

บทที่ 3

การจัดการแฟ้มข้อมูล

บางครั้งข้อมูลที่สร้างขึ้นเป็นแฟ้มข้อมูลแล้วยังไม่สามารถนำมาวิเคราะห์ได้ตามวัตถุประสงค์ของการวิจัย เนื่องจากอาจมีแฟ้มข้อมูลอยู่หลายแฟ้ม จำเป็นต้องรวมข้อมูลจากทุกแฟ้มก่อน หรืออาจต้องการเพียงบางส่วนของแฟ้มข้อมูลนั้นในการวิเคราะห์ข้อมูล หรืออาจต้องการจัดเรียงข้อมูลตามลำดับของค่าของตัวแปรบางตัวก่อนการวิเคราะห์ การจัดการเกี่ยวกับแฟ้มข้อมูลเหล่านี้ทำได้ดังต่อไปนี้

1. การเรียงลำดับข้อมูลตามค่าของตัวแปร

คำสั่งที่ใช้สำหรับจัดเรียงลำดับข้อมูลตามค่าของตัวแปร คือ คำสั่ง **Sort Cases** โดยปกติเราอาจจะป้อนข้อมูลที่ได้จากแบบสอบถามตามลำดับเลขที่ของแบบสอบถาม ดังนั้นถ้าต้องการเรียงลำดับข้อมูลใหม่ตามค่าของตัวแปรที่สนใจ ตัวอย่างเช่น จากแฟ้มข้อมูล **DataTest1.sav** ต้องการให้เรียงลำดับข้อมูลตามประสบการณ์การทำงาน (**exp**) จากน้อยไปมาก และเรียงตามจำนวนเงินเดือน (**salary**) จากน้อยไปมากด้วยมีขั้นตอนดังนี้

1. ไปที่เมนูบาร์ คลิกที่ **Data , Sort Cases** จะได้นหน้าต่าง **Sort Cases**
2. ที่หน้าต่าง **Sort Cases** คลิกที่ตัวแปรที่ต้องการเรียงลำดับก่อน ในที่นี้คือ ตัวแปร **exp** แล้วคลิกเลือกที่ **Ascending** ใน **Sort Order** เพราะต้องการเรียงข้อมูลจากน้อยไปมาก แล้วคลิกที่หัวลูกศร > หน้าช่อง **Sort by :** ผลที่ได้คือ **experiance [exp] – Ascending** อยู่ในช่อง **Sort by :**

3. คลิกที่ตัวแปรที่ต้องการเรียงลำดับต่อมา ในที่นี้คือ ตัวแปร **salary** ทำเช่นเดียวกับข้อ 2 ดังภาพที่ 3.1 แล้วคลิกที่ **OK** ผลลัพธ์ที่ได้คือ ข้อมูลที่จัดเรียงใหม่ตามที่ต้องการ ดังภาพที่ 3.2

4. ไปที่เมนูบาร์ คลิกที่ **File , Save As...** ตั้งชื่อ **DataSort** ในช่อง **File name :** และที่ช่อง **Save as type :** เป็น **SPSS (*.sav)** แล้วคลิกที่ปุ่ม

ภาพที่ 3.2

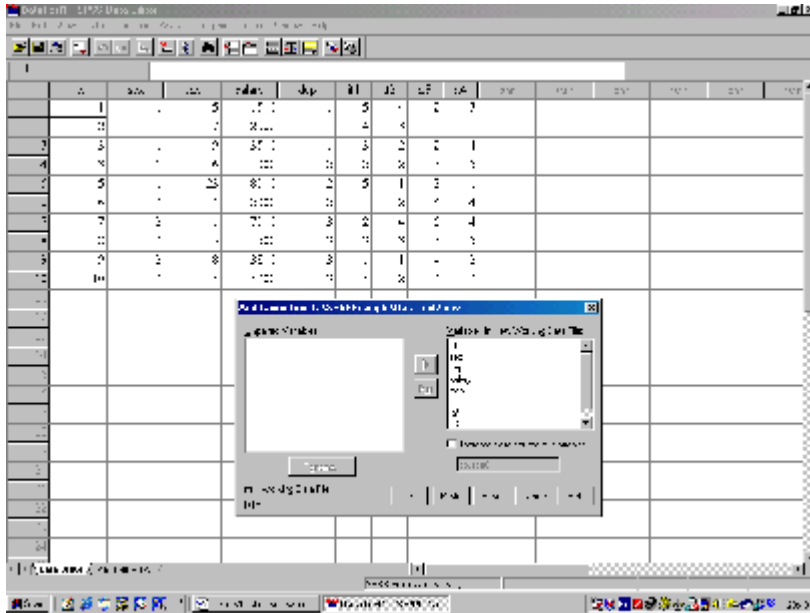
2. การรวมข้อมูลจากเพิ่มข้อมูลตั้งแต่ 2 เพิ่มขึ้นไป

กรณีที่มีตัวอย่างจำนวนมากจำเป็นต้องให้พนักงานหลายคนช่วยกันป้อนข้อมูล โดยแต่ละคนก็จะเก็บ (save) ข้อมูลเป็นชื่อเพิ่มข้อมูลต่าง ๆ ก่อนจะนำข้อมูลไปวิเคราะห์จำเป็นต้องรวมเพิ่มข้อมูลต่าง ๆ นั้นให้เป็นเพิ่มข้อมูลเดียว การรวมเพิ่มข้อมูลอาจทำได้ 2 แบบ คือ แบบแรกรวมเพิ่มข้อมูลที่มีตัวแปรเหมือนกัน ลักษณะนี้เป็นการต่อตัวอย่างจนครบจำนวนตัวอย่างทั้งหมดที่ใช้ในการวิเคราะห์ข้อมูล โดยใช้คำสั่ง **Merge Files** แบบ **Add Cases** แบบที่สองกรณีที่แบบสอบถาม 1 ฉบับมีความยาวมาก แบ่งออกได้หลายตอน ซึ่งในแต่ละตอนก็มีหลายตัวแปร เราอาจต้องแบ่งแบบสอบถามฉบับหนึ่งออกเป็นหลายส่วน แต่ละส่วนประกอบด้วยตัวแปรบางตัวแปรเท่านั้นของตัวอย่างทั้งหมด ให้พนักงานป้อนข้อมูลรับผิดชอบในการป้อนข้อมูลคนละส่วนนี้แล้วเก็บ (save) ข้อมูลเป็นชื่อเพิ่มข้อมูลต่าง ๆ ดังนั้นก่อนจะนำข้อมูลไปวิเคราะห์จำเป็นต้องรวมตัวแปรทั้งหมดจากเพิ่มข้อมูลต่าง ๆ นั้น โดยใช้คำสั่ง **Merge Files** แบบ **Add Variables**

2.1 การใช้คำสั่ง Merge Files แบบ Add Cases

สมมติว่ามีเพิ่มข้อมูล 2 เพิ่มคือ DataTest1.sav มี 10 ตัวอย่าง และ DataTest2.sav มี 10 ตัวอย่าง ต้องการรวมเพิ่มข้อมูล 2 เพิ่มนี้ ซึ่งทั้ง 2 เพิ่มนี้มีตัวแปรเหมือนกัน ทำได้ดังนี้

1. เปิดเพิ่มข้อมูล DataTest1.sav
2. ไปที่เมนูบาร์ คลิกที่ Data , Merge Files , Add Cases... ผลที่ได้คือหน้าต่าง Add Cases : Read File ดังภาพที่ 3.3
3. ที่หน้าต่าง Add Cases : Read File เลือกเพิ่มข้อมูล DataTest2 ซึ่งอยู่ในโฟลเดอร์ C:\SPSSExample แล้วคลิกปุ่ม Open ผลที่ได้คือ หน้าต่าง Add Cases from C:\SPSSExample\DataTest2.sav ดังภาพที่ 3.4 แล้วคลิกปุ่ม OK ผลที่ได้คือ ข้อมูลในเพิ่ม DataTest2 จะไปต่อท้ายเพิ่มข้อมูล DataTest1 ได้จำนวนตัวอย่างทั้งหมดเท่ากับ 20 ตัวอย่าง



ภาพที่ 3.4

2.2 การใช้คำสั่ง Merge Files แบบ Add Variables

สมมติว่ามีเพิ่มข้อมูล 2 เพิ่มคือ DataTest3.sav ที่มีตัวแปร id , sex , exp , และ salary ส่วนเพิ่มข้อมูล DataTest4.sav มีตัวแปร id it1 , it2 , it3 , และ it4 การรวม 2 เพิ่มนี้ทำได้ดังนี้

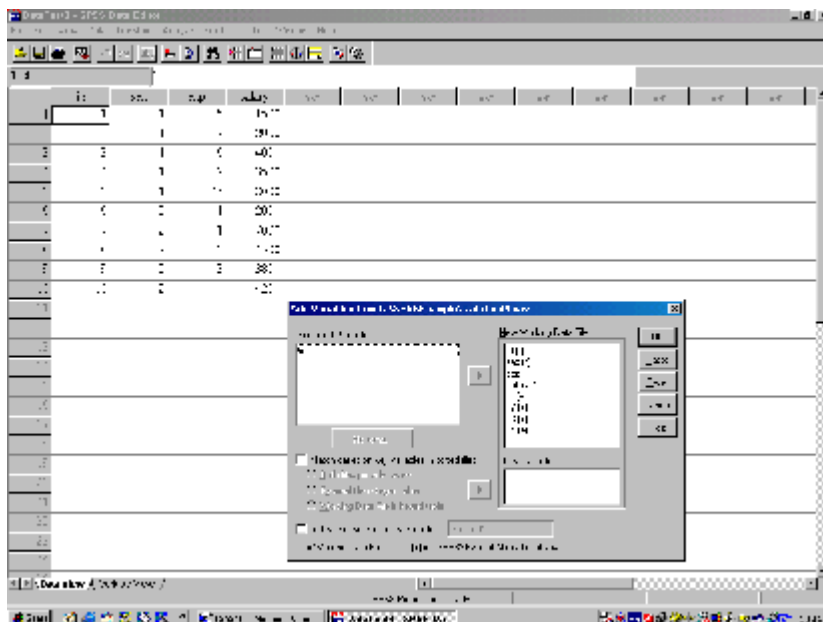
1. เปิดเพิ่มข้อมูล DataTest3.sav ที่อยู่ในโฟลเดอร์ C:\SPSS Example

2. ไปที่เมนูบาร์ คลิกที่ Data , Merge Files , Add Variables... จะได้นหน้าต่าง Add Variables : Read File เลือกเพิ่มข้อมูล DataTest4 แล้วคลิกปุ่ม Open จะได้นหน้าต่าง Add Variables from C:\SPSS Example\DataTest4.sav ดังภาพที่ 3.5

3. หน้าต่าง Add Variables ในช่องภายใต้ New Working Data File : จะปรากฏชื่อตัวแปรของทั้ง 2 เพิ่มข้อมูลนั้น ส่วนภายใต้ช่อง Excluded Variables : จะปรากฏชื่อตัวแปรของเพิ่มข้อมูล DataTest4 ที่ซ้ำกับตัวแปรของเพิ่มข้อมูล DataTest3 คือตัวแปร id ซึ่งจะไม่นำมารวมกับเพิ่มข้อมูล DataTest3

แล้วคลิกที่ Match cases on key variables in sorted files เพื่อเลือกวิธีการจับคู่ตัวอย่าง

4. การเลือกวิธีการจับคู่ตัวอย่างแบบ Both files provide cases : เป็นวิธีการจับคู่ตัวอย่างต่อตัวอย่างที่มีค่าของตัวแปรหลักเหมือนกัน โดยปกติโปรแกรมจะเลือกให้จับคู่ด้วยวิธีนี้อยู่แล้ว ซึ่งการกำหนดตัวแปรหลักทำได้โดย คลิกที่ตัวแปร id ในช่อง Excluded Variables : แล้วคลิกที่หัวลูกศร > หน้าช่อง Key Variables : แล้วคลิกปุ่ม OK จะมีคำเตือนว่า ถ้าข้อมูลไม่ได้เรียงลำดับตามตัวแปรหลักจากน้อยไปมากจะจับคู่ตัวอย่างต่อตัวอย่างไม่สำเร็จ แล้วคลิกปุ่ม OK ผลลัพธ์ที่ได้คือ เพิ่มข้อมูลใหม่ที่รวมตัวแปรของทั้ง 2 แฟ้มข้อมูลเข้าด้วยกันแล้ว



ภาพที่ 3.5

3. การเลือกข้อมูลมาวิเคราะห์เพียงบางส่วน

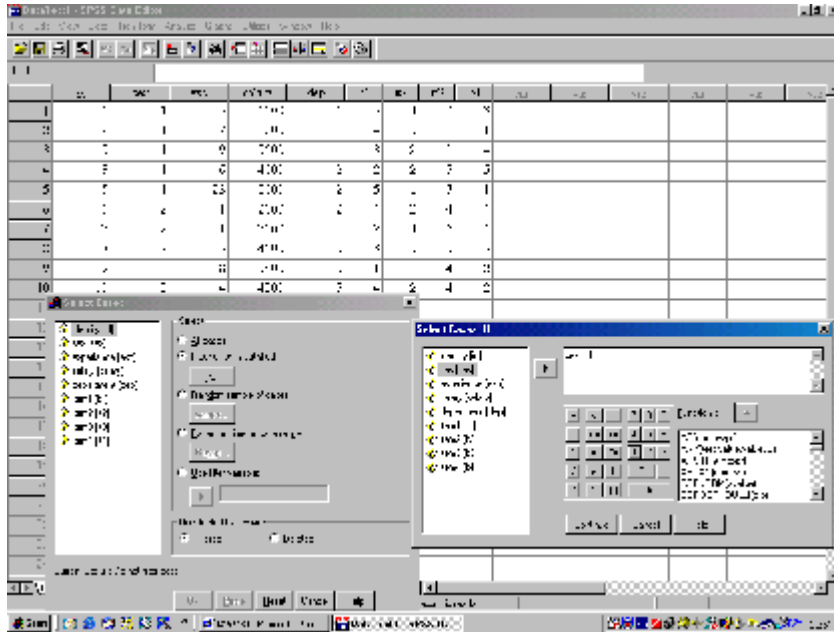
การเลือกข้อมูลมาบางส่วนจากข้อมูลทั้งหมดทำได้โดยคำสั่ง **Select Cases** ถ้าต้องการเลือกข้อมูลเฉพาะบางค่าของตัวแปรหนึ่ง

สมมติว่าต้องการเลือกเฉพาะข้อมูลของพนักงานเพศชาย ในที่นี้คือ ตัวแปร **sex** ที่มีค่าเป็น 1 ของเพิ่มข้อมูล **DataTest1.sav**

1. ไปที่เมนูบาร์ คลิกที่ **Data , Select Cases...** จะได้นหน้าต่าง **Select Cases**

2. ที่หน้าต่าง **Select Cases** ภายในกรอบ **Select** คลิกเลือกที่ **If condition is satisfied** แล้วคลิกปุ่ม **If** เพื่อกำหนดเงื่อนไข ผลที่ได้คือ หน้าต่าง **Select Cases : If** ดังภาพที่ 3.6

3. ที่หน้าต่าง **Select Cases : If** กำหนดเงื่อนไขที่ต้องการคือ **sex = 1** โดยพิมพ์ลงในช่องว่างด้านขวาบน หรืออาจใช้วิธีคลิกตัวแปร **sex** แล้วคลิกหัวลูกศร **>** , คลิกปุ่ม **=** , คลิกปุ่ม **1** เสร็จแล้วคลิกปุ่ม **Continue** จะกลับมาสู่หน้าต่าง **Select Cases** แล้วคลิกปุ่ม **OK** ผลที่ได้คือ เพิ่มข้อมูลที่คัดเลือกเฉพาะข้อมูลของพนักงานเพศชาย



ภาพที่ 3.6

4. การแบ่งข้อมูลออกเป็นกลุ่ม ๆ ตามค่าต่าง ๆ ของตัวแปรหนึ่ง

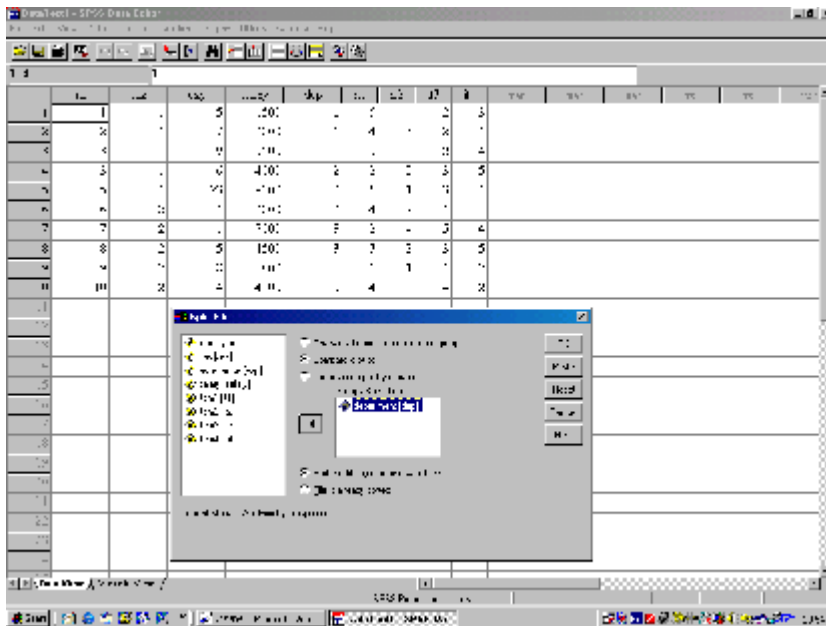
บางครั้งการวิเคราะห์ข้อมูลอาจต้องการวิเคราะห์แยกเป็นกลุ่ม ๆ ตามค่าต่าง ๆ ของตัวแปรหนึ่ง ตัวอย่างเช่น จากแฟ้มข้อมูล DataTest1.sav ต้องการแสดงการแจกแจงความถี่ของตัวแปร sex ในแต่ละค่าของตัวแปร dep โดยใช้คำสั่ง Split File เพื่อแบ่งข้อมูลออกเป็นกลุ่มตามค่าของ dep ก่อนแล้วใช้คำสั่ง Frequencies เพื่อแสดงการแจกแจงความถี่ของ sex ทำได้ดังนี้

1. ไปที่เมนูบาร์ คลิกที่ Data , Split File... จะได้นหน้าต่าง Split File
2. ที่หน้าต่าง Split File คลิกเลือกตัวแปร dep ในช่องซ้ายมือ , คลิกเลือก Compare groups แล้วคลิกที่หัวลูกศร > ผลที่ได้คือ ตัวแปร dep จะย้ายเข้าไปอยู่ในช่องภายใต้ Groups Based on : ดังภาพที่ 3.7 แล้วคลิกปุ่ม OK ผลที่ได้คือแฟ้มข้อมูล DataTest1 ถูกจัดเรียงลำดับใหม่ตามค่าของตัวแปร dep จากน้อยไปมาก

3. ไปที่เมนูบาร์ คลิกที่ **Analyze** , **Descriptive Statistics** , **Frequencies...** จะได้หน้าต่าง **Frequencies**

4. ที่หน้าต่าง **Frequencies** คลิกที่ **sex** , **dep** แล้วคลิกหัวลูกศร > ผลที่ได้คือ **sex** , **dep** ถูกย้ายเข้าไปอยู่ในช่องขวามือของ **Variable(s)** : แล้วคลิกปุ่ม **OK** จะได้หน้าจอ **Output1 - SPSS Viewer** ดังภาพที่ 3.8

ข้อควรระวัง : ถ้าต้องการวิเคราะห์ข้อมูลทั้งหมดอีกครั้ง โดยไม่มีการแบ่งข้อมูลออกเป็นกลุ่ม ๆ ตามค่าต่าง ๆ ของตัวแปรหนึ่ง ต้องกลับมาใช้คำสั่ง **Split File** อีกครั้ง เพื่อกลับสู่สภาพเดิมของข้อมูล โดยคลิกที่ **Analyze all cases , do not create groups** แล้วคลิกปุ่ม **OK**



ภาพที่ 3.7

Frequencies

Warnings

Frequency tables are not produced for the following variables because they are split variables: department.

Statistics

department			sex	department
Dept of Personal	N	Valid	3	3
		Missing	0	0
Dept of Finance	N	Valid	3	3
		Missing	0	0
Dept of Marketting	N	Valid	4	4
		Missing	0	0

Frequency Table

sex

department			Frequency	Percent	Valid Percent	Cumulative Percent
Dept of Personal	Valid	male	3	100.0	100.0	100.0
		female	1	33.3	33.3	100.0
Dept of Finance	Valid	male	2	66.7	66.7	66.7
		female	1	33.3	33.3	100.0
Total			3	100.0	100.0	
Dept of Marketting	Valid	female	4	100.0	100.0	100.0

ภาพที่ 3.8

5. การเปลี่ยนหน่วยของการวิเคราะห์จากระดับ **Micro** เป็นระดับ **Macro**

ตัวอย่างเช่น เพิ่มข้อมูล DataTest1.sav ได้ถูกเตรียมขึ้นแบบ cases ก็เป็นข้อมูลของแต่ละคนหรือแต่ละหน่วยตัวอย่าง แต่ต้องการวิเคราะห์โดยเปลี่ยนหน่วยตัวอย่างจากคน (cases) เป็นแผนก (dep) จึงจำเป็นต้องมีการรวมหน่วยของการวิเคราะห์จากระดับคน (cases) เป็นระดับแผนก (dep) โดยต้องรวมตัวแปรอื่น ๆ ให้มีหน่วยของการวิเคราะห์เป็นระดับแผนกด้วย ดังนี้

1. ไปที่เมนูบาร์ คลิกที่ Data , Aggregate... จะได้นหน้าต่าง Aggregate Data

2. ในหน้าต่าง Aggregate Data คลิกที่ตัวแปร dep ในช่องซ้ายมือ แล้วคลิกที่หัวลูกศร > หน้าช่อง Break Variable (S) : ตัวแปร dep จะย้ายมาอยู่ในช่องนี้

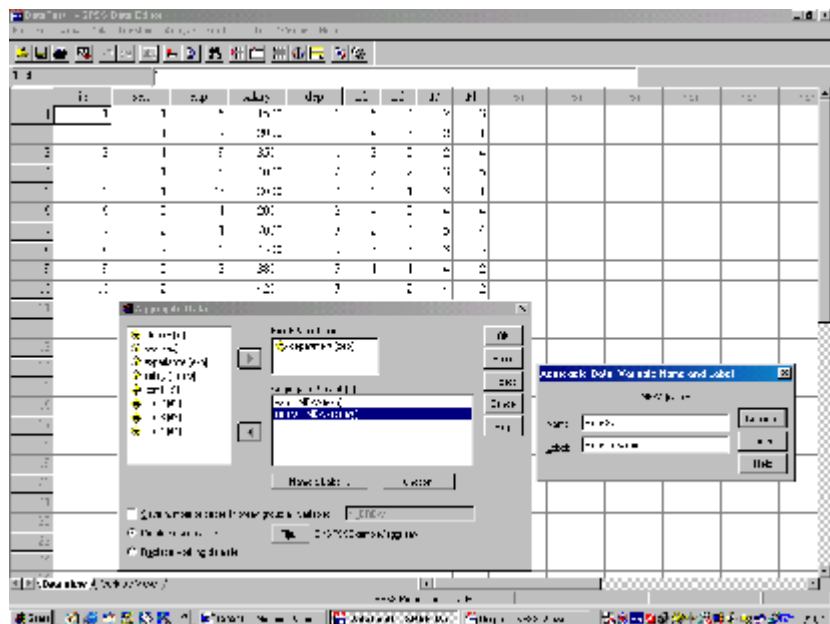
คลิกที่ตัวแปร exp และ salary ในช่องซ้ายมือแล้วคลิกที่หัวลูกศร > หน้าช่อง Aggregate Variable (S) : ในช่องนี้จะปรากฏตัวแปรใหม่คือ $exp = MEAN(exp)$ และ $salary = MEAN(salary)$ แล้วคลิกปุ่ม Name & Label จะได้หน้าต่าง Aggregate Data : Variable Name and Label

3. ในหน้าต่าง Aggregate Data : Variable Name and Label .. เพื่อตั้งชื่อตัวแปรใหม่ในช่อง Name : พิมพ์คำว่า MeanSal ส่วนในช่อง Label : พิมพ์คำว่า Mean of salary ดังภาพที่ 3.9 แล้วคลิกปุ่ม Continue

4. ในหน้าต่าง Aggregate Data ภายในช่อง Aggregate Variable(S) : คลิกที่ตัวแปร $exp = MEAN exp$ แล้วคลิกปุ่ม Name & Label .. เพื่อตั้งชื่อตัวแปรใหม่ในช่อง Name : พิมพ์คำว่า MeanExp_ ส่วนในช่อง Label : พิมพ์คำว่า Mean of exp แล้วคลิกปุ่ม Continue

5. ในหน้าต่าง Aggregate Data คลิกที่ Save number of cases in break group as variable :

6. โปรแกรมจะสร้างเพิ่มข้อมูลขึ้นมาใหม่ชื่อ Aggr.sav ซึ่งมีตัวแปร dep , MeanSal , MeanExp , และ N_Break แล้วคลิกปุ่ม OK



ภาพที่ 3.9