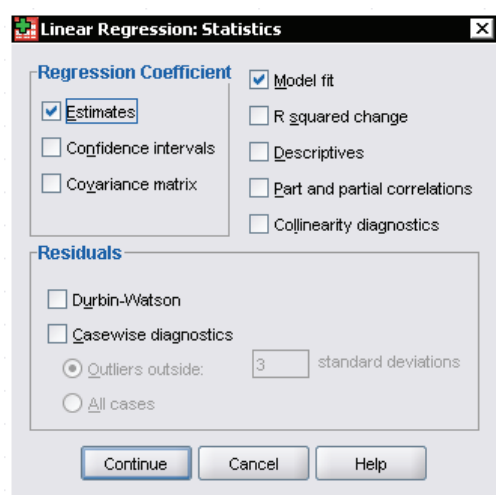
ขั้นที่ 5.3 คลิกที่ Linear จะได้เมนูของคำสั่งดังนี้



ขั้นที่ 5.4 เลือกตัวแปร x เป็นตัวแปรอิสระไปไว้ที่ช่อง Independent(s)
เลือกตัวแปร y เป็นตัวแปรตามไปไว้ที่ช่อง Dependent



ขั้นที่ 5.5 คลิกที่ปุ่ม Statistics จะได้เมนูย่อยเป็น



ขั้นที่ 5.6 คลิกในกรอบสี่เหลี่ยมหน้าช่อง Confidence intervals

ขั้นที่ 5.7 คลิก Continue และ OK ตามลำดับ
จะได้ผลการคำนวณดังนี้

Regression

[DataSet1] C:\Documents and Settings\dumrong\My Documents\datspps16\example16.sav

Variables Entered/Removed^b

Model	Variables Entered	Variables Removed	Method
1	x ^a	.	Enter

a. All requested variables entered.

b. Dependent Variable: y

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	0.99108867 ^a	0.98225675	0.97972200	0.53876550

a. Predictors: (Constant), x

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	112.484	1	112.484	387.516	0.0000002181 ^a
	Residual	2.032	7	.290		
	Total	114.516	8			

a. Predictors: (Constant), x

b. Dependent Variable: y

Coefficients^a

		Model	
		1	
		(Constant)	x
Unstandardized Coefficients	B	.257	2.930
	Std. Error	.532	.149
Standardized Coefficients	Beta		.991
t		.483	19.685
Sig.		0.64406363	0.00000022
95% Confidence Interval for B	Lower Bound	-1.002	2.578
	Upper Bound	1.516	3.282

a. Dependent Variable: y

ขั้นที่ 6. ไม่มีการเปิดตารางสถิติเพื่อหาค่าวิกฤต

ขั้นที่ 7. ช่วงความเชื่อมั่น 95% ของ β คือ (2.578, 3.282)

เพราะว่า 2.5 ไม่อยู่ในช่วงความเชื่อมั่น 95% ของ β เพราะฉะนั้นปฏิเสธ H_0

การทดสอบสมมติฐาน $H_0 : \alpha = 0$ (ระยะตัดแกน = 0)

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \alpha = 0$

กำหนดสมมติฐานอื่น $H_1 : \alpha \neq 0$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. ทำการสุ่มตัวอย่างขนาด n และ คำนวณค่า a , s , S_{xx}

ขั้นที่ 4. เลือกค่าสถิติที่

ขั้นที่ 5. คำนวณค่า $t_{\text{คำนวณ}} = \frac{a}{s \sqrt{\frac{\sum_{i=1}^n x_i^2}{n S_{xx}}}}$, $df = n - 2$ และ $\text{Sig} = P(t > |t_{\text{คำนวณ}}|)$

ขั้นที่ 6. เปิดตารางสถิติหาค่าวิกฤต $-t_{\frac{\alpha}{2}}$ และ $t_{\frac{\alpha}{2}}$ บริเวณวิกฤตคือ $t < -t_{\frac{\alpha}{2}}$ หรือ $t > t_{\frac{\alpha}{2}}$

ขั้นที่ 7. สรุปผล แบบที่ 1 โดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต

ถ้า $t_{\text{คำนวณ}} < -t_{\frac{\alpha}{2}}$ หรือ $t_{\text{คำนวณ}} > t_{\frac{\alpha}{2}}$ แล้วปฏิเสธ H_0

แบบที่ 2 ถ้า $\text{Sig} < \frac{\alpha}{2}$ แล้วปฏิเสธ H_0

หมายเหตุ กรณีของการทดสอบแบบ 1 ข้าง

	$H_1 : \alpha < 0$	$H_1 : \alpha > 0$
ค่าวิกฤต	$-t_{\alpha, n-2}$	$t_{\alpha, n-2}$
บริเวณวิกฤต	$t < -t_{\alpha, n-2}$	$t > t_{\alpha, n-2}$
ปฏิเสธ H_0	ถ้า $t_{\text{คำนวณ}} < 0$ และ $\text{Sig} < \alpha$	ถ้า $t_{\text{คำนวณ}} > 0$ และ $\text{Sig} < \alpha$

จากตัวอย่าง 8.2.1 ในการศึกษาความสัมพันธ์ระหว่างปริมาณอากาศเป็นพิษที่ถูกกำจัดออกไป (y) กับ ปริมาณน้ำฝน (x) จงทดสอบว่า สมการถดถอย $y = \alpha + \beta x$ มีระยะตัดแกน $\alpha = 0$ ที่ระดับนัยสำคัญ 0.05

วิธีทำ

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \alpha = 0$

กำหนดสมมติฐานอื่น $H_1 : \alpha \neq 0$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. ทำการสุ่มตัวอย่างขนาด n และ คำนวณค่า a, s, S_{xx}

ขั้นที่ 4. เลือกค่าสถิติที่

ขั้นที่ 5. คำนวณค่าสถิติที่เลือกจากข้อมูลตัวอย่าง $t_{\text{คำนวณ}} = \frac{a}{s \sqrt{\frac{\sum_{i=1}^n x_i^2}{n S_{xx}}}}$, $df = n - 2$

จากตาราง Coefficient จะได้ $t_{\text{คำนวณ}} = \frac{a}{s \sqrt{\frac{\sum_{i=1}^n x_i^2}{n S_{xx}}}} = 58.583$, $df = 7$ และ $\text{Sig} = 0.0000000001$

ขั้นที่ 6. ค่าวิกฤตคือ $-t_{0.025} = -2.365$, $t_{0.025} = 2.365$ บริเวณวิกฤตคือ $t < -2.365$ หรือ $t > 2.365$

ขั้นที่ 7. สรุปผล แบบที่ 1 เพราะว่า $t_{\text{คำนวณ}}$ อยู่ในบริเวณวิกฤต เพราะฉะนั้นปฏิเสธ H_0

แบบที่ 2 เพราะว่า $\text{Sig} = 0.0000000001 < \frac{\alpha}{2} = 0.025$ เพราะฉะนั้นปฏิเสธ H_0

เพราะฉะนั้นสมการถดถอย $y = \alpha + \beta x$ มีค่าสัมประสิทธิ์การถดถอย $\alpha \neq 0$ ที่ระดับนัยสำคัญ 0.05

หลักการและขั้นตอนการทดสอบสมมติฐานกรณี α_0 ไม่เท่ากับ 0

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \alpha = \alpha_0$

กำหนดสมมติฐานอื่น $H_1 : \alpha \neq \alpha_0$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. ทำการสุ่มตัวอย่างขนาด n และ คำนวณค่าโดยการนำข้อมูลเข้าสู่โปรแกรม SPSS

ขั้นที่ 4. เพราะว่าผลการคำนวณของ SPSS ไม่ให้ค่า $t = \frac{a - \alpha_0}{s \sqrt{\frac{\sum_{i=1}^n x_i^2}{n S_{xx}}}}$ ออกมาโดยตรง

เพราะฉะนั้นเราจึงใช้ช่วงความเชื่อมั่นของ α ช่วยในการสรุปสมมติฐาน

ขั้นที่ 5. ให้หาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่า α

ขั้นที่ 6. ไม่มีการเปิดตารางสถิติเพื่อหาค่าวิกฤต

ขั้นที่ 7. สรุปผลโดยการดูว่า α_0 อยู่ในช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่า α ที่หาได้หรือไม่ ถ้า α_0 อยู่ในช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่า α ที่หาได้แล้ว ไม่ปฏิเสธ H_0

จากตัวอย่างข้อมูล

X	1.50	1.80	2.40	3.00	3.50	3.90	4.40	4.80	5.00
Y	4.80	5.70	7.00	8.30	10.90	12.40	13.10	13.60	15.30

สมมติความสัมพันธ์ของตัวแปรในรูปแบบเชิงเส้นคือ $y = \alpha + \beta x$

งทดสอบสมมติฐานว่า $\alpha = 150$ แยกกับ $\alpha \neq 150$ ที่ระดับนัยสำคัญ 0.05

วิธีทำ กำหนดสมมติฐานหลัก $H_0 : \alpha = 150$

กำหนดสมมติฐานอื่น $H_1 : \alpha \neq 150$

กำหนดระดับนัยสำคัญ 0.05

จากตาราง Coefficients ช่วงความเชื่อมั่น 95% ของ α คือ (146.993, 159.358)

เพราะว่า 150 อยู่ในช่วงความเชื่อมั่น 95% ของ α เพราะฉะนั้นไม่ปฏิเสธ H_0

8.5 การเลือกรูปแบบความสัมพันธ์ที่เหมาะสมกับข้อมูล

ความสัมพันธ์แบบเชิงเดียว ซึ่งประกอบด้วยตัวแปรอิสระ 1 ตัว และตัวแปรตาม 1 ตัว

รูปแบบของสมการความสัมพันธ์เชิงเส้นเชิงเดียวอาจมีรูปแบบเป็น

1. $y = a + bx$

2. $\ln y = a + b \ln x$

3. $y = a + b \ln x$

4. $\ln y = a + bx$

เมื่อเรามีข้อมูลและต้องการรู้ว่ารูปแบบใดเหมาะสมกับข้อมูล

สามารถใช้โปรแกรม Mathcad ช่วยในการเขียนกราฟและคำนวณค่าสหสัมพันธ์ได้ จากตัวอย่าง 8.2.1

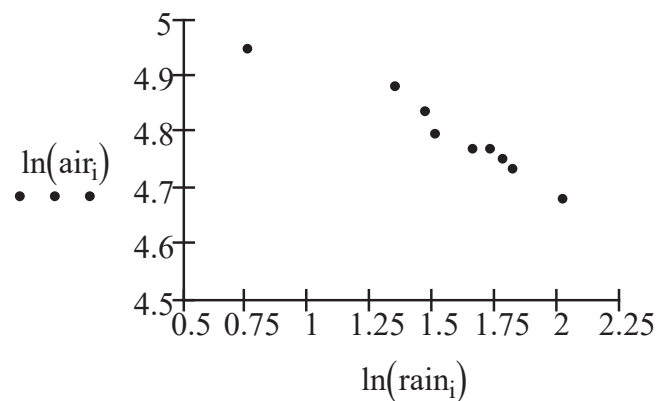
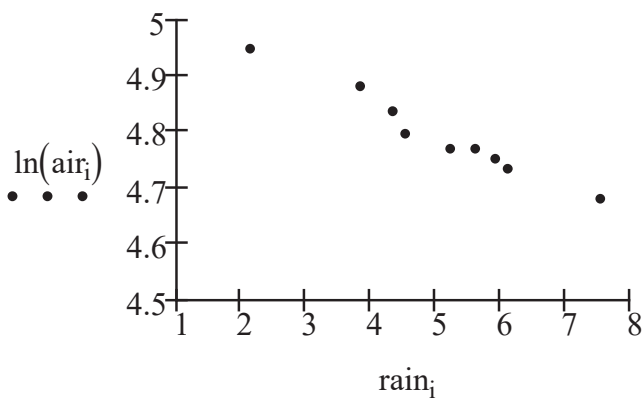
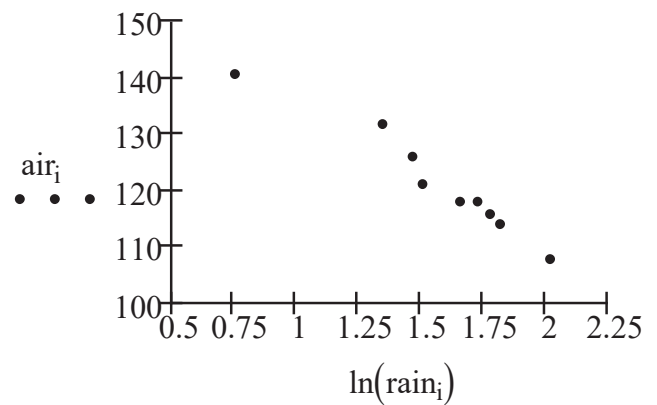
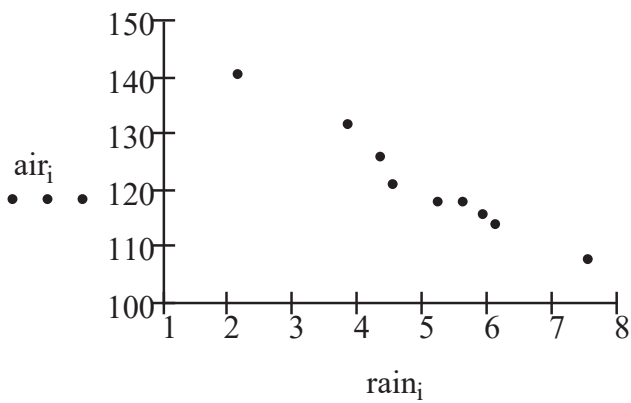
เราสามารถเขียนแผนภาพการกระจาย 4 รูปแบบและคำนวณค่าสหสัมพันธ์ได้ดังนี้

การคำนวณด้วย Mathcad

$$\text{rain} := \begin{pmatrix} 4.30 \\ 4.50 \\ 5.90 \\ 5.60 \\ 6.10 \\ 5.20 \\ 3.80 \\ 2.10 \\ 7.50 \end{pmatrix} \quad \text{air} := \begin{pmatrix} 126.00 \\ 121.00 \\ 116.00 \\ 118.00 \\ 114.00 \\ 118.00 \\ 132.00 \\ 141.00 \\ 108.00 \end{pmatrix} \quad \text{ORIGIN}:=1 \quad i:=1..10$$

$$\text{corr}(\text{rain}, \text{air}) = -0.978658 \quad \text{corr}(\overrightarrow{\ln(\text{rain})}, \text{air}) = -0.976417$$

$$\text{corr}(\text{rain}, \overrightarrow{\ln(\text{air})}) = -0.982471 \quad \text{corr}(\overrightarrow{\ln(\text{rain})}, \overrightarrow{\ln(\text{air})}) = -0.971364$$



$$\text{slope}(\text{rain}, \text{air}) = -6.323988$$

$$\text{intercept}(\text{rain}, \text{air}) = 153.175493$$

$$\text{slope}(\overrightarrow{\text{rain}}, \ln(\text{air})) = -0.051097$$

$$\text{intercept}(\overrightarrow{\text{rain}}, \ln(\text{air})) = 5.052922$$

$$\text{slope}(\overrightarrow{\ln(\text{rain})}, \text{air}) = -26.616482$$

$$\text{intercept}(\overrightarrow{\ln(\text{rain})}, \text{air}) = 162.986181$$

$$\text{slope}(\overrightarrow{\ln(\text{rain})}, \overrightarrow{\ln(\text{air})}) = -0.213116$$

$$\text{intercept}(\overrightarrow{\ln(\text{rain})}, \overrightarrow{\ln(\text{air})}) = 5.129167$$

ผลการคำนวณของ Mathcad จะได้ค่าสัมประสิทธิ์สหสัมพันธ์ของรูปแบบต่างๆ เป็นดังนี้

รูปแบบ	ชื่อรูปแบบใน SPSS	สมการแบบเชิงเส้น	ค่าสัมประสิทธิ์สหสัมพันธ์
1	Linear	$air = a + b \text{rain}$	-0.978658
2	Power	$\ln(\text{air}) = a + b \ln(\text{rain})$	-0.971364
3	Logarithmic	$air = a + b \ln(\text{rain})$	-0.976417
4	Exponential	$\ln(\text{air}) = a + b \text{rain}$	-0.982471

สมการในรูปแบบทั้ง 4 รูปแบบคือ

รูปแบบ	สมการแบบเชิงเส้น	สมการความสัมพันธ์
1	$air = 153.175493 - 6.323988 \text{rain}$	$air = 153.175493 - 6.323988 \text{rain}$
2	$\ln(\text{air}) = 5.129167 - 0.213116 \ln(\text{rain})$	$air = 168.876385 \text{rain}^{-0.213116}$
3	$air = 162.986181 - 26.616482 \ln(\text{rain})$	$air = 162.986181 - 26.616482 \ln(\text{rain})$
4	$\ln(\text{air}) = 5.052922 - 0.051097 \text{rain}$	$air = 156.479029 e^{-0.051097 \text{rain}}$

การตัดสินใจทางด้านสถิติเราเลือกรูปแบบที่มีค่าสัมประสิทธิ์สหสัมพันธ์ที่มีค่าสัมบูรณ์มากที่สุด

เพราะฉะนั้นเลือกรูปแบบความสัมพันธ์ที่ 4 คือ $\ln(\text{air}) = 5.052922 - 0.051097 \text{rain}$

จัดรูปเป็น $air = e^{5.052922 - 0.051097 \text{rain}} = e^{5.052922} e^{-0.051097 \text{rain}} = 156.479029 e^{-0.051097 \text{rain}}$

สรุปสมการแสดงความสัมพันธ์ที่เหมาะสมกับข้อมูลคือ $air = 156.479029 e^{-0.051097 \text{rain}}$

การหาความสัมพันธ์ในรูปแบบต่าง ๆ ด้วย SPSS for Windows

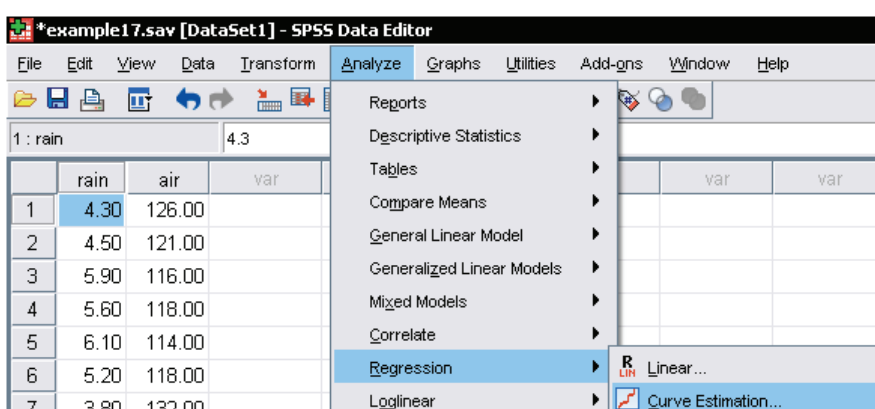
โปรแกรม SPSS for Windows มีคำสั่งที่ช่วยในการหาความสัมพันธ์ในรูปแบบอื่นๆ เช่น

Linear	$y = a + bx = b_0 + b_1 x$ เมื่อ $b_0 = a, b_1 = b$
Power	$\ln(y) = a + b \ln(x)$ หรือ $y = e^a x^b = b_0 x^{b_1}$ เมื่อ $b_0 = e^a, b_1 = b$
Logarithmic	$y = a + b \ln(x) = b_0 + b_1 \ln(x)$ เมื่อ $b_0 = a, b_1 = b$
Exponential	$\ln(y) = a + bx$ หรือ $y = e^a e^{bx} = b_0 e^{b_1 x}$ เมื่อ $b_0 = e^a, b_1 = b$
Quadratic	$y = b_0 + b_1 x + b_2 x^2$

การวิเคราะห์ความสัมพันธ์ของ SPSS มีขั้นตอนดังนี้

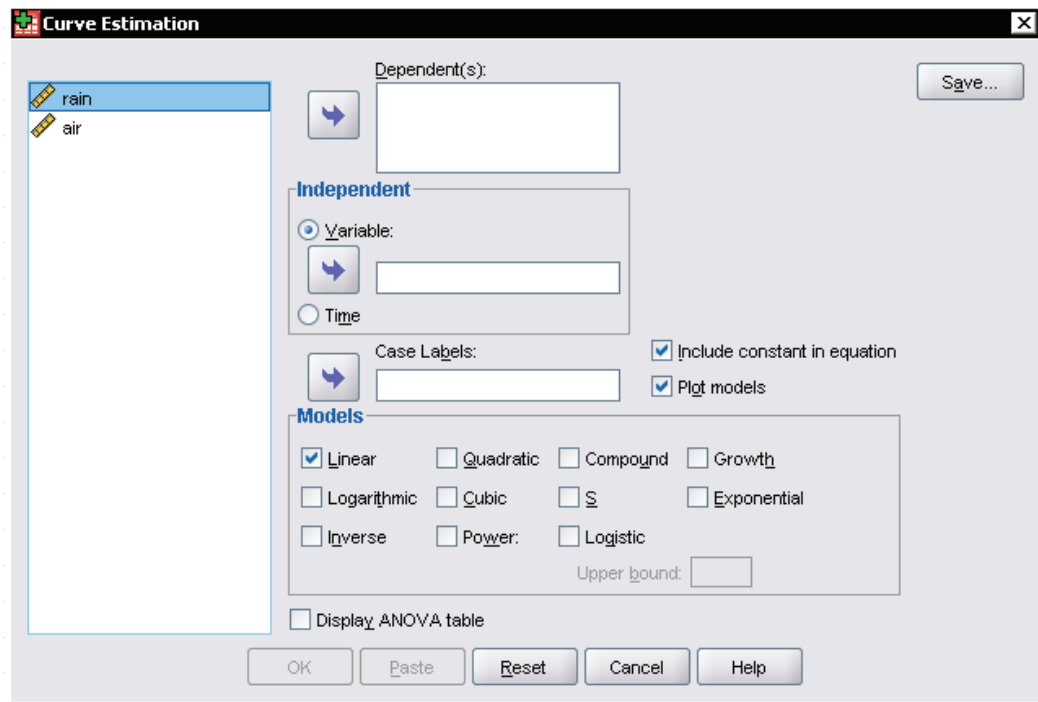
ขั้นที่ 1. นำแฟ้มข้อมูล example17.sav เข้ามาใน SPSS Data Editor

ขั้นที่ 2. เลือกคำสั่ง Analyze / Regression / Curve Estimation



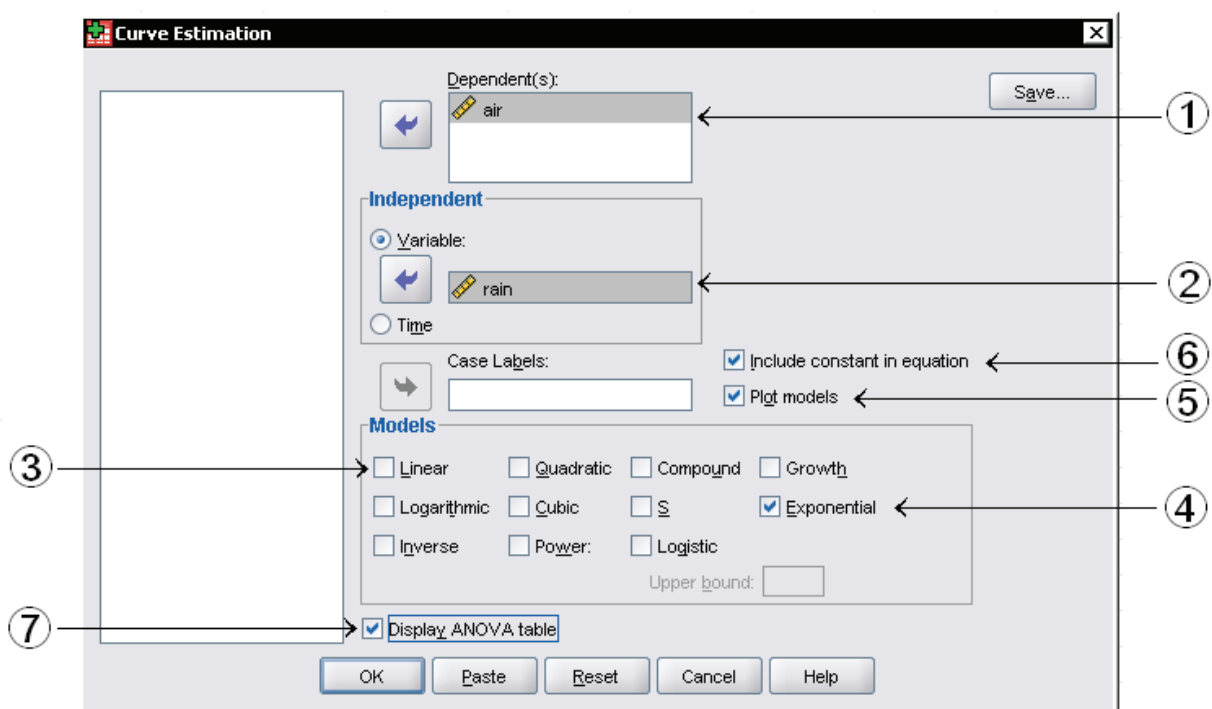
	rain	air
1	4.30	126.00
2	4.50	121.00
3	5.90	116.00
4	5.60	118.00
5	6.10	114.00
6	5.20	118.00
7	3.80	132.00
8	2.10	141.00
9	7.50	108.00

ขั้นที่ 3. คลิกคำสั่ง Curve Estimation จะได้เมนูย่อยเป็น



ขั้นที่ 4. เลือกข้อกำหนดต่าง ๆ ของการวิเคราะห์ข้อมูลเช่น

1. เลือกตัวแปรตาม air
2. เลือกตัวแปรอิสระ rain
3. ยกเลิกการหารูปแบบ Linear
4. เลือกรูปแบบ Exponential (หรือรูปแบบอื่นเช่น Power, Logarithmic)
5. ให้เขียนกราฟของข้อมูลและสมการแสดงความสัมพันธ์ที่คำนวณได้
6. การหาสมการต้องการให้มีพจน์ของค่าคงตัวด้วย
7. ให้แสดงผลตาราง ANOVA



ขั้นที่ 5. คลิก OK จะได้ผลการคำนวณเป็นดังนี้

Curve Fit

[DataSet1] C:\Documents and Settings\dumrong\My Documents\dataspss16\example17.sav

Model Description

Model Name	MOD_2
Dependent Variable	1 air
Equation	1 Exponential ^a
Independent Variable	rain
Constant	Included
Variable Whose Values Label Observations in Plots	Unspecified

a. The model requires all non-missing values to be positive.

Case Processing Summary

	N
Total Cases	9
Excluded Cases ^a	0
Forecasted Cases	0
Newly Created Cases	0

a. Cases with a missing value in any variable are excluded from the analysis.

Variable Processing Summary

	Variables	
	Dependent	Independent
	air	rain
Number of Positive Values	9	9
Number of Zeros	0	0
Number of Negative Values	0	0
Number of Missing Values	User-Missing System-Missing	0 0

air

Exponential

Model Summary

R	R Square	Adjusted R Square	Std. Error of the Estimate
0.98247104	0.96524934	0.96028496	0.01608195

The independent variable is rain.

ANOVA

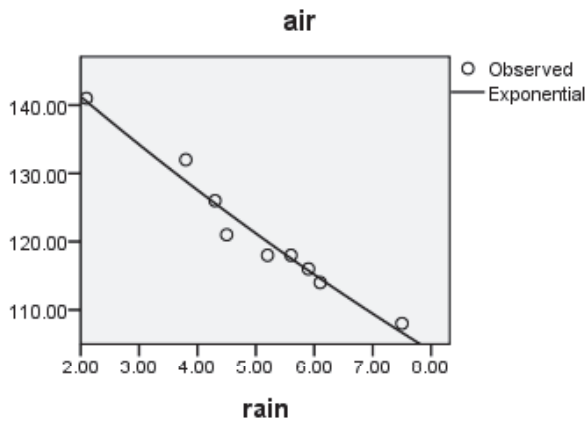
	Sum of Squares	df	Mean Square	F	Sig.
Regression	0.05028653	1	0.05028653	194.43501657	0.00000231
Residual	0.00181040	7	0.00025863		
Total	0.05209693	8			

The independent variable is rain.

Coefficients

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
rain	-0.05109727	0.00366446	-0.98247104	-13.94399572	0.00000231
(Constant)	156.47907058	2.98725029		52.38230998	0.00000000

The dependent variable is ln(air).



การแปลความหมายจากผลการวิเคราะห์ของ SPSS for Windows

1. สมการแสดงความสัมพันธ์ในรูปแบบ Exponential คือ $air = 156.47907058 e^{-0.05109727 \text{ rain}}$
2. สัมประสิทธิ์สหสัมพันธ์ของตัวแปร rain, air ในรูปแบบ Exponential มีค่าเท่ากับ -0.98247104
หมายเหตุ เครื่องหมายบวกหรือลบ ของค่า R ให้ดูจากสัมประสิทธิ์ B ของตัวแปร rain

3. การทดสอบสมมติฐาน

- ขั้นที่ 1. กำหนดสมมติฐาน $H_0 : \rho = 0$ (ตัวแปร rain, air ไม่มีความสัมพันธ์กันในรูปแบบ Exponential)
 $H_1 : \rho \neq 0$ (ตัวแปร rain, air มีความสัมพันธ์กันในรูปแบบ Exponential)

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

ขั้นที่ 3. สุ่มตัวอย่าง

ขั้นที่ 4. เลือกใช้ค่าสถิติเอฟ

ขั้นที่ 5. คำนวณค่าสถิติเอฟจากตาราง ANOVA , $f_{\text{คำนวณ}} = 194.43501657$

ขั้นที่ 6. ค่าวิกฤตคือ $f_{0.05, (1, 7)} = 5.59$ บริเวณวิกฤตคือ $F > 5.59$

ขั้นที่ 7. สรุปผล แบบที่ 1 เพราะว่า $f_{\text{คำนวณ}} = 194.43502 > 5.59$ เพราะฉะนั้นปฏิเสธ H_0
 แบบที่ 2 ถ้า $\text{Sig} = 0.0000 < \alpha = 0.05$ เพราะฉะนั้นปฏิเสธ H_0

เพราะฉะนั้นตัวแปร rain, air มีความสัมพันธ์กันในรูปแบบ Exponential

8.6 การคำนวณค่าสัมประสิทธิ์สหสัมพันธ์และสมการถดถอยพหุคูณ

8.6.1 การคำนวณค่าสัมประสิทธิ์สหสัมพันธ์ของตัวแปรหลายๆ คู่

ในกรณีมีตัวแปรหลายคู่ที่ต้องการหาค่าสัมประสิทธิ์สหสัมพันธ์ เช่น น้ำหนัก (x_1), ความสูง (x_2), อายุ (x_3)

x_1	x_2	x_3
64.00	57.00	8.00
71.00	59.00	10.00
53.00	49.00	6.00
67.00	62.00	11.00
55.00	51.00	8.00
58.00	50.00	7.00
77.00	55.00	10.00
57.00	48.00	9.00
56.00	52.00	10.00
51.00	42.00	6.00
76.00	61.00	12.00
68.00	57.00	9.00

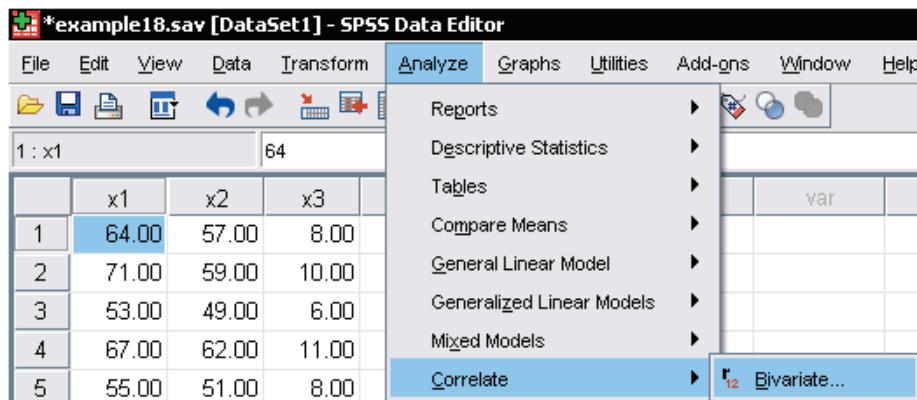
การหาค่าสัมประสิทธิ์สหสัมพันธ์ของตัวแปรหลายคู่พร้อมกันด้วย SPSS for Windows

ขั้นที่ 1. สร้างข้อมูลประกอบด้วยตัวแปร x_1, x_2, x_3

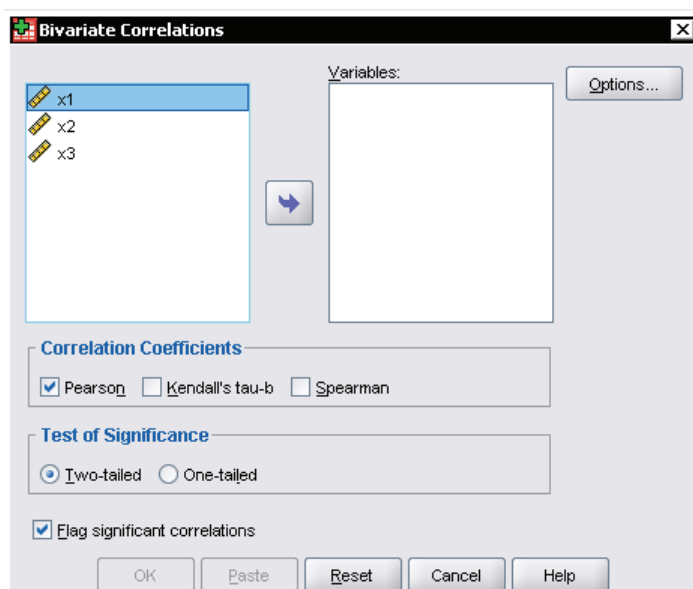
ใน SPSS Data Editor แล้ว Save เพิ่มข้อมูลชื่อ example18.sav

ขั้นที่ 2. เลือกคำสั่ง Analyze / Correlate / Bivariate

	x_1	x_2	x_3
1	64.00	57.00	8.00
2	71.00	59.00	10.00
3	53.00	49.00	6.00
4	67.00	62.00	11.00
5	55.00	51.00	8.00
6	58.00	50.00	7.00
7	77.00	55.00	10.00
8	57.00	48.00	9.00
9	56.00	52.00	10.00
10	51.00	42.00	6.00
11	76.00	61.00	12.00
12	68.00	57.00	9.00



คลิกที่ Bivariate
จะได้เมนูย่อยเป็น

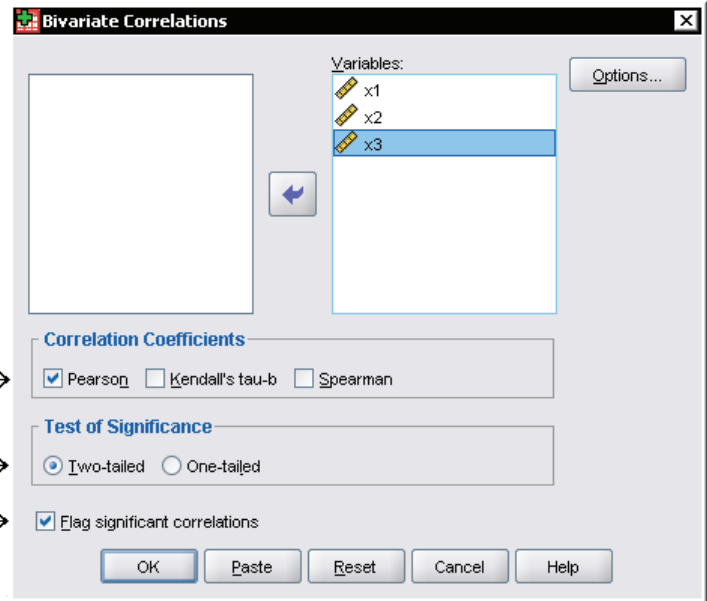


ขั้นที่ 3. เลือกตัวแปร x1, x2 และ x3 มาไว้ที่ช่อง Variables

1. เลือกคำนวณสหสัมพันธ์ Pearson
2. เลือกคำนวณ

Two-tailed Significant

3. ให้แสดงเครื่องหมาย * สำหรับคู่ที่มีค่าความสัมพันธ์แตกต่างกันอย่างมีนัยสำคัญ



- ① →
- ② →
- ③ →

หมายเหตุ

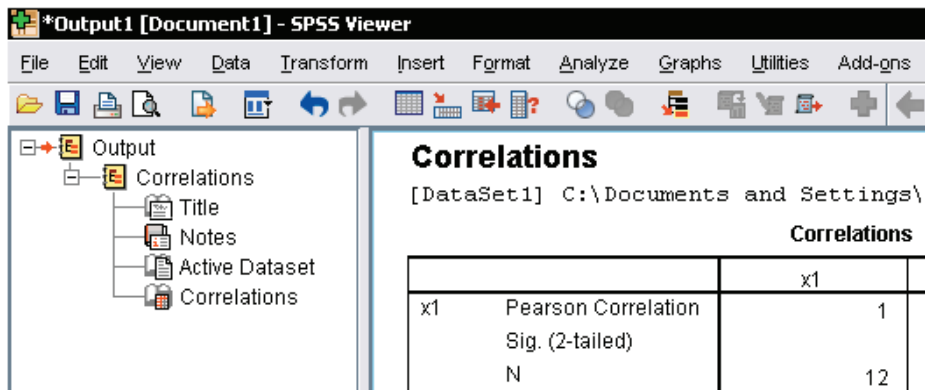
1. สำหรับข้อมูลเชิงปริมาณ และมีการแจกแจงปกติ

ควรใช้ Pearson Correlation Coefficient

2. สำหรับข้อมูลแบบอันดับที่หรือลำดับตำแหน่ง

ควรใช้ Kendall's tau-b หรือ Spearman Correlation Coefficient

ขั้นที่ 4. คลิก OK จะได้ผลการคำนวณเป็น



ผลการคำนวณทั้งหมดคือ

Correlations

[DataSet1] C:\Documents and Settings\dumrong\My Documents\dataspss16\example18.sav

		x1	x2	x3
x1	Pearson Correlation	1	0.81964508**	0.76981680**
	Sig. (2-tailed)		0.00110072	0.00340655
	N	12	12	12
x2	Pearson Correlation	0.81964508**	1	0.79840746**
	Sig. (2-tailed)	0.00110072		0.00184860
	N	12	12	12
x3	Pearson Correlation	0.76981680**	0.79840746**	1
	Sig. (2-tailed)	0.00340655	0.00184860	
	N	12	12	12

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

การแปลความหมายของผลการวิเคราะห์ข้อมูล

ค่าสหสัมพันธ์ของ น้ำหนัก (x_1), ความสูง (x_2) เท่ากับ 0.81964508

ค่าสหสัมพันธ์ของ น้ำหนัก (x_1), อายุ (x_3) เท่ากับ 0.76981680

ค่าสหสัมพันธ์ของ ความสูง (x_2), อายุ (x_3) เท่ากับ 0.79840746

เพราะฉะนั้น น้ำหนัก (x_1), ความสูง (x_2) มีความสัมพันธ์กันมากที่สุด

การทดสอบสมมติฐานเกี่ยวกับความสัมพันธ์ในรูปแบบเชิงเส้น กำหนดนัยสำคัญ $\alpha = 0.05$

สมมติฐานหลัก H_0 : น้ำหนัก (x_1), ความสูง (x_2) ไม่มีความสัมพันธ์ในรูปแบบเชิงเส้น

สมมติฐานอื่น H_1 : น้ำหนัก (x_1), ความสูง (x_2) มีความสัมพันธ์ในรูปแบบเชิงเส้น

จากตาราง Correlations ค่าสัมประสิทธิ์สหสัมพันธ์ = 0.81964508

และ Sig. (2-tailed) = 0.00110072

เพราะว่า Sig. (2-tailed) = 0.00110072 น้อยกว่า 0.05 เพราะฉะนั้นปฏิเสธ H_0

เพราะฉะนั้น น้ำหนัก (x_1), ความสูง (x_2) มีความสัมพันธ์กันในรูปแบบเชิงเส้น

ในทำนองเดียวกัน น้ำหนัก (x_1), อายุ (x_3) มีความสัมพันธ์ในรูปแบบเชิงเส้น

และ ความสูง (x_2), อายุ (x_3) มีความสัมพันธ์ในรูปแบบเชิงเส้น

8.6.2 การคำนวณค่าสัมประสิทธิ์สหสัมพันธ์พหุคูณและสมการถดถอยพหุคูณ

การหาสมการถดถอยพหุคูณเชิงเส้นของ X_1 เทียบกับ X_2 และ X_3

$X_1 = b_{1.23} + b_{12.3} X_2 + b_{13.2} X_3$ ของ X_1 เทียบกับ X_2 และ X_3

สมการปกติคือ $\sum X_1 = b_{1.23} n + b_{12.3} \sum X_2 + b_{13.2} \sum X_3$

$$\sum X_1 X_2 = b_{1.23} \sum X_2 + b_{12.3} \sum X_2^2 + b_{13.2} \sum X_2 X_3$$

$$\sum X_1 X_3 = b_{1.23} \sum X_3 + b_{12.3} \sum X_2 X_3 + b_{13.2} \sum X_3^2$$

จากข้อมูล $\sum X_1 = 753$, $\sum X_2 = 643$, $\sum X_3 = 106$,

$$\sum X_1^2 = 48139, \sum X_2^2 = 34843, \sum X_3^2 = 976,$$

$$\sum X_1 X_2 = 40830, \sum X_1 X_3 = 6796, \sum X_2 X_3 = 5779$$

แทนค่าต่าง ๆ ลงในสมการปกติ จะได้

$$12 b_{1.23} + 643 b_{12.3} + 106 b_{13.2} = 753$$

$$643 b_{1.23} + 34843 b_{12.3} + 5779 b_{13.2} = 40830$$

$$106 b_{1.23} + 5779 b_{12.3} + 976 b_{13.2} = 6796$$

โดยการแก้สมการจะได้

$$b_{1.23} = 3.6512, b_{12.3} = 0.8546 \text{ และ } b_{13.2} = 1.5063$$

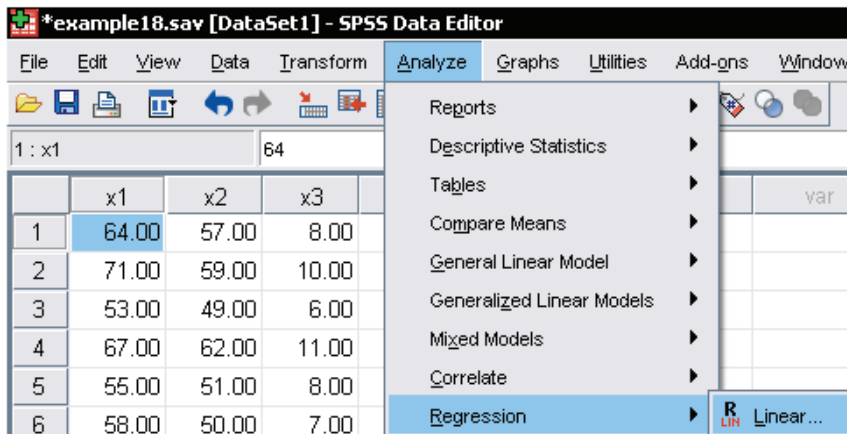
เพราะฉะนั้นสมการถดถอยคือ $X_1 = 3.6512 + 0.8546 X_2 + 1.5063 X_3$

การหาสมการ $X_1 = b_{1.23} + b_{12.3} X_2 + b_{13.2} X_3$ ด้วย SPSS for Windows

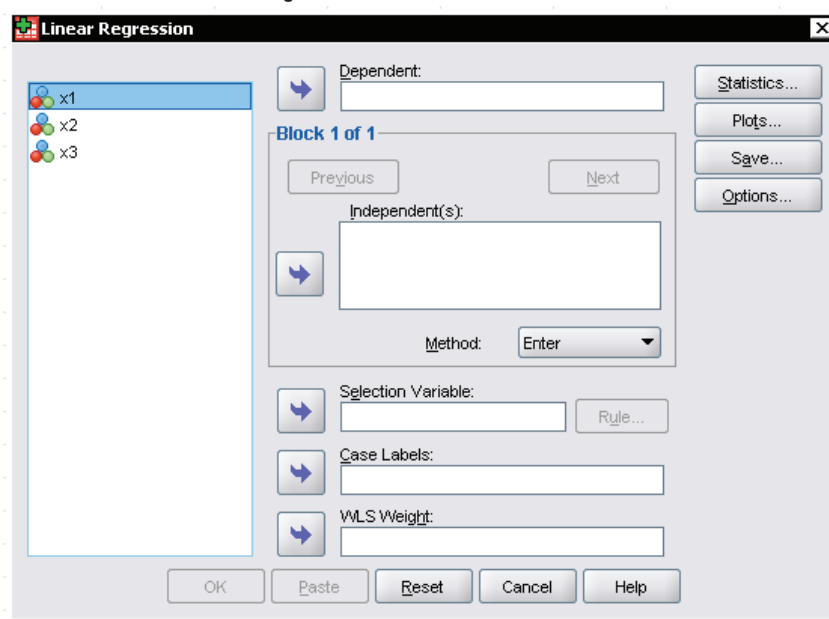
ขั้นที่ 1. นำข้อมูลเข้าสู่ SPSS

1: x1	x2	x3
64.00	57.00	8.00
71.00	59.00	10.00
53.00	49.00	6.00
67.00	62.00	11.00
55.00	51.00	8.00
58.00	50.00	7.00
77.00	55.00	10.00
57.00	48.00	9.00
56.00	52.00	10.00
51.00	42.00	6.00
76.00	61.00	12.00
68.00	57.00	9.00

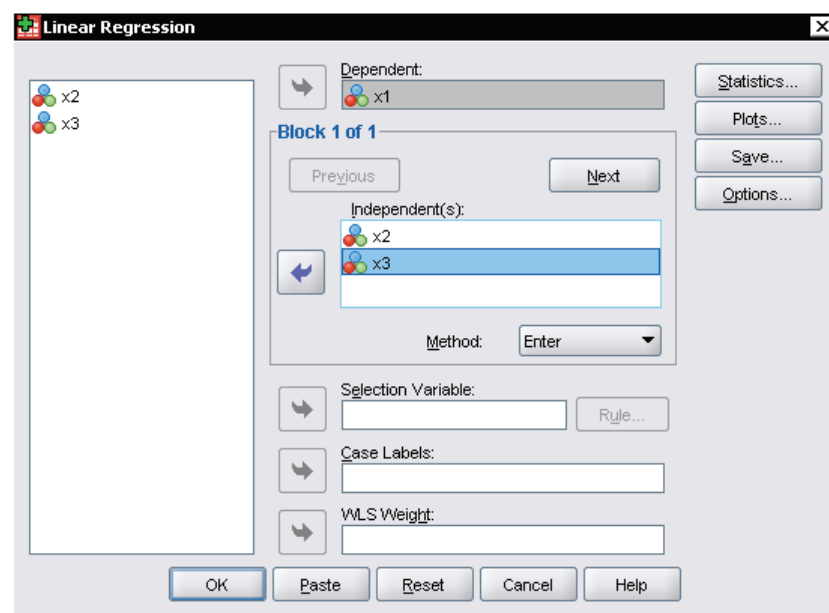
ขั้นที่ 2. เลือกคำสั่ง Analyze / Regression / Linear



ขั้นที่ 3. คลิกที่ Linear จะได้เมนูย่อยเป็น



ขั้นที่ 4. เลือกตัวแปร x1 ไปไว้ที่ช่อง dependent และ เลือกตัวแปร x2, x3 ไปไว้ที่ช่อง Independent[s]



ขั้นที่ 5. คลิก OK จะได้ผลการคำนวณทั้งหมดคือ

Regression

[DataSet1] C:\Documents and Settings\dumrong\My Documents\dataspss16\example18.sav

Variables Entered/Removed^b

Model	Variables Entered	Variables Removed	Method
1	x3, x2 ^a	.	Enter

a. All requested variables entered.

b. Dependent Variable: x1

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	0.84175667 ^a	0.70855430	0.64378858	5.36321469

a. Predictors: (Constant), x3, x2

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	629.373354	2	314.686677	10.940269	0.00389502 ^a
	Residual	258.876646	9	28.764072		
	Total	888.250000	11			

a. Predictors: (Constant), x3, x2

b. Dependent Variable: x1

Coefficients^a

		Model		
		1		
		(Constant)	x2	x3
Unstandardized Coefficients	B	3.6512	0.8546	1.5063
	Std. Error	16.1678	0.4517	1.4143
Standardized Coefficients	Beta		0.5655	0.3183
t		0.2258	1.8921	1.0651
Sig.		0.8264	0.0910	0.3146

a. Dependent Variable: x1

จากตาราง Coefficients จะได้ $b_{1.23} = 3.6512$, $b_{12.3} = 0.8546$, $b_{13.2} = 1.5063$

ความหมายของผลการคำนวณที่ได้คือ $X_1 = b_{1.23} + b_{12.3} X_2 + b_{13.2} X_3$

เพราะฉะนั้นสมการถดถอยคือ $X_1 = 3.6512 + 0.8546 X_2 + 1.5063 X_3$

บทที่ 9

การวิเคราะห์ความแปรปรวน



การวิเคราะห์ความแปรปรวน เป็นการทดสอบสมมติฐานว่าค่าเฉลี่ยของประชากรตั้งแต่ 2 ชุดขึ้นไปเท่ากันหรือไม่ ซึ่งการทดสอบจะสามารถทำได้กับข้อมูลที่เก็บมาจากข้อมูลตัวอย่างที่จำแนกเป็นกลุ่มๆ และทำการทดสอบด้วยการวิเคราะห์ความแปรปรวน โดยมีรูปแบบการวิเคราะห์ความแปรปรวนที่สำคัญคือ

9.1 การวิเคราะห์ความแปรปรวนแบบจำแนกทางเดียว (One-Way ANOVA, Simple-Factor ANOVA)

กลุ่มตัวอย่างขนาด n_1, n_2, \dots, n_k จากประชากรชุดที่ 1, 2, \dots , k ต่างๆ กัน จะจำแนกประชากร k ชุดต่างๆ กันนี้ออกตาม **วิธีการปฏิบัติ** (Treatments) โดยทั่วไปคำว่า วิธีการปฏิบัติจะใช้กับการจำแนกต่างๆ กัน เช่น การอัดแรงคอนกรีตชนิดต่างๆ ปุ๋ยชนิดต่างๆ หรือ บริเวณต่างๆ ของประเทศ ฯลฯ ตัวอย่างเช่น วิศวกรผู้หนึ่งสนใจในการหาค่าเฉลี่ยของการดูดความชื้นในคอนกรีตอัดแรง 5 ชนิด ได้นำตัวอย่างมาทดลองชนิดละ 6 หน่วย นาน 48 ชั่วโมง ได้ข้อมูลดังนี้

ชนิดของคอนกรีต(การดูดความชื้น หน่วยน้ำหนัก %)

1	2	3	4	5
551	595	639	417	563
457	580	615	449	631
450	508	511	517	522
731	583	573	438	613
499	633	648	415	656
632	517	677	555	679

ในกรณีทั่วไป สมมติว่าประชากร k ชุด เป็นอิสระต่อกันและมีการแจกแจงปกติ โดยมีค่าเฉลี่ย $\mu_1, \mu_2, \dots, \mu_k$ ตามลำดับและความแปรปรวนเท่ากันคือ σ^2

การทดสอบสมมติฐาน $H_0 : \mu_1 = \mu_2 = \dots = \mu_k$

แย้งกับ $H_1 : \text{อย่างน้อยที่สุดค่าเฉลี่ย 2 ค่าไม่เท่ากัน}$

จากรูปแบบทั่วไปของข้อมูล

ลำดับที่	วิธีการปฏิบัติ (Treatment)						
	1	2	...	j	...	k	
1	x_{11}	x_{12}	...	x_{1j}	...	x_{1k}	
2	x_{21}	x_{22}	...	x_{2j}	...	x_{2k}	
:	:	:		:		:	
i	x_{i1}	x_{i2}	...	x_{ij}	...	x_{ik}	
:	:	:		:		:	
	x_{n11}	x_{n22}		x_{nj}	...	x_{nk}	
จำนวน	n_1	n_2		n_j		n_k	
รวม	$T_{.1}$	$T_{.2}$		$T_{.j}$		$T_{.k}$	$T_{..}$
ค่าเฉลี่ย	$\bar{x}_{.1}$	$\bar{x}_{.2}$		$\bar{x}_{.j}$		$\bar{x}_{.k}$	$\bar{x}_{..}$

หมายเหตุ จำนวนตัวอย่างในแต่ละ Treatment ไม่จำเป็นต้องมีขนาดเท่ากัน

$$N = \text{จำนวนข้อมูลทั้งหมด} = n_1 + n_2 + \dots + n_k$$

$$x_{ij} = \text{ค่าสังเกตตัวที่ } i \text{ ซึ่งเลือกจากวิธีการปฏิบัติที่ } j$$

$$T_{.j} = \text{ผลรวมของค่าสังเกตทั้งหมดจากตัวอย่างที่มีวิธีการปฏิบัติที่ } j$$

$$\bar{x}_{.j} = \text{ค่าเฉลี่ยของค่าสังเกตจากตัวอย่างที่มีวิธีการปฏิบัติที่ } j$$

$$T_{..} = \text{ผลรวมทั้งหมดของค่าสังเกต}$$

$$\bar{x}_{..} = \frac{T_{..}}{N} = \text{ค่าเฉลี่ยของค่าสังเกตทั้งหมด}$$

การแปรผันของข้อมูลจำแนกเป็น 3 ส่วนคือ

1. การแปรผันโดยรวมทั้งหมด (SST : TOTAL SUM OF SQUARE, Total Sum Square)

$$SST = \sum_{j=1}^k \sum_{i=1}^{n_j} x_{ij}^2 - \frac{T_{..}^2}{N}$$

2. การแปรผันที่เกิดจาก Treatment ต่างกัน (SSTR : TREATMENT SUM OF SQUARE, Between-

Groups Sum Square),
$$SSTR = \sum_{j=1}^k \frac{T_{.j}^2}{n_j} - \frac{T_{..}^2}{N}$$

3. การแปรผันที่เกิดภายในกลุ่ม Treatment เดียวกัน (SSE : ERROR SUM OF SQUARE, Within-Groups Sum Square, Residual Sum Square) $SSE = SST - SSTR$

นำค่า SST, SSTR, SSE มาสร้างเป็นตารางวิเคราะห์ความแปรปรวน (ANalysis Of VAriance : ANOVA)

ตาราง ANOVA ในตำราสถิติ

แหล่งการแปรผัน	ผลบวกกำลังสอง	ระดับชั้นความเสรี	ค่าเฉลี่ยของผลบวกกำลังสอง	f _{คำนวณ}
วิธีการปฏิบัติ (Treatment)	SSTR	k - 1	$MSTR = \frac{SSA}{k-1}$	$f_{คำนวณ} = \frac{MSTR}{MSE}$
ความคลาดเคลื่อน (Error)	SSE	N - k	$MSE = \frac{SSE}{N-k}$	
ทั้งหมด (Total)	SST	N - 1		

รูปแบบของตาราง ANOVA ที่คำนวณโดยใช้โปรแกรม SPSS for Windows

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	SSTR	k - 1	MSTR	f _{คำนวณ}	P(F > f _{คำนวณ})
Within Groups	SSE	N - k	MSE		
Total	SST	N - 1			

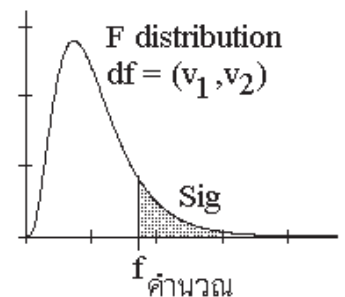
หมายเหตุ ความหมายของ Sig. ของค่าสถิติเอฟ f_{คำนวณ}

จากค่าสถิติเอฟ f_{คำนวณ} และ ระดับชั้นความเสรี $v_1 = k - 1, v_2 = N - k$ ที่คำนวณได้จากตัวอย่าง

$Sig = P(F > f_{คำนวณ})$

= พื้นที่ใต้โค้งของการแจกแจงเอฟทางทางด้านขวาที่ระยะ f_{คำนวณ}

$$= \int_{f_{คำนวณ}}^{\infty} \frac{\Gamma(\frac{v_1+v_2}{2}) (\frac{v_1}{v_2})^{\frac{v_1}{2}} f^{\frac{v_1}{2}-1}}{\Gamma(\frac{v_1}{2}) \Gamma(\frac{v_2}{2}) (1 + \frac{v_1}{v_2} f)^{\frac{v_1+v_2}{2}}} df$$



ขั้นตอนการวิเคราะห์ความแปรปรวนแบบจำแนกทางเดียว

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$
 กำหนดสมมติฐานอื่น $H_1 : \mu_1 \neq \mu_2 \neq \mu_3 \neq \dots \neq \mu_k$ (ค่าเฉลี่ยอย่างน้อย 2 ชุดต่างกัน)

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. ทำการสุ่มตัวอย่าง

ขั้นที่ 4. เลือกค่าสถิติ F

ขั้นที่ 5. คำนวณค่าสถิติ f_{คำนวณ} โดยการสร้างตาราง ANOVA

ขั้นที่ 6. เปิดตารางหาค่าวิกฤต f_α ระดับชั้นความเสรี $v_1 = k - 1, v_2 = N - k$ บริเวณวิกฤตคือ $F > f_\alpha$

ขั้นที่ 7. สรุปผล แบบที่ 1. ถ้า $f_{คำนวณ} > f_\alpha$ แล้วปฏิเสธ H_0

แบบที่ 2 ถ้า $Sig < \alpha$ แล้วปฏิเสธ H_0

การทดสอบเพื่อหาประชากรที่มีค่าเฉลี่ยไม่เท่ากัน (MULTIPLE COMPARISONS)

ในกรณีปฏิเสธ H_0 เราสามารถหาค่าเฉลี่ยประชากรคู่ที่ต่างกัน วิธีที่นิยมใช้ในการทดสอบหาคู่ของประชากรที่มีค่าเฉลี่ยเท่ากัน หรือคู่ของประชากรที่มีค่าเฉลี่ยต่างกัน มีหลายวิธี เช่นการทดสอบโดยใช้วิธีของ SCHEFFE' METHOD (S METHOD) ซึ่งมีขั้นตอนการทำงานดังนี้

การทดสอบสมมติฐาน $H_0 : \mu_L - \mu_M = 0$

แย้งกับ $H_1 : \mu_L - \mu_M \neq 0$ ที่ระดับนัยสำคัญ α

การสรุปผล ถ้า $|\bar{x}_L - \bar{x}_M| > \sqrt{\text{MSE}\left(\frac{1}{n_L} + \frac{1}{n_M}\right)} \sqrt{(k-1)f_{\alpha,(v_1,v_2)}}$ แล้วปฏิเสธ H_0

MSE หมายถึง MSE จากตารางวิเคราะห์ความแปรปรวน

n_L หมายถึง ขนาดตัวอย่างชุดที่ L

n_M หมายถึง ขนาดตัวอย่างชุดที่ M

$f_{\alpha,(v_1,v_2)}$ หมายถึงค่าเอฟทำให้ $P(F > f_{\alpha,(v_1,v_2)}) = \alpha$, $v_1 = k - 1$, $v_2 = N - k$

ตัวอย่าง 9.1.1 ข้อมูลการดูความชื้นของคอนกรีต 5 ชนิดเป็นดังนี้

ชนิดที่ 1	ชนิดที่ 2	ชนิดที่ 3	ชนิดที่ 4	ชนิดที่ 5
551	595	639	417	563
457	580	615	449	631
450	508	511	517	522
731	583	573	438	613
499	633	648	415	656
632	517	677	555	679

กำหนด $\mu_1, \mu_2, \mu_3, \mu_4, \mu_5$ เป็นค่าเฉลี่ยประชากรของการดูความชื้นของคอนกรีตชนิดที่ 1, 2, 3, 4, 5 ตามลำดับ จงทดสอบสมมติฐานว่า $\mu_1, \mu_2, \mu_3, \mu_4, \mu_5$ เท่ากัน ที่ระดับนัยสำคัญ 0.05

วิธีทำ

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$

กำหนดสมมติฐานอื่น $H_1 : \mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4 \neq \mu_5$ (ค่าเฉลี่ยอย่างน้อย 2 ชุดต่างกัน)

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

ขั้นที่ 3. ทำการสุ่มตัวอย่าง

ขั้นที่ 4. เลือกค่าสถิติ F

ขั้นที่ 5. คำนวณค่าสถิติ F จากตัวอย่างและสร้างตาราง ANOVA

คำนวณค่า $N = 30$

$$SST = \sum_{j=1}^5 \sum_{i=1}^6 x_{ij}^2 - \frac{T_{..}^2}{N} = 551^2 + 457^2 + \dots + 679^2 - \frac{16854^2}{30} = 209376.800$$

$$SSTR = \sum_{j=1}^5 \frac{T_{.j}^2}{6} - \frac{T_{..}^2}{N} = \frac{3320^2 + 3416^2 + 3663^2 + 2791^2 + 3664^2}{6} - \frac{16854^2}{30} = 85356.4667$$

$$SSE = 209377.800 - 85356.4667 = 124021.333$$

ตาราง ANOVA

แหล่งการแปรผัน SOV	ผลบวกกำลังสอง SS	ระดับชั้นความเสรี DF	ค่าเฉลี่ยของผล บวกกำลังสอง (MS)	f _{คำนวณ}
วิธีการปฏิบัติ (Treatment)	SSTR = 85365	k - 1 = 4	21339	4.30
ความคลาดเคลื่อน (Error)	SSE = 124021	N - k = 25	4961	
ทั้งหมด (Total)	SST = 209377	N - 1 = 29		

ขั้นที่ 6. ค่าวิกฤต $f_{0.05} = 2.76$ ระดับชั้นความเสรี $v_1 = 4$ $v_2 = 25$ บริเวณวิกฤตคือบริเวณ $F > 2.76$

ขั้นที่ 7. สรุปผลเพราะว่า $f_{คำนวณ} = 4.30 > 2.76$ เพราะฉะนั้นปฏิเสธ H_0

การวิเคราะห์ความแปรปรวนแบบจำแนกทางเดียวด้วย SPSS for Windows

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$
กำหนดสมมติฐานอื่น $H_1 : \mu_1 \neq \mu_2 \neq \mu_3 \neq \dots \neq \mu_k$ (ค่าเฉลี่ยอย่างน้อย 2 ชุดต่างกัน)

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. ทำการสุ่มตัวอย่าง

ขั้นที่ 4. เลือกค่าสถิติ F

ขั้นที่ 5. คำนวณค่าสถิติ F จากตัวอย่าง (สร้างตาราง ANOVA)

5.1 คำนวณค่า $f_{คำนวณ}$ ไปใช้ในการสรุปผล ตามวิธีทางทฤษฎีข้างต้น

5.2 คำนวณค่า Sig ของค่าสถิติ F ที่คำนวณได้ไปใช้ในการสรุปผล

ขั้นที่ 6. 6.1 หากค่าวิกฤตคือ f_α ระดับชั้นความเสรี $v_1 = k - 1$, $v_2 = N - k$ บริเวณวิกฤตคือ $F > f_\alpha$

6.2 ใช้ค่า Sig ในการสรุปผล

ขั้นที่ 7. สรุปผล แบบที่ 1 ถ้า $f_{คำนวณ} > f_\alpha$ แล้วปฏิเสธ H_0

แบบที่ 2 ถ้า $Sig < \alpha$ แล้วปฏิเสธ H_0

จากข้อมูลตัวอย่าง 9.1.1

การวิเคราะห์ความแปรปรวนแบบจำแนกทางเดียวด้วย SPSS for Windows

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$
กำหนดสมมติฐานอื่น $H_1 : \mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4 \neq \mu_5$ (ค่าเฉลี่ยอย่างน้อย 2 ชุดต่างกัน)

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

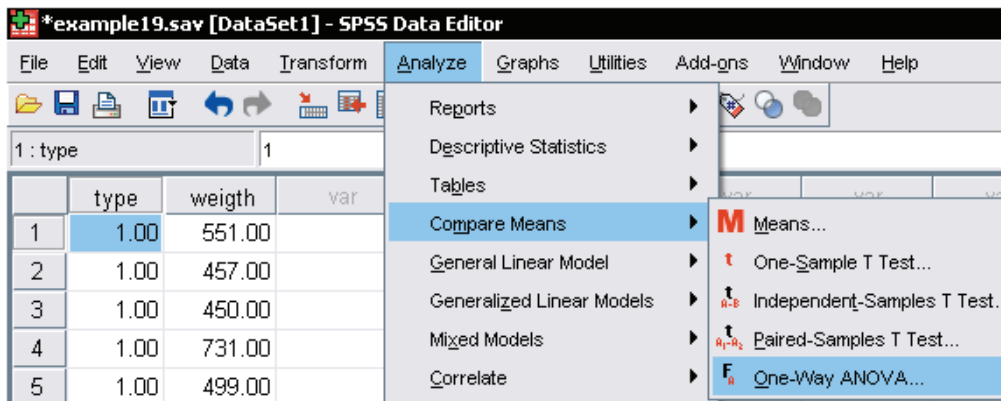
ขั้นที่ 3. ทำการสุ่มตัวอย่างและสร้างแฟ้มข้อมูล

ขั้นที่ 3.1 สร้างแฟ้มข้อมูลโดยกำหนดให้ ตัวแปร type เป็นตัวแปรจำแนกกลุ่ม และ weight เป็นตัวแปรน้ำหนักการวัดความถี่ของคอนกรีต

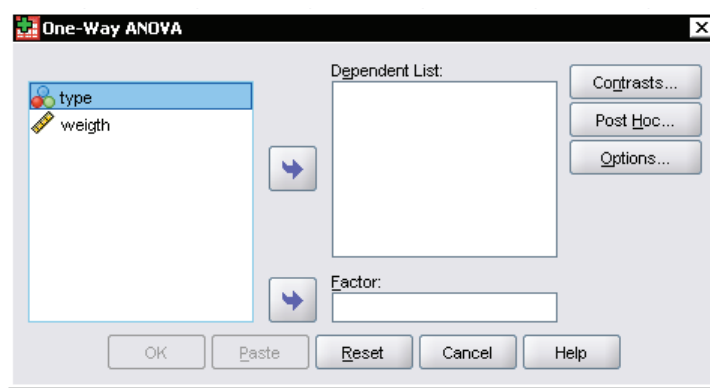
	type	weight
1	1.00	551.00
2	1.00	457.00

บันทึกเป็นแฟ้มข้อมูลชื่อ example19.sav

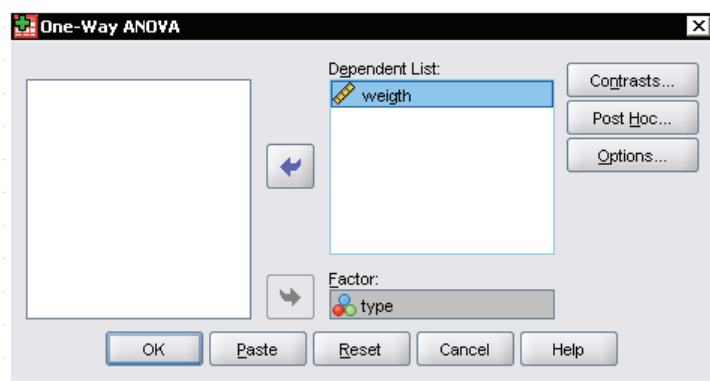
ขั้นที่ 3.2 ใช้คำสั่ง Analyze / Compare Means / One-Way ANOVA



ขั้นที่ 3.3 คลิกที่ One-Way ANOVA จะได้เมนูย่อยเป็น



ขั้นที่ 3.4 เลือกตัวแปร type ไปไว้ที่ช่อง Factor และ เลือกตัวแปร weighth ไปไว้ที่ช่อง Dependent List



ขั้นที่ 3.5 คลิก OK จะได้ผลการคำนวณดังนี้

Oneway
[DataSet1] C:\Documents and Settings\dumrong\My Documents\data\spss16\ex

ANOVA

weighth	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	85356.4667	4	21339.1167	4.3015	0.00875
Within Groups	124020.3333	25	4960.8133		
Total	209376.8000	29			

ผลการคำนวณทั้งหมดคือ

Oneway

[DataSet1] C:\Documents and Settings\dumrong\My Documents\datspss16\example19.sav

ANOVA

weigth					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	85356.4667	4	21339.1167	4.3015	0.00875
Within Groups	124020.3333	25	4960.8133		
Total	209376.8000	29			

- ขั้นที่ 4. สรุปผลโดยเปรียบเทียบค่า Sig กับ α หรือ เปรียบเทียบค่า $f_{\text{คำนวณ}}$ กับค่าวิกฤต
- ขั้นที่ 5. จากตาราง ANOVA จะได้ $f_{\text{คำนวณ}} = 4.3015$ และ $\text{Sig} = 0.00875$
- ขั้นที่ 6. ค่าวิกฤตคือ $f_{0.05} = 2.76$ ระดับชั้นความเสรี $v_1 = 4, v_2 = 25$ บริเวณวิกฤตคือบริเวณ $F > 2.76$
- ขั้นที่ 7. สรุปผล แบบที่ 1 เพราะว่า $f_{\text{คำนวณ}} = 4.3015 > 2.76$ เพราะฉะนั้นปฏิเสธ H_0
หรือ แบบที่ 2 เพราะว่า $\text{Sig} = 0.00875 < 0.05$ เพราะฉะนั้นปฏิเสธ H_0

- หมายเหตุ
1. ในทางปฏิบัติการสรุปผลโดยดูค่า Sig มีความสะดวกมากกว่า
 2. ที่มาของค่า Sig คือ โดยการคำนวณด้วย MATHCAD

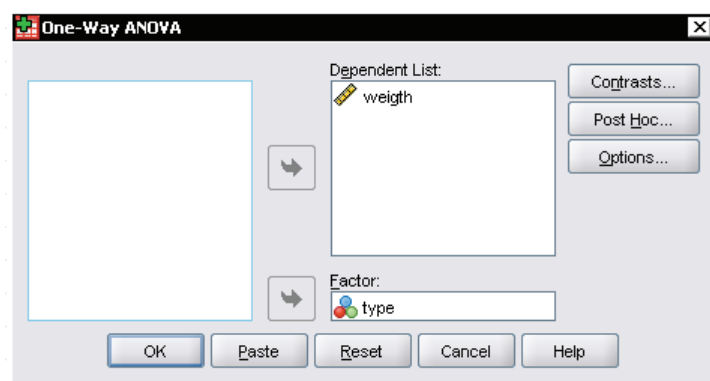
$$v1 := 4 \quad v2 := 25$$

$$h(f) := \frac{\Gamma\left(\frac{v1 + v2}{2}\right) \cdot \left(\frac{v1}{v2}\right)^{\frac{v1}{2}} \cdot f^{\left(\frac{v1}{2}\right) - 1}}{\Gamma\left(\frac{v1}{2}\right) \cdot \Gamma\left(\frac{v2}{2}\right) \cdot \left[1 + \left(\frac{v1}{v2}\right) \cdot f\right]^{\frac{v1 + v2}{2}}} \quad \text{Sig} := 1 - \int_0^{4.302} h(f) df \quad \text{Sig} = 0.008747$$

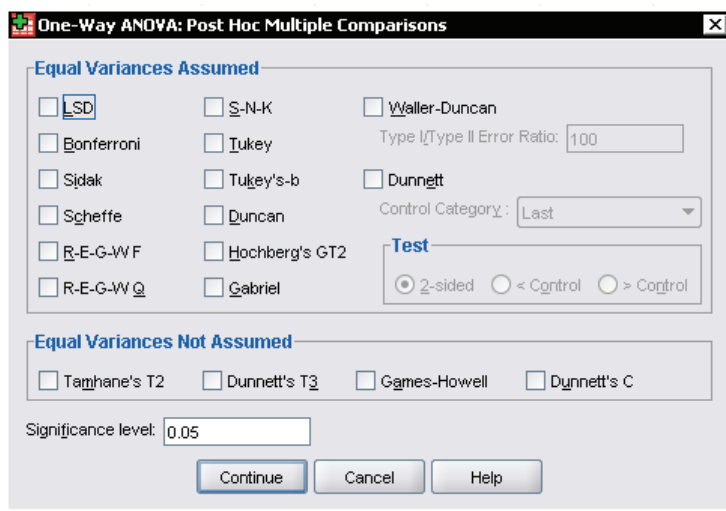
$$\text{Significant} := 1 - pF(4.302, 4, 25) \quad \text{Significant} = 0.008747$$

การทดสอบเพื่อหาประชากรที่มีค่าเฉลี่ยไม่เท่ากัน (MULTIPLE COMPARISONS)

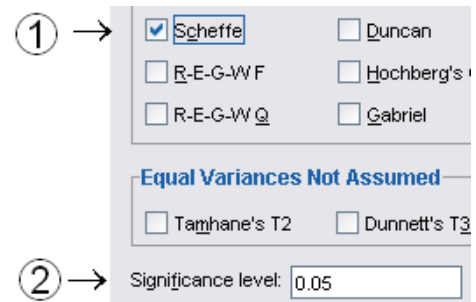
จากขั้นที่ 3.5



ขั้นที่ 3.6 คลิกที่ปุ่ม Post Hoc จะได้เมนูย่อยดังนี้



ขั้นที่ 3.7 1. คลิกเลือกวิธี Scheffe
2. เลือกระดับนัยสำคัญ 0.05



ขั้นที่ 3.8 คลิก Continue และ OK ตามลำดับ จะได้ผลการคำนวณเพิ่มเติมจากเดิมดังนี้

Multiple Comparisons

weigth Scheffe		Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
(I) type	(J) type				Lower Bound	Upper Bound
1	2	-16.00000	40.66454	.997	-151.0824	119.0824
	3	-57.16667	40.66454	.740	-192.2491	77.9158
	4	88.16667	40.66454	.346	-46.9158	223.2491
	5	-57.33333	40.66454	.738	-192.4158	77.7491
2	1	16.00000	40.66454	.997	-119.0824	151.0824
	3	-41.16667	40.66454	.903	-176.2491	93.9158
	4	104.16667	40.66454	.195	-30.9158	239.2491
	5	-41.33333	40.66454	.902	-176.4158	93.7491
3	1	57.16667	40.66454	.740	-77.9158	192.2491
	2	41.16667	40.66454	.903	-93.9158	176.2491
	4	145.33333*	40.66454	.030	10.2509	280.4158
	5	-.16667	40.66454	1.000	-135.2491	134.9158
4	1	-88.16667	40.66454	.346	-223.2491	46.9158
	2	-104.16667	40.66454	.195	-239.2491	30.9158
	3	-145.33333*	40.66454	.030	-280.4158	-10.2509
	5	-145.50000*	40.66454	.030	-280.5824	-10.4176
5	1	57.33333	40.66454	.738	-77.7491	192.4158
	2	41.33333	40.66454	.902	-93.7491	176.4158
	3	.16667	40.66454	1.000	-134.9158	135.2491
	4	145.50000*	40.66454	.030	10.4176	280.5824

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

หมายเหตุ คู่ที่มี * เป็นคู่ประชากรที่มีค่าเฉลี่ยแตกต่างกัน ที่ระดับนัยสำคัญ 0.05
ผลการวิเคราะห์ข้อมูล ค่าเฉลี่ยประชากรคู่ที่ 3 และ 4 แตกต่างกัน ที่ระดับนัยสำคัญ 0.05
ค่าเฉลี่ยประชากรคู่ที่ 4 และ 5 แตกต่างกัน ที่ระดับนัยสำคัญ 0.05

9.2 การวิเคราะห์ความแปรปรวนแบบที่มีการสุ่มอย่างสมบูรณ์ในแต่ละกลุ่ม (Randomized Complete Block Designs, Multiple-Factor ANOVA)

การวิเคราะห์ความแปรปรวนแบบที่มีการสุ่มอย่างสมบูรณ์ในแต่ละกลุ่มหรือแบบจำแนกสองทาง เป็นการวิเคราะห์สำหรับการทดลองแบบ Randomized Block Design ซึ่งแบ่งการทดลองที่มีลักษณะแบบเดียวกัน ออกเป็น กลุ่ม (Block) ในแต่ละกลุ่มมีจำนวนหน่วยของการทดลองเท่ากันเรียกว่า วิธีการปฏิบัติ (Treatment) เพื่อให้ข้อมูลมีการสุ่มอย่างสมบูรณ์ในแต่ละกลุ่ม ตัวอย่างเช่น ในการเปรียบเทียบประสิทธิภาพของเครื่องจักร 4 ชนิด ที่ใช้ผลิตสินค้าอย่างหนึ่งได้จัดเจ้าหน้าที่ 5 คน สลับการควบคุมเครื่องจักร ข้อมูลคือจำนวนวินาทีทั้งหมดที่ใช้ในการผลิตจนสำเร็จ ข้อมูลที่ได้คือ

เจ้าหน้าที่	เครื่องจักร			
	1	2	3	4
1	44	38	47	36
2	46	40	52	43
3	34	36	44	32
4	43	38	46	33
5	38	42	49	39

ลักษณะของข้อมูลจะเห็นได้ว่า ค่าสังเกต 1 ค่ามีผลมาจาก การควบคุมของเจ้าหน้าที่ และ เครื่องจักรต่าง ๆ กัน ในทางสถิติถือว่าเป็นค่าสังเกตที่มีการแปรผันมาจาก 2 ปัจจัย (Factor)

9.2.1 กรณีข้อมูล 1 Treatment และ 1 Block มีค่าสังเกต 1 ค่า ลักษณะข้อมูลโดยทั่วไป

	วิธีการปฏิบัติ (Treatment)						รวม	ค่าเฉลี่ย
	1	2	...	j	...	k		
กลุ่มที่ 1	x_{11}	x_{12}	...	x_{1j}	...	x_{1k}	$T_{1.}$	$\bar{x}_{1.}$
กลุ่มที่ 2	x_{21}	x_{22}	...	x_{2j}	...	x_{2k}	$T_{2.}$	$\bar{x}_{2.}$
กลุ่มที่ 3	x_{31}	x_{32}	...	x_{3j}	...	x_{3k}	$T_{3.}$	$\bar{x}_{3.}$
:	:	:		:		:	:	:
กลุ่มที่ i	x_{i1}	x_{i2}	...	x_{ij}	...	x_{ik}	$T_{i.}$	$\bar{x}_{i.}$
:	:	:		:		:	:	:
กลุ่มที่ b	x_{b1}	x_{b2}	...	x_{bj}	...	x_{bk}	$T_{b.}$	$\bar{x}_{b.}$
รวม	$T_{.1}$	$T_{.2}$...	$T_{.j}$...	$T_{.k}$	$T_{..}$	
ค่าเฉลี่ย	$\bar{x}_{.1}$	$\bar{x}_{.2}$...	$\bar{x}_{.j}$...	$\bar{x}_{.k}$		$\bar{x}_{..}$

N คือจำนวนค่าสังเกตทั้งหมด

x_{ij} คือค่าสังเกตของวิธีการปฏิบัติที่ j ในกลุ่มที่ i, $j = 1, 2, \dots, k$ และ $i = 1, 2, \dots, b$

$\bar{x}_{.j}$ = ค่าเฉลี่ยของวิธีการปฏิบัติที่ j (j^{th} treatment mean)

$\bar{x}_{..}$ = ค่าเฉลี่ยทั้งหมด (grand mean หรือ overall mean)

\bar{x}_i = ค่าเฉลี่ยของกลุ่มที่ i (i^{th} block mean)

$T_{.j}$ = ผลรวมของค่าสังเกตที่ได้จากวิธีการปฏิบัติที่ j (j^{th} treatment total)

T_i = ผลรวมของค่าสังเกตที่ได้จากกลุ่มที่ i (i^{th} block total)

$T_{..}$ = ผลรวมทั้งหมด (grand total หรือ overall total)

การแปรผันของข้อมูลจำแนกเป็น 4 ส่วนคือ

1. การแปรผันโดยรวมทั้งหมด (SST : TOTAL SUM OF SQUARE, Total Sum Square)

$$SST = \sum_{j=1}^k \sum_{i=1}^b (x_{ij} - \bar{x}_{..})^2 = \sum_{j=1}^k \sum_{i=1}^b x_{ij}^2 - \frac{T_{..}^2}{N}$$

2. การแปรผันที่เกิดจาก Treatment ต่างกัน (SSTR : TREATMENT SUM OF SQUARE, Between-Treatment Sum Square)

$$SSTR = b \sum_{j=1}^k (\bar{x}_{.j} - \bar{x}_{..})^2 = \frac{\sum_{j=1}^k T_{.j}^2}{b} - \frac{T_{..}^2}{N}$$

3. การแปรผันที่เกิดจาก Block ต่างกัน (SSBL : BLOCK SUM OF SQUARE, Between-Groups Sum Square)

$$SSBL = k \sum_{i=1}^b (\bar{x}_i - \bar{x}_{..})^2 = \frac{\sum_{i=1}^b T_i^2}{k} - \frac{T_{..}^2}{N}$$

4. การแปรผันที่เกิดร่วมกันระหว่าง Treatment และ Block (SSE : ERROR SUM OF SQUARE, Within-Groups Sum Square, Residual Sum Square)

$$SSE = SST - SSTR - SSBL$$

นำค่า SST, SSTR, SSBL, SSE มาสร้างเป็นตารางวิเคราะห์ความแปรปรวน

ตาราง ANOVA ในตำราสถิติ

แหล่งการแปรผัน	ผลบวก กำลังสอง	ระดับชั้น ความเสรี	ค่าเฉลี่ยของผล บวกกำลังสอง	f คำนวณ
วิธีการปฏิบัติ (Treatment)	SSTR	$k - 1$	$MSTR = \frac{SSTR}{k - 1}$	$f_{\text{treatment}} = \frac{MSTR}{MSE}$
กลุ่ม (Block)	SSBL	$b - 1$	$MSBL = \frac{SSBL}{b - 1}$	$f_{\text{block}} = \frac{MSBL}{MSE}$
ความคลาดเคลื่อน (Error)	SSE	$(b - 1)(k - 1)$	$MSE = \frac{SSE}{(b - 1)(k - 1)}$	
ทั้งหมด (Total)	SST	$N - 1$		

รูปแบบตาราง ANOVA ที่ได้จาก SPSS for Windows

Source	Sum of Squares (SS)	df	Mean Square $= \frac{SS}{df}$	F $= \frac{MS^*}{MSE}$	Sig.
Corrected Model	SSTR + SSBL	$(b - 1) + (k - 1)$	MS_Corrected	$f_{correct}$	$P(F > f_{correct})$
Intercept	SST - SS(Total)	1	MS_intercept	$f_{intercept}$	$P(F > f_{intercept})$
Block	SSBL	$b - 1$	MSBL	f_{block}	$P(F > f_{block})$
Treatment	SSTR	$k - 1$	MSTR	$f_{treatment}$	$P(F > f_{treatment})$
Error	SSE	$(k - 1)(b - 1)$	MSE		
Total	SS(Total)	N			
Corrected Total	SST	$N - 1$			

การทดสอบเพื่อหาประชากรที่มีค่าเฉลี่ยไม่เท่ากัน (MULTIPLE COMPARISONS)

ภายหลังการสรุปผล หากเราต้องปฏิเสธ H_0 ในการหาค่าเฉลี่ยของวิธีการปฏิบัติ (Treatment) หรือค่าเฉลี่ยของกลุ่ม (Block) ที่ต่างกัน วิธีที่นิยมใช้ในการทดสอบหาคู่ที่มีค่าเฉลี่ยเท่ากันหรือคู่ที่ค่าเฉลี่ยต่างกัน เช่นวิธี Fisher's LSD ของ เซอร์ โรนัลด์ เอ ฟิชเชอ (Sir Ronald A. Fisher ค.ศ. 1890 - 1962)

LSD หมายถึง LEAST SIGNIFICANT DIFFERENCE

การทดสอบหาวิธีการปฏิบัติ (Treatment) ที่มีค่าเฉลี่ยแตกต่างกัน

กำหนดสมมติฐาน $H_0' : \mu_L = \mu_M$

แย้งกับ $H_1' : \mu_L \neq \mu_M$

กำหนดระดับนัยสำคัญ α

การสรุปผล ถ้า $|\bar{x}_{.L} - \bar{x}_{.M}| > t_{\frac{\alpha}{2}} \sqrt{MSE \left(\frac{2}{b}\right)}$ แล้ว ปฏิเสธ H_0'

การทดสอบหากลุ่ม (Block) ที่มีค่าเฉลี่ยแตกต่างกัน

กำหนดสมมติฐาน $H_0'' : \mu_L = \mu_M$

แย้งกับ $H_1'' : \mu_L \neq \mu_M$

กำหนดระดับนัยสำคัญ α

การสรุปผล ถ้า $|\bar{x}_{L.} - \bar{x}_{M.}| > t_{\frac{\alpha}{2}} \sqrt{MSE \left(\frac{2}{k}\right)}$ แล้ว ปฏิเสธ H_0''

โดยที่ MSE หมายถึง $\frac{SSE}{(b-1)(k-1)}$ จาก ANOVA TABLE , b หมายถึง จำนวน Blocks

k หมายถึงจำนวน Treatments และ $t_{\frac{\alpha}{2}}$ หมายถึง ค่า t อ่านจากตาราง t ระดับชั้นความเสรี $(b - 1)(k - 1)$

ขั้นตอนการทดสอบสมมติฐานด้วยการวิเคราะห์ความแปรปรวนแบบจำแนกสองทาง

ขั้นที่ 1. กำหนดสมมติฐานเกี่ยวกับวิธีการปฏิบัติ (Treatment)

$H_0 : \mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$

$H_1 : \mu_1 \neq \mu_2 \neq \mu_3 \neq \dots \neq \mu_k$ (ค่าเฉลี่ยอย่างน้อย 2 ชุดต่างกัน)

สมมติฐานเกี่ยวกับความแตกต่างระหว่างกลุ่ม (Block)

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \dots = \mu_b.$$

$$H_1 : \mu_1 \neq \mu_2 \neq \mu_3 \neq \dots \neq \mu_b. \text{ (ค่าเฉลี่ยอย่างน้อย 2 ชุดต่างกัน)}$$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. ทำการสุ่มตัวอย่าง

ขั้นที่ 4. เลือกค่าสถิติ F

ขั้นที่ 5. คำนวณค่าสถิติ F จากตัวอย่างโดยการสร้างตาราง ANOVA

ขั้นที่ 6. เปิดตารางหาค่าวิกฤต

6.1 ค่าวิกฤตของการสรุปผลเกี่ยวกับ Treatment คือ f_α

โดยมีค่าระดับชั้นความเสรี $v_1 = k - 1$, $v_2 = (b - 1)(k - 1)$ บริเวณวิกฤตคือ $F > f_\alpha$

6.2 ค่าวิกฤตของการสรุปผลเกี่ยวกับ Block คือ f_α

โดยมีค่าระดับชั้นความเสรี $v_1 = b - 1$, $v_2 = (b - 1)(k - 1)$ บริเวณวิกฤตคือ $F > f_\alpha$

ขั้นที่ 7. สรุปผล

7.1 การสรุปผลเกี่ยวกับ Treatment 7.1.1 ถ้า $f_{\text{treatment}} > f_\alpha$ ของ Treatment แล้ว ปฏิเสธ H_0

7.1.2 ถ้า Sig ของค่าสถิติ $f_{\text{treatment}} < \alpha$ แล้ว ปฏิเสธ H_0

7.2 การสรุปผลเกี่ยวกับ Block

7.2.1 ถ้า $f_{\text{block}} > f_\alpha$ ของ Block แล้ว ปฏิเสธ H_0

7.2.2 ถ้า Sig ของค่าสถิติ $f_{\text{block}} < \alpha$ แล้ว ปฏิเสธ H_0

ตัวอย่าง 9.2.1 ในการเปรียบเทียบประสิทธิภาพของเครื่องจักร 4 ชนิด และความสามารถของคนที่คุณเครื่องจักร 5 คน ข้อมูลจากการสุ่มตัวอย่างคือ

	เครื่องจักร 1	เครื่องจักร 2	เครื่องจักร 3	เครื่องจักร 4
เจ้าหน้าที่คนที่ 1	44	38	47	36
เจ้าหน้าที่คนที่ 2	46	40	52	43
เจ้าหน้าที่คนที่ 3	34	36	44	32
เจ้าหน้าที่คนที่ 4	43	38	46	33
เจ้าหน้าที่คนที่ 5	38	42	49	39

จงทดสอบสมมติฐานที่ระดับนัยสำคัญ 0.05 ว่า เครื่องจักร 4 เครื่องมีอัตราเร็วเท่ากัน

จงทดสอบสมมติฐานที่ระดับนัยสำคัญ 0.05 ว่า เจ้าหน้าที่ 5 คน ปฏิบัติการด้วยอัตราเร็วเท่ากัน

วิธีทำ ขั้นที่ 1. สมมติฐานเกี่ยวกับวิธีการปฏิบัติ (เครื่องจักร Treatment)

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$$

$$H_1 : \mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4 \text{ (ค่าเฉลี่ยอย่างน้อย 1 คู่แตกต่างกัน)}$$

สมมติฐานเกี่ยวกับความแตกต่างระหว่างกลุ่ม(เจ้าหน้าที่ Block)

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5.$$

$$H_1 : \mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4 \neq \mu_5. \text{ (ค่าเฉลี่ยอย่างน้อย 1 คู่แตกต่างกัน)}$$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

ขั้นที่ 3. นำข้อมูลเข้าสู่ SPSS Data Editor

ขั้นที่ 3.1 การสร้างแฟ้มข้อมูล

กำหนดตัวแปร man เป็นตัวแปรจำแนกคน

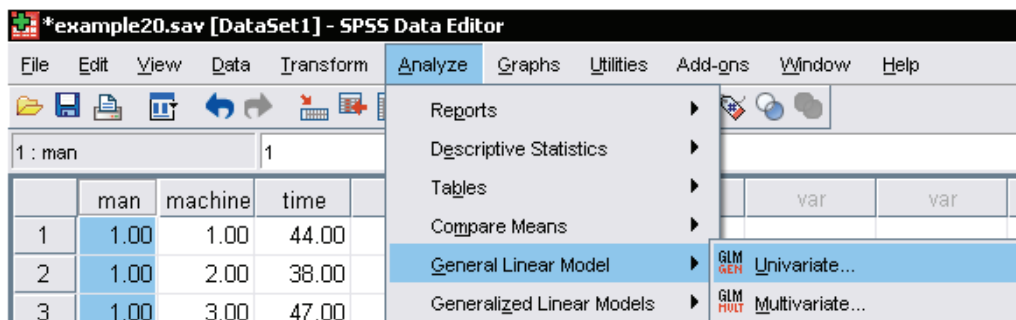
ตัวแปร machine เป็นตัวแปรจำแนกเครื่องจักร

ตัวแปร time เป็นตัวแปรเก็บข้อมูลที่ต้องการวิเคราะห์

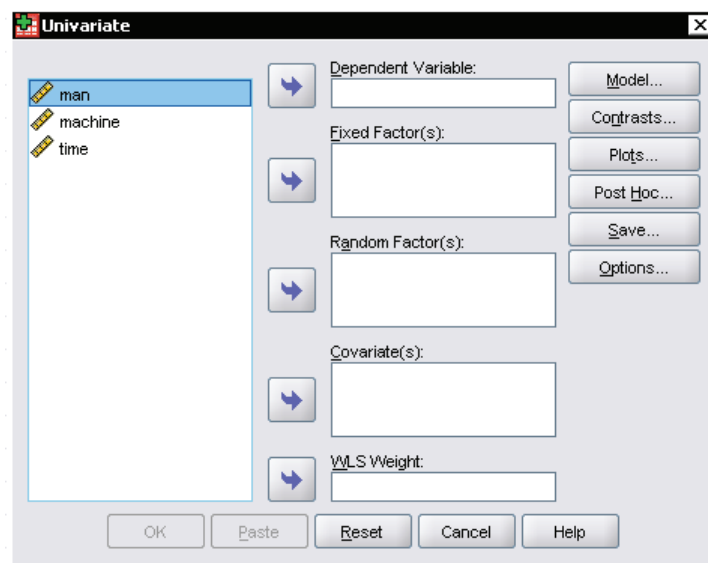
เสร็จแล้ว Save ลงแฟ้มข้อมูลชื่อ example20.sav

ขั้นที่ 3.2 เลือกใช้คำสั่ง Analyze / General Linear Model / Univariate

	man	machine	time
1	1.00	1.00	44.00
2	1.00	2.00	38.00
3	1.00	3.00	47.00
4	1.00	4.00	36.00
5	2.00	1.00	46.00
6	2.00	2.00	40.00



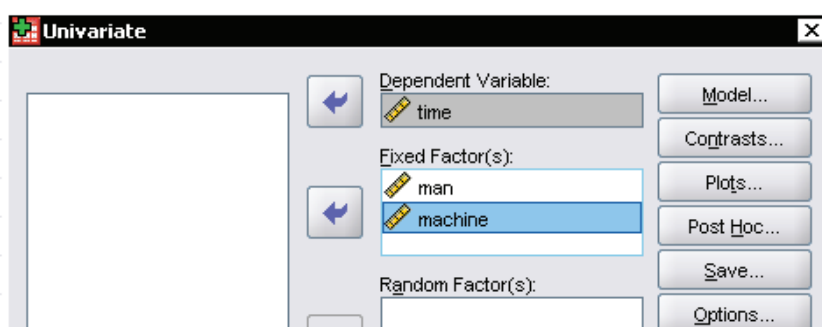
ขั้นที่ 3.3 เลือกคำสั่ง Univariate จะได้เมนูย่อยดังนี้



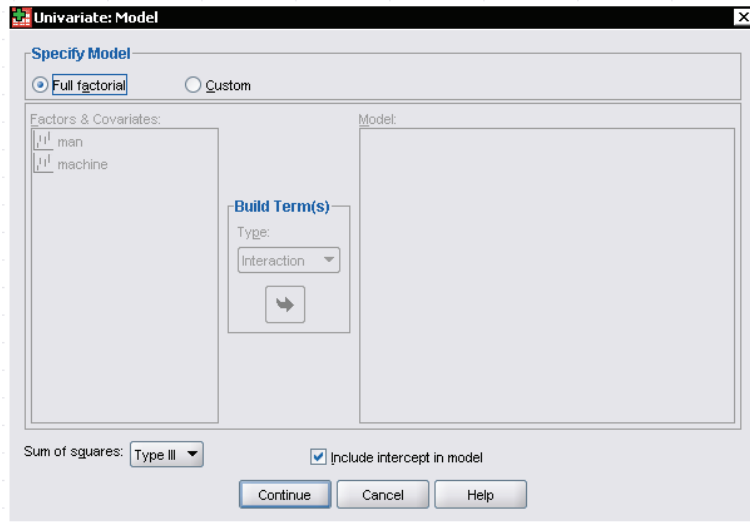
ขั้นที่ 3.4 เลือกตัวแปร time ไปที่ช่อง Dependent Variable

เลือกตัวแปร machine ไปที่ช่อง Fixed Factor(s)

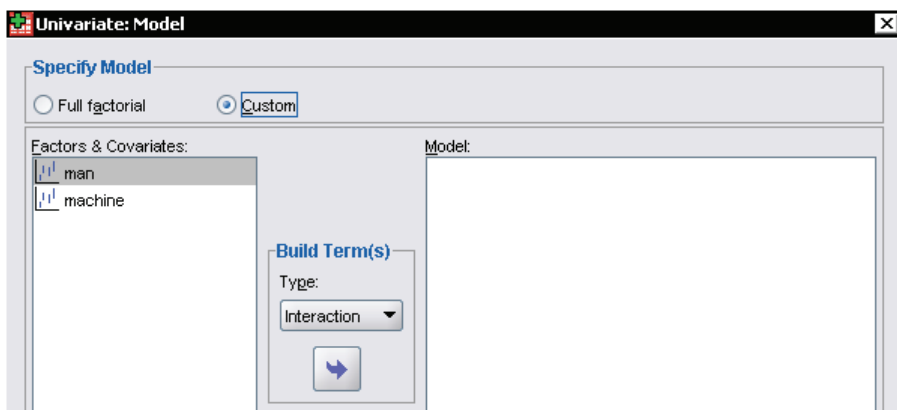
เลือกตัวแปร man ไปที่ช่อง Fixed Factor(s)



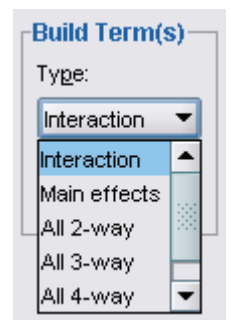
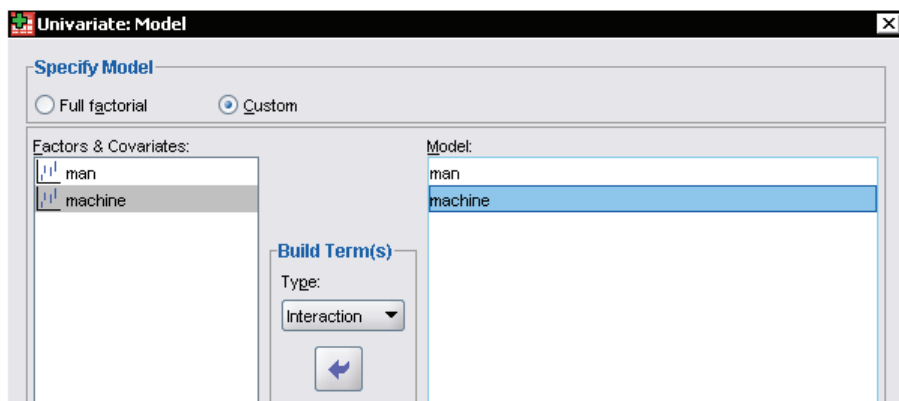
ขั้นที่ 3.5 คลิกที่ปุ่ม Model จะได้เมนูย่อย



ขั้นที่ 3.6 ตรงตำแหน่ง Specify Model ให้เลือก • Custom

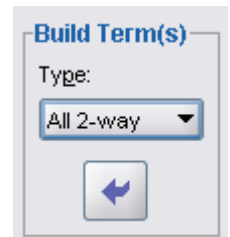


ขั้นที่ 3.7 เลือกตัวแปร machine จากช่อง Factor & Covariates แล้วคลิก  เพื่อย้ายมาไว้ที่ช่อง Model
เลือกตัวแปร man จากช่อง Factor & Covariates แล้วคลิก  เพื่อย้ายมาไว้ที่ช่อง Model

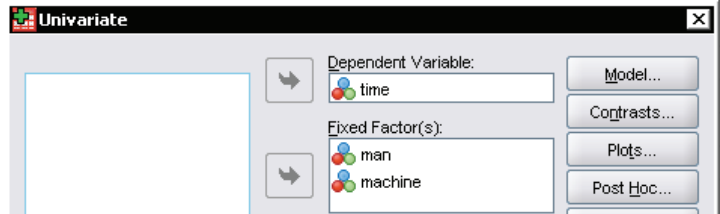


ขั้นที่ 3.8 ในกรอบของ Build Term(s) คลิกที่ช่อง Interaction จะได้ แถบเมนูย่อย

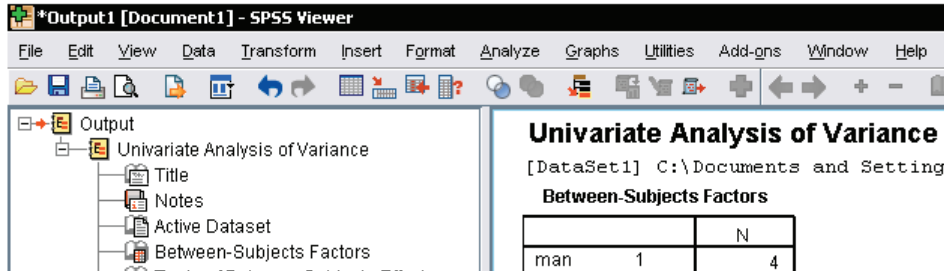
ขั้นที่ 3.9 ให้เลือก All 2-way



ขั้นที่ 3.10 คลิกที่ Continue
จะกลับไปเมนูย่อย Univariate



ขั้นที่ 3.11 คลิก OK จะได้ผลการคำนวณดังนี้



ผลการคำนวณทั้งหมดคือ

Univariate Analysis of Variance

[DataSet1] C:\Documents and Settings\dumrong\My Documents\datspss16\example20.sav

Between-Subjects Factors

		N
man	1	4
	2	4
	3	4
	4	4
	5	4
machine	1	5
	2	5
	3	5
	4	5

Tests of Between-Subjects Effects

Dependent Variable: time

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	500.3 ^a	7	71.471	11.637	0.0001669961
Intercept	33620.0	1	33620.000	5474.084	0.0000000000
man	161.5	4	40.375	6.574	0.0048466575
machine	338.8	3	112.933	18.388	0.0000877779
Error	73.7	12	6.142		
Total	34194.0	20			
Corrected Total	574.0	19			

a. R Squared = .872 (Adjusted R Squared = .797)

จากผลการคำนวณของ SPSS จะได้ว่า

$$SSTR(\text{machine}) = 338.8$$

$$MSTR(\text{machine}) = 112.933$$

$$SSBL(\text{man}) = 161.5$$

$$MSBL(\text{man}) = 40.375$$

$$SSE = 73.7$$

$$MSE = 6.14$$

$$SST = 574.0$$

ขั้นที่ 4. เลือกค่าสถิติ F

ขั้นที่ 5. คำนวณค่าสถิติ F จากตัวอย่าง

ตาราง ANOVA

แหล่งการแปรผัน	ผลบวก กำลังสอง	ระดับชั้น ความเสรี	ค่าเฉลี่ยของผล บวกกำลังสอง	f _{คำนวณ}
วิธีการปฏิบัติ(Treatment) (machine)	338.800	3	112.933	$f_{\text{treatment}} = 18.39$
กลุ่ม(Block) (man)	161.500	4	40.38	$f_{\text{block}} = 6.58$
ความคลาดเคลื่อน (Error)	73.700	12	6.14	
ทั้งหมด (Total)	574.000	19		

ขั้นที่ 6. เปิดตารางหาค่าวิกฤต

ค่าวิกฤตของการสรุปผลเกี่ยวกับ machine มีค่าเป็น $f_{0.05,(3,12)} = 3.49$

ค่าวิกฤตของการสรุปผลเกี่ยวกับ man มีค่าเป็น $f_{0.05,(4,12)} = 3.26$

ขั้นที่ 7. สรุปผล

การสรุปผลเกี่ยวกับ machine (Treatment)

1. เพราะว่า F คำนวณของ machine = 18.39 > 3.49 เพราะฉะนั้นปฏิเสธ H_0
- หรือ 2. เพราะว่า Sig = 0.0000877779 < 0.05 เพราะฉะนั้นปฏิเสธ H_0

การสรุปผลเกี่ยวกับ man (Block)

1. เพราะว่า F คำนวณของ man = 6.58 > 3.26 เพราะฉะนั้นปฏิเสธ H_0
- หรือ 2. เพราะว่า Sig = 0.0048466575 < 0.05 เพราะฉะนั้นปฏิเสธ H_0

การทดสอบหาเครื่องจักรที่มีประสิทธิภาพต่างกัน (เครื่องจักรที่ 1 และ เครื่องจักรที่ 2)

กำหนดสมมติฐาน $H_0 : \mu_1 = \mu_2$

แย้งกับ $H_1 : \mu_1 \neq \mu_2$

กำหนดระดับนัยสำคัญ $\alpha = 0.05$

$\bar{x}_{.1} = 41.0$, $\bar{x}_{.2} = 38.8$, $|\bar{x}_{.1} - \bar{x}_{.2}| = 2.2$, $t_{0.025,12} = 2.179$,

$$\sqrt{\text{MSE}\left(\frac{2}{b}\right)} = \sqrt{6.142\left(\frac{2}{5}\right)} = 1.5674 \text{ และ } t_{\frac{\alpha}{2}} \sqrt{\text{MSE}\left(\frac{2}{b}\right)} = 2.179 \sqrt{6.142\left(\frac{2}{5}\right)} = 3.4153$$

เพราะว่า $|\bar{x}_{.1} - \bar{x}_{.2}| = 2.2 < 3.4153$ เพราะฉะนั้นไม่ปฏิเสธ H_0

เพราะฉะนั้นยอมรับว่าเครื่องจักรที่ 1 และ เครื่องจักรที่ 2 มีความสามารถเท่ากัน

การทดสอบหากกลุ่มที่มีค่าเฉลี่ยแตกต่างกัน (เจ้าหน้าที่คนที่ 1 กับ เจ้าหน้าที่คนที่ 2)

กำหนดสมมติฐาน $H_0 : \mu_1 = \mu_2$.

แย้งกับ $H_1 : \mu_1 \neq \mu_2$.

กำหนดระดับนัยสำคัญ $\alpha = 0.05$

$\bar{x}_{1.} = 41.25$, $\bar{x}_{2.} = 45.25$, $|\bar{x}_{1.} - \bar{x}_{2.}| = 4$, $t_{0.025,12} = 2.179$,

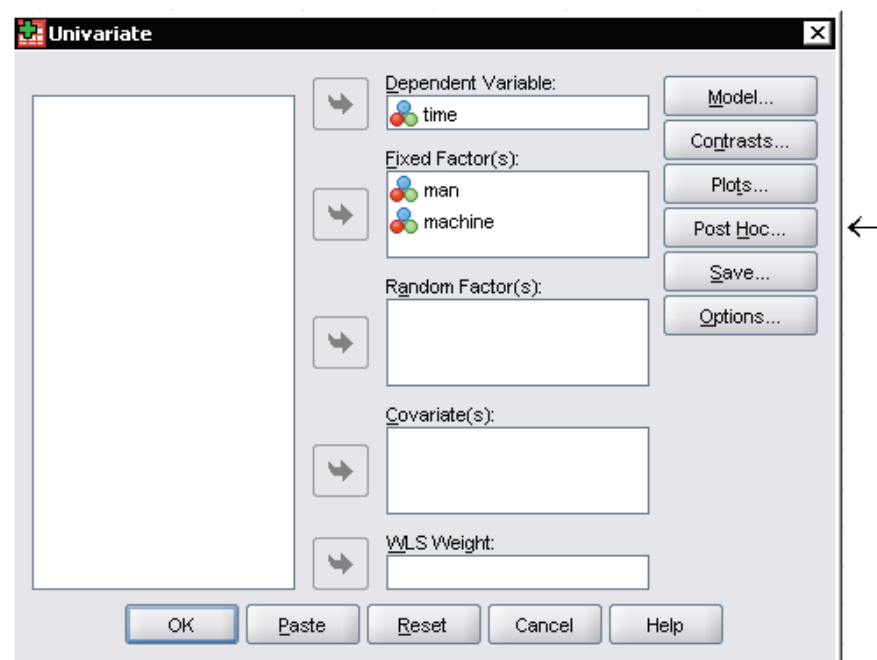
$$\sqrt{\text{MSE}\left(\frac{2}{b}\right)} = \sqrt{6.142\left(\frac{2}{4}\right)} = 1.7524 \text{ และ } t_{\frac{\alpha}{2}} \sqrt{\text{MSE}\left(\frac{2}{k}\right)} = 2.179 \sqrt{6.142\left(\frac{2}{4}\right)} = 3.8184$$

เพราะว่า $|\bar{x}_{1.} - \bar{x}_{2.}| = 4 > 3.8184$ เพราะฉะนั้นปฏิเสธ H_0

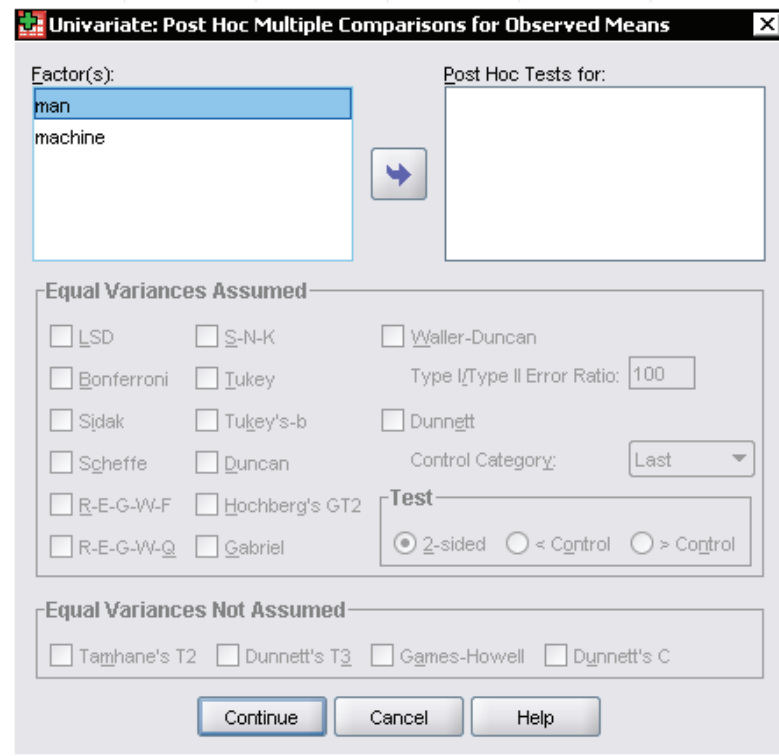
เพราะฉะนั้นเจ้าหน้าที่คนที่ 1 กับ เจ้าหน้าที่คนที่ 2 มีความสามารถแตกต่างกัน

การคำนวณ Multiple Comparisons โดยใช้ SPSS for Windows

จากขั้นที่ 3.10 มีเมนูย่อยเป็น

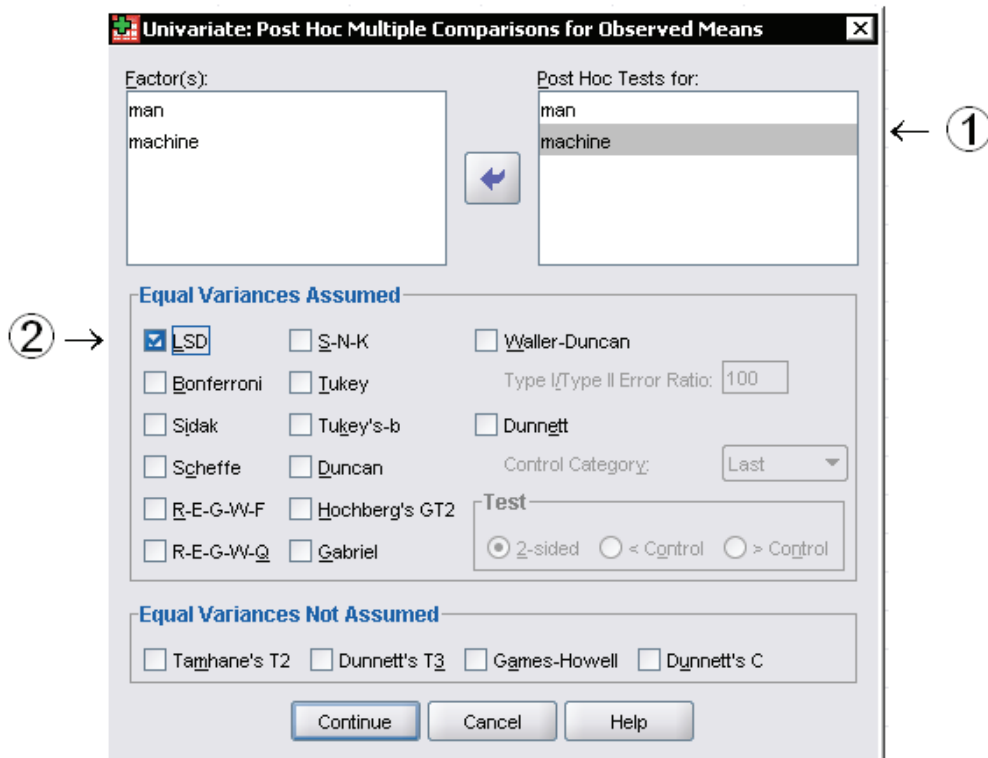


ก่อนที่จะคลิก OK ให้คลิกปุ่ม Post Hoc จะได้เมนูย่อยดังนี้



ขั้นที่ 3.12

1. เลือกตัวแปร man, machine มาไว้ที่ช่อง Post Hoc Tests for
2. เลือกการทดสอบแบบ LSD



ขั้นที่ 3.13 คลิก Continue และ OK ตามลำดับ

จะได้ผลการคำนวณเพิ่มเติมจากเดิมเป็นส่วนของการทดสอบ Multiple comparison ดังนี้

Post Hoc Tests man

Multiple Comparisons

time
LSD

(I) man	(J) man	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1	2	-4.0000 [*]	1.75238	.041	-7.8181	-.1819
	3	4.7500 [*]	1.75238	.019	.9319	8.5681
	4	1.2500	1.75238	.489	-2.5681	5.0681
	5	-.7500	1.75238	.676	-4.5681	3.0681
2	1	4.0000 [*]	1.75238	.041	.1819	7.8181
	3	8.7500 [*]	1.75238	.000	4.9319	12.5681
	4	5.2500 [*]	1.75238	.011	1.4319	9.0681
	5	3.2500	1.75238	.088	-.5681	7.0681
3	1	-4.7500 [*]	1.75238	.019	-8.5681	-.9319
	2	-8.7500 [*]	1.75238	.000	-12.5681	-4.9319
	4	-3.5000	1.75238	.069	-7.3181	.3181
	5	-5.5000 [*]	1.75238	.009	-9.3181	-1.6819
4	1	-1.2500	1.75238	.489	-5.0681	2.5681
	2	-5.2500 [*]	1.75238	.011	-9.0681	-1.4319
	3	3.5000	1.75238	.069	-.3181	7.3181
	5	-2.0000	1.75238	.276	-5.8181	1.8181
5	1	.7500	1.75238	.676	-3.0681	4.5681
	2	-3.2500	1.75238	.088	-7.0681	.5681
	3	5.5000 [*]	1.75238	.009	1.6819	9.3181
	4	2.0000	1.75238	.276	-1.8181	5.8181

Based on observed means.

The error term is Mean Square(Error) = 6.142.

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

หมายเหตุ คู่ที่มีเครื่องหมาย * คือคู่ประชากรที่มีค่าเฉลี่ยแตกต่างกันโดยวิธี LSD โดยมีระดับนัยสำคัญ 0.05

การสรุปผลที่ระดับนัยสำคัญ 0.05 เป็นดังนี้

เจ้าหน้าที่คนที่มีความสามารถแตกต่างกันคือ

- (เจ้าหน้าที่คนที่ 1 , เจ้าหน้าที่คนที่ 2) ,
- (เจ้าหน้าที่คนที่ 1 , เจ้าหน้าที่คนที่ 3) ,
- (เจ้าหน้าที่คนที่ 2 , เจ้าหน้าที่คนที่ 3) ,
- (เจ้าหน้าที่คนที่ 2 , เจ้าหน้าที่คนที่ 4) ,
- (เจ้าหน้าที่คนที่ 3 , เจ้าหน้าที่คนที่ 5)

machine

Multiple Comparisons

time
LSD

(I) machine	(J) machine	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1	2	2.2000	1.56738	.186	-1.2150	5.6150
	3	-6.6000*	1.56738	.001	-10.0150	-3.1850
	4	4.4000*	1.56738	.016	.9850	7.8150
2	1	-2.2000	1.56738	.186	-5.6150	1.2150
	3	-8.8000*	1.56738	.000	-12.2150	-5.3850
	4	2.2000	1.56738	.186	-1.2150	5.6150
3	1	6.6000*	1.56738	.001	3.1850	10.0150
	2	8.8000*	1.56738	.000	5.3850	12.2150
	4	11.0000*	1.56738	.000	7.5850	14.4150
4	1	-4.4000*	1.56738	.016	-7.8150	-.9850
	2	-2.2000	1.56738	.186	-5.6150	1.2150
	3	-11.0000*	1.56738	.000	-14.4150	-7.5850

Based on observed means.

The error term is Mean Square(Error) = 6.142.

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

หมายเหตุ คู่ที่มีเครื่องหมาย * คือคู่ประชากรที่มีค่าเฉลี่ยแตกต่างกันโดยวิธี LSD โดยมีระดับนัยสำคัญ 0.05 การสรุปผลที่ระดับนัยสำคัญ 0.05 เป็นดังนี้

เพราะฉะนั้นเครื่องจักรที่มีความสามารถแตกต่างกันคือ (เครื่องจักรที่ 1 , เครื่องจักรที่ 3) , (เครื่องจักรที่ 1 , เครื่องจักรที่ 4) , (เครื่องจักรที่ 2 , เครื่องจักรที่ 3) , (เครื่องจักรที่ 3 , เครื่องจักรที่ 4)

9.2.2 กรณีข้อมูล 1 Treatment และ 1 Block มีค่าสังเกตมากกว่า 1 ค่า

ในกรณี 1 Treatment 1 Block มีค่าสังเกตมากกว่าหนึ่งค่าเช่นจากตัวอย่าง 9.2.1 หากเราให้เจ้าหน้าที่ทำงานกับเครื่องจักรซ้ำ 5 ครั้งเพื่อตรวจสอบว่าผลของความชำนาญกับเครื่องจักรจะทำให้เวลาเฉลี่ยในการทำงานของเจ้าหน้าที่และเครื่องจักรมีค่าเฉลี่ยเท่ากันหรือไม่ ดังนั้นเราจึงเลือกสมมติฐานเพื่อทำการทดสอบ 3 แบบคือ สมมติฐานเกี่ยวกับวิธีการปฏิบัติ (Treatment)

H_0 : ไม่มีความแตกต่างระหว่างประชากรที่มี Treatment แตกต่างกัน

H_1 : มีความแตกต่างระหว่างประชากรที่มี Treatment แตกต่างกันอย่างน้อยหนึ่งคู่ของ Treatment

สมมติฐานเกี่ยวกับความแตกต่างระหว่างกลุ่ม (Block)

H_0 : ไม่มีความแตกต่างระหว่างประชากรที่มี Block แตกต่างกัน

H_1 : มีความแตกต่างระหว่างประชากรที่มี Block แตกต่างกันอย่างน้อยหนึ่งคู่ของ Block

สมมติฐานเกี่ยวกับ Treatment และ Block

H_0 : ไม่มีความแตกต่างระหว่างประชากรที่มี Treatment และ Block ร่วมกัน

H_1 : มีความแตกต่างระหว่างประชากรที่มี Treatment และ Block ร่วมกันอย่างน้อยหนึ่งคู่

การแปรผันของข้อมูลจำแนกเป็น 5 ส่วนคือ

1. การแปรผันโดยรวมทั้งหมด (SST : TOTAL SUM OF SQUARE, Total Sum Square)

$$SST = \sum_{j=1}^k \sum_{i=1}^b \sum_{s=1}^m x_{ij}^2 - \frac{T_{..}^2}{N}, m = \text{จำนวนค่าสังเกตของแต่ละ treatment และ block}$$

2. การแปรผันที่เกิดจาก Treatment ต่างกัน (SSTR : TREATMENT SUM OF SQUARE, Between-Treatment Sum Square)

$$SSTR = \frac{\sum_{j=1}^k T_{.j}^2}{bm} - \frac{T_{..}^2}{N}$$

3. การแปรผันที่เกิดจาก Block ต่างกัน (SSBL : BLOCK SUM OF SQUARE, Between-Groups Sum Square)

$$SSBL = \frac{\sum_{i=1}^b T_{i.}^2}{km} - \frac{T_{..}^2}{N}$$

4. การแปรผันที่เกิดร่วมกันของ Treatment และ Block (SSTRBL : Sum of Square Interactions TR*BL)

$$SSTRBL = \sum_{j=1}^k \sum_{i=1}^b \frac{\left(\sum_{s=1}^m x_{ij} \right)^2}{m} - \frac{T_{..}^2}{N}$$

5. การแปรผันในส่วนอื่นๆ ของ Treatment และ Block (SSE : ERROR SUM OF SQUARE, Within-Groups Sum Square, Residual Sum Square)

$$SSE = SST - SSTR - SSBL - SSTRBL$$

นำค่า SST, SSTR, SSBL, SSTRBL, SSE มาสร้างเป็นตารางวิเคราะห์ความแปรปรวน

ตาราง ANOVA ในตำราสถิติ

แหล่งการแปรผัน	ผลบวกกำลังสอง	ระดับชั้นความเสรี	ค่าเฉลี่ยของผลบวกกำลังสอง	f _{คำนวณ}
Treatment	SSTR	k - 1	MSTR = $\frac{SSTR}{k-1}$	f _{treatment} = $\frac{MSTR}{MSE}$
Block	SSBL	b - 1	MSBL = $\frac{SSBL}{b-1}$	f _{block} = $\frac{MSBL}{MSE}$
Interactions Treatment and Block	SSTRBL	(k - 1)(b - 1)	MSTRBL = $\frac{SSTRBL}{(k-1)(b-1)}$	f _{interactions} = $\frac{MSTRBL}{MSE}$
Error	SSE	bk(m - 1)	MSE = $\frac{SSE}{kb(m-1)}$	
Total	SST	N - 1		

รูปแบบตาราง ANOVA ที่ได้จาก SPSS for Windows

Source	Sum of Squares	df	Mean Square = $\frac{SS}{df}$	F = $\frac{MS}{MSE}$	Sig.
Corrected Model	SSTR + SSBL + SSTRBL	kb - 1			
Intercept	SS Intercept	1			
TREATMENT	SSTR	k - 1	MSTR	$f_{\text{treatment}}$	$P(F > f_{\text{treatment}})$
BLOCK	SSBL	b - 1	MSBL	f_{block}	$P(F > f_{\text{block}})$
Interaction TR*BL	SSTRBL	$(k - 1)(b - 1)$	MSTRBL	$f_{\text{interactions}}$	$P(F > f_{\text{interactions}})$
Error	SSE	$kb(m - 1)$	MSE		
Total	SS Intercept + SST	N			
Corrected Total	SST	N - 1			

การสรุปผลการทดสอบที่ระดับนัยสำคัญ α

สมมติฐานเกี่ยวกับวิธีการปฏิบัติ (Treatment)

H_0 : ไม่มีความแตกต่างระหว่างประชากรที่มี Treatment แตกต่างกัน

H_1 : มีความแตกต่างระหว่างประชากรที่มี Treatment แตกต่างกันอย่างน้อยหนึ่งคู่ของ Treatment

การสรุปผล แบบที่ 1 ถ้า $\text{Sig} = P(F > f_{\text{treatment}}) < \alpha$ แล้วปฏิเสธ H_0

แบบที่ 2 ถ้า $f_{\text{treatment}} > f_{\alpha, (k-1, kb(m-1))}$ แล้วปฏิเสธ H_0

สมมติฐานเกี่ยวกับความแตกต่างระหว่างกลุ่ม (Block)

H_0 : ไม่มีความแตกต่างระหว่างประชากรที่มี Block แตกต่างกัน

H_1 : มีความแตกต่างระหว่างประชากรที่มี Block แตกต่างกันอย่างน้อยหนึ่งคู่ของ Block

การสรุปผล แบบที่ 1 ถ้า $\text{Sig} = P(F > f_{\text{block}}) < \alpha$ แล้วปฏิเสธ H_0

แบบที่ 2 ถ้า $f_{\text{block}} > f_{\alpha, (b-1, kb(m-1))}$ แล้วปฏิเสธ H_0

สมมติฐานเกี่ยวกับ Treatment และ Block

H_0 : ไม่มีความแตกต่างระหว่างประชากรที่มี Treatment และ Block ร่วมกัน

H_1 : มีความแตกต่างระหว่างประชากรที่มี Treatment และ Block ร่วมกันอย่างน้อยหนึ่งคู่

การสรุปผล แบบที่ 1 ถ้า $\text{Sig} = P(F > f_{\text{interactions}}) < \alpha$ แล้วปฏิเสธ H_0

แบบที่ 2 ถ้า $f_{\text{interactions}} > f_{\alpha, ((k-1)(b-1), kb(m-1))}$ แล้วปฏิเสธ H_0

ตัวอย่าง 9.2.2 ในการเปรียบเทียบประสิทธิภาพของเครื่องจักร 4 ชนิด และความสามารถของคนที่คุมเครื่องจักร 5 คน โดยให้ทำการควบคุมเครื่องจักรเครื่องเดียวซ้ำคนละ 5 ครั้ง ได้ข้อมูลจากการสุ่มตัวอย่างคือ

	เครื่องจักร 1	เครื่องจักร 2	เครื่องจักร 3	เครื่องจักร 4
เจ้าหน้าที่คนที่ 1	44	38	47	36
	35	35	35	36
	35	32	39	36
	32	35	32	35
	37	36	37	36
เจ้าหน้าที่คนที่ 2	46	40	52	43
	35	37	42	33
	39	36	32	32
	39	35	35	35
	38	36	38	31
เจ้าหน้าที่คนที่ 3	34	36	44	32
	34	35	40	36
	35	34	39	36
	34	34	38	35
	33	36	39	34
เจ้าหน้าที่คนที่ 4	43	38	46	33
	35	35	39	35
	38	36	39	35
	38	32	32	37
	32	35	38	37
เจ้าหน้าที่คนที่ 5	38	36	35	39
	39	34	32	35
	39	32	36	36
	40	32	35	36
	35	32	32	32

จงทดสอบสมมติฐานที่ระดับนัยสำคัญ 0.01

1. ไม่มีความแตกต่างของค่าเฉลี่ยระหว่าง เครื่องจักร (Treatment) ที่แตกต่างกัน
2. ไม่มีความแตกต่างของค่าเฉลี่ยระหว่าง เจ้าหน้าที่ (Block) ที่แตกต่างกัน
3. ไม่มีความแตกต่างของค่าเฉลี่ยระหว่าง เครื่องจักร (Treatment) และ เจ้าหน้าที่ (Block) ที่แตกต่างกัน

วิธีทำ

ขั้นที่ 1. สมมติฐานเกี่ยวกับวิธีการปฏิบัติ (เครื่องจักร Treatment)

H_0 : ไม่มีความแตกต่างระหว่างค่าเฉลี่ยของเครื่องจักร (Treatment) ที่แตกต่างกัน

H_1 : มีความแตกต่างระหว่างค่าเฉลี่ยของเครื่องจักรอย่างน้อยหนึ่งคู่

สมมติฐานเกี่ยวกับความแตกต่างระหว่างกลุ่ม (เจ้าหน้าที่ Block)

H_0 : ไม่มีความแตกต่างระหว่างค่าเฉลี่ยของเจ้าหน้าที่ (Block) ที่แตกต่างกัน

H_1 : มีความแตกต่างระหว่างค่าเฉลี่ยของเจ้าหน้าที่อย่างน้อยหนึ่งคู่

สมมติฐานเกี่ยวกับ Treatment และ Block

H_0 : ไม่มีความแตกต่างระหว่างค่าเฉลี่ยของเครื่องจักร (Treatment) และ เจ้าหน้าที่ (Block)

H_1 : มีความแตกต่างระหว่างค่าเฉลี่ยของเครื่องจักร (Treatment) และ เจ้าหน้าที่ (Block) อย่างน้อยหนึ่งคู่

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.01$

ขั้นที่ 3. เข้าสู่ SPSS Data Editor

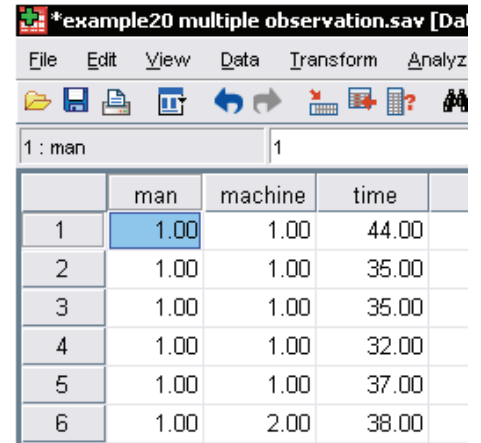
ขั้นที่ 3.1 การสร้างแฟ้มข้อมูล กำหนดให้

ตัวแปร man เป็นตัวแปรจำแนกคน

ตัวแปร machine เป็นตัวแปรจำแนกเครื่องจักร

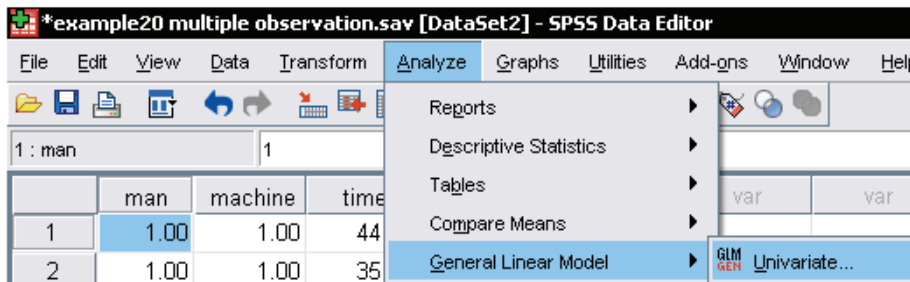
ตัวแปร time เป็นตัวแปรเก็บข้อมูลที่ต้องการวิเคราะห์

เสร็จแล้ว Save ลงแฟ้มข้อมูลชื่อ example20 multiple observation.sav

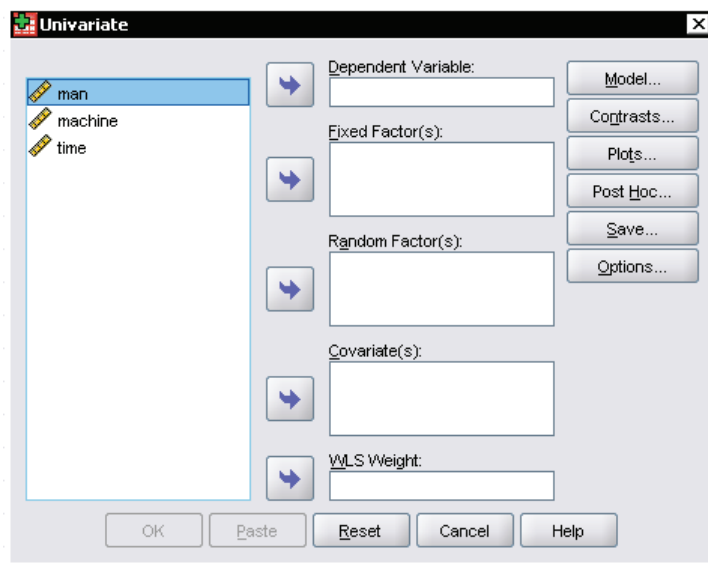


	man	machine	time
1	1.00	1.00	44.00
2	1.00	1.00	35.00
3	1.00	1.00	35.00
4	1.00	1.00	32.00
5	1.00	1.00	37.00
6	1.00	2.00	38.00

ขั้นที่ 3.2 เลือกใช้คำสั่ง Analyze / General Linear Model / Univariate

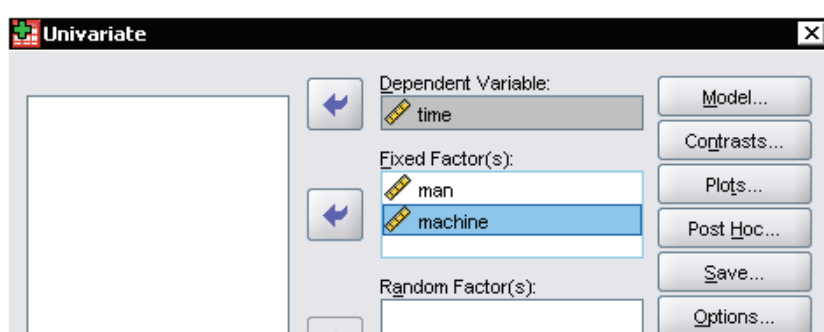


ขั้นที่ 3.3 เลือกคำสั่ง Univariate จะได้เมนูย่อยดังนี้



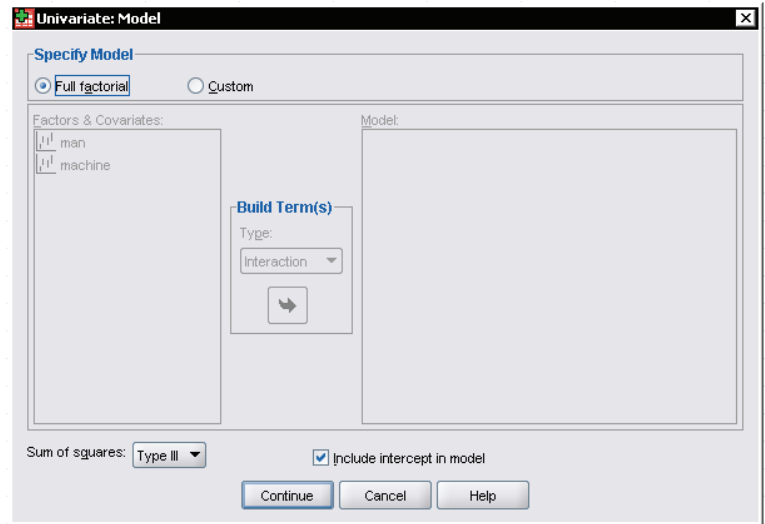
ขั้นที่ 3.4 เลือกตัวแปร time ไปช่อง Dependent Variable

เลือกตัวแปร machine และ man ไปช่อง Fixed Factor(s)



ขั้นที่ 3.5

คลิกที่ปุ่ม Model จะได้เมนูย่อย

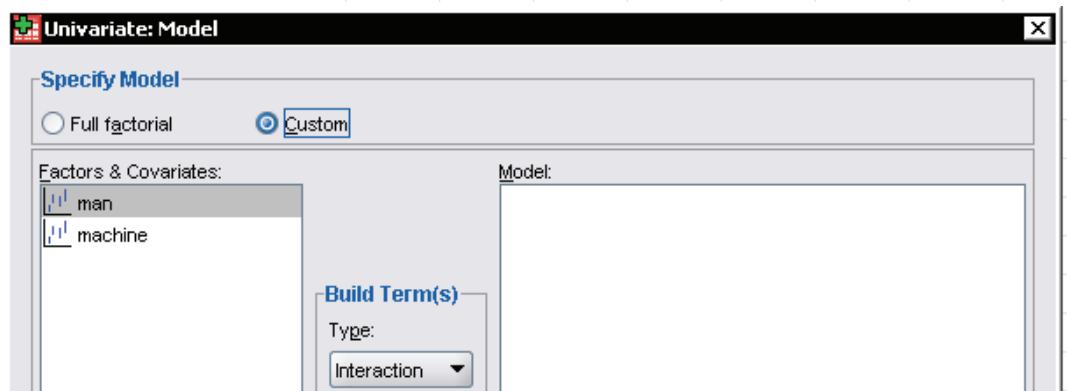


ขั้นที่ 3.6

ตรงตำแหน่ง

Specify Model

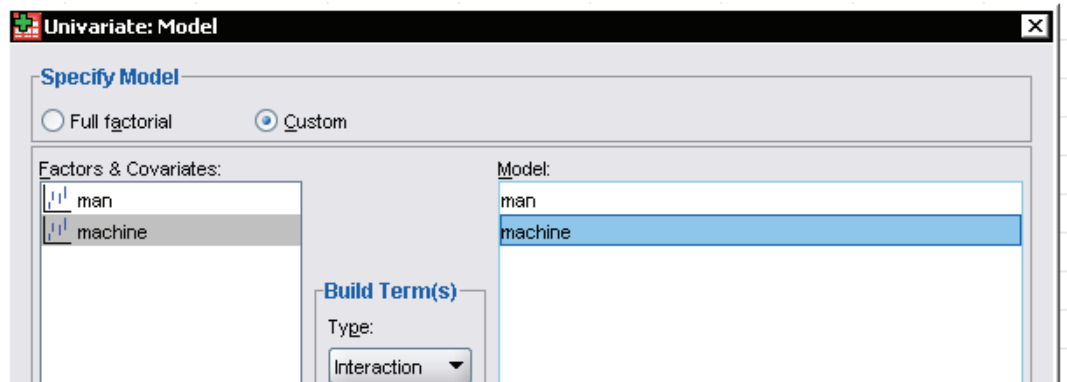
ให้เลือก • Custom



ขั้นที่ 3.7

เลือกตัวแปร machine(F) จากช่อง Factor & Covariates มาไว้ที่ช่อง Model

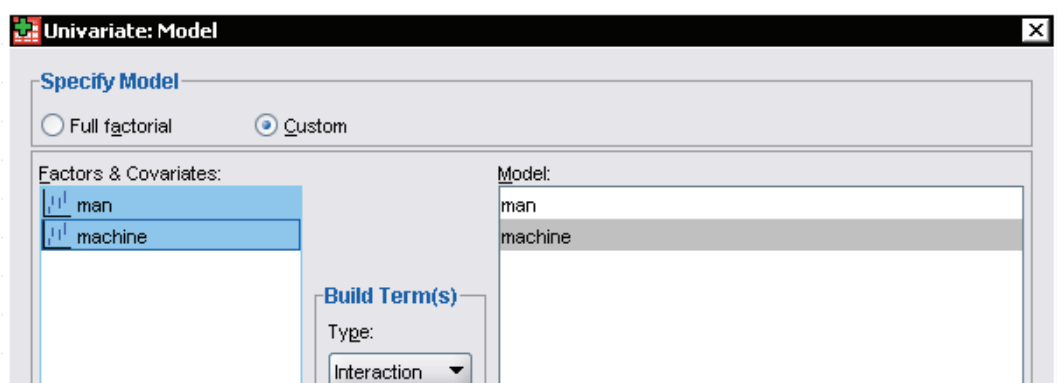
เลือกตัวแปร man(F) จากช่อง Factor & Covariates มาไว้ที่ช่อง Model



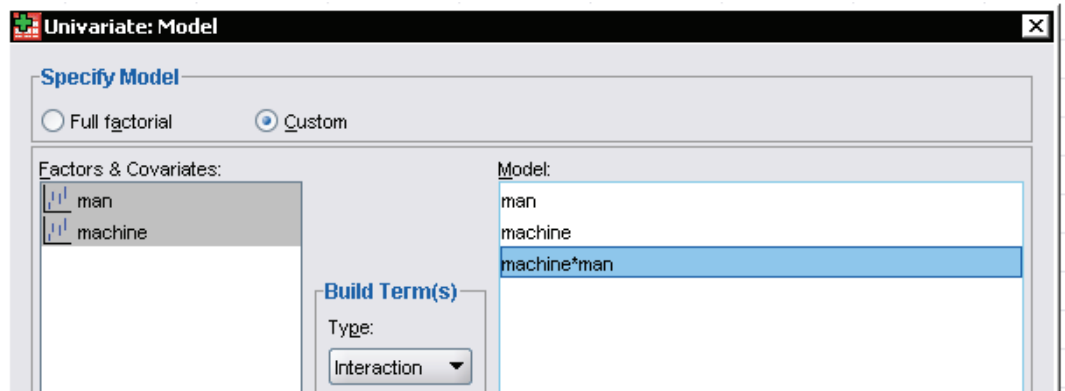
ขั้นที่ 3.8

เลือกตัวแปร man และ machine พร้อมกัน ทำได้ดังนี้

คลิกที่ตัวแปร man และ กด Shift ค้างไว้ แล้วคลิกที่ตัวแปร machine จะได้ผลบนจอภาพเป็น



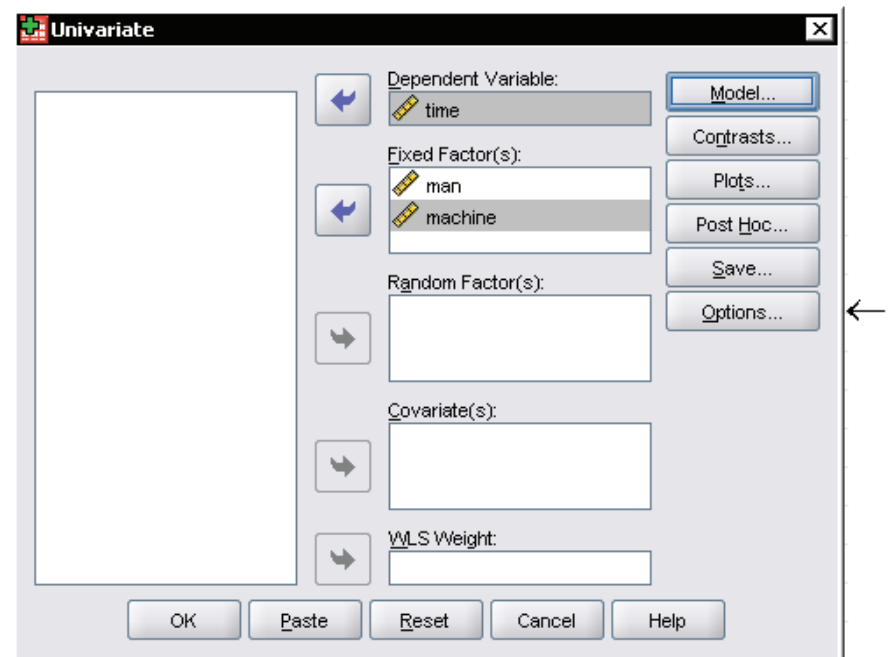
ขั้นที่ 3.9

คลิกที่ จะได้ผลในเมนูย่อย
เป็นดังนี้

ขั้นที่ 3.10

คลิก Continue

จะกลับไปเมนูย่อย Univariate

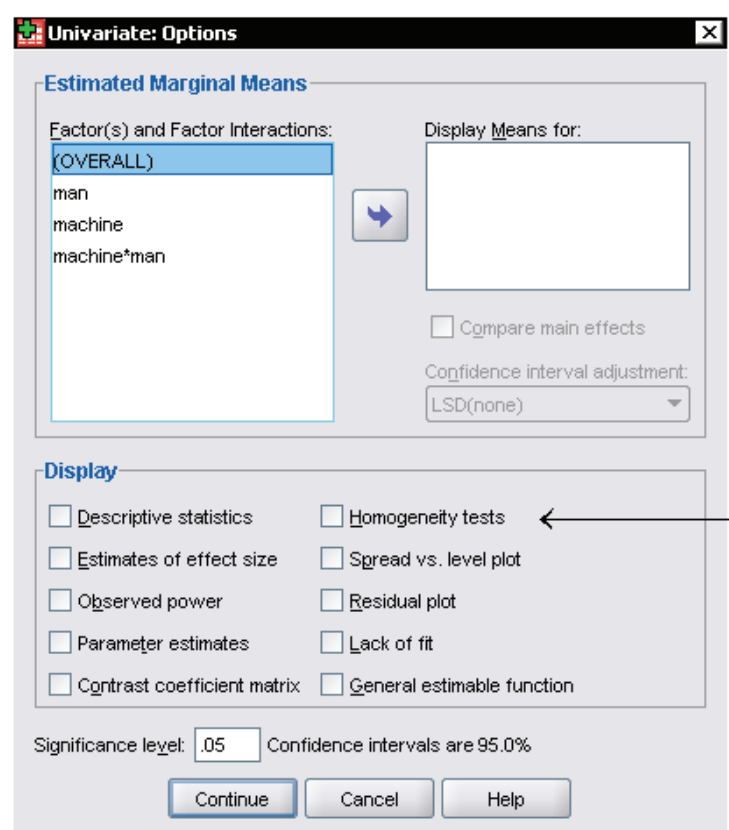


ขั้นที่ 3.11

ถ้าต้องการทดสอบว่าความแปรปรวน
ของข้อมูลเท่ากันจริงหรือไม่

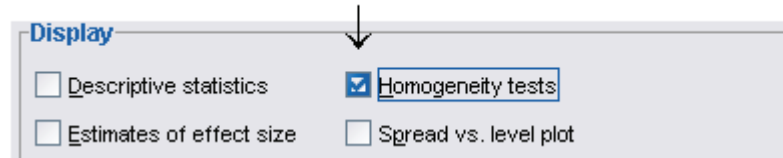
ให้คลิกปุ่ม Options

จะได้เมนูย่อยเป็น



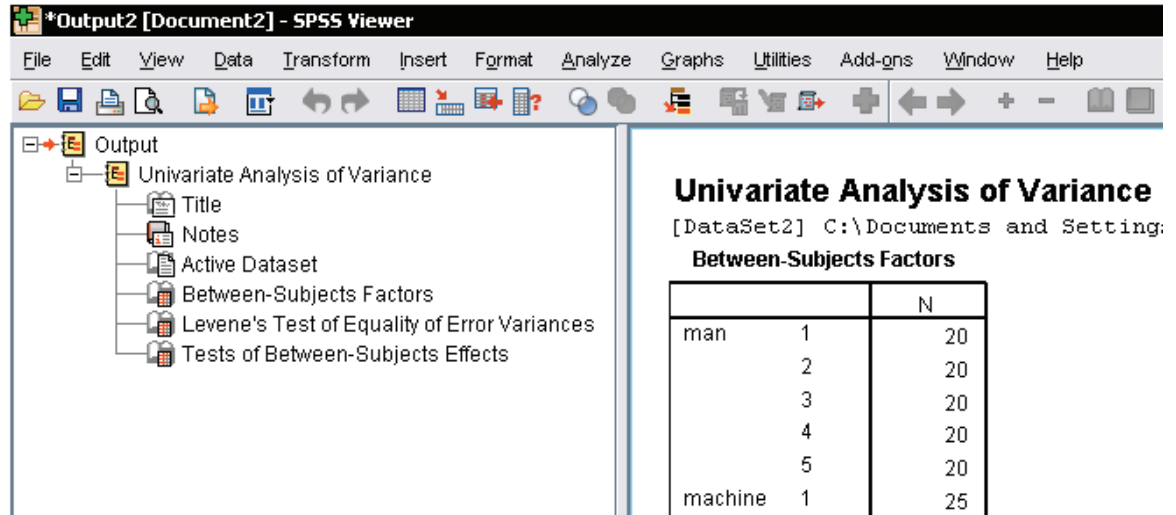
ขั้นที่ 3.12

เลือก Homogeneity tests



ขั้นที่ 3.13

คลิก Continue กลับไปเมนูย่อย Univariate และคลิก OK ตามลำดับจะได้ผลการคำนวณดังนี้



ผลการคำนวณทั้งหมดคือ

Univariate Analysis of Variance

[DataSet2] C:\Documents and Settings\dumrong\My Documents\datasps16\example20 multiple observation.sav

Between-Subjects Factors

		N
man	1	20
	2	20
	3	20
	4	20
	5	20
machine	1	25
	2	25
	3	25
	4	25

Levene's Test of Equality of Error Variances^a

Dependent Variable: time

F	df1	df2	Sig.
1.979700	19	80	0.018697

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + man + machine + man * machine

Tests of Between-Subjects Effects

Dependent Variable: time

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	401.5600 ^a	19	21.1347	1.7911	0.038191
Intercept	132350.4400	1	132350.4400	11216.1390	0.000000
man	66.4600	4	16.6150	1.4081	0.238851
machine	162.6800	3	54.2267	4.5955	0.005096
man * machine	172.4200	12	14.3683	1.2177	0.285890
Error	944.0000	80	11.8000		
Total	133696.0000	100			
Corrected Total	1345.5600	99			

a. R Squared = .298 (Adjusted R Squared = .132)

การสรุปผลเกี่ยวกับความแปรปรวน

จากตาราง Levene's Test of Equality of Error Variance

จะได้ว่า $\text{Sig} = P(F > 1.979700) = 0.018697 > \alpha = 0.01$

เพราะฉะนั้นยอมรับว่าค่าความแปรปรวนของประชากรทุกกลุ่มมีค่าเท่ากันที่ระดับนัยสำคัญ 0.01

การสรุปผลเกี่ยวกับค่าเฉลี่ย

สมมติฐานเกี่ยวกับวิธีการปฏิบัติ (เครื่องจักร, Treatment = ตัวแปร machine)

เพราะว่า $\text{Sig} = P(F > f_{\text{treatment}}) = P(F > 4.5955) = 0.005096 < \alpha = 0.01$ เพราะฉะนั้นปฏิเสธ H_0

เพราะฉะนั้นมีความแตกต่างระหว่างค่าเฉลี่ยของเครื่องจักร(Treatment) อย่างน้อยหนึ่งคู่

สมมติฐานเกี่ยวกับกลุ่ม (เจ้าหน้าที่, Block = ตัวแปร man)

เพราะว่า $\text{Sig} = P(F > f_{\text{block}}) = P(F > 1.4081) = 0.238851 > \alpha = 0.01$ เพราะฉะนั้นไม่ปฏิเสธ H_0

เพราะฉะนั้นไม่มีความแตกต่างระหว่างค่าเฉลี่ยของเจ้าหน้าที่ (Block) ที่แตกต่างกัน

สมมติฐานเกี่ยวกับวิธีการปฏิบัติและกลุ่ม

เพราะว่า $\text{Sig} = P(F > f_{\text{interactions}}) = P(F > 1.2177) = 0.285890 > \alpha = 0.01$

เพราะฉะนั้นไม่ปฏิเสธ H_0

เพราะฉะนั้นไม่มีความแตกต่างระหว่างค่าเฉลี่ยของเครื่องจักร(Treatment) และ เจ้าหน้าที่(Block)

บทที่ 10

การทดสอบสมมติฐาน แบบนอนพาราเมตริก



ในกรณีที่เราไม่ทราบการแจกแจงของประชากรและเราต้องการทดสอบสมมติฐานเกี่ยวกับลักษณะบางอย่างของประชากร เราจะทำการทดสอบสมมติฐาน แบบนอนพาราเมตริก (Nonparametric Test) การทดสอบแบบนอนพาราเมตริกที่สำคัญได้แก่

- การทดสอบว่าตัวอย่างที่เราเลือกมาเป็นไปโดยสุ่มหรือไม่
- การทดสอบว่าประชากรมีการแจกแจงตามที่เราคาดไว้หรือไม่
- การทดสอบว่าประชากร 2 กลุ่มมีความสัมพันธ์กันหรือไม่
- การทดสอบว่าค่าเฉลี่ยของประชากร k กลุ่มตัวอย่างเท่ากันหรือไม่

10.1 การทดสอบว่าตัวอย่างที่เราเลือกมาเป็นไปโดยสุ่มหรือไม่

การทดสอบว่าข้อมูลตัวอย่างที่เราเก็บรวบรวมมาได้มีลักษณะของการเกิดเป็นไปโดยสุ่มหรือไม่ สามารถทำการทดสอบได้โดยใช้วิธี ทดสอบรันส์ (Runs Test)

การทดสอบสมมติฐานโดยใช้ Runs Test ของ SPSS for windows

ตัวอย่าง 10.1.1 ข้อมูลจำนวนคนที่อยู่ในแถวเพื่อรอถอนเงินจากเครื่อง ATM ที่เก็บมาในช่วงเวลา 40 วัน ต่อเนื่องกันเป็นดังนี้

6	7	5	6	8	6	8	6	6	4
3	2	4	4	3	4	7	5	6	8
6	6	3	5	2	5	4	4	3	7
5	5	4	3	7	4	6	5	2	8

จงทดสอบว่าจำนวนคนที่อยู่ในแถวเป็นไปอย่างสุ่ม กำหนดระดับนัยสำคัญ 0.05

วิธีทำ ขั้นที่ 1. กำหนดสมมติฐานหลัก H_0 : จำนวนคนที่อยู่ในแถวเป็นไปอย่างสุ่ม
กำหนดสมมติฐานอื่น H_1 : จำนวนคนที่อยู่ในแถวไม่เป็นไปอย่างสุ่ม

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

ขั้นที่ 3. ทำการสุ่มตัวอย่างและทำการคำนวณ

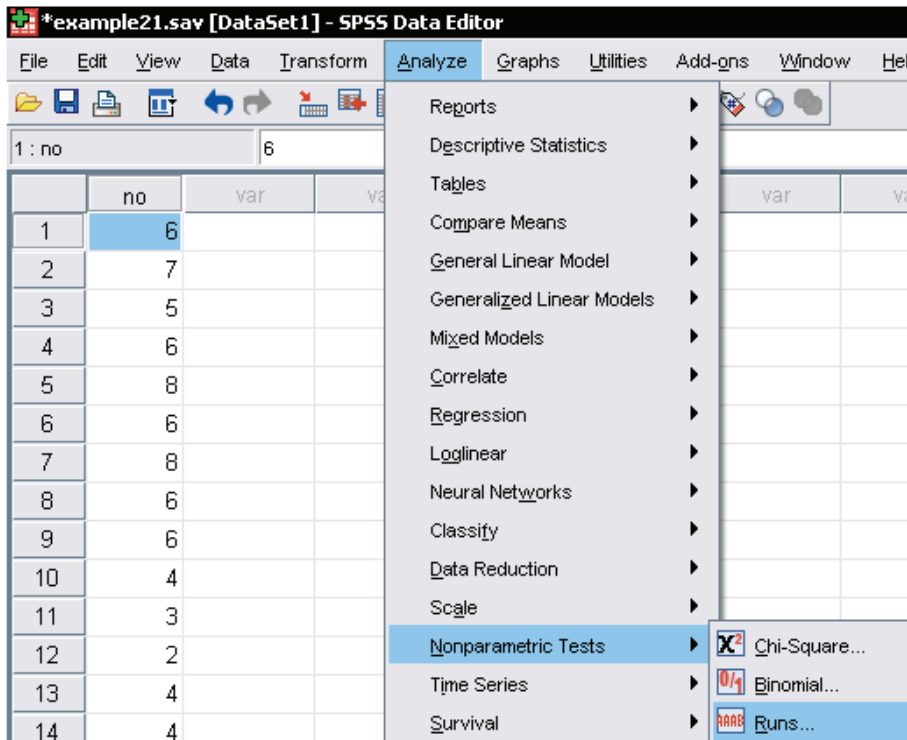
ขั้นที่ 3.1 สร้างแฟ้มข้อมูลประกอบด้วยตัวแปร จำนวนคน (no)

เสร็จแล้ว Save ไว้ที่แฟ้มข้อมูลชื่อ example21.sav

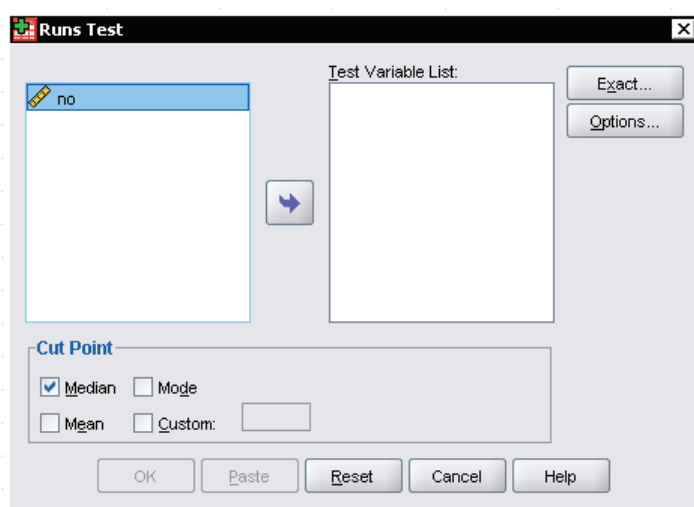
ขั้นที่ 3.2 เลือกคำสั่ง Analyze / Nonparametric Tests / Runs

*example21.sav [DataSet1]

	no	var
1	6	
2	7	
39	2	
40	8	

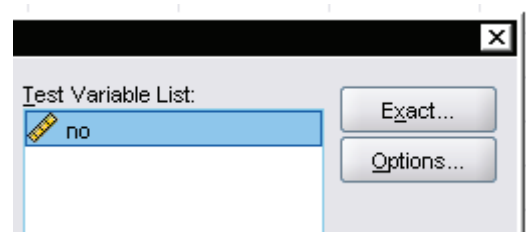


ขั้นที่ 3.3 คลิกที่คำสั่ง Runs จะได้เมนูย่อย

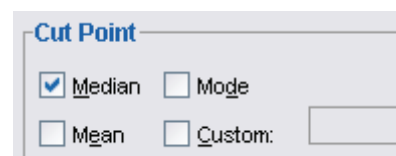


ขั้นที่ 3.4 เลือกตัวแปร no

ไปไว้ที่ช่อง Test Variable List



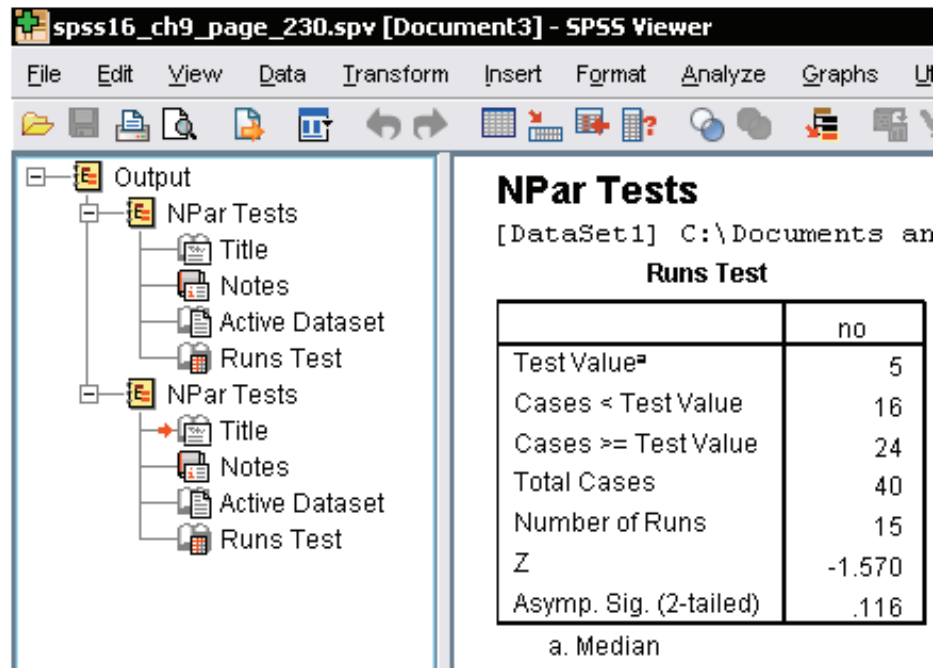
ขั้นที่ 3.5 เลือก Cut Point เป็น Median



หมายเหตุ การทำ Runs Test

สามารถทำการทดสอบโดยการเปรียบเทียบกับค่า Mean Mode หรือค่าอื่น ๆ ที่กำหนดเอง Custom ได้

ขั้นที่ 3.6 คลิก OK
จะได้ผลการคำนวณ
บนจอภาพเป็นดังนี้



ขั้นที่ 4. เลือกค่าสถิติ Z

ขั้นที่ 5. $z_{\text{คำนวณ}} = -1.570$ และ $\text{Asymp. Sig. (2-tailed)} = 0.116$

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤตและบริเวณวิกฤต

กรณีใช้ค่า Z ค่าวิกฤตคือ $-z_{\frac{\alpha}{2}}$ และ $z_{\frac{\alpha}{2}}$ บริเวณวิกฤตคือ $Z < -z_{\frac{\alpha}{2}}$ หรือ $Z > z_{\frac{\alpha}{2}}$

เพราะฉะนั้นค่าวิกฤตคือ -1.96 และ 1.96 บริเวณวิกฤตคือ $Z < -1.96$ หรือ $Z > 1.96$

ขั้นที่ 7. สรุปผล

แบบที่ 1 โดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต

โดยมีเกณฑ์การสรุปผล ถ้า $z_{\text{คำนวณ}} < -z_{\frac{\alpha}{2}}$ หรือ $z_{\text{คำนวณ}} > z_{\frac{\alpha}{2}}$ แล้วปฏิเสธ H_0

เพราะว่า $z_{\text{คำนวณ}} = -1.570$ ไม่อยู่ในบริเวณวิกฤต เพราะฉะนั้นไม่ปฏิเสธ H_0

หรือ แบบที่ 2 โดยการเปรียบเทียบ Sig กับค่า α เกณฑ์การสรุปผล ถ้า $\text{Sig} < \alpha$ แล้วปฏิเสธ H_0

เพราะว่า $\text{Sig} = 0.116$ มากกว่า 0.05 เพราะฉะนั้นไม่ปฏิเสธ H_0

หมายเหตุ การสรุปผลโดยใช้ค่าการเปรียบเทียบ Sig กับค่า α มีความสะดวกมากกว่า

การทดสอบว่าประชากรมีค่าเฉลี่ยตามที่เราคาดไว้หรือไม่

ตัวอย่าง 10.1.2 ข้อมูลของจำนวนชัลเฟอร์ออกไซด์ที่ออกมาจากโรงงานอุตสาหกรรมในแต่ละวัน ที่เก็บมาได้ในช่วง 60 วัน เป็นดังนี้

17	15	20	29	19	18	22	25	27	9
24	20	17	6	24	14	15	23	24	26
19	23	28	19	16	22	24	17	20	13
19	10	23	18	31	13	20	17	24	14
28	19	16	22	24	17	20	13	19	10
23	18	17	15	20	29	19	18	22	25

จงทดสอบสมมติฐานว่า ค่าเฉลี่ยของชัลเฟอร์ออกไซด์เท่ากับ 20 ที่ระดับนัยสำคัญ 0.05

วิธีทำ ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \mu = 20$
กำหนดสมมติฐานอื่น $H_1 : \mu \neq 20$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

ขั้นที่ 3. ทำการสุ่มตัวอย่างและทำการคำนวณค่าสถิติ Z

ขั้นที่ 3.1 สร้างแฟ้มข้อมูลประกอบด้วย

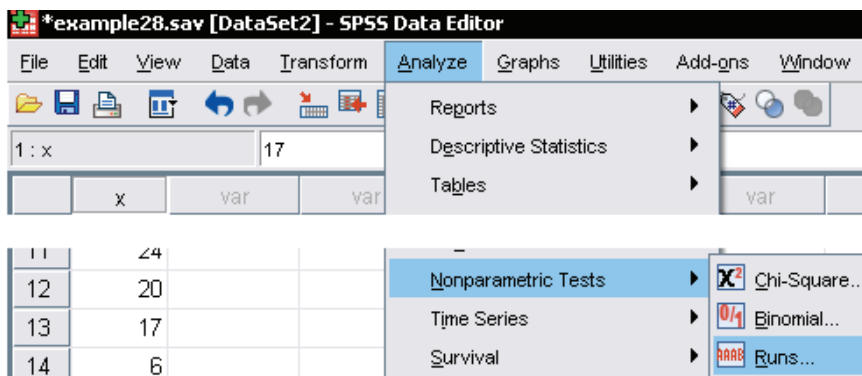
ตัวแปร x แทนจำนวนซัลเฟอร์ออกไซด์

เสร็จแล้ว Save ในแฟ้มข้อมูลชื่อ example28.sav

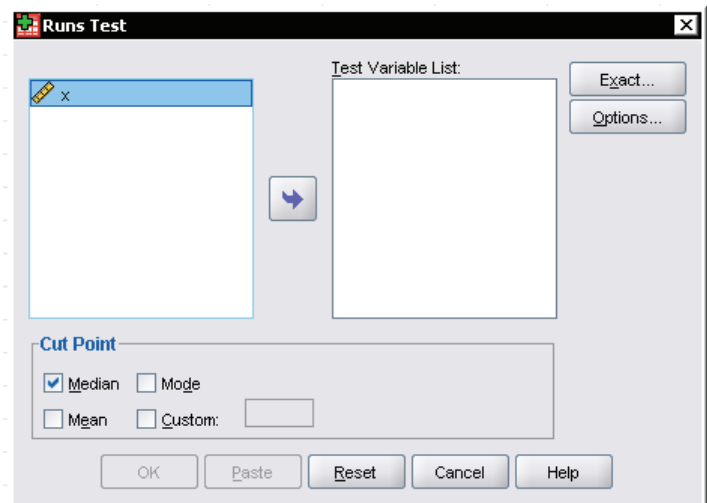
ขั้นที่ 3.2 เลือกคำสั่ง Analyze / Nonparametric Tests / Runs

*example28.sav [DataSe

	x	var
1	17	
2	15	
59	22	
60	25	



ขั้นที่ 3.3 คลิกที่ Runs จะได้เมนูย่อย



ขั้นที่ 3.4

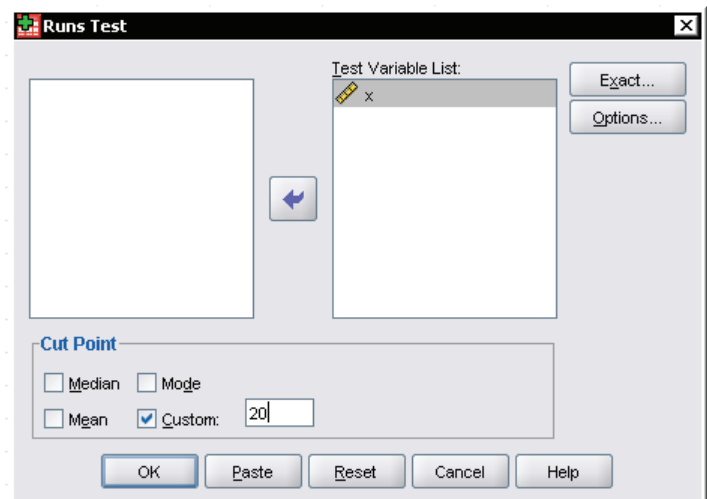
1. เลือกตัวแปร x ไปไว้ที่ช่อง Test Variable List

2. คลิกที่ช่อง Median

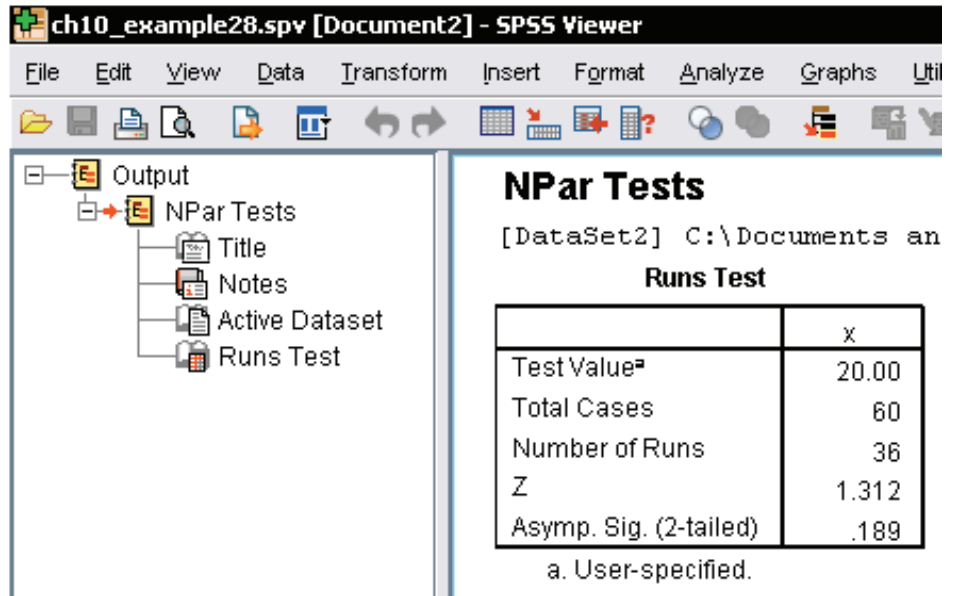
เพื่อยกเลิกการทดสอบเทียบกับค่า Median

3. คลิกที่ช่อง Custom

และพิมพ์ค่า 20 ในช่อง Custom



ขั้นที่ 3.6 คลิก OK จะได้ผลการคำนวณเป็น



ขั้นที่ 4. เลือกค่าสถิติ Z

ขั้นที่ 5. $z_{\text{คำนวณ}} = 1.312$ และ $\text{Asymp. Sig. (2-tailed)} = 0.189$

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤตและบริเวณวิกฤต

กรณีใช้ค่า Z ค่าวิกฤตคือ $-z_{\frac{\alpha}{2}}$ และ $z_{\frac{\alpha}{2}}$ บริเวณวิกฤตคือ $Z < -z_{\frac{\alpha}{2}}$ หรือ $Z > z_{\frac{\alpha}{2}}$

เพราะฉะนั้นค่าวิกฤตคือ -1.96 และ 1.96 บริเวณวิกฤตคือ $Z < -1.96$ หรือ $Z > 1.96$

ขั้นที่ 7. สรุปผล แบบที่ 1 โดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต

โดยมีเกณฑ์การสรุปผลว่า ถ้า $z_{\text{คำนวณ}} < -z_{\frac{\alpha}{2}}$ หรือ $z_{\text{คำนวณ}} > z_{\frac{\alpha}{2}}$ แล้วปฏิเสธ H_0

เพราะว่า $z_{\text{คำนวณ}} = 1.312$ ไม่อยู่ในบริเวณวิกฤต เพราะฉะนั้น ไม่ปฏิเสธ H_0

หรือ แบบที่ 2 โดยการเปรียบเทียบ Sig กับค่า α โดยมีเกณฑ์การสรุปผลว่า ถ้า $\text{Sig} < \alpha$ แล้วปฏิเสธ H_0

เพราะว่า $\text{Sig} = 0.189 > 0.05$ เพราะฉะนั้น ไม่ปฏิเสธ H_0

10.2 การทดสอบว่าประชากรมีการแจกแจงตามที่เราคาดไว้หรือไม่

การทดสอบว่าประชากรที่เราสนใจมีการแจกแจงปกติจริงหรือไม่ ประชากรที่เราสนใจมีการแจกแจง uniform จริงหรือไม่ ประชากรที่เราสนใจมีการแจกแจงปัวส์ซองจริงหรือไม่ เราสามารถทำการทดสอบแบบ Nonparametric Test ได้

ตัวอย่าง 10.2.1 การทดสอบว่าน้ำหนักของนักเรียนมีการแจกแจงปกติจริงหรือไม่ จึงทำการสุ่มตัวอย่างน้ำหนักนักเรียนมา 50 คน ได้ข้อมูลดังนี้

50	69	108	85	132	67	121	80	59	64
148	61	50	103	110	66	95	55	128	101
137	145	103	96	136	127	149	111	76	134
87	117	50	77	108	133	98	124	95	124
109	123	107	65	92	101	125	66	90	110

กำหนดระดับนัยสำคัญ 0.05

วิธีทำ ขั้นที่ 1. กำหนดสมมติฐานหลัก H_0 : ข้อมูลน้ำหนักมีการแจกแจงปกติ
 กำหนดสมมติฐานอื่น H_1 : ข้อมูลน้ำหนักไม่มีการแจกแจงปกติ

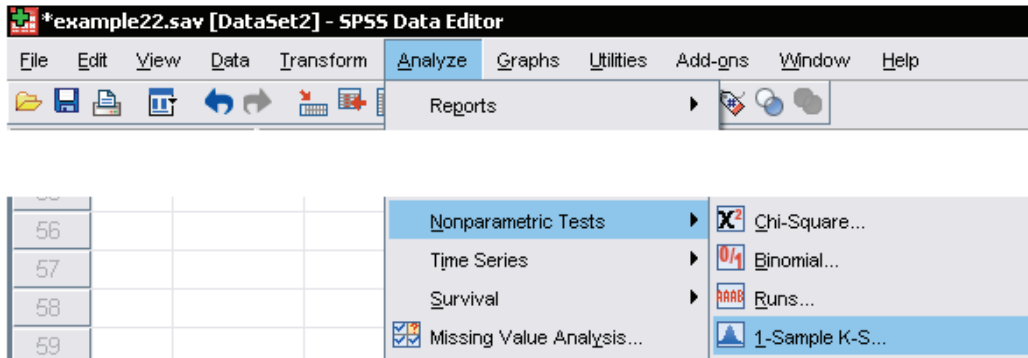
ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

ขั้นที่ 3. ทำการสุ่มตัวอย่าง และ ทำการคำนวณค่าสถิติ Z

ขั้นที่ 3.1 สร้างแฟ้มข้อมูล

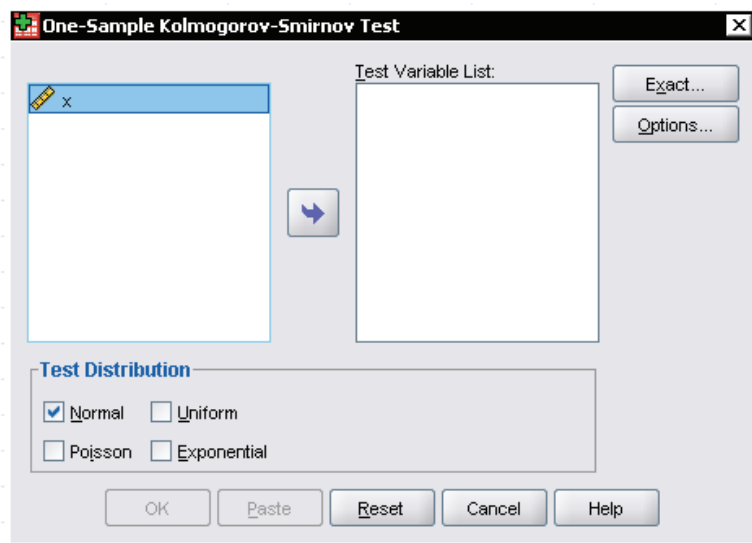
และ Save ในชื่อ example22.sav

ขั้นที่ 3.2 เลือกคำสั่ง Analyze / Nonparametric Tests / 1-Sample K-S



*example22.sa		
File	Edit	View
1 : x		
	x	
1	59	
2	69	
49	90	
50	110	

ขั้นที่ 3.3 คลิกที่คำสั่ง 1-Sample K-S จะได้เมนูย่อย

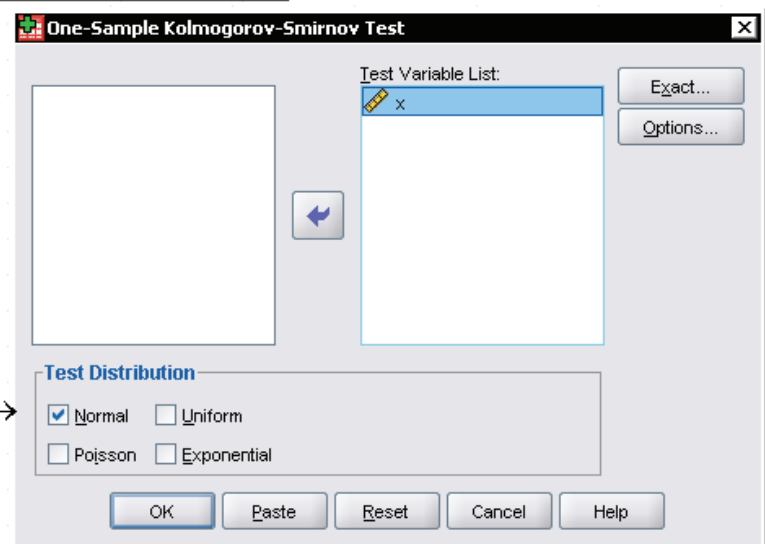
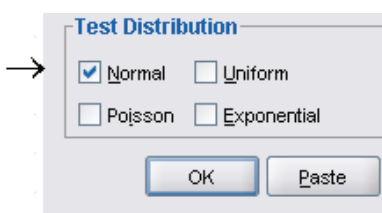


ขั้นที่ 3.4 เลือกตัวแปร x

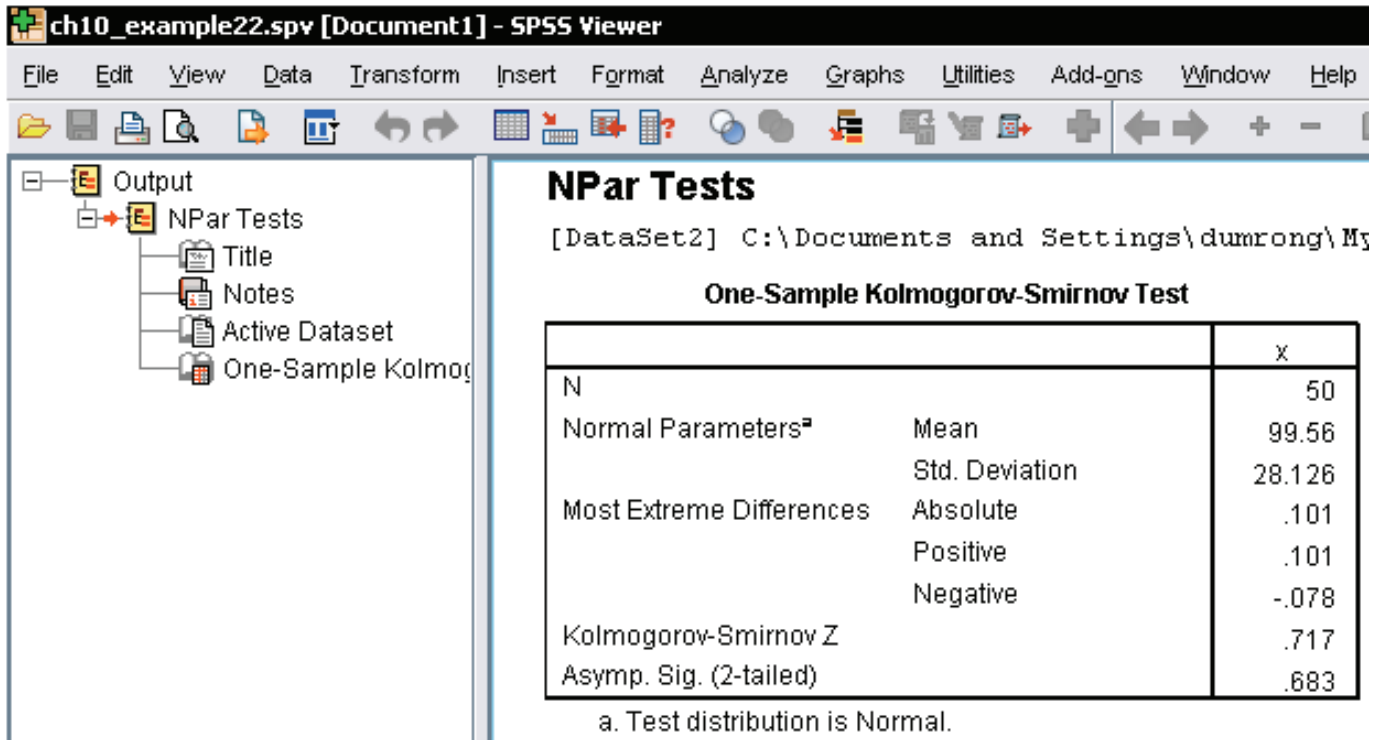
ไปที่ช่อง Test Variable List

เลือกการทดสอบ

Test Distribution เป็น Normal



ขั้นที่ 3.5 คลิก OK จะได้ผลการคำนวณเป็น



ขั้นที่ 4. เลือกค่าสถิติ Z (Kolmogorov-Smirnov Z)

ขั้นที่ 5. $z_{\text{คำนวณ}} = 0.717$ และ $\text{Asymp. Sig. (2-tailed)} = 0.683$

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤตและบริเวณวิกฤต

กรณีใช้ค่า Z ค่าวิกฤตคือ $-z_{\frac{\alpha}{2}}$ และ $z_{\frac{\alpha}{2}}$ บริเวณวิกฤตคือ $Z < -z_{\frac{\alpha}{2}}$ หรือ $Z > z_{\frac{\alpha}{2}}$

เพราะฉะนั้นค่าวิกฤตคือ -1.96 และ 1.96 บริเวณวิกฤตคือ $Z < -1.96$ หรือ $Z > 1.96$

ขั้นที่ 7. สรุปผล

แบบที่ 1 โดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต

โดยมีเกณฑ์การสรุปผลว่า ถ้า $z_{\text{คำนวณ}} < -z_{\frac{\alpha}{2}}$ หรือ $z_{\text{คำนวณ}} > z_{\frac{\alpha}{2}}$ แล้วปฏิเสธ H_0

เพราะว่า $z_{\text{คำนวณ}} = 0.717$ ไม่อยู่ในบริเวณวิกฤต เพราะฉะนั้นไม่ปฏิเสธ H_0

หรือ แบบที่ 2 โดยการเปรียบเทียบ Sig กับค่า α โดยมีเกณฑ์การสรุปผลว่า ถ้า $\text{Sig} < \alpha$ แล้วปฏิเสธ H_0

เพราะว่า $\text{Sig} = 0.683 > 0.05$ เพราะฉะนั้น ไม่ปฏิเสธ H_0

หมายเหตุ การสรุปผลโดยใช้ค่าการเปรียบเทียบ Sig กับค่า α มีความสะดวกมากกว่า

10.3 การทดสอบว่าประชากร 2 กลุ่มมีค่าเฉลี่ยเท่ากันหรือไม่

10.3.1 ประชากร 2 ชุดไม่เป็นอิสระต่อกัน

ในกรณีที่ประชากร 2 ชุดไม่เป็นอิสระต่อกัน และไม่ทราบการแจกแจงของประชากร เราสามารถทำการทดสอบสมมติฐานว่าค่าเฉลี่ยของประชากร 2 ชุดเท่ากันหรือไม่ โดยใช้วิธีทดสอบ Wilcoxon Signed Ranks Test

ตัวอย่าง 10.3.1 การทดสอบว่าโปรแกรมการควบคุมน้ำหนักโดยใช้เวลา 40 วันจะมีผลทำให้น้ำหนักลดลงได้ทำการเก็บข้อมูลน้ำหนักของชาย 40 คนได้ข้อมูลดังนี้

คนที่	น้ำหนักก่อน เข้าโปรแกรม	น้ำหนักหลัง เข้าโปรแกรม
1	147.00	137.90
2	183.50	176.20
3	232.10	219.00
4	161.60	163.80
5	197.50	193.50
6	206.30	201.40
7	177.00	180.60
8	215.40	203.20
9	147.70	149.00
10	208.10	195.40
11	137.90	140.00
12	176.20	170.00
13	219.00	210.00
14	163.80	160.00
15	137.90	140.00
16	176.20	170.00
17	219.00	210.00
18	163.80	165.00
19	193.50	195.00
20	201.40	205.00

คนที่	น้ำหนักก่อน เข้าโปรแกรม	น้ำหนักหลัง เข้าโปรแกรม
21	180.60	185.00
22	203.20	195.00
23	137.90	140.00
24	176.20	170.00
25	219.00	200.00
26	163.80	155.00
27	193.50	190.00
28	201.40	200.00
29	180.60	170.00
30	137.90	140.00
31	176.20	177.00
32	219.00	211.00
33	163.80	174.00
34	193.50	195.00
35	201.40	200.00
36	180.60	180.00
37	203.20	203.00
38	149.00	150.00
39	195.40	185.00
40	145.00	150.00

จงทดสอบว่าโปรแกรมการควบคุมน้ำหนักไม่ทำให้น้ำหนักเปลี่ยนแปลง กำหนดระดับนัยสำคัญ 0.05

วิธีทำ **ขั้นที่ 1.** กำหนดสมมติฐานหลัก H_0 : ค่าเฉลี่ยของน้ำหนักก่อนและหลังเข้าโปรแกรมเท่ากัน

กำหนดสมมติฐานอื่น H_1 : ค่าเฉลี่ยของน้ำหนักก่อนและหลังเข้าโปรแกรมไม่เท่ากัน

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

ขั้นที่ 3. ทำการสุ่มตัวอย่างและคำนวณค่าสถิติ Z

ขั้นที่ 3.1 สร้างเพิ่มข้อมูล

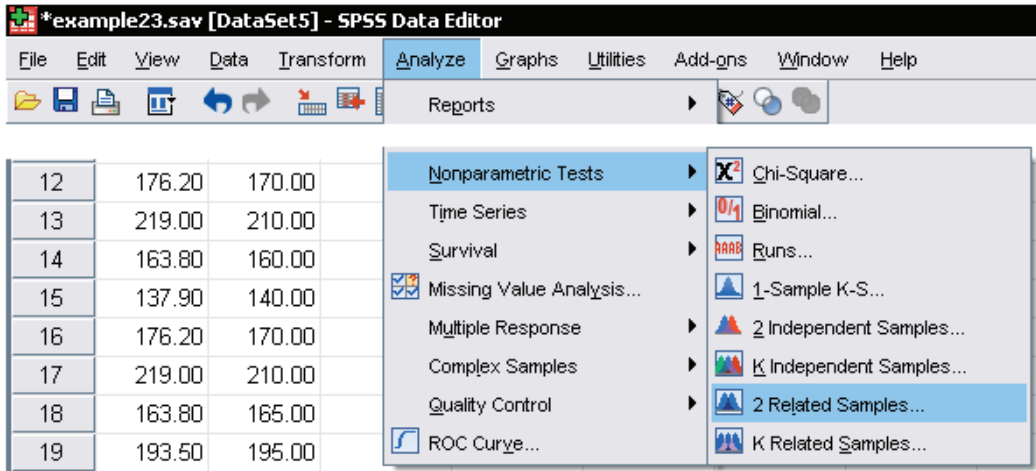
ให้ x เป็นน้ำหนักก่อนเข้าโปรแกรม

และ y เป็นน้ำหนักหลังเข้าโปรแกรม

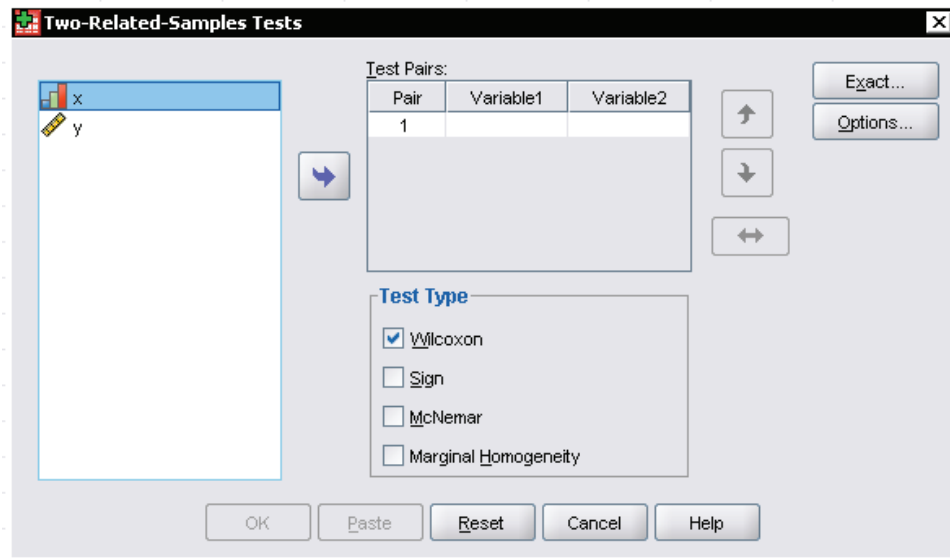
เสร็จแล้ว Save ข้อมูลในแฟ้มชื่อ example23.sav

	x	y	var
1	147.00	137.90	
2	183.50	176.20	
3	232.10	219.00	
4	161.60	163.80	

ขั้นที่ 3.2 เลือกคำสั่ง Analyze / Nonparametric Tests / 2 Related Samples

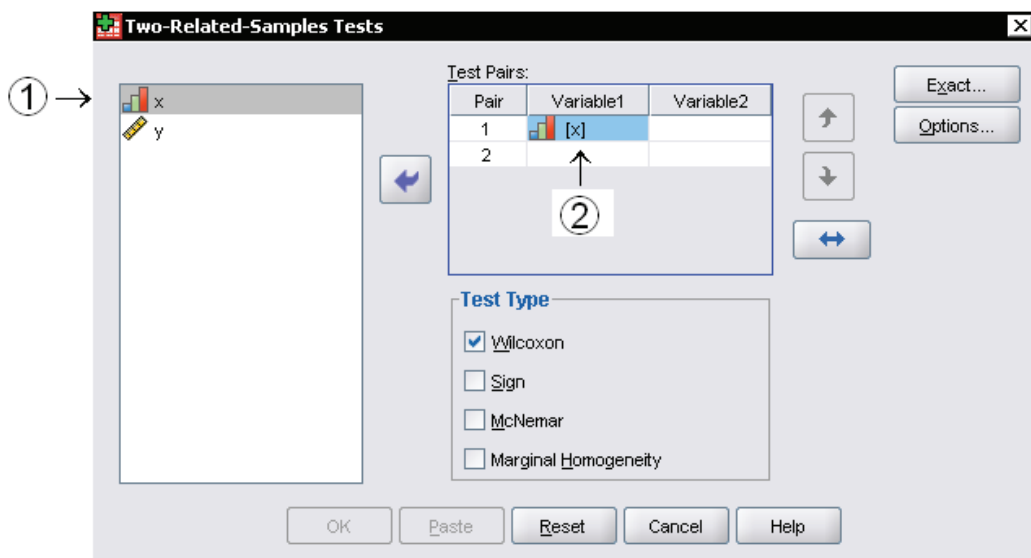


ขั้นที่ 3.3 คลิกที่คำสั่ง 2 Related Samples จะได้เมนูย่อย



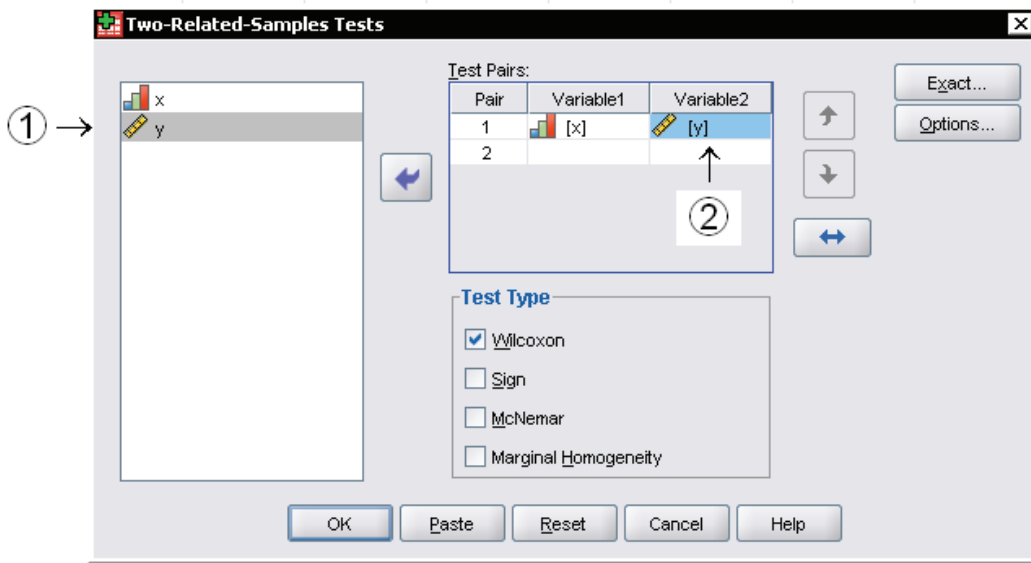
ขั้นที่ 3.4 1. เลือกตัวแปร x

2. แล้วกด  ตัวแปร x จะย้ายมาที่ช่อง Variable1 ที่กรอบ Test Pairs

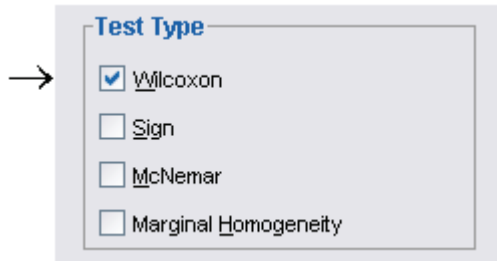


ขั้นที่ 3.5 1. เลือกตัวแปร y

2. แล้วกด  ตัวแปร y จะย้ายมาที่ช่อง Variable2 ที่กรอบ Test Pairs



ขั้นที่ 3.6 เลือกวิธีทดสอบ Wilcoxon



ขั้นที่ 3.7 คลิก OK จะได้ผลการคำนวณเป็น

ch10_example23.spv [Document2] - SPSS Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Add-ons Window Help

Output

- Output
 - NParr Tests
 - Title
 - Notes
 - Active Dataset
 - Wilcoxon Signed Ranks Test
 - Title
 - Ranks
 - Test Statistics

NPar Tests

[DataSet5] C:\Documents and Settings\dumrong\My Doc

Wilcoxon Signed Ranks

Ranks

		N	Mean Rank	Sum of Ranks
y - x	Negative Ranks	24 ^a	25.21	605.00
	Positive Ranks	16 ^b	13.44	215.00
	Ties	0 ^c		
	Total	40		

a. $y < x$
 b. $y > x$
 c. $y = x$

Test Statistics^b

	y - x
Z	-2.622 ^a
Asymp. Sig. (2-tailed)	.009

a. Based on positive ranks.
 b. Wilcoxon Signed Ranks Test

ขั้นที่ 4. เลือกค่าสถิติ Z

ขั้นที่ 5. $z_{\text{คำนวณ}} = -2.622$ และ Asymp. Sig. (2-tailed) = 0.009

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤตและบริเวณวิกฤต

กรณีใช้ค่า Z ค่าวิกฤตคือ $-z_{\frac{\alpha}{2}}$ และ $z_{\frac{\alpha}{2}}$ บริเวณวิกฤตคือ $Z < -z_{\frac{\alpha}{2}}$ หรือ $Z > z_{\frac{\alpha}{2}}$

เพราะฉะนั้นค่าวิกฤตคือ -1.96 และ 1.96 บริเวณวิกฤตคือ $Z < -1.96$ หรือ $Z > 1.96$

ขั้นที่ 7. สรุปผล แบบที่ 1 โดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต

โดยมีเกณฑ์การสรุปผลว่า ถ้า $z_{\text{คำนวณ}} < -z_{\frac{\alpha}{2}}$ หรือ $z_{\text{คำนวณ}} > z_{\frac{\alpha}{2}}$ แล้วปฏิเสธ H_0

เพราะว่า $z_{\text{คำนวณ}} = -2.622 < -1.96$ เพราะฉะนั้น ปฏิเสธ H_0

หรือ แบบที่ 2 โดยการเปรียบเทียบ Sig กับค่า α โดยมีเกณฑ์การสรุปผลว่า ถ้า Sig < α แล้วปฏิเสธ H_0

เพราะว่า Sig = 0.009 < 0.05 เพราะฉะนั้น ปฏิเสธ H_0

หมายเหตุ การสรุปผลโดยใช้ค่าการเปรียบเทียบ Sig กับค่า α มีความสะดวกมากกว่า

10.3.2 ประชากร 2 ชุดเป็นอิสระต่อกัน

ตัวอย่าง 10.3.2 ปริมาณของนิโคตินที่มีในบุหรี่ยี่ห้อคือ

ยี่ห้อ A	2.1	4.0	6.3	5.4	4.8	3.7	6.1	3.3		
ยี่ห้อ B	4.1	0.6	3.1	2.5	4.0	6.2	1.6	2.2	1.9	5.4

จงทดสอบว่าปริมาณของนิโคตินที่มีในบุหรี่ยี่ห้อเท่ากัน กำหนดระดับนัยสำคัญ 0.05

วิธีทำ ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \mu_1 = \mu_2$

กำหนดสมมติฐานอื่น $H_1 : \mu_1 \neq \mu_2$

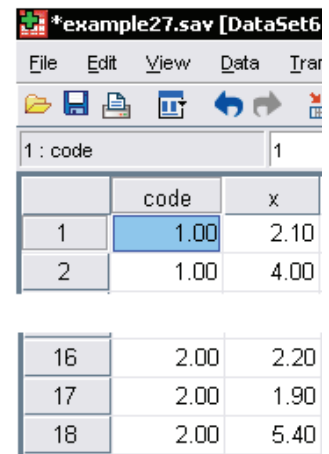
ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

ขั้นที่ 3. ทำการสุ่มตัวอย่างและทำการคำนวณค่าสถิติ Z

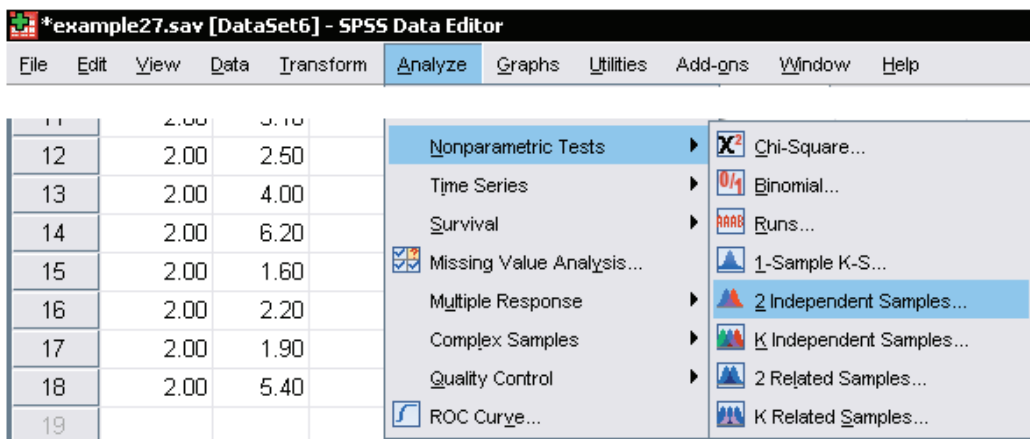
ขั้นที่ 3.1 สร้างแฟ้มข้อมูล code เป็นตัวแปรจำแนกกลุ่ม

x เป็นตัวแปรปริมาณนิโคติน และ Save แฟ้มข้อมูลชื่อ example27.sav

ขั้นที่ 3.2 เลือกคำสั่ง



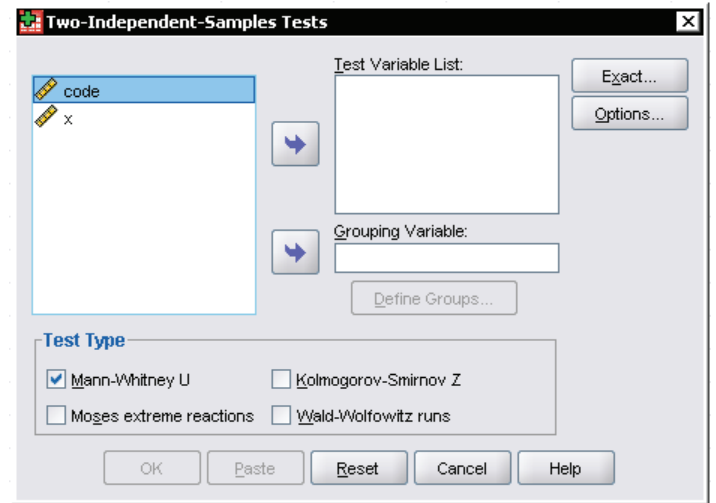
Analyze / Nonparametric Tests / 2 Independent Samples



ขั้นที่ 3.3 คลิกที่คำสั่งย่อย

2 Independent Samples

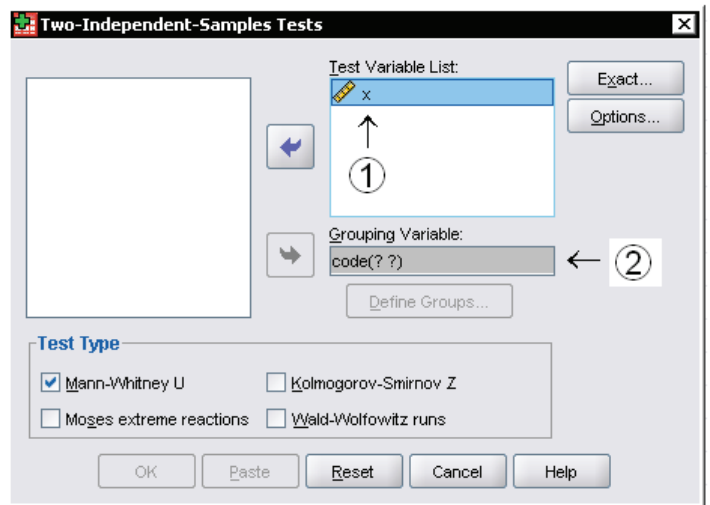
จะได้เมนูย่อย



ขั้นที่ 3.4

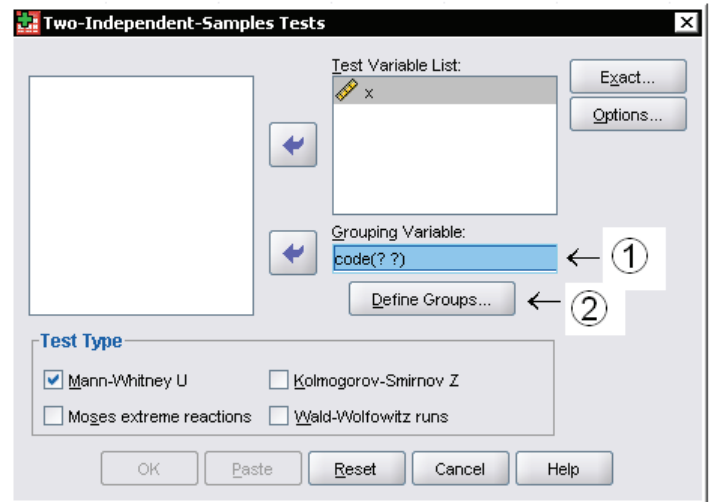
1. เลือกตัวแปร x มาที่ช่อง Test Variable List
2. เลือกตัวแปร code มาที่ช่อง Grouping Variable
3. เลือก Test Type เป็น Mann-Whitney

③ →



ขั้นที่ 3.5

1. คลิกที่ code(? ?) จะทำให้ปุ่ม Define Groups เป็นสีเข้มขึ้นเพื่อกำหนดกลุ่มทดสอบ
2. คลิกที่ Define Groups จะได้เมนูย่อย

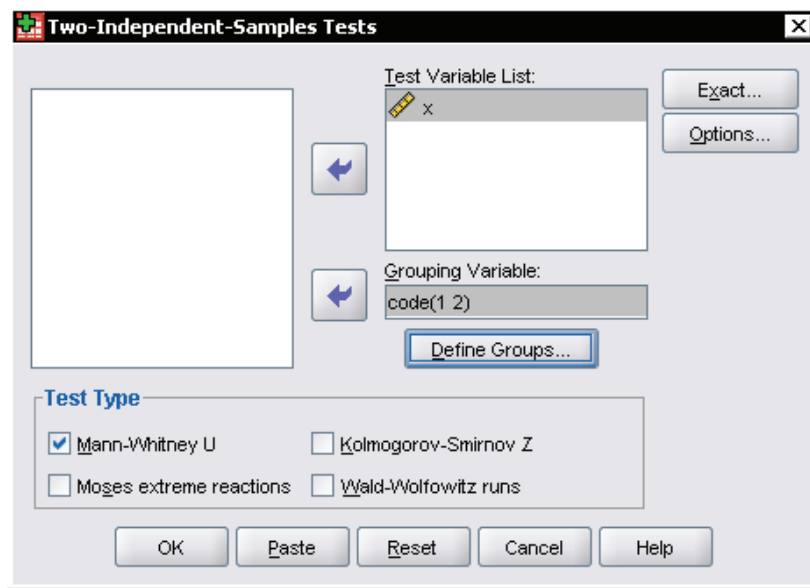


ขั้นที่ 3.6

พิมพ์ 1 ในช่อง Group 1
พิมพ์ 2 ในช่อง Group 2
บนจอภาพจะเป็นดังนี้



ขั้นที่ 3.7 คลิก Continue จะกลับมาที่เมนูย่อย



ขั้นที่ 3.8 คลิก OK จะได้ผลการคำนวณเป็น

ch10_example27.spv [Document2] - SPSS Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Add-ons Window

Output

- Output
 - NPar Tests
 - Title
 - Notes
 - Active Dataset
 - Mann-Whitney Test
 - Title
 - Ranks
 - Test Statistics

NPar Tests

[DataSet6] C:\Documents and Settings\dumron

Mann-Whitney

Ranks

	code	N	Mean Rank	Sum of Ranks
x	1	8	11.62	93.00
	2	10	7.80	78.00
	Total	18		

Test Statistics^b

	x
Mann-Whitney U	23.000
Wilcoxon W	78.000
Z	-1.512
Asymp. Sig. (2-tailed)	.131
Exact Sig. [2*(1-tailed Sig.)]	.146 ^a

a. Not corrected for ties.
b. Grouping Variable: code

ขั้นที่ 4. เลือกค่าสถิติ Z

ขั้นที่ 5. $z_{\text{คำนวณ}} = -1.512$ และ Asymp. Sig. (2-tailed) = 0.131

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤตและบริเวณวิกฤต

กรณีใช้ค่า Z ค่าวิกฤตคือ $-z_{\frac{\alpha}{2}}$ และ $z_{\frac{\alpha}{2}}$ บริเวณวิกฤตคือ $Z < -z_{\frac{\alpha}{2}}$ หรือ $Z > z_{\frac{\alpha}{2}}$

เพราะฉะนั้นค่าวิกฤตคือ -1.96 และ 1.96 บริเวณวิกฤตคือ $Z < -1.96$ หรือ $Z > 1.96$

ขั้นที่ 7. สรุปผล

แบบที่ 1 โดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต

โดยมีเกณฑ์การสรุปผลว่า ถ้า $z_{\text{คำนวณ}} < -z_{\frac{\alpha}{2}}$ หรือ $z_{\text{คำนวณ}} > z_{\frac{\alpha}{2}}$ แล้วปฏิเสธ H_0

เพราะว่า $z_{\text{คำนวณ}} = -1.512$ ไม่อยู่ในบริเวณวิกฤต เพราะฉะนั้นไม่ปฏิเสธ H_0

หรือ แบบที่ 2 โดยการเปรียบเทียบ Sig กับค่า α

โดยมีเกณฑ์การสรุปผลว่า ถ้า Sig < α แล้วปฏิเสธ H_0

เพราะว่า Sig = 0.131 > 0.05 เพราะฉะนั้นไม่ปฏิเสธ H_0

หมายเหตุ การสรุปผลโดยใช้ค่าการเปรียบเทียบ Sig กับค่า α มีความสะดวกมากกว่า

10.4 การทดสอบว่าประชากร k กลุ่มมีค่าเฉลี่ยเท่ากันหรือไม่

10.4.1 ประชากร k กลุ่มเป็นอิสระต่อกัน

ในกรณีที่ประชากร k ชุดอิสระต่อกัน และ ไม่ทราบการแจกแจงของประชากร เราสามารถทำการทดสอบสมมติฐานว่าค่าเฉลี่ยของประชากร 2 ชุดเท่ากันหรือไม่ โดยใช้วิธีทดสอบ **Kruskal – Wallis Test**

ตัวอย่าง 10.4.1 คะแนนสอบวิชาภาษาอังกฤษของนักเรียน 3 กลุ่มที่มาจากวิธีการสอนที่ต่างกัน

วิธีที่ 1	94	88	91	74	87	97	
วิธีที่ 2	85	82	79	84	63	72	80
วิธีที่ 3	89	67	72	76	69		

จงทดสอบสมมติฐานว่าวิธีการสอนทั้งสามแบบให้ผลเหมือนกัน กำหนดระดับนัยสำคัญ 0.05

วิธีทำ

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \mu_1 = \mu_2 = \mu_3$

กำหนดสมมติฐานอื่น $H_1 : \mu_1 \neq \mu_2 \neq \mu_3$ (ค่าเฉลี่ยอย่างน้อย 1 คู่แตกต่างกัน)

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

ขั้นที่ 3. ทำการสุ่มตัวอย่างและทำการคำนวณค่าสถิติ Z

ขั้นที่ 3.1 สร้างแฟ้มข้อมูล

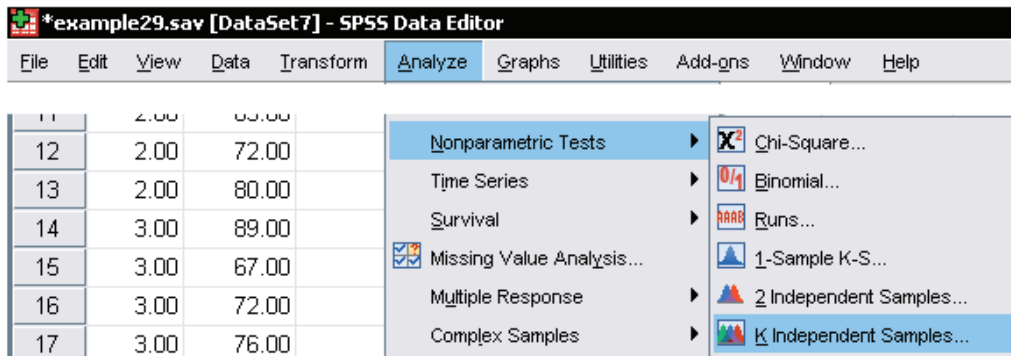
code เป็นตัวแปรจำแนกกลุ่ม

x เป็นตัวแปรเก็บคะแนน

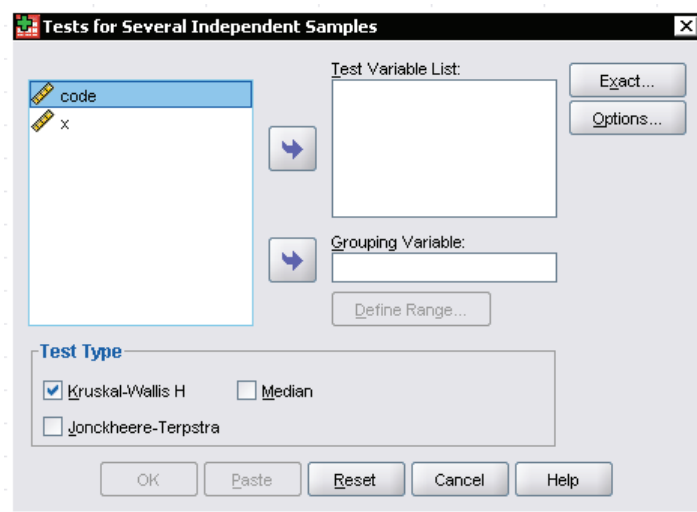
และ Save ลงแฟ้มข้อมูลชื่อ example29.sav

code	x
1	94.00
2	88.00
17	76.00
18	69.00

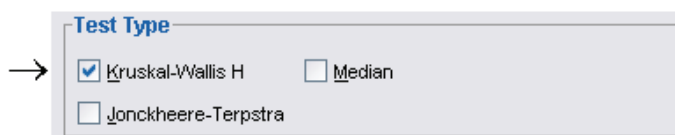
ขั้นที่ 3.2 เลือกคำสั่ง Analyze / Nonparametric Tests / K Independent Samples



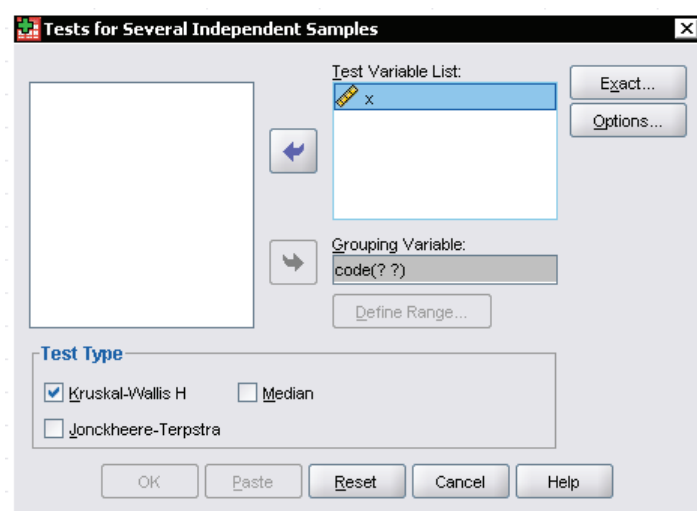
ขั้นที่ 3.3 คลิกที่คำสั่ง K Independent Samples จะได้เมนูย่อย



เลือกการทดสอบ Test Type เป็น Kruskal-Wallis H



ขั้นที่ 3.4 เอาตัวแปร x ไปไว้ที่ช่อง Test Variable List เอาตัวแปร code ไปไว้ที่ช่อง Grouping Variable



ขั้นที่ 3.5 1. คลิกที่ code(? ?)

จะทำให้ปุ่ม Define Range

เป็นสีเข้มขึ้นเพื่อจะทำการกำหนดกลุ่มทดสอบ

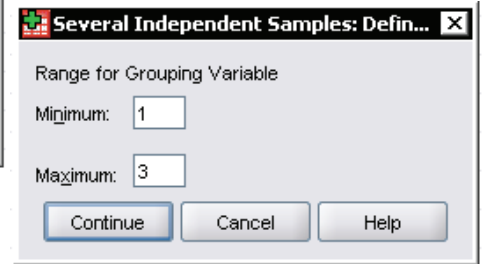
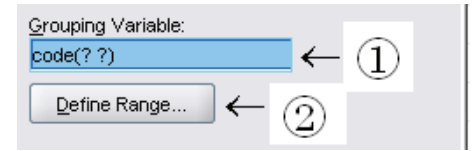
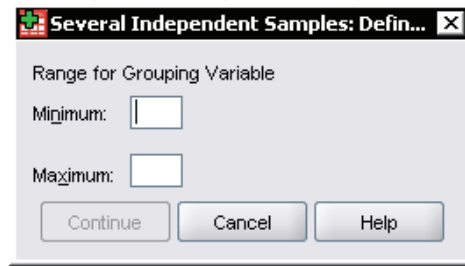
2. คลิกที่ Define Groups

จะได้เมนูย่อย Define Range

3. พิมพ์ 1 ในช่อง Minimum

4. พิมพ์ 3 ในช่อง Maximum

บนจอภาพจะเป็นดังนี้

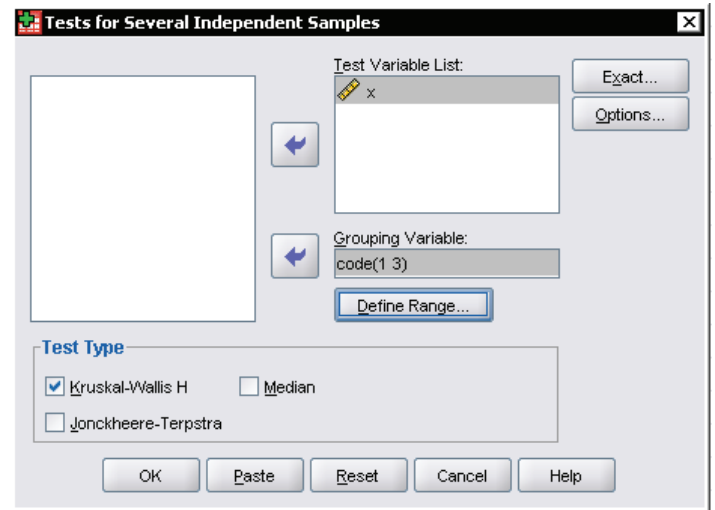


ขั้นที่ 3.6 คลิก Continue

จะได้จอภาพเป็น

ข้อสังเกต จากของเดิม code(? ?)

กลายเป็น code(1 3)



ขั้นที่ 3.7 คลิก OK จะได้ผลการคำนวณเป็น

ch10_example29.spv [Document3] - SPSS Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Add-o

Output

- Output
 - NPAr Tests
 - Title
 - Notes
 - Active Dataset
 - Kruskal-Wallis Test
 - Title
 - Ranks
 - Test Statistics

NPAr Tests
[DataSet7] C:\Documents and Sett

Kruskal-Wallis
Ranks

	code	N	Mean Rank
x	1	6	14.00
	2	7	7.93
	3	5	6.30
	Total	18	

Test Statistics^{a,b}

	x
Chi-Square	6.673
df	2
Asymp. Sig.	.036

a. Kruskal Wallis Test
b. Grouping Variable: code

จากตาราง Test Statistics

Test Statistics^{a,b}

	x
Chi-Square	6.673
df	2
Asymp. Sig.	.036

a. Kruskal Wallis Test

b. Grouping Variable: code

หมายเหตุ Asymp. Sig = 0.036 มาจากค่า $P(\chi^2 > 6.673)$

ขั้นที่ 4. เลือกค่าสถิติ χ^2

ขั้นที่ 5. $\chi^2_{\text{คำนวณ}} = 6.673$ และ Asymp. Sig. = 0.036

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤตและบริเวณวิกฤต

ค่าวิกฤตคือ χ^2_{α} เมื่อ $df = k - 1$ บริเวณวิกฤตคือ $\chi^2 > \chi^2_{\alpha}$

เพราะฉะนั้นค่าวิกฤตคือ $\chi^2_{0.05} = 5.99$ บริเวณวิกฤตคือ $\chi^2 > 5.99$

ขั้นที่ 7. สรุปผล แบบที่ 1 โดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต

โดยมีเกณฑ์การสรุปผลว่า ถ้า $\chi^2_{\text{คำนวณ}} > \chi^2_{\alpha}$ แล้วปฏิเสธ H_0

เพราะว่า $\chi^2_{\text{คำนวณ}} = 6.673 > 5.99$ เพราะฉะนั้นปฏิเสธ H_0

หรือ แบบที่ 2 โดยการเปรียบเทียบ Sig กับค่า α โดยมีเกณฑ์การสรุปผลว่า ถ้า $\text{Sig} < \alpha$ แล้วปฏิเสธ H_0

เพราะว่า $\text{Sig} = 0.036 < 0.05$ เพราะฉะนั้นปฏิเสธ H_0

หมายเหตุ การสรุปผลโดยใช้ค่าการเปรียบเทียบ Sig กับค่า α มีความสะดวกมากกว่า

10.4.2 ประชากร 2 กลุ่มไม่เป็นอิสระต่อกัน

ตัวอย่าง 10.4.2 เครื่องมือ 3 แบบสำหรับวัดปริมาณของซัลเฟอร์มอน็อกไซด์ในบรรยากาศได้ข้อมูลเป็นดังนี้

ปริมาณของซัลเฟอร์มอน็อกไซด์ที่วัดได้ในแต่ละวัน

วันที่	เครื่องมือแบบ A	เครื่องมือแบบ B	เครื่องมือแบบ C
1.	0.96	0.87	0.76
2.	0.82	0.74	0.85
3.	0.75	0.63	0.74
4.	0.61	0.55	0.46
5.	0.89	0.76	0.78
6.	0.64	0.70	0.81
7.	0.81	0.69	0.72
8.	0.68	0.57	0.56
9.	0.65	0.53	0.56
10.	0.84	0.88	0.74
11.	0.59	0.51	0.62
12.	0.94	0.79	0.68

จงทดสอบที่ระดับนัยสำคัญ 0.05 ว่าผลการวัดของเครื่องมือทั้ง 3 แบบมีผลไม่แตกต่างกัน

วิธีทำ **ขั้นที่ 1.** กำหนดสมมติฐานหลัก $H_0 : \mu_1 = \mu_2 = \mu_3$
 กำหนดสมมติฐานอื่น $H_1 : \mu_1 \neq \mu_2 \neq \mu_3$ (ค่าเฉลี่ยอย่างน้อย 1 คู่แตกต่าง)

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

ขั้นที่ 3. ทำการสุ่มตัวอย่างและทำการคำนวณค่าสถิติ Z

ขั้นที่ 3.1 สร้างแฟ้มข้อมูลประกอบด้วยตัวแปร 3 ตัว

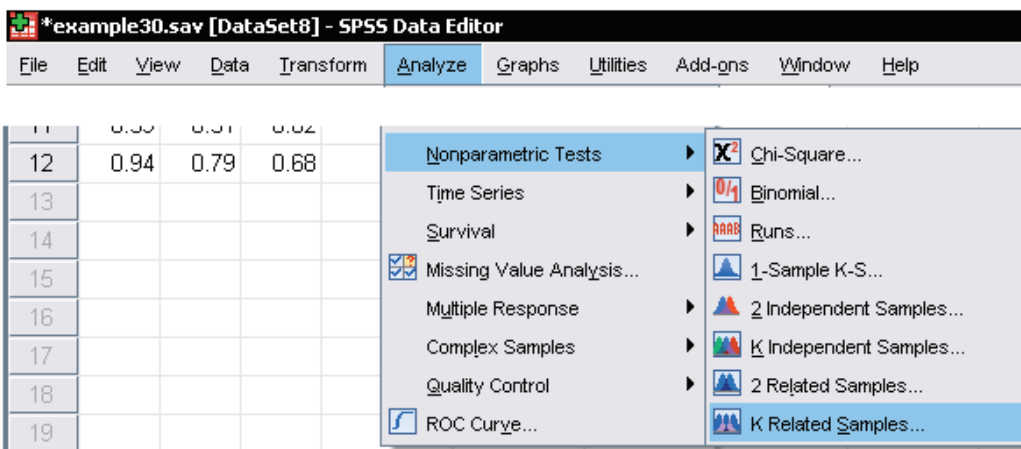
คือ a , b , c เป็นปริมาณของซัลเฟอร์มอนีอกไซด์

ในบรรยากาศได้ด้วยเครื่องมือแบบ A , B , C ตามลำดับ

และ Save ลงแฟ้มข้อมูลชื่อ example30.sav

ขั้นที่ 3.2 เลือกคำสั่ง Analyze / Nonparametric Tests / K Related Samples

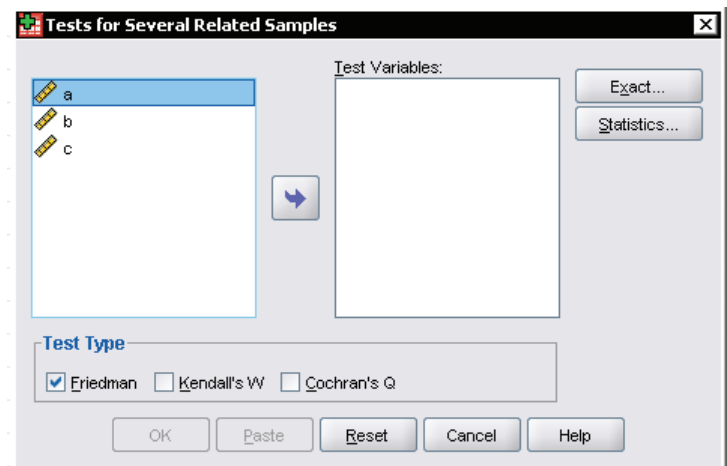
	a	b	c
1	0.96	0.87	0.76
2	0.82	0.74	0.85
3	0.75	0.63	0.74
4	0.61	0.55	0.46
5	0.89	0.76	0.78
6	0.64	0.70	0.81
7	0.81	0.69	0.72
8	0.68	0.57	0.56
9	0.65	0.53	0.56
10	0.84	0.88	0.74
11	0.59	0.51	0.62
12	0.94	0.79	0.68



ขั้นที่ 3.3

คลิกที่ K Related Samples

จะได้เมนูย่อย



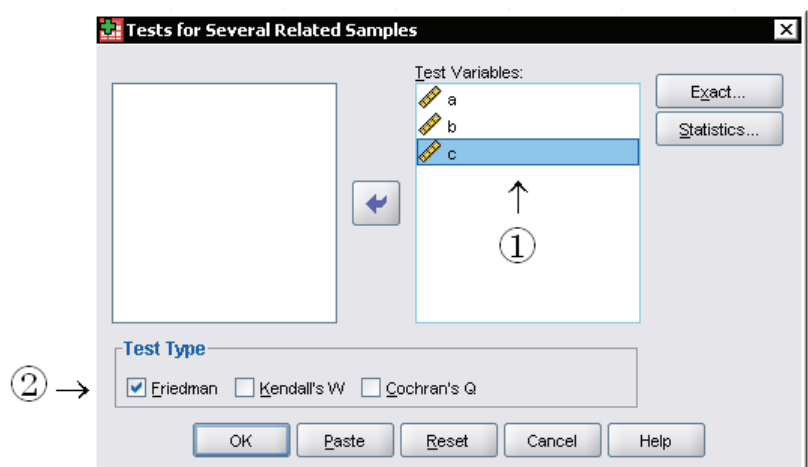
ขั้นที่ 3.4

1. เลือกตัวแปร a , b , c

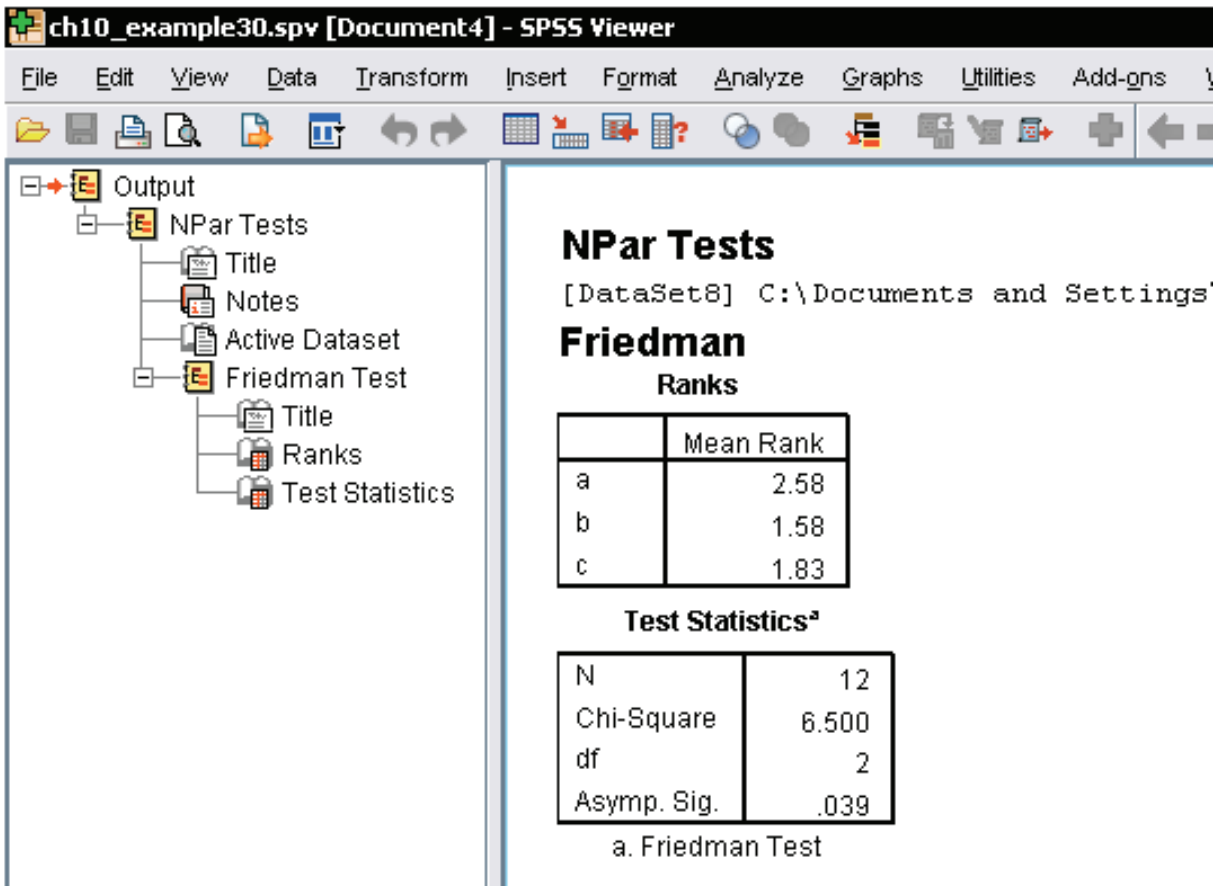
มาที่ช่อง Test Variables

2. เลือกวิธีทดสอบ Test Type

เป็น Friedman



ขั้นที่ 3.5 คลิก OK จะได้ผลการคำนวณเป็น



หมายเหตุ Asymp. Sig = 0.039 มาจากค่า $P(\chi^2 > 6.5)$

ขั้นที่ 4. เลือกค่าสถิติ χ^2

ขั้นที่ 5. คำนวณค่าสถิติจากตัวอย่างได้ $\chi^2_{\text{คำนวณ}} = 6.500$ และ Asymp. Sig. = 0.039

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤตและบริเวณวิกฤต

ค่าวิกฤตคือ χ^2_{α} เมื่อ $df = k - 1$ และบริเวณวิกฤตคือ $\chi^2 > \chi^2_{\alpha}$

เพราะฉะนั้นค่าวิกฤตคือ $\chi^2_{0.05} = 5.99$ บริเวณวิกฤตคือ $\chi^2 > 5.99$

ขั้นที่ 7. สรุปผล

แบบที่ 1 โดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต

โดยมีเกณฑ์การสรุปผลว่า ถ้า $\chi^2_{\text{คำนวณ}} > \chi^2_{\alpha}$ แล้วปฏิเสธ H_0

เพราะว่า $\chi^2_{\text{คำนวณ}} = 6.5 > 5.99$ เพราะฉะนั้นปฏิเสธ H_0

หรือ แบบที่ 2 โดยการเปรียบเทียบ Sig กับ α โดยมีเกณฑ์การสรุปผลว่า ถ้า $\text{Sig} < \alpha$ แล้วปฏิเสธ H_0

เพราะว่า $\text{Sig} = 0.039 < 0.05$ เพราะฉะนั้นปฏิเสธ H_0

หมายเหตุ การสรุปผลโดยใช้ค่าการเปรียบเทียบ Sig กับค่า α มีความสะดวกมากกว่า

10.5 การหาสัมประสิทธิ์สหสัมพันธ์ตำแหน่งที่ (Rank Correlation Coefficient)

ในกรณีที่ข้อมูล 2 ชุด เป็นข้อมูลแบบลำดับที่ ตำแหน่งที่ หรือข้อมูลที่มีความหมายในลักษณะของการเรียงลำดับ กับข้อมูลเชิงปริมาณ การวัดความสัมพันธ์นิยมใช้สัมประสิทธิ์สหสัมพันธ์ตำแหน่งที่ของสเปียร์แมน การหาสัมประสิทธิ์สหสัมพันธ์ตำแหน่งที่ของสเปียร์แมน (Spearman Rank Correlation Coefficient)

ตัวอย่าง 10.5.1 ข้อมูลของจำนวนชั่วโมงดูหนังสือและคะแนนสอบที่นักเรียนทำได้ เป็นดังนี้

นักเรียนคนที่	1	2	3	4	5	6	7	8	9	10
เวลาดูหนังสือ(หน่วย ชม.)	8	5	11	13	10	5	18	15	2	8
คะแนน	56	44	79	72	70	54	95	85	33	65

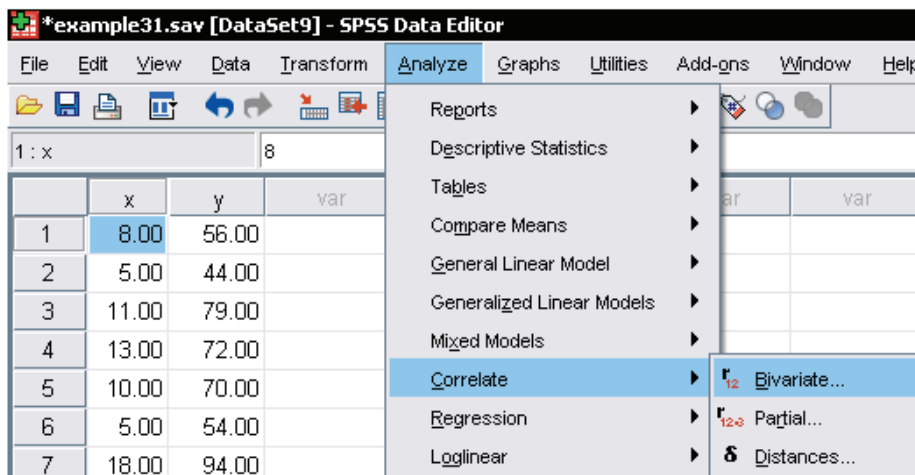
จงหาสัมประสิทธิ์สหสัมพันธ์ตำแหน่งที่ของสเปียร์แมน และอธิบายความสัมพันธ์ที่ได้

วิธีทำ

ขั้นที่ 1. สร้างแฟ้มข้อมูลโดยมีตัวแปร x แทนจำนวนชั่วโมงที่ดูหนังสือ และ ตัวแปร y แทนคะแนนที่ได้

เสร็จแล้ว Save ลงแฟ้มข้อมูลชื่อ example31.sav

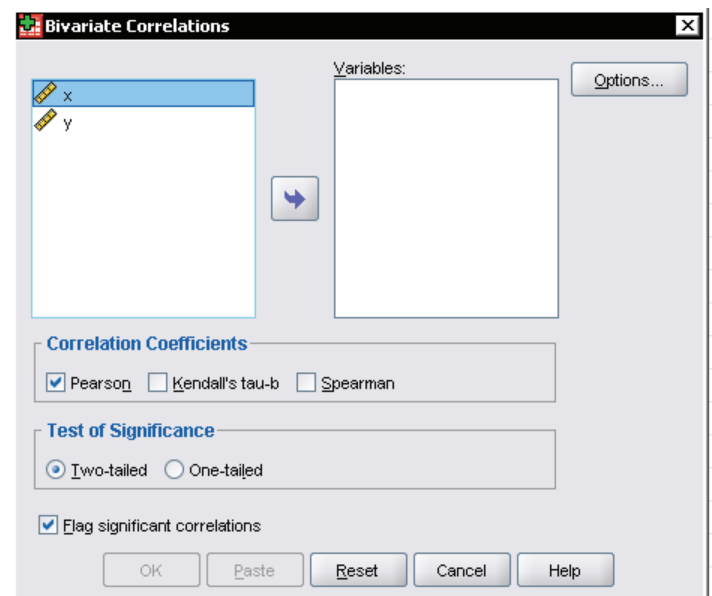
ขั้นที่ 2. เลือกคำสั่ง Analyze / Correlate / Bivariate



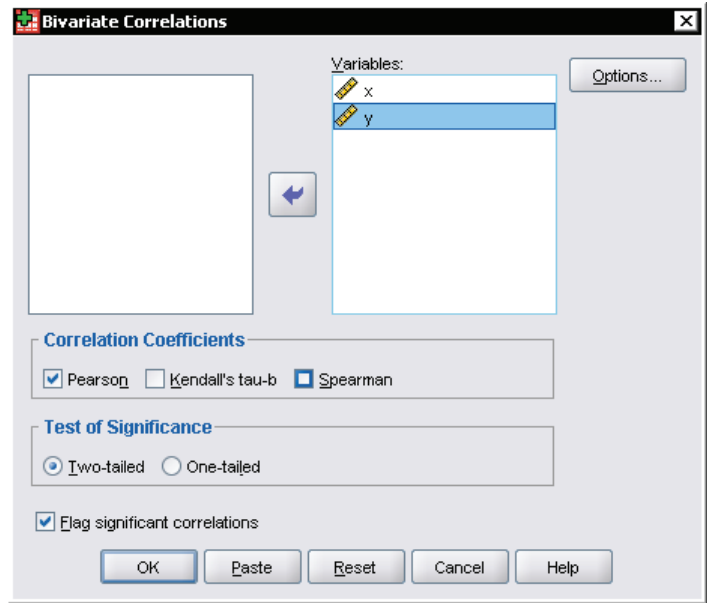
The screenshot shows the data table in SPSS Data Editor. The table has two columns: 'x' and 'y'. The data points are as follows:

	x	y
1	8.00	56.00
2	5.00	44.00
3	11.00	79.00
4	13.00	72.00
5	10.00	70.00
6	5.00	54.00
7	18.00	94.00
8	15.00	85.00
9	2.00	33.00
10	8.00	65.00

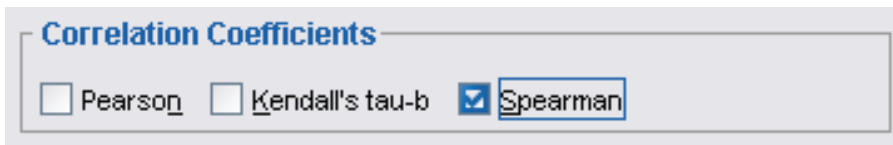
ขั้นที่ 3. คลิกที่คำสั่ง Bivariate จะได้เมนูย่อย



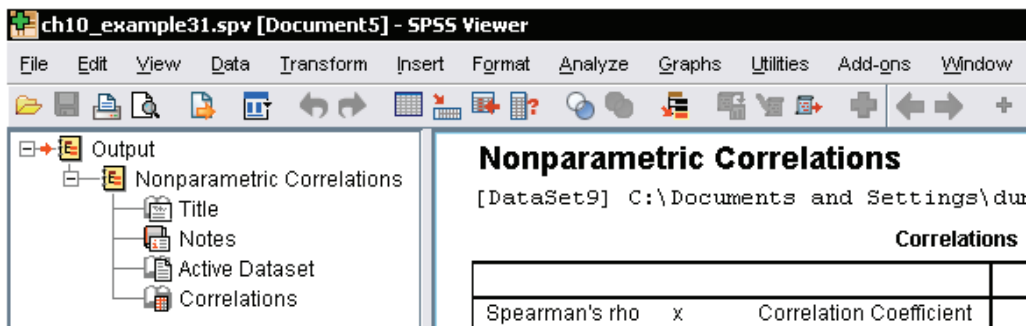
ขั้นที่ 4. นำตัวแปร x และ ตัวแปร y มาไว้ที่ช่อง Variables



ขั้นที่ 5. คลิกที่ Pearson เพื่อยกเลิก และ คลิกที่ Spearman เพื่อเลือกค่าความสัมพันธ์อันดับต่ำแหน่งที่ของสเปียร์แมน



ขั้นที่ 6. คลิก OK จะได้ผลการคำนวณเป็นดังนี้



ผลการคำนวณทั้งหมดคือ

Nonparametric Correlations

[DataSet9] C:\Documents and Settings\dumrong\My Documents\dataspss16\example31.sav

Correlations

			x	y
Spearman's rho	x	Correlation Coefficient	1.00000000**	0.98172557
		Sig. (2-tailed)	.	0.00000048
		N	10	10
y	y	Correlation Coefficient	0.98172557**	1.00000000
		Sig. (2-tailed)	0.00000048	.
		N	10	10

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

จากตาราง Correlations ค่าสัมประสิทธิ์สหสัมพันธ์ตำแหน่งที่ของสเปียร์แมน = 0.98172557

การทดสอบว่าจำนวนชั่วโมงดูหนังสือ และคะแนนสอบที่นักเรียนทำได้มีความสัมพันธ์

กำหนดระดับนัยสำคัญ 0.05

วิธีทำ

ขั้นที่ 1. กำหนดสมมติฐานหลัก H_0 : จำนวนชั่วโมงดูหนังสือ และคะแนนสอบมีความสัมพันธ์
กำหนดสมมติฐานอื่น H_1 : จำนวนชั่วโมงดูหนังสือ และคะแนนสอบ ไม่มีความสัมพันธ์

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

ขั้นที่ 3. ทำการสุ่มตัวอย่าง

ขั้นที่ 4. เลือกค่าสถิติ Spearman's rho Correlation Coefficient และค่า Sig. (2-tailed)

ขั้นที่ 5. จากตารางวิเคราะห์ข้อมูลจะได้

Spearman's rho Correlation Coefficient = 0.98172557 และค่า Sig. (2-tailed) = 0.00000048

ขั้นที่ 6. ไม่ต้องเปิดตารางหาค่าวิกฤต

ขั้นที่ 7. สรุปผลโดยการเปรียบเทียบ Sig กับค่า α

โดยมีเกณฑ์การสรุปผลว่า ถ้า Sig. (2-tailed) < α แล้วปฏิเสธ H_0

เพราะว่า Sig. (2-tailed) = 0.00000048 < 0.05 เพราะฉะนั้นปฏิเสธ H_0

เพราะฉะนั้น จำนวนชั่วโมงดูหนังสือ และคะแนนสอบที่นักเรียนทำได้มีความสัมพันธ์กัน

ภาคผนวกที่ 1

การคำนวณค่า Significant ของค่าสถิติ

ผลการวิเคราะห์ข้อมูลทางสถิติของ SPSS for Windows ส่วนใหญ่ผลการวิเคราะห์ที่จะมีการแสดงค่าของ Significant เช่น Sig. (1-tailed), Sig. (2-tailed) ซึ่งที่มาของค่า Significant คำนวณมาจากค่าสถิติ Z, t, F, Chi-Square ตัวอย่างเช่น

ค่าสถิติ $z = 1.312$	มีค่า Sig. (2-tailed) = 0.18952014
ค่าสถิติ $t = 1.581, df = 4$	มีค่า Sig. (2-tailed) = 0.189034
ค่าสถิติ Chi-Square = 3.822, $df = 3$	มีค่า Asymp. Sig. = 0.281338
ค่าสถิติ $F = 4.302, v_1 = 4, v_2 = 25$	มีค่า Sig. = 0.009

1. การหาค่า Significant ของค่าสถิติ Z

ค่า Significant ของค่าสถิติ $Z = k$ คำนวณมาจากค่าของความน่าจะเป็นหรือพื้นที่ใต้โค้งทางทางด้านขวาของโค้งการแจกแจงความน่าจะเป็นของตัวแปรสุ่ม Z ตั้งแต่ $z = |k|$ ถึง $z = \infty$

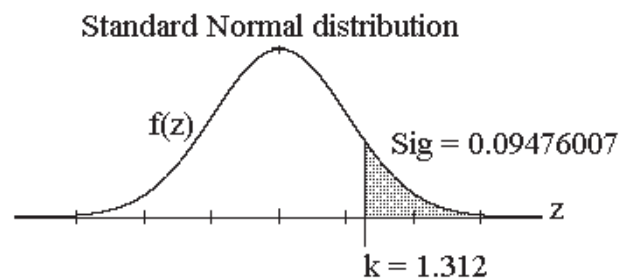
หมายเหตุ เราใช้สัญลักษณ์ Sig. แทนค่า Significant ของค่าสถิติ $Z = k$ และ $\text{Sig. (2-tailed)} = 2 \text{ Sig.}$

$$\text{Sig.} = P(|k| < Z < \infty)$$

$$= \int_{|k|}^{\infty} f(z) dz = 0.5 - \int_0^{|k|} f(z) dz$$

ตัวอย่างการคำนวณเช่น

$$f(z) := \frac{1}{\sqrt{2 \cdot \pi}} \cdot e^{-\frac{z^2}{2}} \quad \text{Sig.} := 0.5 - \int_0^{1.312} f(z) dz \quad \text{Sig.} = 0.09476007$$



คำสั่งของ Mathcad : $\text{cnorm}(k) = P(-\infty < Z < k) = \int_{-\infty}^k f(z) dz$

เพราะฉะนั้น Significant ของค่าสถิติ $z = k$ จะมีค่าเท่ากับ $1 - \text{cnorm}(|k|)$

คำสั่งที่ช่วยในการคำนวณค่าสถิติ Z ของ Mathcad

$\text{dnorm}(z, 0, 1)$	ฟังก์ชันการแจกแจงความน่าจะเป็นของ Z
$\text{pnorm}(k, 0, 1), \text{cnorm}(k)$	$\text{pnorm}(k) = \text{cnorm}(k) = P(-\infty < Z < k)$
$\text{qnorm}(A, 0, 1)$	ระยะบนแกน $Z = k$ ที่ทำให้ $P(-\infty < Z < k) = A$

หมายเหตุ ค่าวิกฤต z_α มีค่าเท่ากับ $qnorm(1 - \alpha, 0, 1)$ ตัวอย่างการคำนวณ

$$\begin{aligned} pnorm(1.312, 0, 1) &= 0.905240 & qnorm(1 - 0.025, 0, 1) &= 1.959964 \\ 1 - pnorm(1.312, 0, 1) &= 0.094760 & qnorm(1 - 0.005, 0, 1) &= 2.575829 \\ qnorm(0.094760, 0, 1) &= -1.312000 & qnorm(1 - 0.05, 0, 1) &= 1.644854 \end{aligned}$$

เพราะฉะนั้น $z_{0.025} = 1.96, z_{0.05} = 1.645, z_{0.005} = 2.58$

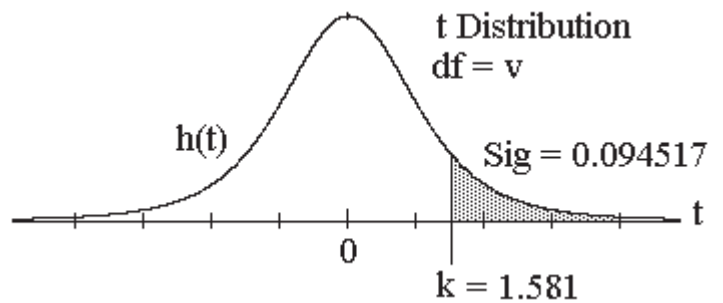
2. การหาค่า Significant ของค่าสถิติ t

ค่า Significant ของค่าสถิติ $t = k$ คำนวณมาจากค่าของความน่าจะเป็นหรือพื้นที่ใต้โค้งทางทางด้านขวาของโค้งการแจกแจงความน่าจะเป็นของตัวแปรสุ่ม t ตั้งแต่ $t = |k|$ ถึง $t = \infty$

t distribution

$$v := 4$$

$$h(t) := \left(\frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right) \cdot \sqrt{\pi \cdot v}} \right) \cdot \left[1 + \left(\frac{t^2}{v} \right) \right]^{-\frac{v+1}{2}}$$



หมายเหตุ เราใช้สัญลักษณ์ Sig. แทนค่า Significant ของค่าสถิติ $t = k$ และ $Sig. (2-tailed) = 2 Sig.$

$$Sig. = P(|k| < t < \infty) = \int_{t=|k|}^{\infty} h(t) dt = 0.5 - \int_0^{t=|k|} h(t) dt$$

ตัวอย่างการคำนวณเช่น

t distribution

$$v := 4$$

$$h(t) := \left(\frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right) \cdot \sqrt{\pi \cdot v}} \right) \cdot \left[1 + \left(\frac{t^2}{v} \right) \right]^{-\frac{v+1}{2}}$$

$$Sig := 0.5 - \int_0^{1.581} h(t) dt$$

$$Sig = 0.094517$$

$$\text{คำสั่งของ Mathcad : } pt(k, df) = P(-\infty < t < k) = \int_{-\infty}^k h(t) dt$$

เพราะฉะนั้น Significant ของค่าสถิติ $t = k$ จะมีค่าเท่ากับ $1 - pt(|k|, df)$

คำสั่งที่ช่วยในการคำนวณค่าสถิติ t ของ Mathcad

dt(t, df)	ฟังก์ชันการแจกแจงความน่าจะเป็นของตัวแปรสุ่ม t
pt(k, df)	$pt(k, df) = P(-\infty < t < k)$
qt(A, df)	ระยะบนแกน $t = k$ ที่ทำให้ $P(0 < t < k) = A$

หมายเหตุ ค่าวิกฤต $t_{\alpha, df}$ มีค่าเท่ากับ $qt(1 - \alpha, df)$ ตัวอย่างการคำนวณ

$$\begin{aligned}
 \text{pt}(1.581, 4) &= 0.905483 & \text{pt}(-1.581, 4) &= 0.094517 & 1 - \text{pt}(1.581, 4) &= 0.094517 \\
 \text{qt}(0.905482704) &= 1.581000 & \text{qt}(0.094517304) &= -1.581000 & & \\
 \text{qt}(1 - 0.05, 4) &= 2.131847 & \text{qt}(1 - 0.094517, 4) &= 1.581003 & \text{qt}(1 - 0.025, 4) &= 2.776445
 \end{aligned}$$

เพราะฉะนั้น $t_{0.05, 4} = 2.131847$, $t_{0.025, 4} = 2.776445$

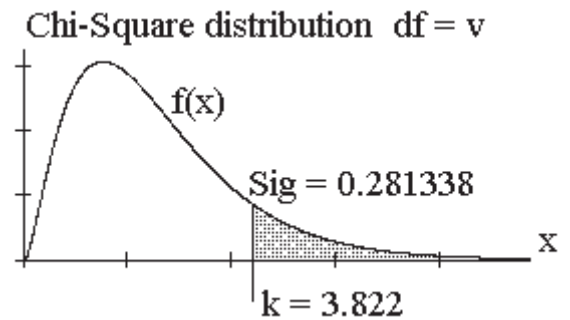
3. การหาค่า Significant ของค่าสถิติไคสแควร์

ค่า Significant ของค่าสถิติ $\chi^2 = k$ คำนวณมาจากค่าของความน่าจะเป็นหรือพื้นที่ใต้โค้งทางทางด้านขวาของโค้งการแจกแจงความน่าจะเป็นของตัวแปรสุ่มไคสแควร์ตั้งแต่ $\chi^2 = k$ ถึง $\chi^2 = \infty$

Chi-square distribution

$$v := 3$$

$$f(x) := \left(\frac{1}{2^{\frac{v}{2}} \cdot \Gamma\left(\frac{v}{2}\right)} \right) \cdot x^{\left(\frac{v}{2}\right)-1} \cdot e^{-\frac{x}{2}}$$



หมายเหตุ เราใช้สัญลักษณ์ Sig. แทนค่า Significant ค่าสถิติไคสแควร์ $\chi^2 = k$

$$\text{Sig.} = P(k < \chi^2 < \infty) = \int_k^\infty f(x) dx = 1 - \int_0^k f(x) dx$$

ตัวอย่างการคำนวณเช่น $v := 3$ $\text{TOL} := 0.000001$

$$\begin{aligned}
 f(x) &:= \left(\frac{1}{2^{\frac{v}{2}} \cdot \Gamma\left(\frac{v}{2}\right)} \right) \cdot x^{\left(\frac{v}{2}\right)-1} \cdot e^{-\frac{x}{2}} & \text{Sig} &:= 1 - \int_0^{3.822} f(x) dx \\
 & & \text{Sig} &= 0.281338
 \end{aligned}$$

หมายเหตุ Asymp. Sig. ของค่าสถิติไคสแควร์ = Sig. ของค่าสถิติไคสแควร์

คำสั่งที่ช่วยในการคำนวณค่าสถิติไคสแควร์ของ Mathcad

dchisq(x, df)	ฟังก์ชันการแจกแจงความน่าจะเป็นของตัวแปรสุ่มไคสแควร์
pchisq(k, df)	$\text{pchisq}(k, df) = P(k < \chi^2 < \infty)$
qchisq(A, df)	ระยะบนแกน $\chi^2 = k$ ที่ทำให้ $P(0 < \chi^2 < k) P(0 < \chi^2 < k) = A$

หมายเหตุ ค่าวิกฤต $\chi^2_{\alpha, df}$ มีค่าเท่ากับ $\text{qchisq}(1 - \alpha, df)$ ตัวอย่างการคำนวณ

$$\text{pchisq}(3.822, 3) = 0.718662 \quad 1 - \text{pchisq}(3.822, 3) = 0.281338 \quad \text{qchisq}(0.718662, 3) = 3.821996$$

$$\text{qchisq}(1 - 0.95, 3) = 0.3518 \quad \text{qchisq}(1 - 0.025, 3) = 9.3484 \quad \text{qchisq}(1 - 0.05, 3) = 7.8147$$

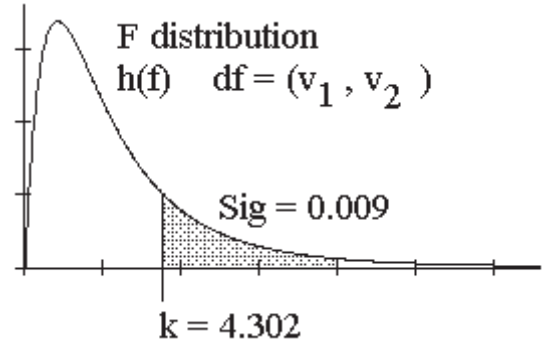
เพราะฉะนั้น $\chi^2_{0.95, 3} = 0.3528$, $\chi^2_{0.025, 3} = 9.3484$, $\chi^2_{0.05, 3} = 7.8147$

4. การหาค่า Significant ของค่าสถิติ F

ค่า Significant ของค่าสถิติ $F = k$ คำนวณมาจากค่าของความน่าจะเป็นหรือพื้นที่ใต้โค้งทางทางด้านขวาของโค้งการแจกแจงความน่าจะเป็นของตัวแปรสุ่มเอฟตั้งแต่ $F = k$ ถึง $F = \infty$

$v_1 := 4 \quad v_2 := 10$

$$h(f) := \frac{\Gamma\left(\frac{v_1 + v_2}{2}\right) \cdot \left(\frac{v_1}{v_2}\right)^{\frac{v_1}{2}} \cdot f^{\left(\frac{v_1}{2}\right)-1}}{\Gamma\left(\frac{v_1}{2}\right) \cdot \Gamma\left(\frac{v_2}{2}\right) \cdot \left[1 + \left(\frac{v_1}{v_2}\right) \cdot f\right]^{\frac{v_1+v_2}{2}}}$$



หมายเหตุ เราใช้สัญลักษณ์ Sig แทนค่า Significant ของค่าสถิติ $F = k$

$$\text{Sig.} = P(k < F < \infty) = \int_{F=k}^{\infty} h(f) df = 1 - \int_0^{F=k} h(f) df$$

ตัวอย่างการคำนวณเช่น

$v_1 := 4 \quad v_2 := 25$

$$h(f) := \frac{\Gamma\left(\frac{v_1 + v_2}{2}\right) \cdot \left(\frac{v_1}{v_2}\right)^{\frac{v_1}{2}} \cdot f^{\left(\frac{v_1}{2}\right)-1}}{\Gamma\left(\frac{v_1}{2}\right) \cdot \Gamma\left(\frac{v_2}{2}\right) \cdot \left[1 + \left(\frac{v_1}{v_2}\right) \cdot f\right]^{\frac{v_1+v_2}{2}}}$$

$$\text{Sig} := 1 - \int_0^{4.302} h(f) df \quad \text{Sig} = 0.008747$$

คำสั่งที่ช่วยในการคำนวณค่าสถิติ F ของ Mathcad

$dF(x, df1, df2)$	ฟังก์ชันการแจกแจงความน่าจะเป็นของตัวแปรสุ่ม F
$pF(k, df1, df2)$	$pF(k, df1, df2) = P(k < F < k)$
$qF(A, df1, df2)$	ระยะบนแกน $F = k$ ที่ทำให้ $P(k < F < k) = A$

หมายเหตุ ค่าวิกฤต $f_{\alpha, (df1, df2)}$ มีค่าเท่ากับ $qF(1-\alpha, df1, df2)$ ตัวอย่างการคำนวณ

$pF(4.302, 4, 25) = 0.991253 \quad 1 - pF(4.302, 4, 25) = 0.008747 \quad qF(0.991253, 4, 25) = 4.302032$

$qF(1 - 0.05, 4, 24) = 2.776289 \quad qF(1 - 0.025, 4, 24) = 3.379359 \quad qF(1 - 0.01, 4, 24) = 4.218445$

เพราะฉะนั้น $f_{0.05, (4, 24)} = 2.776289, f_{0.025, (4, 24)} = 3.379359, f_{0.01, (4, 24)} = 4.218445$

ภาคผนวกที่ 2

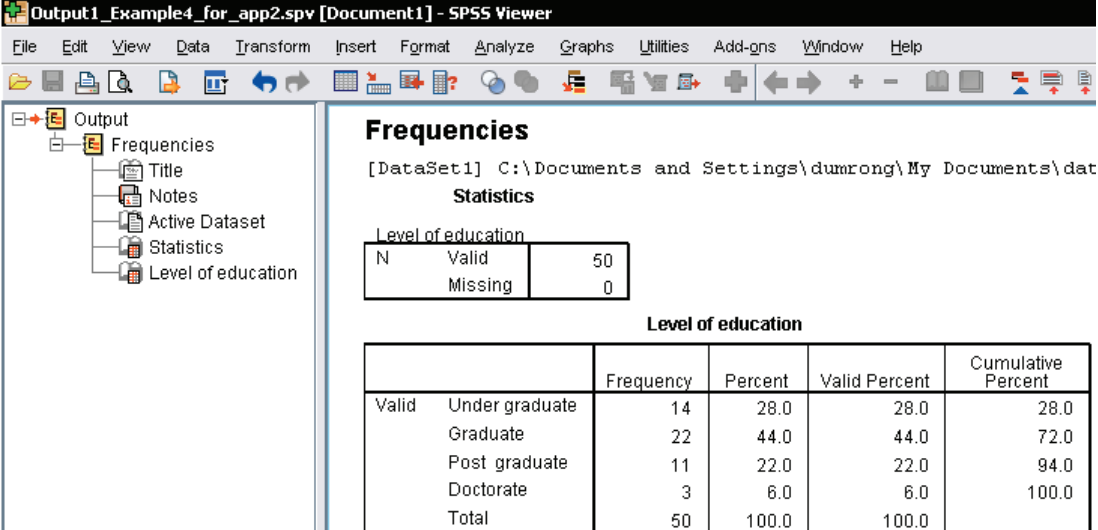
การเชื่อมโยงข้อมูล SPSS for Windows กับ Microsoft Word

จากผลการวิเคราะห์ข้อมูลของ SPSS for Windows เราสามารถนำผลการคำนวณไปไว้ที่ Microsoft Word ได้ทั้งในรูปแบบข้อความ ตาราง และ ข้อมูล

1. การนำตารางการวิเคราะห์จาก SPSS ไป Word

1.1 การ Copy ตารางจาก SPSS Viewer มาไว้ที่ Word ในรูปแบบ Object

เปิดแฟ้มข้อมูล example4.sav ใน SPSS Data Editor วิเคราะห์ข้อมูลด้วยคำสั่ง Analyze / Descriptive Statistics / Frequencies เพื่อทำการแจกแจงความถี่ของตัวแปร educ ได้ผลบนจอภาพ SPSS Viewer ดังนี้

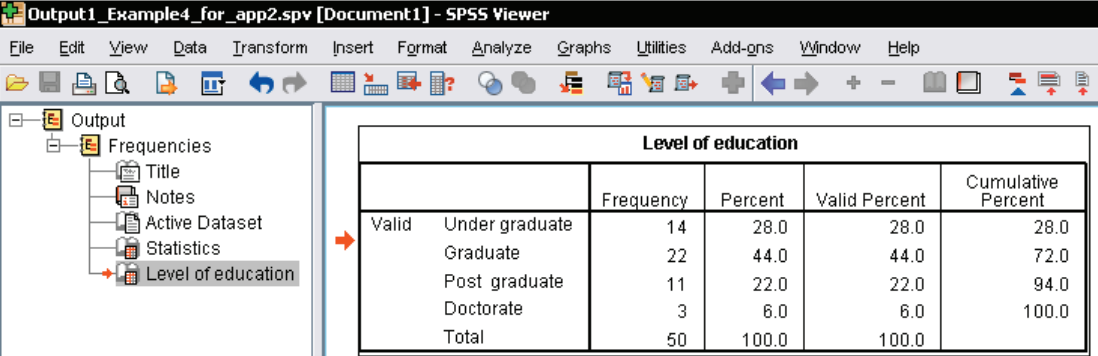


The screenshot shows the SPSS Viewer interface with the 'Frequencies' output for the variable 'Level of education'. The output includes a summary table and a detailed table with columns for Frequency, Percent, Valid Percent, and Cumulative Percent.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Under graduate	14	28.0	28.0	28.0
	Graduate	22	44.0	44.0	72.0
	Post graduate	11	22.0	22.0	94.0
	Doctorate	3	6.0	6.0	100.0
	Total	50	100.0	100.0	

เราต้องการตารางการแจกแจงความถี่ของตัวแปร ระดับการศึกษา (educ) ไปไว้ที่ Microsoft Word

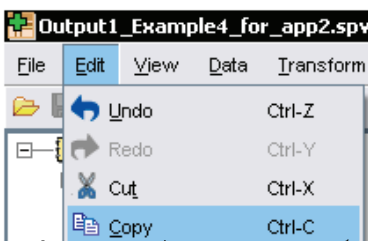
ขั้นที่ 1. คลิกที่ตารางผลการวิเคราะห์ความถี่ของตัวแปรระดับการศึกษา



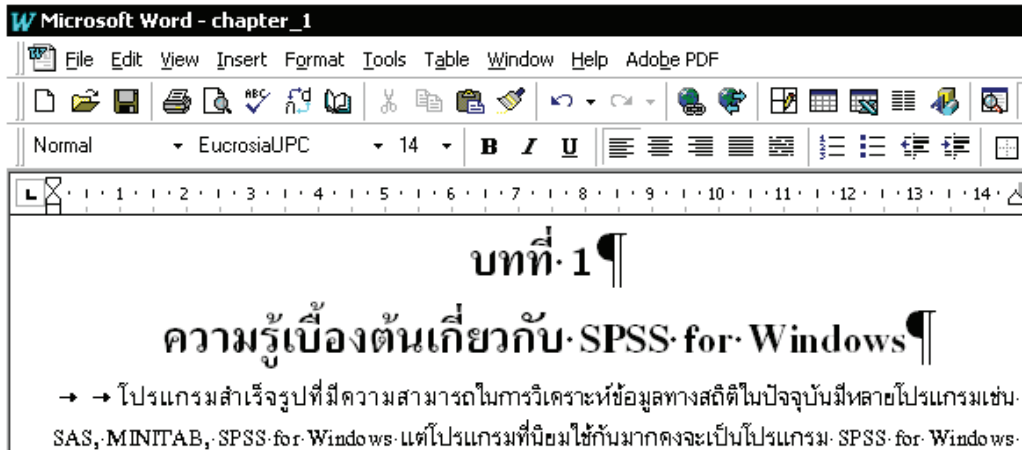
The screenshot shows the same SPSS Viewer interface, but with the 'Level of education' table selected in the output window. A red arrow points to the table, indicating it is the target for copying.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Under graduate	14	28.0	28.0	28.0
	Graduate	22	44.0	44.0	72.0
	Post graduate	11	22.0	22.0	94.0
	Doctorate	3	6.0	6.0	100.0
	Total	50	100.0	100.0	

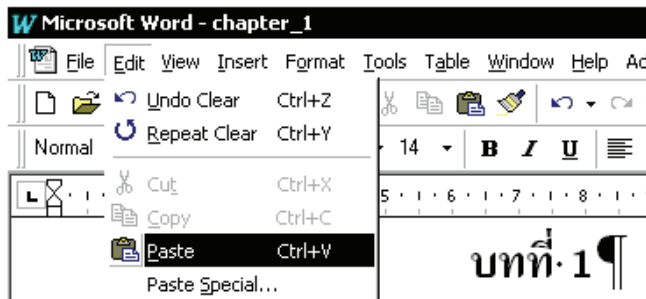
ขั้นที่ 2. คลิกที่ Edit และ Copy



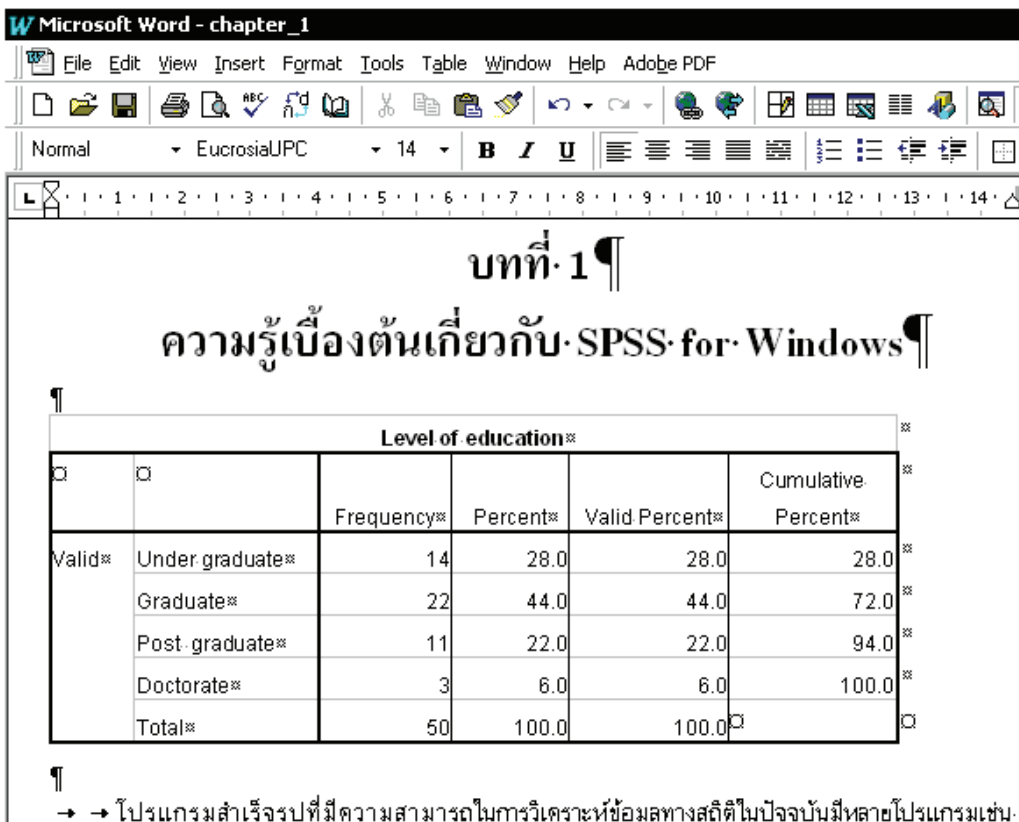
ขั้นที่ 3. ไปที่ Window การทำงานของ Microsoft Word



ขั้นที่ 4. ขณะที่อยู่ใน Microsoft Word ให้คลิกเมาส์ตรงตำแหน่งที่ต้องการแทรกตาราง ใช้คำสั่ง Edit และ Paste หรือ กด Ctrl + V

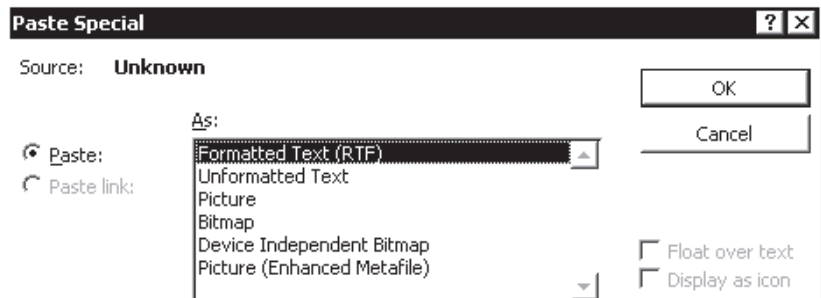


จะได้ตารางของการวิเคราะห์หม่าอยู่ใน Window ของ Microsoft Word ที่เราทำขณะนั้น โดยรูปแบบตารางจะเป็น Text หรือ Table ในรูปแบบของ Text



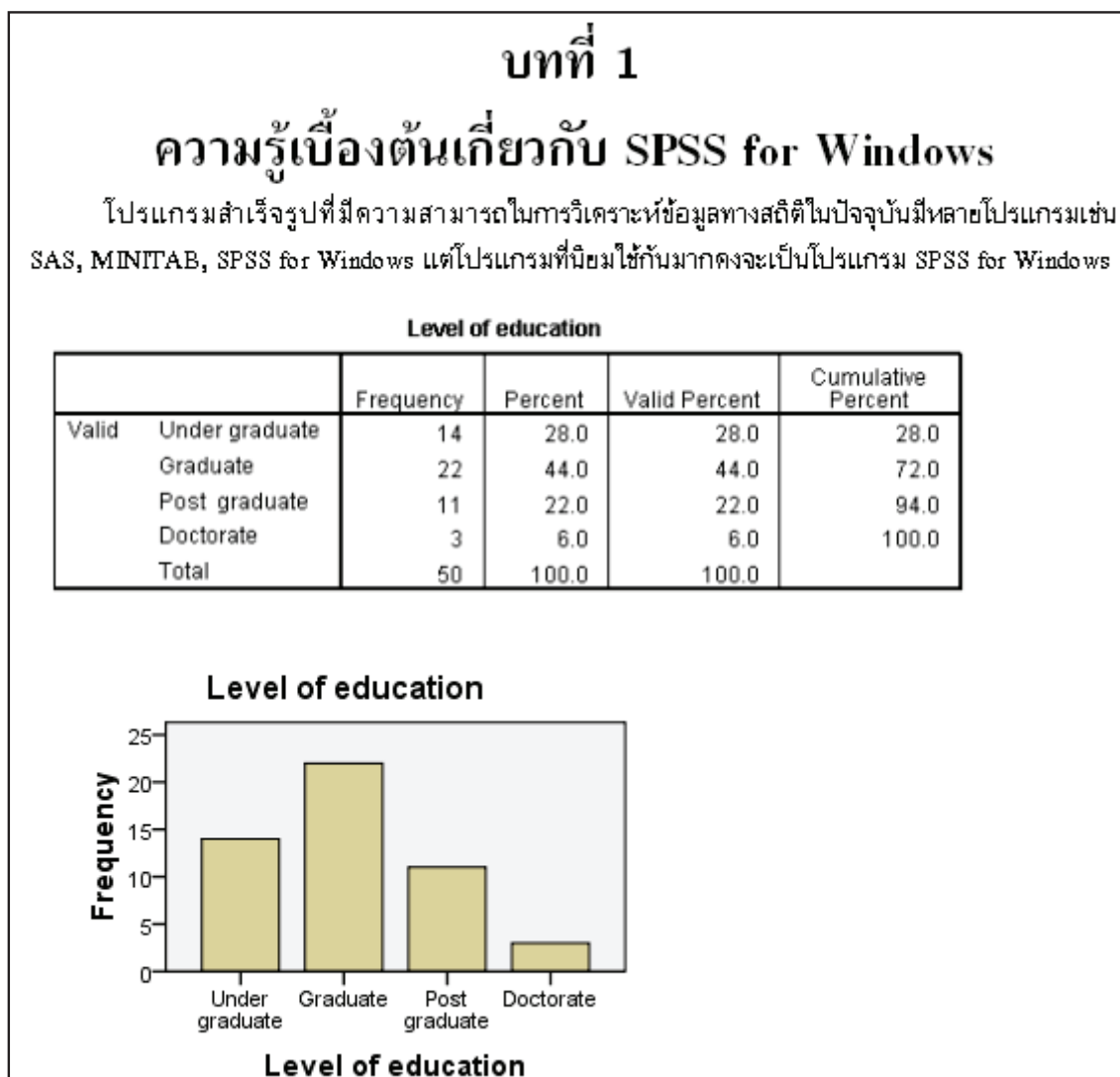
1.2 การ Copy ตารางจาก SPSS Viewer มาไว้ที่ Word ในรูปแบบ Object

- ไปที่ SPSS Viewer **ขั้นที่ 1.** คลิกที่ตารางผลการวิเคราะห์ความถี่ ใน SPSS Viewer
- ขั้นที่ 2.** คลิกที่ Edit และ Copy
- ไปที่ Microsoft Word **ขั้นที่ 3.** ไปหน้าของ Word ที่ต้องการแทรกตาราง
- ขั้นที่ 4.** คลิกเมาส์ตรงที่ต้องการแทรกตาราง ใช้คำสั่ง Edit และ Paste Spacial
จะได้เมนูย่อยเป็น



- ขั้นที่ 5.** เลือก Picture หรือ Bitmap จะได้ตารางในรูปแบบ Object มาแทรกกับข้อความของ Word

หมายเหตุ การ Copy รูปภาพจาก SPSS View มาไว้ที่ Word ทำได้ในทำนองเดียวกัน
ตัวอย่างการ Copy ตาราง และ กราฟแท่ง จาก SPSS Viewer มาไว้ที่ Microsoft Word

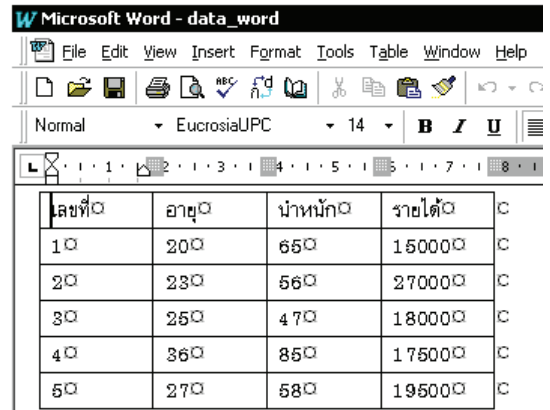
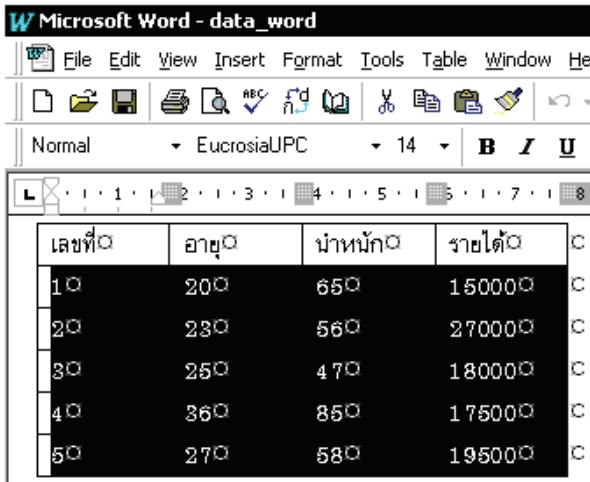


2. การนำข้อมูลจากตารางของ Word ไปเป็นข้อมูลของ SPSS

เราสามารถนำข้อมูลจากตารางของ Word ไปเป็นข้อมูลของ SPSS ได้ตามขั้นตอนดังนี้ ตัวอย่างเช่นเราต้องการหา อายุเฉลี่ย น้ำหนักเฉลี่ย และรายได้เฉลี่ย ของคน 5 คนจากตารางใน Word

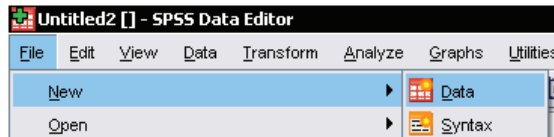
ขั้นที่ 1. เลือกข้อมูลจากตารางของ Word

แล้วกด Ctrl + C (ทำการ Copy)

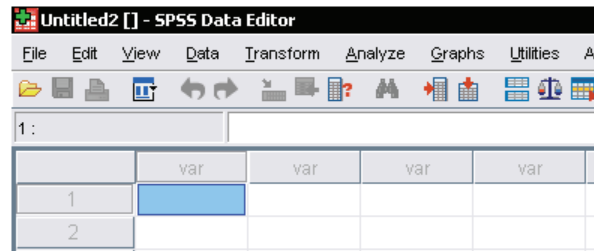


ขั้นที่ 2. ไปที่ SPSS Data Editor

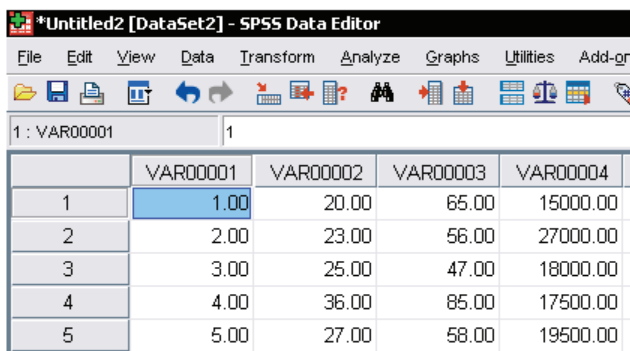
เลือกเมนู File / New / Data และ คลิก Data



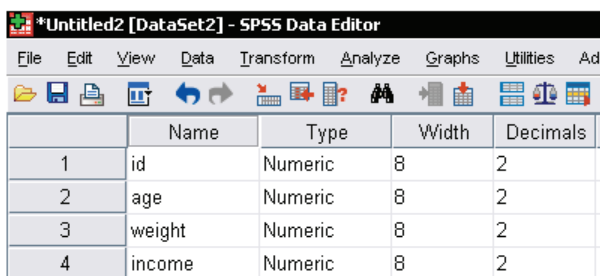
จะได้จอภาพใน SPSS Data Editor เป็นดังนี้



ขั้นที่ 3. กด Ctrl+V เพื่อเอาข้อมูลที่เร Copy มาจาก Word วางลงใน SPSS



ขั้นที่ 4. เปลี่ยนชื่อตัวแปรใน Variable View ให้เหมาะสมกับข้อมูล



กลับมาที่ Data View จะได้ผลดังนี้

	id	age	weight	income	vs
1	1.00	20.00	65.00	15000.00	
2	2.00	23.00	56.00	27000.00	
3	3.00	25.00	47.00	18000.00	
4	4.00	36.00	85.00	17500.00	
5	5.00	27.00	58.00	19500.00	

ขั้นที่ 5. วิเคราะห์ข้อมูลหาค่าเฉลี่ยของ age, weight และ income

Descriptives
[DataSet2] C:\Documents and Settings\dumrong\My Documents\datas

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
age	5	20.00	36.00	26.2000	6.05805
weight	5	47.00	85.00	62.2000	14.27235
income	5	15000.00	27000.00	19400.0000	4546.97702
Valid N (listwise)	5				

ขั้นที่ 6. Copy ตารางจาก SPSS Viewer ด้วยคำสั่ง Copy Objects ที่วิเคราะห์ได้มาไว้ที่ Microsoft Word จะได้ผลดังนี้

Microsoft Word - data_word

เลขที่	อายุ	น้ำหนัก	รายได้
1	20	65	15000
2	23	56	27000
3	25	47	18000
4	36	85	17500
5	27	58	19500

Descriptive Statistics**

	N*	Minimum*	Maximum*	Mean*	Std. Deviation*
age**	5	20.00	36.00	26.2000	6.05805
weight**	5	47.00	85.00	62.2000	14.27235
income**	5	15000.00	27000.00	19400.0000	4546.97702
Valid N (listwise)**	5				

3. การนำข้อมูล Data จาก SPSS Data Editor มาทำงานที่ Microsoft Word

การนำข้อมูลที่เป็น Data จาก SPSS Data Editor มาที่ Word ทำได้ดังนี้

จากแฟ้มข้อมูล example4.sav ใน SPSS Data Editor

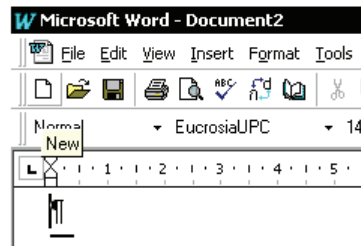
	id	sex	age	educ	status	income
1	1	1	37	2	4	5500
2	2	2	29	3	1	4100
3	3	2	48	1	2	5400

ขั้นที่ 1. เลือกบริเวณของข้อมูลที่ต้องการ

	id	sex	age	educ	status	income	grade
1	1	1	37	2	4	5500	3.78
2	2	2	29	3	1	4100	3.89
3	3	2	48	1	2	5400	3.67
4	4	1	99	1	2	9999	2.78

ขั้นที่ 2. กด Ctrl + C เพื่อ Copy

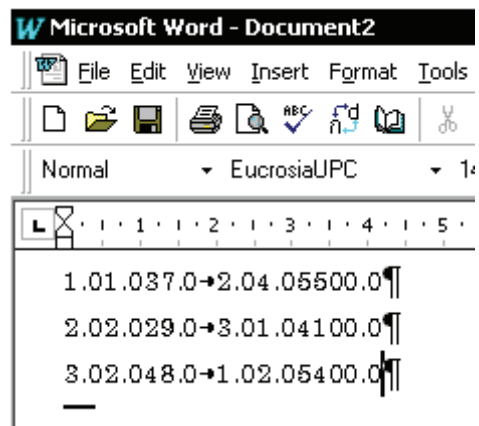
หรือใช้คำสั่ง Edit / Copy



ขั้นที่ 3. ไปที่โปรแกรม Microsoft Word

ขั้นที่ 4. แล้วกด Ctrl + V เพื่อเอาข้อมูลที่ Copy

ไว้มาทำงานต่อใน Microsoft Word



ขั้นที่ 5. หากต้องการใช้ข้อมูลในรูปแบบตาราง

ให้ใช้ความสามารถของ Microsoft Word

โดยใช้คำสั่ง Convert Texts to Table

1.0	1.0	37.0	2.0	4.0	5500.0
2.0	2.0	29.0	3.0	1.0	4100.0
3.0	2.0	48.0	1.0	2.0	5400.0

ภาคผนวกที่ 3

การเชื่อมโยงข้อมูล SPSS for Windows กับ Excel

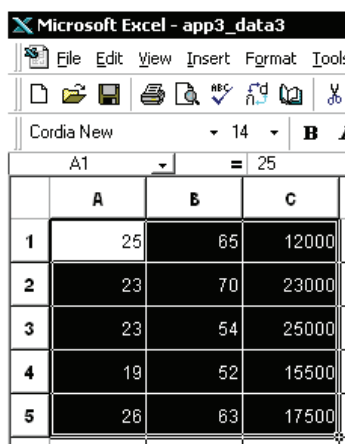
เนื่องจากโปรแกรม Excel เป็นโปรแกรมที่มีผู้นิยมใช้กันมากและอาจเก็บข้อมูลที่ต้องการวิเคราะห์ในรูปแบบของ Excel Sheet และถึงแม้ว่าโปรแกรม Excel จะมีความสามารถในการวิเคราะห์ข้อมูลทางด้านสถิติ แต่ถ้าเราสามารถเชื่อมโยง ข้อมูลและผลลัพธ์การวิเคราะห์ข้อมูล ระหว่าง SPSS for Windows กับ Excel ได้ ก็จะเป็นประโยชน์อย่างมากในการทำงาน

1. การนำข้อมูลจาก Excel ไปเป็นข้อมูลของ SPSS for Windows

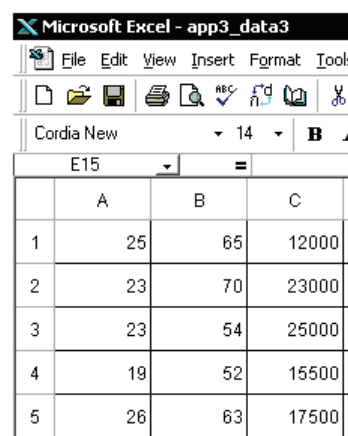
ข้อมูลทางสถิติเราสามารถนำข้อมูลที่สร้างไว้ด้วยโปรแกรม Excel มาวิเคราะห์ด้วย SPSS for Windows ได้ ตัวอย่างเช่น เรามีข้อมูลใน Excel ดังนี้ เราต้องการนำข้อมูลจาก Excel ไปวิเคราะห์ที่ SPSS

ขั้นที่ 1.

เลือกบริเวณที่ต้องการ Copy ใน Excel



	A	B	C
1	25	65	12000
2	23	70	23000
3	23	54	25000
4	19	52	15500
5	26	63	17500



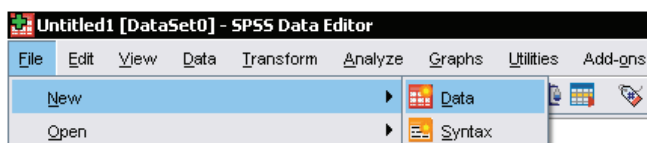
	A	B	C
1	25	65	12000
2	23	70	23000
3	23	54	25000
4	19	52	15500
5	26	63	17500

ขั้นที่ 2. Copy ใน Excel

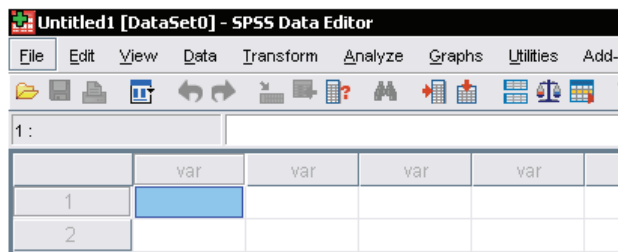
โดยการกด Ctrl + C

ขั้นที่ 3. ไปที่ SPSS Data Editor

เลือกเมนู File / New / Data



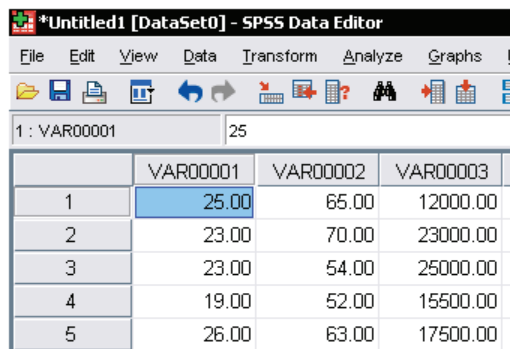
คลิกที่ Data จะได้จอภาพเป็น



	var	var	var	var
1				
2				

ขั้นที่ 4. กด Ctrl + V เพื่อนำข้อมูลที่

Copy มาจาก Excel วางลงในตารางของ SPSS



	VAR00001	VAR00002	VAR00003
1	25.00	65.00	12000.00
2	23.00	70.00	23000.00
3	23.00	54.00	25000.00
4	19.00	52.00	15500.00
5	26.00	63.00	17500.00

คลิก Variable View

	Name	Type	Width	Decimals
1	VAR00001	Numeric	8	2
2	VAR00002	Numeric	8	2
3	VAR00003	Numeric	8	2

ขั้นที่ 5.

เปลี่ยนชื่อตัวแปรตามความเหมาะสม

	Name	Type	Width	Decimals
1	age	Numeric	9	0
2	weight	Numeric	4	2
3	income	Numeric	9	2

คลิก Data View เพื่อกลับจอบทภาพของ Data
จะได้ผลบจนจอบทภาพดังนี้

	age	weight	income
1	25	65.00	12000.00
2	23	70.00	23000.00
3	23	54.00	25000.00
4	19	52.00	15500.00
5	26	63.00	17500.00

ขั้นที่ 6. วิเคราะห์ข้อมูลเพื่อหาค่าเฉลี่ยของทุกตัวแปร

Descriptives
[DataSet0]

	N	Minimum	Maximum	Mean	Std. Deviation
age	5	19	26	23.20	2.683
weight	5	52.00	70.00	60.8000	7.59605
income	5	12000.00	25000.00	18600.0000	5354.90429
Valid N (listwise)	5				

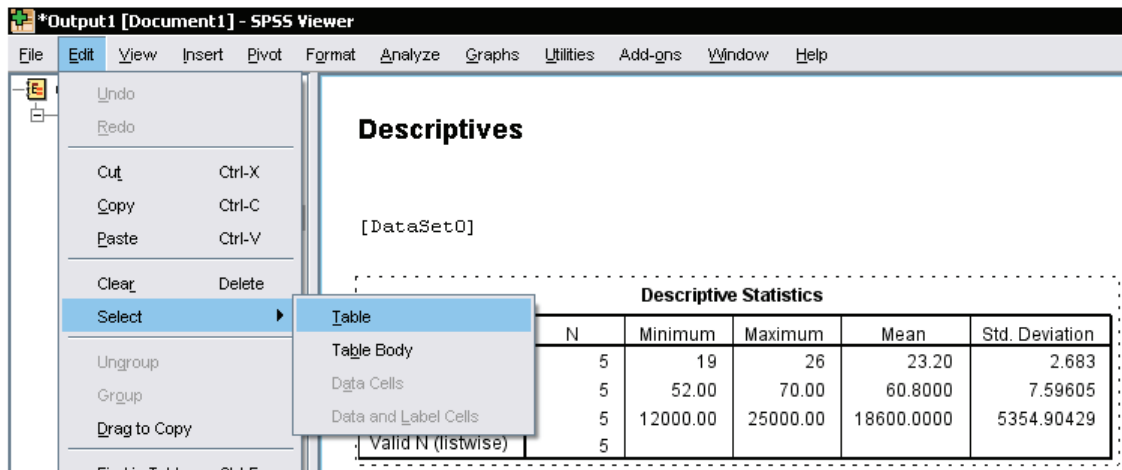
2. การนำผลการวิเคราะห์ข้อมูลของ SPSS กลับไปที่ Excel

ขั้นที่ 1. ใน SPSS Viewer ไปที่ตารางที่ต้องการ แล้ว กดดับเบิลคลิก

Descriptives
[DataSet0]

	N	Minimum	Maximum	Mean	Std. Deviation
age	5	19	26	23.20	2.683
weight	5	52.00	70.00	60.8000	7.59605
income	5	12000.00	25000.00	18600.0000	5354.90429
Valid N (listwise)	5				

ขั้นที่ 2. เลือกเมนู Edit / Select / Table

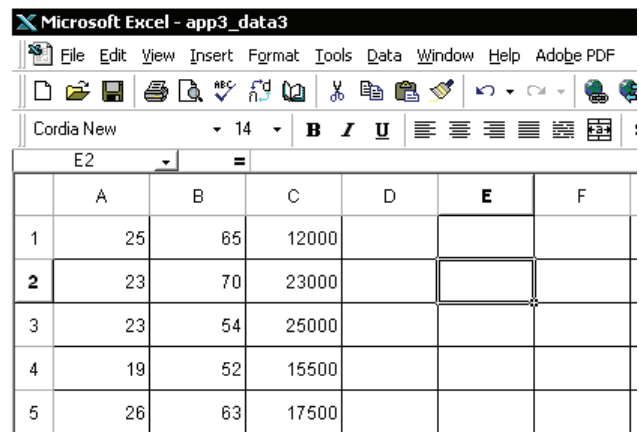


จะได้ผลที่ตารางที่เลือกไว้เป็นดังนี้

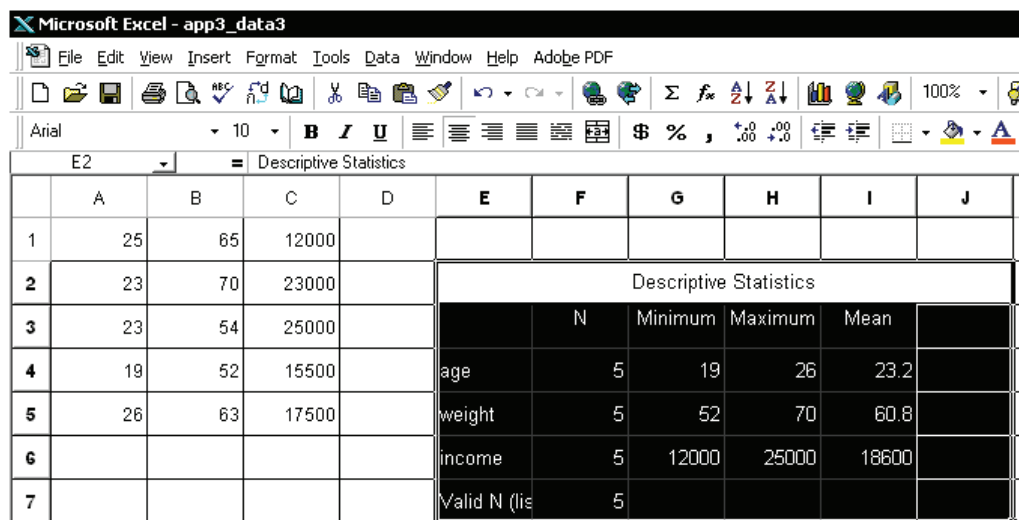


ขั้นที่ 3. กด Ctrl + C เพื่อ Copy ตาราง

ขั้นที่ 4. ไปที่ Excel และหาดำแหน่งที่เหมาะสมที่จะวางตารางการวิเคราะห์ข้อมูลที่เรา Copy ไว้แล้ว เช่นที่ตำแหน่ง Cell E2



ขั้นที่ 5. คลิกคำสั่ง Edit / Paste จะได้ผลบนจอภาพ Excel เป็นดังนี้



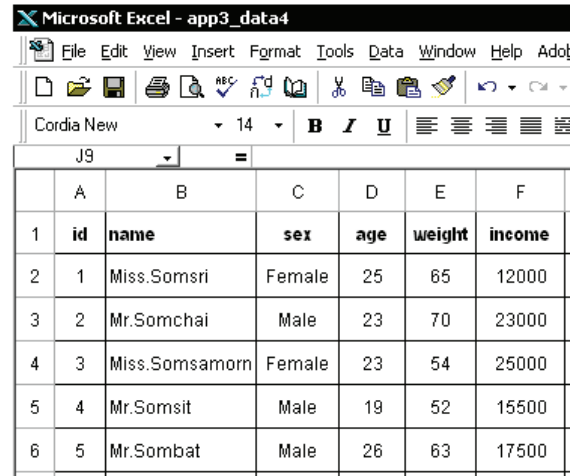
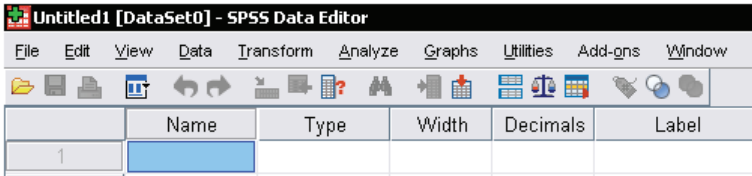
3. การนำข้อมูลแบบ String และ Number จาก Excel ไปไว้ที่ SPSS for Windows

ขั้นที่ 1. สร้างตัวอย่างข้อมูลที่ Excel เป็นดังนี้

จากข้อมูลใน Excel ที่มีทั้ง String และ Numerics
การนำข้อมูลจาก Excel ไปที่ SPSS for Windows
เราต้องเตรียมตัวแปรที่ SPSS ให้เหมาะสมกับตัวแปร

ขั้นที่ 2. ไปที่ SPSS Data Editor

คลิกที่ Variable View



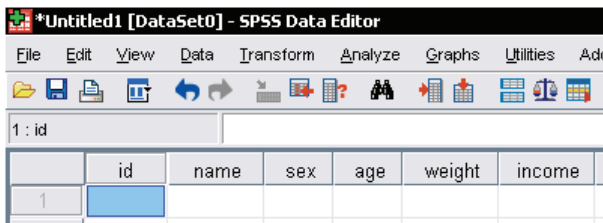
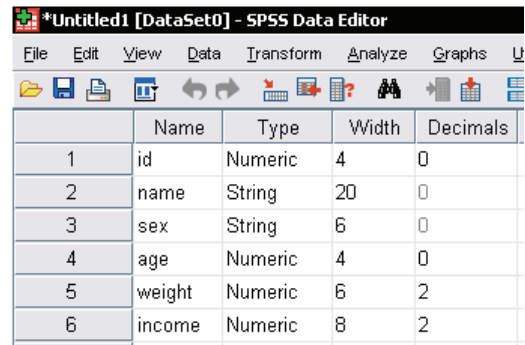
ขั้นที่ 3. กำหนดตัวแปรให้

เหมาะสมกับข้อมูล

เช่นชื่ออาจต้องกำหนดความกว้าง

width = 20

ขั้นที่ 4. คลิก Data View



ขั้นที่ 5. เพื่อถ่ายต่อความเข้าใจขอให้ทำการ

Copy ตัวแปรครั้งละตัวจาก Excel มา Paste ที่ SPSS

ขั้นที่ 5.1 Copy ที่ column ของตัวแปร id ดูจากที่แรก

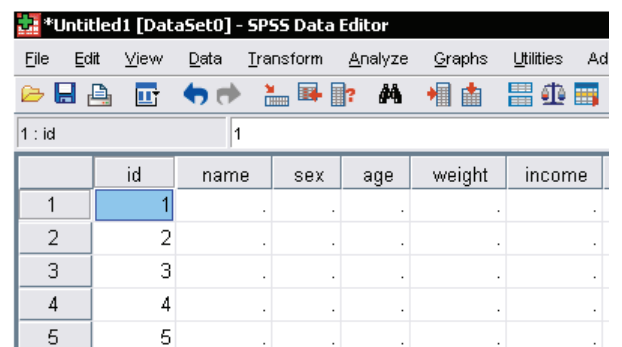
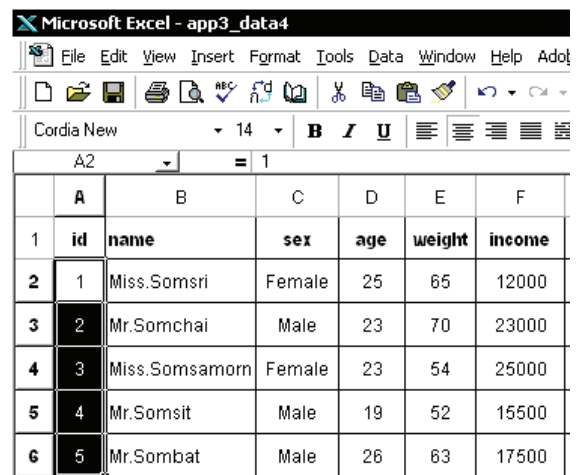
ใน Excel

ขั้นที่ 5.2 ไปที่ SPSS Data Editor

เลื่อนเมาส์ไปที่ตำแหน่งค่าสังเกตแรกของตัวแปร id

ใช้คำสั่ง Edit / Paste

จะได้ผลบนจอภาพดังนี้



หมายเหตุ ข้อมูลประเภทเดียวกันเราสามารถ Copy พร้อมกันจาก Excel แล้วมา Paste ที่ SPSS ได้

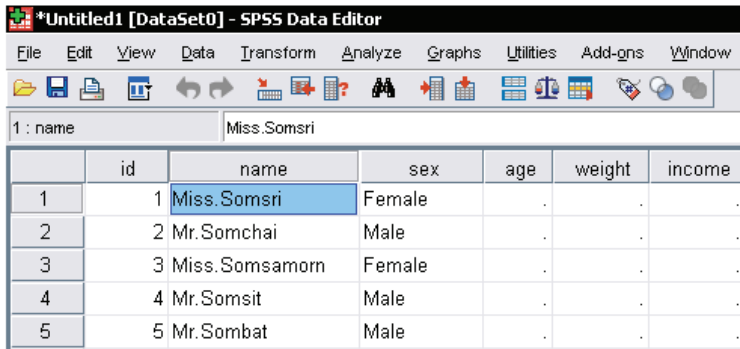
ขั้นที่ 5.3 กลับไปที่ Excel

Copy ที่ข้อมูลตัวแปร name และ sex ดูจากที่แรก

ขั้นที่ 5.4 ไปที่ SPSS Data Editor

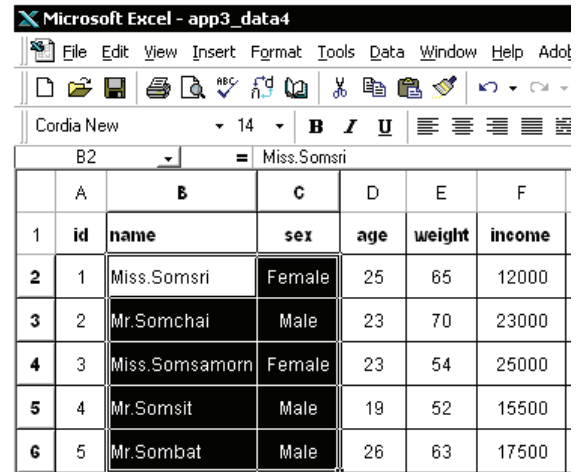
เลื่อนเมาส์ไปที่ตำแหน่งค่าสังเกตแรกของตัวแปร name

ใช้คำสั่ง Edit / Paste จะได้ผลบจนจภาพดังนี้



*Untitled1 [DataSet0] - SPSS Data Editor

	id	name	sex	age	weight	income
1	1	Miss.Somsri	Female			
2	2	Mr.Somchai	Male			
3	3	Miss.Somsamorn	Female			
4	4	Mr.Somsit	Male			
5	5	Mr.Sombat	Male			



Microsoft Excel - app3_data4

	A	B	C	D	E	F
1	id	name	sex	age	weight	income
2	1	Miss.Somsri	Female	25	65	12000
3	2	Mr.Somchai	Male	23	70	23000
4	3	Miss.Somsamorn	Female	23	54	25000
5	4	Mr.Somsit	Male	19	52	15500
6	5	Mr.Sombat	Male	26	63	17500

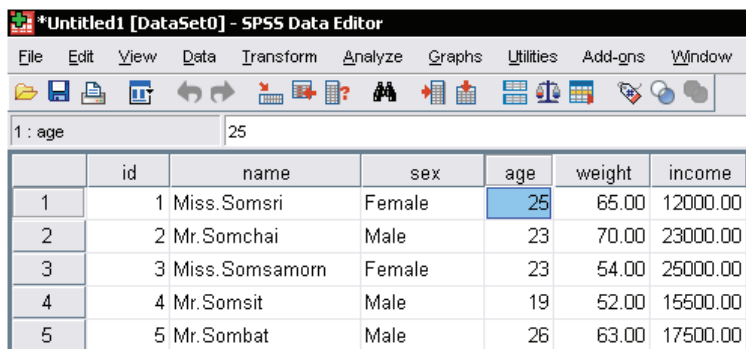
ขั้นที่ 5.5 กลับไปที่ Excel

Copy ที่ข้อมูลของตัวแปร age, weight, income ดูจากที่แรก

ขั้นที่ 5.6 ไปที่ SPSS Data Editor

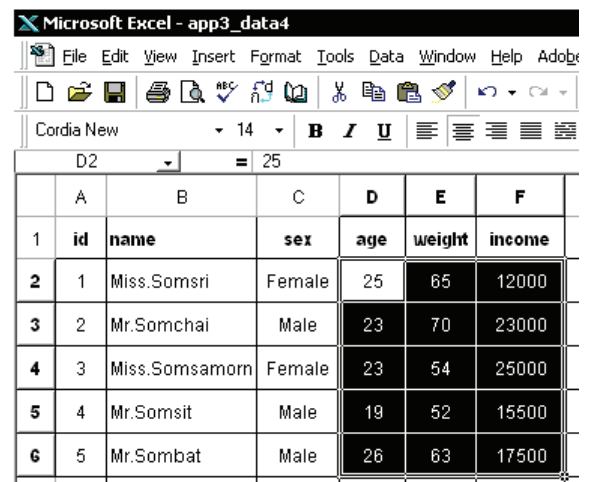
เลื่อนเมาส์ไปที่ตำแหน่งค่าสังเกตแรกของตัวแปร age

ใช้คำสั่ง Edit / Paste จะได้ผลบจนจภาพดังนี้



*Untitled1 [DataSet0] - SPSS Data Editor

	id	name	sex	age	weight	income
1	1	Miss.Somsri	Female	25	65.00	12000.00
2	2	Mr.Somchai	Male	23	70.00	23000.00
3	3	Miss.Somsamorn	Female	23	54.00	25000.00
4	4	Mr.Somsit	Male	19	52.00	15500.00
5	5	Mr.Sombat	Male	26	63.00	17500.00



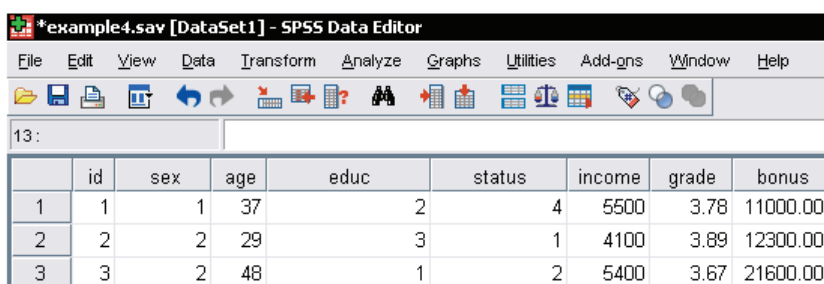
Microsoft Excel - app3_data4

	A	B	C	D	E	F
1	id	name	sex	age	weight	income
2	1	Miss.Somsri	Female	25	65	12000
3	2	Mr.Somchai	Male	23	70	23000
4	3	Miss.Somsamorn	Female	23	54	25000
5	4	Mr.Somsit	Male	19	52	15500
6	5	Mr.Sombat	Male	26	63	17500

ขณะนี้เรามีข้อมูลที่พร้อมจะวิเคราะห์ข้อมูลด้วย SPSS แล้ว

4. การนำข้อมูลจาก SPSS for Windows ไปไว้ที่ Excel

จากจอภาพ SPSS Data Editor



*example4.sav [DataSet1] - SPSS Data Editor

	id	sex	age	educ	status	income	grade	bonus
1	1	1	37	2	4	5500	3.78	11000.00
2	2	2	29	3	1	4100	3.89	12300.00
3	3	2	48	1	2	5400	3.67	21600.00

ขั้นที่ 1. เลือกบริเวณที่
ต้องการใน SPSS

	id	sex	age	educ	status	income	grade	bonus
1	1	1	37	2	4	5500	3.78	11000.00
2	2	2	29	3	1	4100	3.89	12300.00
3	3	2	48	1	2	5400	3.67	21600.00
4	4	1	99	1	2	9999	2.78	19998.00
5	5	2	33	2	9	9999	3.00	29997.00

ขั้นที่ 2. ไปที่ Excel เลือกตำแหน่งที่ต้องการ Paste ข้อมูลที่ Copy มาจาก SPSS
เช่นเลือกตำแหน่ง Cell A1

	A	B	C	D	E	F	G
1							
2							

ขั้นที่ 3. คลิกคำสั่ง Edit / Paste จะได้ผลดังนี้

	A	B	C	D	E	F	G	H
1	1	1	37	2	4	5500	3.78	11000
2	2	2	29	3	1	4100	3.89	12300
3	3	2	48	1	2	5400	3.67	21600
4	4	1	99	1	2	9999	2.78	19998
5	5	2	33	2	9	9999	3	29997

จัดรูปแบบให้สวยงามด้วยความสามารถของ Excel ตัวอย่างเช่น

	A	B	C	D	E	F	G	H
1	1	1	37	2	4	5500	3.78	11000
2	2	2	29	3	1	4100	3.89	12300
3	3	2	48	1	2	5400	3.67	21600
4	4	1	99	1	2	9999	2.78	19998
5	5	2	33	2	9	9999	3	29997

ภาคผนวกที่ 4

SPSS Syntax Editor กับ โปรแกรมภาษา SPSS

โปรแกรมภาษา SPSS เป็นเรื่องที่มีประโยชน์มากสำหรับการวิเคราะห์ข้อมูลที่มีขั้นตอนการสั่งงานหลายขั้นตอน หรือ การทำงานที่ต้องมีการทำอยู่เป็นประจำเช่น การตรวจสอบมาตรฐานสินค้า การวิเคราะห์คะแนนสอบ วิเคราะห์ยอดการขายสินค้า ฯลฯ งานต่างๆ เหล่านี้เราสามารถนำโปรแกรม Syntax ของ SPSS มาช่วยในการทำงานได้ นอกจากนั้นในบางหน่วยงานอาจจะมีโปรแกรมภาษา SPSS ที่ยังมีความจำเป็นที่จะต้องใช้งาน และต้องการผู้ที่ใช้โปรแกรมภาษา SPSS เหล่านี้เป็น

1. การบันทึกคำสั่งต่างๆ ของการวิเคราะห์ข้อมูลเป็น Syntax

ใน SPSS รุ่นแรกๆ เช่น SPSS/PC version 3 ผู้ที่จะวิเคราะห์ข้อมูลด้วย SPSS ต้องเขียนโปรแกรมภาษา SPSS ได้บางพอสมควร แต่ในปัจจุบันเราสามารถสั่งให้ SPSS ทำการบันทึกคำสั่งต่างๆ ที่เราสั่งให้โปรแกรม Syntax แล้วพิมพ์ไว้ที่ SPSS Syntax Editor ตัวอย่างเช่นการบันทึกคำสั่งการวิเคราะห์ของตัวแปร age ในข้อมูลของแฟ้ม example4.sav เป็น Syntax

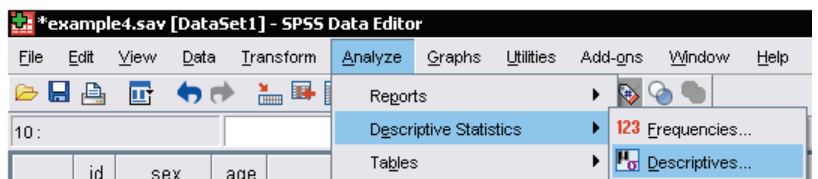
ขั้นที่ 1. เปิดแฟ้ม example4.sav

คลิกเมาส์ที่ Analyze

บนเมนูบาร์จอภาพจะเป็นดังนี้

และ คลิก Descriptive Statistics

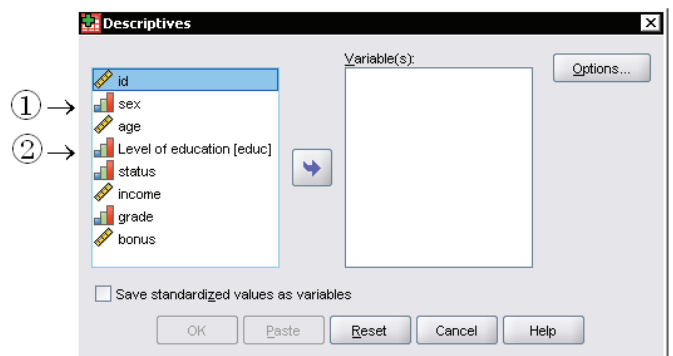
จอภาพเป็นดังนี้



ขั้นที่ 2. คลิก Descriptives จะได้เมนูย่อย

หมายเหตุ

1. เครื่องหมายแสดงชนิดตัวแปรว่าเป็นข้อมูลตัวเลข
2. ตัวแปรที่มี Value Labels จะแสดงค่า Value Label เช่นตัวแปร educ



ขั้นที่ 3. การเลือกตัวแปร age เพื่อทำการคำนวณ

ให้นำเมาส์ไปคลิกที่ตัวแปร age

แล้วคลิกที่ปุ่ม  เพื่อย้ายตัวแปร age ไปทางขวา

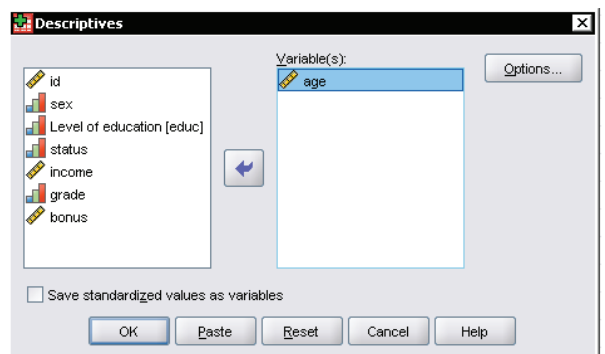
บนจอภาพจะกลายเป็น

หมายเหตุ ถ้าเราคลิก OK ขณะนี้โปรแกรม SPSS

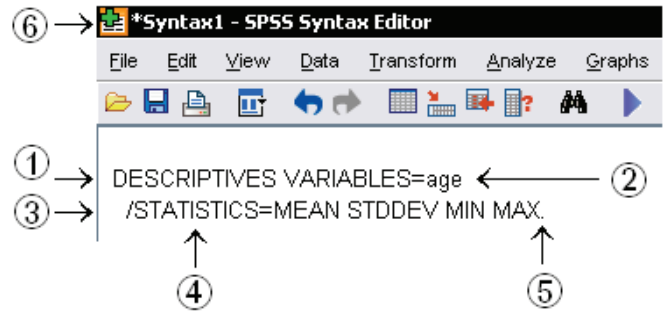
จะทำการวิเคราะห์ข้อมูลให้ทันที

แต่ถ้าเราต้องการขั้นตอนที่เราสั่งมาทั้งหมด

บันทึกเป็น Syntax ให้ทำดังนี้



ขั้นที่ 4. คลิกปุ่ม Paste บนเมนูย่อย Descriptives บนจอภาพจะมี Window ของ SPSS Syntax Editor และโปรแกรมภาษา SPSS ชนิดหนึ่งที่เรียกว่า Syntax ปรากฏอยู่
ข้อสังเกต



1. เป็นผลมาจากการคลิกคำสั่ง

Analyze / Descriptive Statistics / Descriptives จากขั้นที่ 1.

2. เป็นผลมาจากการเลือกตัวแปร age มาวิเคราะห์จากขั้นที่ 2.

3. เป็นผลมาจากค่า Default ที่ SPSS กำหนดให้

4. มาจากรูปแบบภาษาของ Syntax ที่กำหนดว่า VARIABLES, STATISTICS เป็นคำสั่งย่อยของชุดคำสั่ง DESCRIPTIVES ต้องมีการขึ้นด้วยเครื่องหมาย /

5. มาจากรูปแบบภาษาของ Syntax ที่กำหนดว่าเมื่อจบชุดคำสั่งต้องลงท้ายด้วยเครื่องหมาย .

6. Window ของ SPSS Syntax Editor

ความหมายทางด้านโปรแกรม SPSS

DESCRIPTIVES วิเคราะห์ข้อมูลด้วยคำสั่ง Analyze / Descriptive Statistics / Descriptives

VARIABLES=age เลือกตัวแปร age มาทำการวิเคราะห์

/STATISTICS=MEAN STDDEV MIN MAX .

คำนวณค่าสถิติตามที่กำหนดคือ ค่าเฉลี่ย (MEAN) ส่วนเบี่ยงเบนมาตรฐาน (STDDEV) ค่าต่ำสุด (MIN) ค่าสูงสุด (MAX)

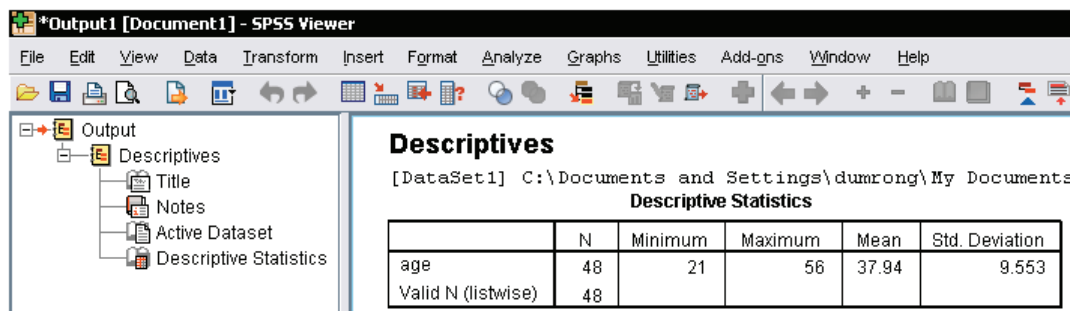
2. การสั่งให้โปรแกรม Syntax ทำงาน

ขั้นที่ 4. สั่งให้โปรแกรม Syntax ทำงาน

คลิกคำสั่ง Run / All



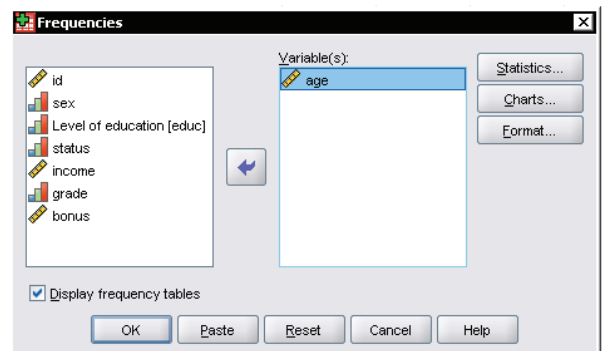
จะได้ผลการคำนวณ
ดังนี้



เมื่อเราวิเคราะห์

เพิ่มเติมด้วย

คำสั่ง Analyze / Descriptive Statistics / Frequencies



และเลือกจำนวนค่าสถิติเพิ่มเติมดังนี้

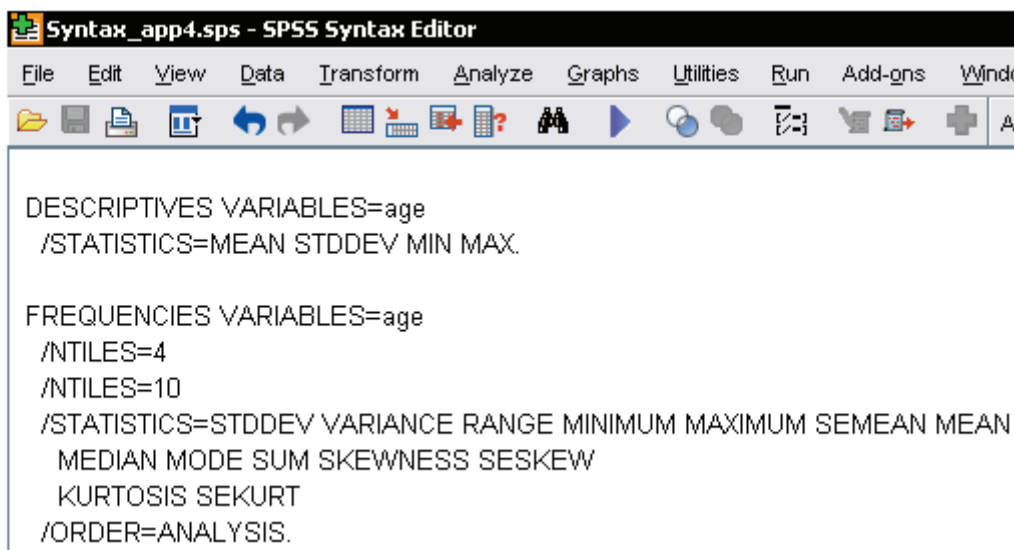
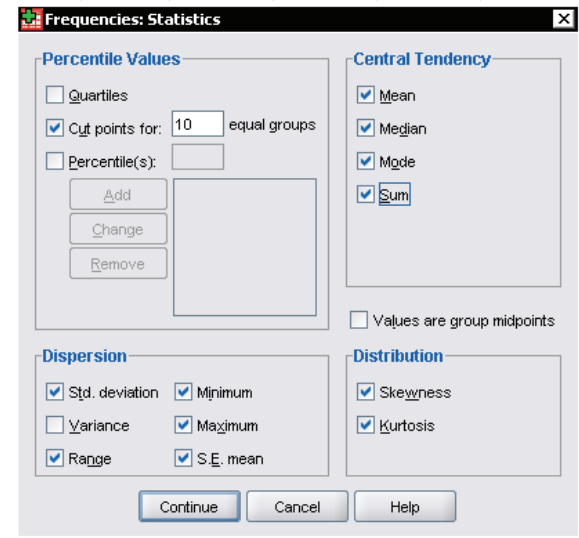
เมื่อคลิก Paste ที่เมนูย่อย

โปรแกรม SPSS จะนำขั้นตอนต่าง ๆ

ที่เราสั่งไว้ล่าสุดไปพิมพ์ต่อใน SPSS Syntax Editor

ผลของโปรแกรมที่ได้ใน SPSS Syntax Editor เป็นดังนี้

1. เป็นโปรแกรมส่วนแรกจาก คำสั่ง Descriptives
2. เป็นโปรแกรมใหม่จาก คำสั่ง Frequencies



ความหมายทางด้านโปรแกรม SPSS

FREQUENCIES

วิเคราะห์ข้อมูลด้วยคำสั่ง Analyze / Descriptive Statistics / Frequencies

VARIABLES=age

เลือกตัวแปร age มาทำการแจกแจงความถี่

/NTILES=4

สั่งให้คำนวณค่าควอไทล์ที่ 1, 2 และ 3

/NTILES=10

สั่งให้คำนวณค่าเปอร์เซ็นต์ไทล์ที่ 10, 20, 30, ... , 80, 90

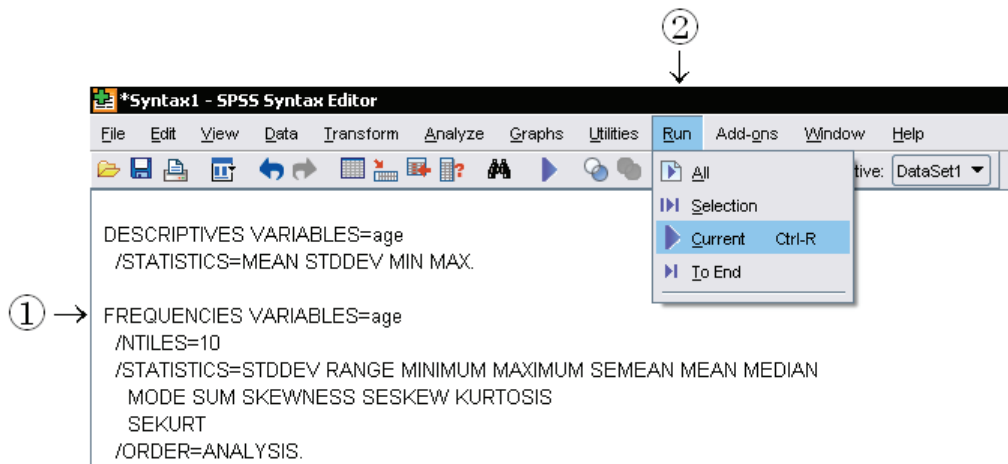
/STATISTICS=STDDEV VARIANCE RANGE MINIMUM MAXIMUM SEMEAN MEAN MEDIAN

MODE

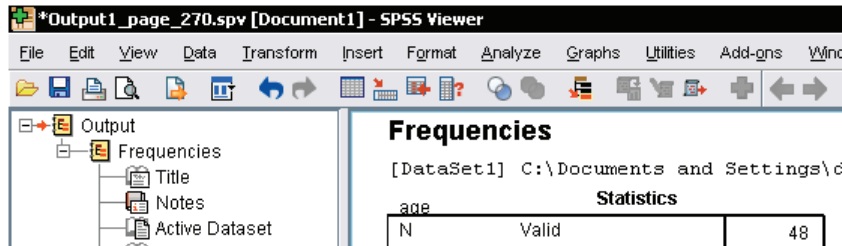
สั่งให้คำนวณค่าสถิติต่างๆ เช่น ส่วนเบี่ยงเบนมาตรฐาน (STDDEV) ความแปรปรวน (VARIANCE) พิสัย (RANGE) ค่าต่ำสุด (MINIMUM) ค่าสูงสุด (MAXIMUM) Standard Error Mean (SEMEAN) ค่าเฉลี่ย (MEAN) มัธยฐาน (MEDIAN) ฐานนิยม (MODE)

3. การสั่งให้บางส่วนของโปรแกรม Syntax ทำงาน

- ขั้นที่ 1. เลื่อนเมาส์ Pointer มาที่คำสั่ง FREQUENCIES ใน SPSS Syntax Editor
- ขั้นที่ 2. คลิกคำสั่ง Run / Current

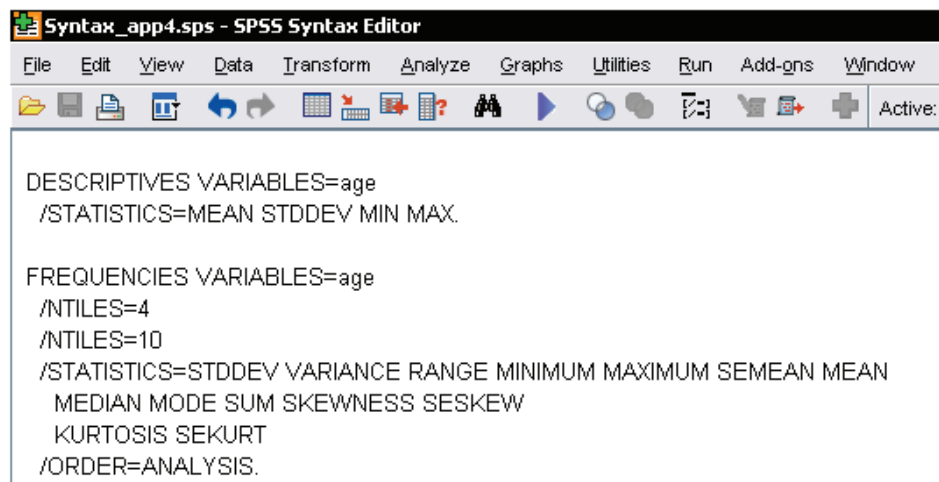


จะได้ผลการคำนวณ ตั้งแต่ตำแหน่งที่เราเลือกไว้ คือคำสั่ง FREQUENCIES



4. การแก้ไขบางส่วนของโปรแกรม Syntax

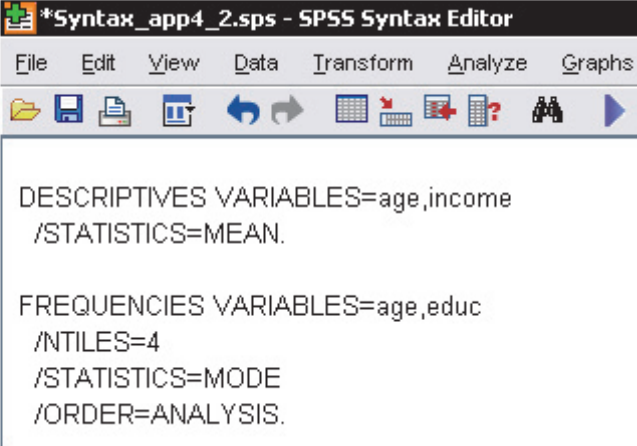
จากโปรแกรม Syntax เดิมเราสามารถเพิ่มเติมและแก้ไขคำสั่งได้ด้วยอย่างเช่น จากโปรแกรม



ตัวอย่างการแก้ไขเช่น

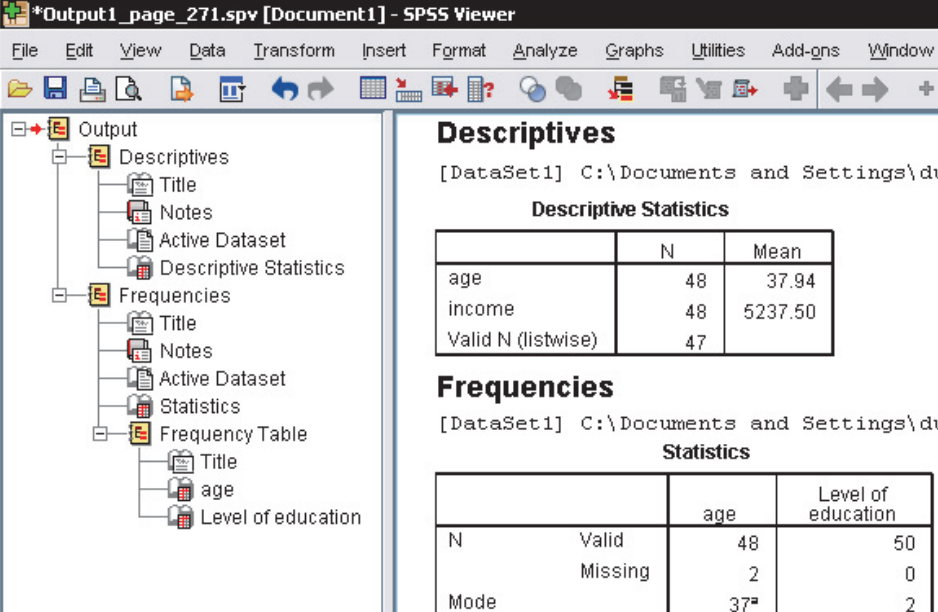
1. เพิ่มตัวแปรของการวิเคราะห์ของคำสั่ง DESCRIPTIVES อีก 1 ตัวคือตัวแปร income
2. ค่าสถิติของคำสั่ง DESCRIPTIVE คือ MEAN ค่าเดียวเท่านั้น
3. เพิ่มตัวแปรของการวิเคราะห์ของคำสั่ง FREQUENCIES อีก 1 ตัวคือตัวแปร educ
4. ยกเลิกการคำนวณเปอร์เซ็นต์ไทล์ 10, 20, 30, ... , 90
5. การวิเคราะห์ด้วยคำสั่ง FREQUENCIES ให้คำนวณค่าสถิติเฉพาะค่าฐานนิยม (MODE) เท่านั้น

จะได้โปรแกรม Syntax ใหม่เป็น



① → DESCRIPTIVES VARIABLES=age,income
 ② → /STATISTICS=MEAN.
 ③ → FREQUENCIES VARIABLES=age,educ
 ④ → /NTILES=4
 ⑤ → /STATISTICS=MODE
 /ORDER=ANALYSIS.

สั่งให้โปรแกรม Syntax ทำงานโดยคลิกคำสั่ง Run / All จะได้ผลดังนี้



Descriptives
 [DataSet1] C:\Documents and Settings\du

Descriptive Statistics

	N	Mean
age	48	37.94
income	48	5237.50
Valid N (listwise)	47	

Frequencies
 [DataSet1] C:\Documents and Settings\du

Statistics

		age	Level of education
N	Valid	48	50
	Missing	2	0
Mode		37 ^a	2

หมายเหตุ เราสามารถบันทึกเพิ่ม Syntax ไว้ทำงานต่อได้ โดยใช้คำสั่ง File / Save

บรรณานุกรม

- Joseph G. Van Matre , Glenn H. Gilbreath , **Statistics for Business and Economics** , Third Edition ,Business Publication,Inc., Homewood, Illinois ,1987
- Ronald E. Walpole , Raymond H. Myers , **Probability and Statistics for Engineers and Scientists** Third Edition , Macmillan Publishing Company , NewYork , 1985.
- SPSS Base 7.5 Application Guide** , SPSS Inc. USA 1997
- SPSS Base 7.5 for Windows User's Guide** , SPSS Inc. USA 1997
- กรรณิกา ทิตาราม สถิติเชิงคณิตศาสตร์ ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์ มหาวิทยาลัย กรุงเทพมหานคร 2528
- คณาจารย์ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย **ความน่าจะเป็นและสถิติ** พิทักษ์ การพิมพ์ กรุงเทพมหานคร 2528
- ดำรงค์ ทิพย์โยธา **การวิเคราะห์ข้อมูลทางสถิติ และความน่าจะเป็นด้วยโปรแกรมสำเร็จรูป SPSS for Windows & Mathcad** โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย กรุงเทพมหานคร 2541
- ดำรงค์ ทิพย์โยธา **การวิเคราะห์ข้อมูลทางสถิติด้วย SPSS for Windows version 9.0** โรงพิมพ์จุฬาลงกรณ์ มหาวิทยาลัย กรุงเทพมหานคร 2543
- ดำรงค์ ทิพย์โยธา **การวิเคราะห์ข้อมูลทางสถิติด้วย SPSS for Windows version 10.0** โรงพิมพ์จุฬาลงกรณ์ มหาวิทยาลัย กรุงเทพมหานคร 2545
- ดำรงค์ ทิพย์โยธา **การวิเคราะห์ข้อมูลทางสถิติด้วย SPSS for Windows version 12.0** โรงพิมพ์จุฬาลงกรณ์ มหาวิทยาลัย กรุงเทพมหานคร 2548
- ดำรงค์ ทิพย์โยธา **ความน่าจะเป็นและสถิติ สรุปเนื้อหา โจทย์แบบฝึกหัดและเฉลย** โรงพิมพ์จุฬาลงกรณ์ มหาวิทยาลัย กรุงเทพมหานคร 2544
- ดำรงค์ ทิพย์โยธา **คู่มือ MATHCAD** โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย กรุงเทพมหานคร 2541
- พรพรรณ แยมกลิ่น , สุพัตตา ปวนะฤทธิ์ เอกสารประกอบคำบรรยาย **วิชาความน่าจะเป็นและสถิติ** ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย กรุงเทพมหานคร 2530
- ศิริชัย พงษ์วิชัย **การวิเคราะห์ข้อมูลทางสถิติด้วยคอมพิวเตอร์ พิมพ์ครั้งที่ 8** สำนักพิมพ์จุฬาลงกรณ์ มหาวิทยาลัย กรุงเทพมหานคร 2539

การหาค่า Significant (Sig.) ของค่าสถิติด้วย Mathcad

คำสั่งของ Mathcad เกี่ยวกับการแจกแจงค่าสถิติ t

$dt(t, v)$ = ฟังก์ชันการแจกแจงความน่าจะเป็นของตัวแปรสุ่ม t

$$dt(t, v) = \frac{\Gamma(\frac{v+1}{2})}{\Gamma(\frac{v}{2})\sqrt{\pi v}} \left(1 + \frac{t^2}{v}\right)^{-\frac{v+1}{2}}, df = v$$

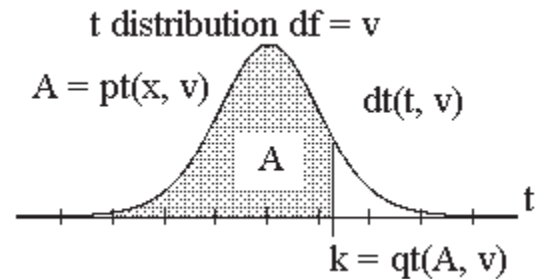
$A = pt(k, v) = P(t < k)$ = พื้นที่ใต้โค้งทางทางด้านขวาที่ระยะ k

$k = qt(A, v)$ = ระยะบนแกนที่

ที่ทำให้พื้นที่ใต้โค้งทางทางด้านขวามีค่าเท่ากับ A

ตัวอย่างเช่น $pt(2.242, 23) = 0.98254978$ และ $qt(0.98254978, 23) = 2.242$

เพราะฉะนั้น Significant ของค่าสถิติ $t = k$ มีค่าเท่ากับ $1 - pt(k, v)$ เมื่อ $df = v$



คำสั่งของ Mathcad เกี่ยวกับการแจกแจงค่าสถิติไคสแควร์

$dchisq(x, v)$ = ฟังก์ชันการแจกแจงความน่าจะเป็นของตัวแปรสุ่มไคสแควร์

$$dchisq(x, v) = \frac{1}{2^{\frac{v}{2}}\Gamma(\frac{v}{2})} x^{\frac{v}{2}-1} e^{-\frac{x}{2}}, df = v$$

$A = pchisq(k, v) = P(\chi^2 < k)$

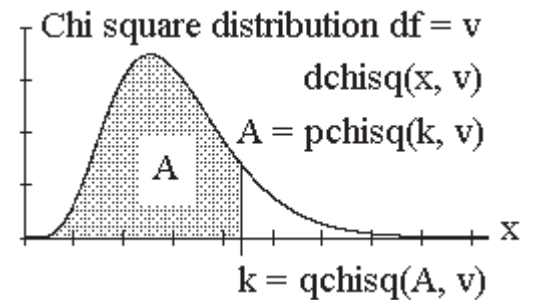
= พื้นที่ใต้โค้งทางทางด้านขวาที่ระยะ k

$k = qchisq(A, v)$ = ระยะบนแกนไคสแควร์

ที่ทำให้พื้นที่ใต้โค้งทางทางด้านขวามีค่าเท่ากับ A

ตัวอย่างเช่น $pchisq(3.822, 3) = 0.71866249$ และ $qchisq(0.71866249, 3) = 3.822$

เพราะฉะนั้น Significant ของค่าสถิติไคสแควร์ k มีค่าเท่ากับ $1 - pchisq(k, v)$ เมื่อ $df = v$



คำสั่งของ Mathcad เกี่ยวกับการแจกแจงค่าสถิติเอฟ

$dF(x, v_1, v_2)$ = ฟังก์ชันการแจกแจงความน่าจะเป็นของตัวแปรสุ่มเอฟ

$$dF(x, v_1, v_2) = \frac{\Gamma(\frac{v_1+v_2}{2})\left(\frac{v_1}{v_2}\right)^{\frac{v_1}{2}} f^{\frac{v_1}{2}-1}}{\Gamma(\frac{v_1}{2})\Gamma(\frac{v_2}{2})\left(1 + \frac{v_1}{v_2} f\right)^{\frac{v_1+v_2}{2}}}, df = v_1, v_2$$

$A = pF(k, v_1, v_2) = P(F < k)$

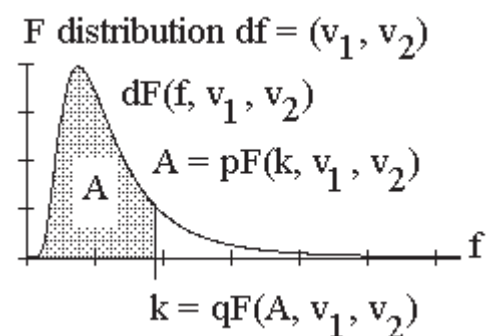
= พื้นที่ใต้โค้งทางทางด้านขวาที่ระยะ k

$k = qF(A, v_1, v_2)$ = ระยะบนแกนเอฟ

ที่ทำให้พื้นที่ใต้โค้งทางทางด้านขวามีค่าเท่ากับ A

ตัวอย่างเช่น $pF(6.574, 4, 12) = 0.99515349$ และ $qF(0.99515349, 4, 12) = 6.57400016$

เพราะฉะนั้น Significant ของค่าสถิติ $f = k$ มีค่าเท่ากับ $1 - pchisq(k, v_1, v_2)$ เมื่อ $df = v_1, v_2$



ความหมายของ Significant (Sig.) ของค่าสถิติ

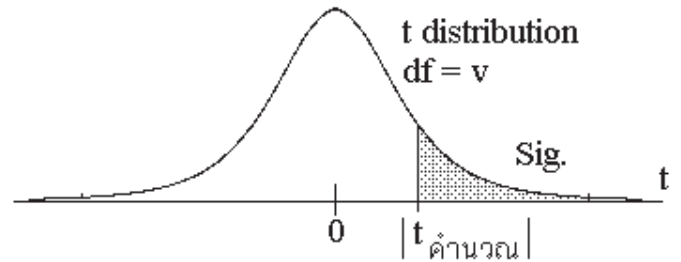
ความหมายของ Sig. (2-tailed) ของค่าสถิติที (t_{คำนวณ})

จากค่าสถิติ t_{คำนวณ} และ ระดับชั้นความเสรี df = v ที่คำนวณได้จากตัวอย่าง

$$P(t > | t_{\text{คำนวณ}} |) = \text{พื้นที่ใต้โค้งของการแจกแจงที}$$

$$\text{ทางทางด้านขวาที่ระยะ } t_{\text{คำนวณ}}$$

$$= \int_{|t_{\text{คำนวณ}}|}^{\infty} \frac{\Gamma(\frac{v+1}{2})}{\Gamma(\frac{v}{2})\sqrt{\pi v}} (1 + \frac{x^2}{v})^{-\frac{v+1}{2}} dx$$



Sig. (1-tailed) ของค่าสถิติ t_{คำนวณ} มีค่าเท่ากับ P(t > | t_{คำนวณ} |)

Sig. (2-tailed) ของค่าสถิติ t_{คำนวณ} มีค่าเท่ากับ 2 P(t > | t_{คำนวณ} |)

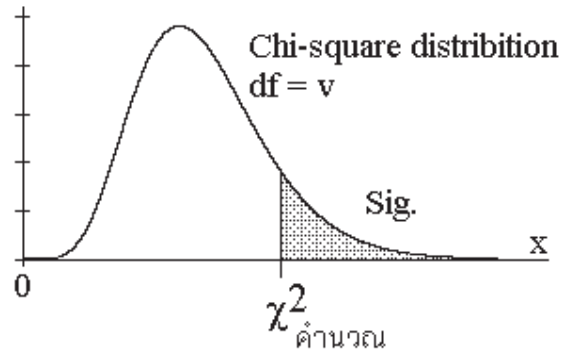
ความหมายของ Sig. ของค่าสถิติไคสแควร์ (χ²_{คำนวณ})

จากค่าสถิติ χ²_{คำนวณ} และ ระดับชั้นความเสรี df = v ที่คำนวณได้จากตัวอย่าง

$$P(x > \chi^2_{\text{คำนวณ}}) = \text{พื้นที่ใต้โค้งของการแจกแจงไคสแควร์}$$

$$\text{ทางทางด้านขวาที่ระยะ } \chi^2_{\text{คำนวณ}}$$

$$= \int_{\chi^2_{\text{คำนวณ}}}^{\infty} \frac{1}{2^{\frac{v}{2}} \Gamma(\frac{v}{2})} x^{\frac{v}{2}-1} e^{-\frac{x}{2}} dx$$



Sig. ของค่าสถิติ χ²_{คำนวณ} มีค่าเท่ากับ P(x > χ²_{คำนวณ})

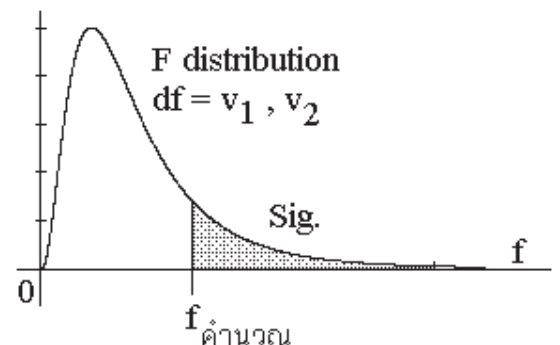
ความหมายของ Sig. ของค่าสถิติเอฟ (f_{คำนวณ})

จากค่าสถิติ f_{คำนวณ} และ ระดับชั้นความเสรี v₁, v₂ ที่คำนวณได้จากตัวอย่าง

$$P(F > f_{\text{คำนวณ}}) = \text{พื้นที่ใต้โค้งของการแจกแจงเอฟ}$$

$$\text{ทางทางด้านขวาที่ระยะ } f_{\text{คำนวณ}}$$

$$= \int_{f_{\text{คำนวณ}}}^{\infty} \frac{\Gamma(\frac{v_1+v_2}{2}) (\frac{v_1}{v_2})^{\frac{v_1}{2}} f^{\frac{v_1}{2}-1}}{\Gamma(\frac{v_1}{2}) \Gamma(\frac{v_2}{2}) (1 + \frac{v_1}{v_2} f)^{\frac{v_1+v_2}{2}}} df$$



Sig. ของค่าสถิติ f_{คำนวณ} มีค่าเท่ากับ P(F > f_{คำนวณ})