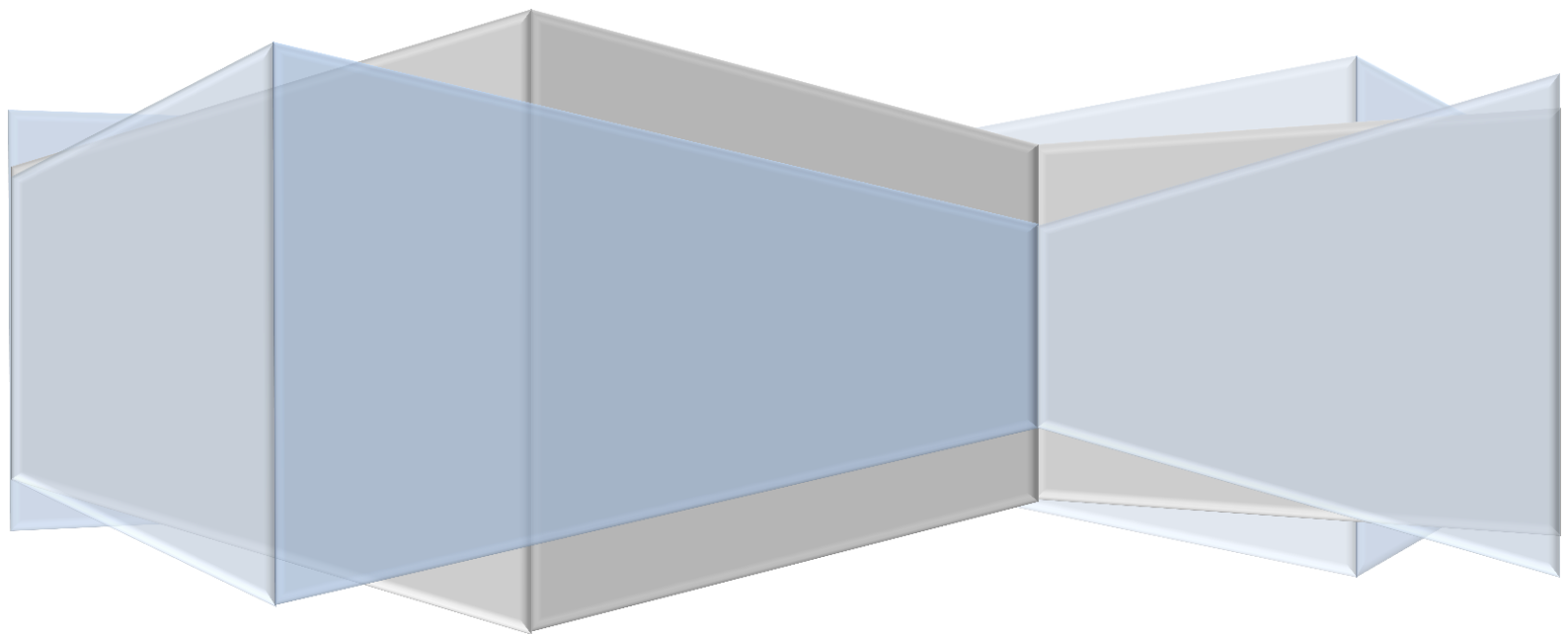




คู่มือการใช้งาน
สำหรับการจัดการฐานข้อมูลตัวอย่างซ้ำ (ORG)
จากข้อมูลภาวะการทำงานของประชากร
โดย
สถาบันวิจัยเพื่อการประเมินและออกแบบนโยบาย
มหาวิทยาลัยหอการค้าไทย



คำนำ

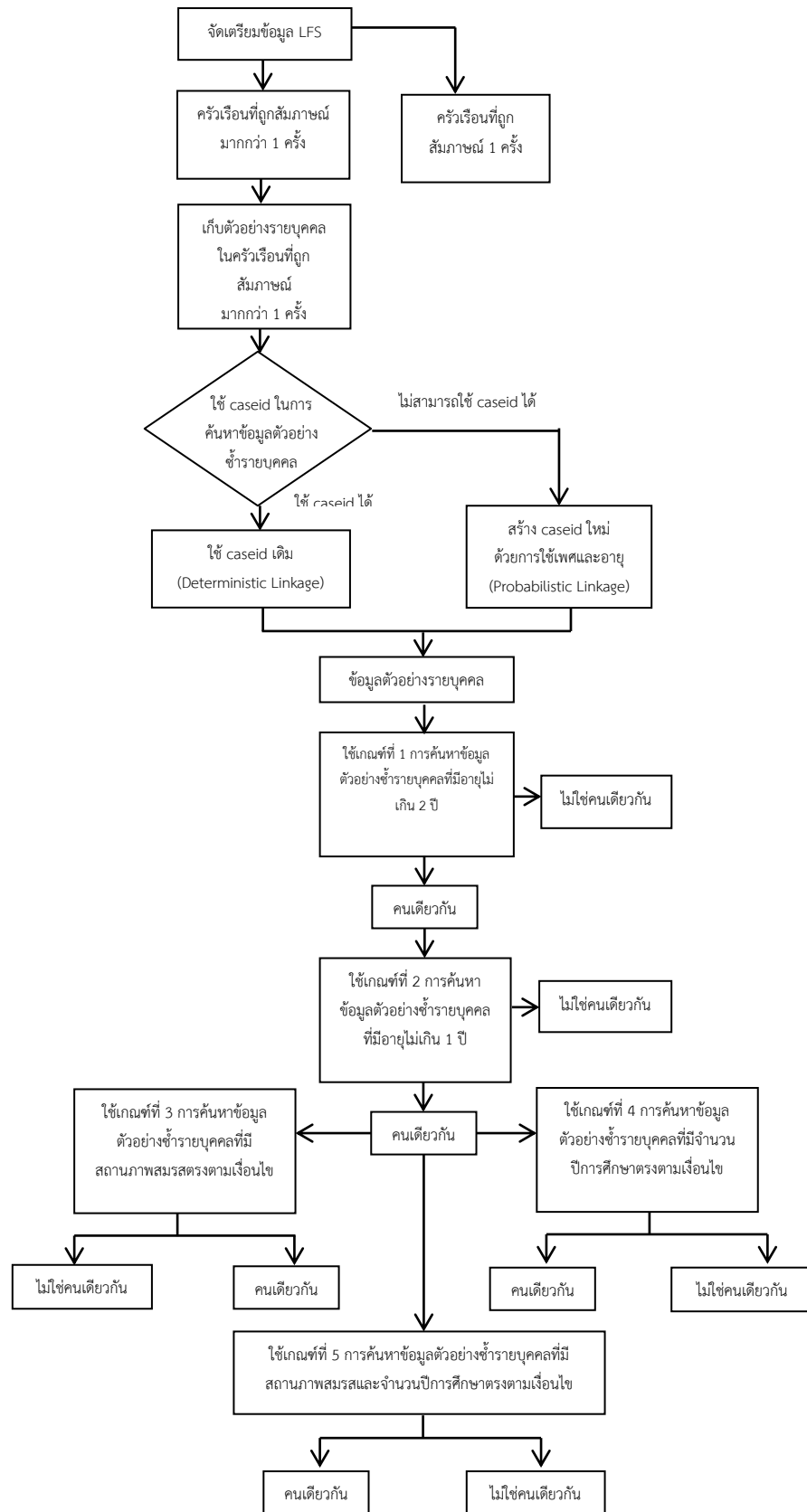
คู่มือการใช้งานชุดคำสั่งจัดทำขึ้นภายใต้โครงการ “การจัดทำฐานข้อมูลแบบตัวอย่างซ้ำจากข้อมูลภาวะการทำงานของประชากร” ซึ่งได้รับทุนจากสำนักงานกองทุนสนับสนุนการวิจัย (สกว.) โดยโครงการจะนำข้อมูลการสำรวจภาวะการทำงานของประชากร (Labor Force Survey: LFS) ของสำนักงานสถิติแห่งชาติ นับตั้งแต่ปี พ.ศ. 2545 – 2556 มาสร้างข้อมูลตัวอย่างซ้ำด้านแรงงานของประชากรไทยในส่วนของข้อมูล ORG โดยจะเผยแพร่ Matching ID และ STATA code พร้อมคู่มือการสร้างข้อมูลตัวอย่างซ้ำในเวปไซต์ เพื่อช่วยให้นักวิจัยที่มีข้อมูลการสำรวจภาวะการทำงานของประชากรสามารถสร้างข้อมูลได้อย่างสะดวกมากขึ้น

คณะนักวิจัย

สารบัญ

เรื่อง	หน้า
คำนำ	1
1. กระบวนการจัดทำข้อมูล ORG	3
2. วิธีที่ 1 ขั้นตอนการใช้ Matching ID เพื่อสร้างข้อมูล ORG	4
3. วิธีที่ 2 ขั้นตอนการใช้โค้ดคำสั่ง เพื่อสร้างข้อมูล ORG	7
4. เกณฑ์ที่ใช้ในการค้นหาข้อมูลตัวอย่างข้าราชการ	17

1. กระบวนการจัดทำข้อมูล ORG



2. วิธีที่ 1 ขั้นตอนการใช้ Matching ID เพื่อสร้างข้อมูล ORG

1. เตรียมข้อมูลภาวะการทำงานของประชากรจากสำนักงานสถิติแห่งชาติ ตั้งแต่ปี 2002-2013 (พ.ศ. 2545-2556) ซึ่งมีทั้งไฟล์ในรูปแบบ .sav (SPSS file) และ .dat (Text file)
2. บันทึกข้อมูลภาวะการทำงานของประชากรที่อยู่ในรูปของไฟล์ .dta (Stata file) ให้อยู่ในโฟลเดอร์ที่ชื่อ data

Name	Date modified	Type	Size
ifs2002q1	25/1/2561 10:03	Stata Dataset	2,182 KB
ifs2002q2	25/1/2561 10:03	Stata Dataset	21 KB
ifs2002q3	25/1/2561 10:03	Stata Dataset	25 KB
ifs2002q4	25/1/2561 10:03	Stata Dataset	2,131 KB
ifs2003q1	25/1/2561 10:03	Stata Dataset	2,130 KB
ifs2003q2	25/1/2561 10:03	Stata Dataset	18 KB
ifs2003q3	25/1/2561 10:03	Stata Dataset	17 KB
ifs2003q4	25/1/2561 10:03	Stata Dataset	2,106 KB

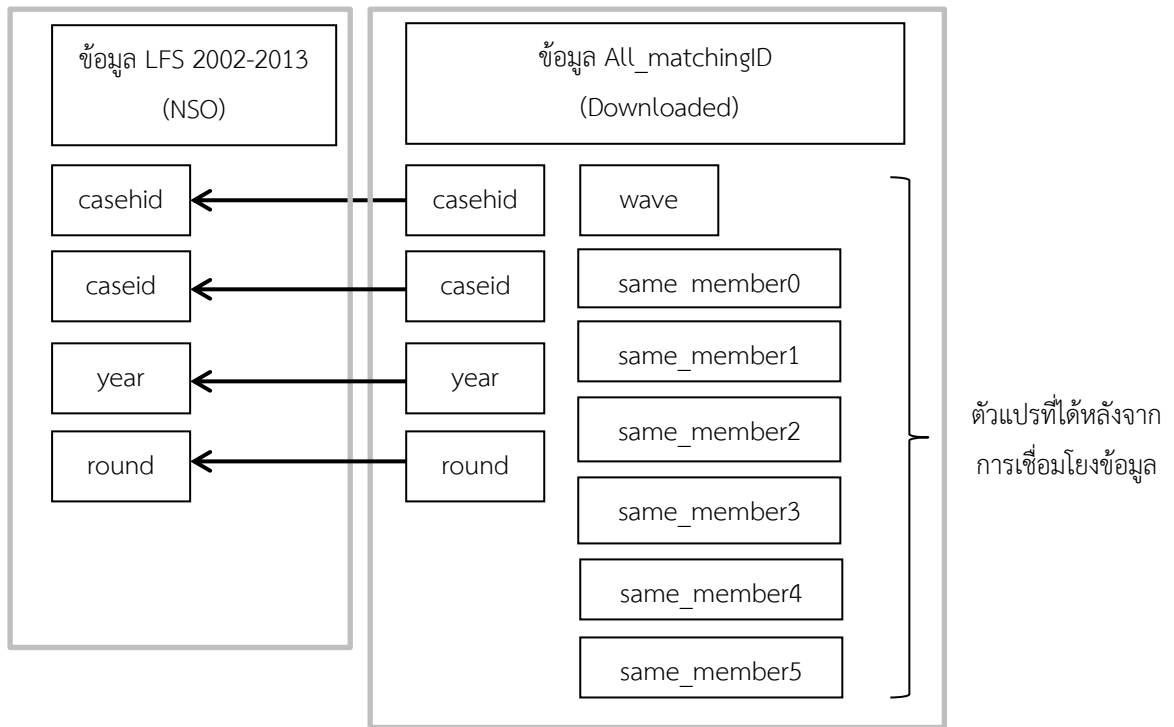
3. ดาวน์โหลดข้อมูล matching ID ที่ชื่อ All_matchingID ใส่ไว้ในโฟลเดอร์ชื่อ data และ ดาวน์โหลด command (Stata dofile) ใส่ไว้ในโฟลเดอร์ชื่อ dofile ผ่านเว็บไซต์ที่เผยแพร่ โดย command ที่โหลดมาพร้อมกับข้อมูล All_matchingID จะใช้สร้างตัวแปร ID เพื่อเชื่อมโยงข้อมูลภาวะการทำงานของประชากร

ข้อมูล matching ID แยกรายชุด

Name	Date modified	Type	Size
All_matchingID_2002-2003_ORG	17/1/2561 15:32	Stata Dataset	220,702 KB
All_matchingID_2004-2006_ORG	17/1/2561 15:33	Stata Dataset	293,803 KB
All_matchingID_2006-2008_ORG	17/1/2561 15:33	Stata Dataset	352,003 KB
All_matchingID_2008-2010_ORG	17/1/2561 15:33	Stata Dataset	344,633 KB
All_matchingID_2010-2011_ORG	17/1/2561 15:33	Stata Dataset	344,633 KB
All_matchingID_2012-2013_ORG	17/1/2561 15:33	Stata Dataset	238,147 KB

Stata dofile แยกรายชุด

do-matchingID_2002-2003	25/1/2561 9:59	Stata Do-file	4 KB
do-matchingID_2004-2006	25/1/2561 12:24	Stata Do-file	4 KB
do-matchingID_2006-2008	25/1/2561 13:21	Stata Do-file	4 KB
do-matchingID_2008-2010	25/1/2561 13:26	Stata Do-file	4 KB
do-matchingID_2010-2011	25/1/2561 13:28	Stata Do-file	4 KB
do-matchingID_2012-2013	25/1/2561 13:29	Stata Do-file	4 KB



แผนภาพการเชื่อมโยงโดยใช้ Matching ID

4. เปิดข้อมูลภาวะการทำงานของประชากร แล้วใช้ Stata dofile ที่ดาวน์โหลดมาในขั้นตอนที่ 2 สร้างตัวแปร ID ครัวเรือน และ ID รายบุคคล เพื่อให้สามารถเชื่อมโยง ID ดังกล่าวเข้ากับ ข้อมูล All_matchingID_2002-2013_ORG โดยใช้ตัวแปร casehid (ID ครัวเรือน), caseid (ID รายบุคคล), year (ปี) และ round (ไตรมาส) ในการเชื่อมโยง ซึ่งตัวแปรที่ใช้สร้าง ID ครัวเรือน และ ID รายบุคคล ประกอบด้วยตัวแปร ดังนี้
 - ID ครัวเรือน ให้สร้างแล้วตั้งชื่อว่า casehid

$$\text{casehid} = \text{region} + \text{province} + \text{areatype} + \text{psunumbr} + \text{samset} + \text{ea_set} + \text{enumdist} + \text{blk_vil} + \text{hhtype} + \text{hhnumber}$$

ตัวอย่าง casehid = 11010001A1060000101

ตัวแปร	คำอธิบาย	ตัวอย่าง	หลักที่
region	ภาค	1	1
province	จังหวัด	10	2-3
areatype	อยู่ในเขตเทศบาล/นอกเขตเทศบาล	1	4
psunumbr	ลำดับที่ชุมนุมอาคาร/หมู่บ้านตัวอย่าง	0001	5-8
samset	ชุดครัวเรือนตัวอย่าง	A	9
ea_set	ลำดับที่ของชุดครัวเรือนตัวอย่าง	1	10
enumdist	เขตที่แจ่งนับ	060	11-13
blk_vil	หมู่บ้าน	000	14-16
hhtype	ประเภทของครัวเรือน	1	17
hhnumber	หมายเลขครัวเรือน	01	18-19

➤ ID รายบุคคล ให้สร้างแล้วตั้งชื่อว่า caseid

casehid = region + province + areatype + psunumbr + samset + ea_set +
enumdist + blk_vil + hhtype + hhnumber + memnumbr

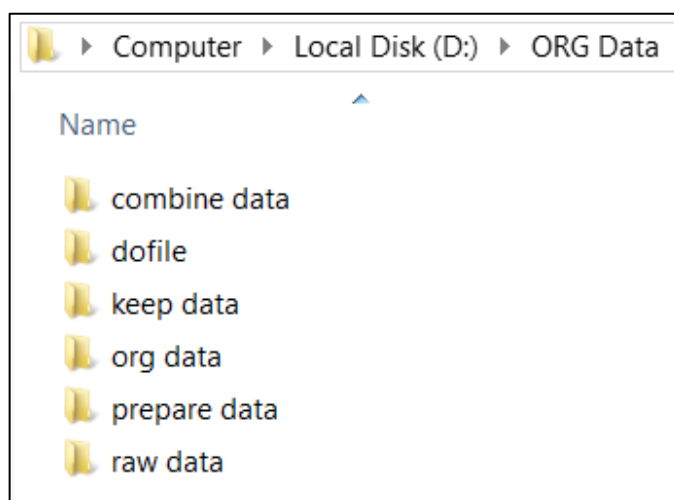
ตัวอย่าง caseid = 11010001A106000010101

ตัวแปร	คำอธิบาย	ตัวอย่าง	หลักที่
region	ภาค	1	1
province	จังหวัด	10	2-3
areatype	อยู่ในเขตเทศบาล/นอกเขตเทศบาล	1	4
psunumbr	ลำดับที่ชุมนุมอาคาร/หมู่บ้านตัวอย่าง	0001	5-8
samset	ชุดครัวเรือนตัวอย่าง	A	9
ea_set	ลำดับที่ของชุดครัวเรือนตัวอย่าง	1	10
enumdist	เขตที่แจ่งนับ	060	11-13
blk_vil	หมู่บ้าน	000	14-16
hhtype	ประเภทของครัวเรือน	1	17
hhnumber	หมายเลขครัวเรือน	01	18-19
memnumbr	หมายเลขสมาชิกครัวเรือน	01	20-21

5. หลังจากเชื่อมโยงข้อมูลเข้าด้วยกัน ก็จะได้ข้อมูลตัวอย่างซ้ำจากข้อมูลภาวะการทำงานของประชากรโดยมีตัวแปรที่ระบุว่า สมาชิกครัวเรือนคนใดที่ถูกสัมภาษณ์มากกว่า 1 ครั้ง ซึ่งประกอบไปด้วยตัวแปรที่ชื่อ same_member0, same_member1, same_member2, same_member3, same_member4 และ same_member5 ตามลำดับ โดยตั้งชื่อหมายเลขตามเกณฑ์ที่นักวิจัยใช้ในการค้นหาข้อมูลตัวอย่างซ้ำรายบุคคล ซึ่งจะอธิบายรายละเอียดของแต่ละเกณฑ์ในส่วนถัดไป

3. วิธีที่ 2 ขั้นตอนการใช้โค้ดคำสั่ง เพื่อสร้างข้อมูล ORG

1. เตรียมข้อมูลภาวะการทำงานของประชากรจากสำนักงานสถิติแห่งชาติ ตั้งแต่ปี 2002-2013 (พ.ศ. 2545-2556) ซึ่งมีทั้งไฟล์ในรูปแบบ .sav (SPSS file) และ .dat (Text file)
2. ดาวน์โหลดโค้ดคำสั่งที่อยู่ในรูปแบบไฟล์ .do ซึ่งต้องเปิดกับโปรแกรม Stata
3. เตรียมไฟล์เตอร์เพื่อจัดเก็บข้อมูล ซึ่งประกอบไปด้วยไฟล์เตอร์ต่างๆ ดังต่อไปนี้
สร้างไฟล์เตอร์หลักตั้งชื่อว่า ORG Data ภายในไฟล์เตอร์หลักจะประกอบไปด้วย 6 ไฟล์เตอร์ย่อย คือ
 - 1) raw data - ใส่ข้อมูล .dat ที่ได้จาก NSO แยกรายปี
 - 2) prepare data
 - 3) combine data
 - 4) keep data
 - 5) org data
 - 6) dofile - ใส่ dofile ที่ load มา



4. ใน folder raw data ให้ใส่ข้อมูลที่เป็นไฟล์ .dat โดยตั้งชื่อไฟล์เตอร์และชื่อ data ดังนี้

Computer > Local Disk (D:) > ORG Data > raw data

Name	Date modified
LFS-2002	1/11/2560 11:15
LFS-2003	1/11/2560 11:15
LFS-2004	1/11/2560 11:15
LFS-2005	1/11/2560 11:15
LFS-2006	1/11/2560 11:15
LFS-2007	1/11/2560 11:15
LFS-2008	1/11/2560 11:15
LFS-2009	1/11/2560 11:15
LFS-2010	1/11/2560 11:15
LFS-2011	1/11/2560 11:15
LFS-2012	1/11/2560 11:15
LFS-2013	1/11/2560 11:15

- โฟลเดอร์ LFS-2002 ใส่ raw data ที่อยู่ในรูปแบบไฟล์ .dat โดยให้ตั้งชื่อเรียงตามไตรมาสที่ 1-4 ดังนี้

Local Disk (D:) > ORG Data > raw data > LFS-2002

Name	Date modified	Type
LFS145	23/4/2545 17:14	DAT File
LFS245	17/7/2545 15:00	DAT File
LFS345	19/11/2545 10:27	DAT File
LFS445	15/1/2546 9:30	DAT File

- โฟลเดอร์ LFS-2003 ใส่ raw data ที่อยู่ในรูปแบบไฟล์ .dat โดยให้ตั้งชื่อเรียงตามไตรมาสที่ 1-4 ดังนี้

Local Disk (D:) > ORG Data > raw data > LFS-2003

Name	Date modified	Type
LFS146	11/4/2546 6:34	DAT File
LFS246	23/7/2546 8:30	DAT File
LFS346	15/10/2546 5:06	DAT File
LFS446	28/1/2547 4:29	DAT File

- โพล์เตอร์ LFS-2004 ใส่ raw data ที่อยู่ในรูปแบบไฟล์ .dat โดยให้ตั้งชื่อเรียงตามไตรมาสที่ 1-4 ดังนี้

Name	Date modified	Type
LFS147	27/4/2547 4:42	DAT File
LFS247	16/7/2547 19:03	DAT File
LFS347	18/11/2547 13:51	DAT File
LFS447	2/3/2548 14:33	DAT File

- โพล์เตอร์ LFS-2005 ใส่ raw data ที่อยู่ในรูปแบบไฟล์ .dat โดยให้ตั้งชื่อเรียงตามไตรมาสที่ 1-4 ดังนี้

Name	Date modified	Type
LFS148	20/4/2548 13:35	DAT File
LFS248	18/7/2548 11:08	DAT File
LFS348	28/10/2548 11:28	DAT File
LFS448	18/1/2549 12:40	DAT File

- โพล์เตอร์ LFS-2006 ใส่ raw data ที่อยู่ในรูปแบบไฟล์ .dat โดยให้ตั้งชื่อเรียงตามไตรมาสที่ 1-4 ดังนี้

Name	Date modified	Type
LFS149	29/5/2549 14:26	DAT File
LFS249	19/7/2549 14:57	DAT File
LFS349	25/10/2549 8:39	DAT File
LFS449	26/1/2550 10:47	DAT File

- โพล์เตอร์ LFS-2007 ใส่ raw data ที่อยู่ในรูปแบบไฟล์ .dat โดยให้ตั้งชื่อเรียงตามไตรมาสที่ 1-4 ดังนี้

Name	Date modified	Type
LFS150	4/5/2550 14:50	DAT File
LFS250	26/7/2550 8:42	DAT File
LFS350	29/10/2550 15:56	DAT File
LFS450	23/1/2551 12:48	DAT File

- โพล์เตอร์ LFS-2008 ใส่ raw data ที่อยู่ในรูปแบบไฟล์ .dat โดยให้ตั้งชื่อเรียงตามไตรมาสที่ 1-4 ดังนี้

Name	Date modified	Type
WGT0406	30/7/2551 9:33	DAT File
WGT0709	27/10/2551 15:59	DAT File
WGT1012	11/2/2552 9:27	DAT File
WGTMA3	21/5/2551 14:51	DAT File

- โพล์เตอร์ LFS-2009 ใส่ raw data ที่อยู่ในรูปแบบไฟล์ .dat โดยให้ตั้งชื่อเรียงตามไตรมาสที่ 1-4 ดังนี้

Name	Date modified	Type
WGT0103	28/4/2552 9:01	DAT File
WGT0406	28/7/2552 13:35	DAT File
WGT0709	30/10/2552 13:24	DAT File
WGT1012	27/1/2553 14:41	DAT File

- โพล์เตอร์ LFS-2010 ใส่ raw data ที่อยู่ในรูปแบบไฟล์ .dat โดยให้ตั้งชื่อเรียงตามไตรมาสที่ 1-4 ดังนี้

Local Disk (D:) > ORG Data > raw data > LFS-2010

Name	Date modified	Type
WGT0103	30/4/2553 13:01	DAT File
WGT0406	23/7/2553 17:07	DAT File
WGT0708	14/10/2553 11:58	DAT File
WGT1012	1/2/2554 10:18	DAT File

- โพล์เตอร์ LFS-2011 ใส่ raw data ที่อยู่ในรูปแบบไฟล์ .dat โดยให้ตั้งชื่อเรียงตามไตรมาสที่ 1-4 ดังนี้

Local Disk (D:) > ORG Data > raw data > LFS-2011

Name	Date modified	Type
LFS54WGTQTR1	28/4/2554 8:28	DAT File
LFS254WGT0406	19/7/2554 16:23	DAT File
lfs354	19/10/2554 16:33	DAT File
lfs454	10/2/2555 10:27	DAT File

- โพล์เตอร์ LFS-2012 ใส่ raw data ที่อยู่ในรูปแบบไฟล์ .dat โดยให้ตั้งชื่อเรียงตามไตรมาสที่ 1-4 ดังนี้

Local Disk (D:) > ORG Data > raw data > LFS-2012

Name	Date modified	Type
lfs155	30/5/2555 10:03	DAT File
lfs255	23/8/2555 14:42	DAT File
Lfs355	26/10/2555 16:21	DAT File
lfs455	9/9/2555 8:09	DAT File

- โพล์เตอร์ LFS-2013 ใส่ raw data ที่อยู่ในรูปแบบไฟล์ .dat โดยให้ตั้งชื่อเรียงตามไตรมาสที่ 1-4 ดังนี้

Name	Date modified	Type
lfs156	24/4/2556 10:35	DAT File
lfs256	30/7/2556 15:11	DAT File
LFS356	14/10/2556 10:06	DAT File
lfs456	6/1/2557 8:21	DAT File

5. ดาวน์โหล้ dofile ทั้งหมด 13 ไฟล์ ไปใส่ไว้ในโพลเตอร์ dofile

Name	Date modified	Type
1.samsetA_ORG_dofile	3/11/2560 14:25	Stata Do-file
2.samsetB_ORG_dofile	3/11/2560 14:46	Stata Do-file
3.samsetC_ORG_dofile	3/11/2560 14:46	Stata Do-file
4.samsetD_ORG_dofile	3/11/2560 14:53	Stata Do-file
5.samsetE_ORG_dofile	3/11/2560 15:33	Stata Do-file
6.samsetF_ORG_dofile	3/11/2560 15:34	Stata Do-file
7.samsetG_ORG_dofile	3/11/2560 15:34	Stata Do-file
8.samsetH_ORG_dofile	3/11/2560 15:34	Stata Do-file
9.samsetI_ORG_dofile	3/11/2560 15:39	Stata Do-file
10.samsetJ_ORG_dofile	3/11/2560 15:40	Stata Do-file
11.rotation1_ORG_dofile	6/11/2560 12:05	Stata Do-file
12.rotation2_ORG_dofile	6/11/2560 12:05	Stata Do-file
13.rotation3_ORG_dofile	6/11/2560 12:06	Stata Do-file

- dofile ที่ใช้ข้อมูล LFS ในปี 2002-2003 คือ
 - 1.samsetA_ORG_dofile.do
 - 2.samsetB_ORG_dofile.do
- dofile ที่ใช้ข้อมูล LFS ในปี 2004-2006 คือ
 - 3.samsetC_ORG_dofile.do
 - 4.samsetD_ORG_dofile.do

- dofile ที่ใช้ข้อมูล LFS ในปี 2006-2008 คือ
 - 5.samsetE_ORG_dofile.do
 - 6.samsetF_ORG_dofile.do
- dofile ที่ใช้ข้อมูล LFS ในปี 2008-2010 คือ
 - 7.samsetG_ORG_dofile.do
 - 8.samsetH_ORG_dofile.do
- dofile ที่ใช้ข้อมูล LFS ในปี 2010-2011 คือ
 - 9.samsetI_ORG_dofile.do
 - 10.samsetJ_ORG_dofile.do
- dofile ที่ใช้ข้อมูล LFS ในปี 2012-2013 คือ
 - 11.rotation1_ORG_dofile.do
 - 12.rotation2_ORG_dofile.do
 - 13.rotation3_ORG_dofile.do

6. เปิด dofile แต่ละชุด รั้นตามหมายเลข 1-13

```

1 ***** RIPPED VERSION DATE 3-11-2017 *****
2
3 clear
4 clear matrix
5 clear mata
6 set more off
7 set mem lg
8
9 //.....Set Path (Automatic: Do not change).....//
10 local fullpath c(pwd) /* display "fullpath" */
11 local splitat = strpos('fullpath','\ORG Data') /* display "splitat" */
12 global path = substr('fullpath',1,'splitat') /* display "${path}" */
13 sysdir set PERSONAL "${path}\ORG Data"
14
15 //...Set Path for this folder (Automatic: Do not change)...//
16
17 global raw "${path}\ORG Data\raw data/"
18 global pre "${path}\ORG Data\prepare data/"
19 global comb "${path}\ORG Data\combine data/"
20 global keep "${path}\ORG Data\keep data/"
21 global org "${path}\ORG Data\org data/"
22
23 *****
24
25
26
27 ***** (0) CUT RAW DATA IN 2002-2003 FOR SAMSET A *****
28
29 **>>>..... LFS 2002 Quarter 1 .....<<<<<
30 //***** (1) cut stata file from law data*****//
31 clear all
32 set more off
33 infix q1-4 ///
34 cwt 2-3 ///
35 area 4 ///
36

```

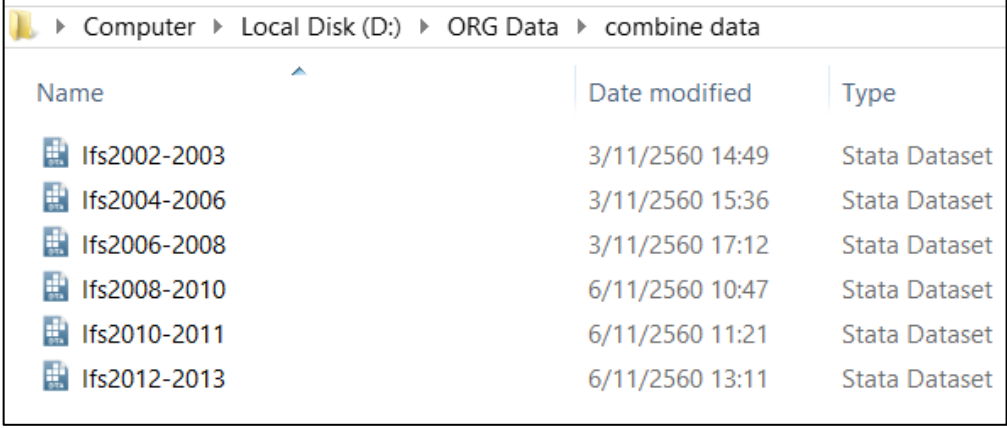
7. ข้อมูล LFS ที่ยังไม่มีการจัดการให้เป็นข้อมูลตัวอย่างซ้ำ ORG แต่จะเป็นข้อมูลที่มีการสร้างชื่อตัวแปรให้เหมือนกันในทุกปีและทุกไตรมาสจะถูกเก็บอยู่ในโฟลเดอร์ prepare data

Name	Date modified	Type
ifs2002q1	3/11/2560 14:48	Stata Dataset
ifs2002q2	3/11/2560 14:48	Stata Dataset
ifs2002q3	3/11/2560 14:48	Stata Dataset
ifs2002q4	3/11/2560 14:48	Stata Dataset
ifs2003q1	3/11/2560 14:49	Stata Dataset
ifs2003q2	3/11/2560 14:49	Stata Dataset
ifs2003q3	3/11/2560 14:49	Stata Dataset
ifs2003q4	3/11/2560 14:49	Stata Dataset
ifs2004q1	3/11/2560 15:33	Stata Dataset
ifs2004q2	3/11/2560 15:33	Stata Dataset
ifs2004q3	3/11/2560 15:33	Stata Dataset
ifs2004q4	3/11/2560 15:34	Stata Dataset

8. ข้อมูล LFS ที่แบ่งตาม samset และ rotation แต่ยังไม่มีการจัดการให้เป็นข้อมูล ORG ตัวอย่างซ้ำ จะถูกเก็บไว้ในโฟลเดอร์ keep data

Name	Date modified	Type
Rotation1_samsetA	6/11/2560 12:08	Stata Dataset
Rotation1_samsetA_keep	6/11/2560 12:11	Stata Dataset
Rotation1_samsetB	6/11/2560 12:10	Stata Dataset
Rotation1_samsetB_keep	6/11/2560 12:11	Stata Dataset
Rotation1_samsetC	6/11/2560 12:10	Stata Dataset
Rotation1_samsetC_keep	6/11/2560 12:11	Stata Dataset
Rotation1_samsetD	6/11/2560 12:11	Stata Dataset
Rotation1_samsetD_keep	6/11/2560 12:12	Stata Dataset

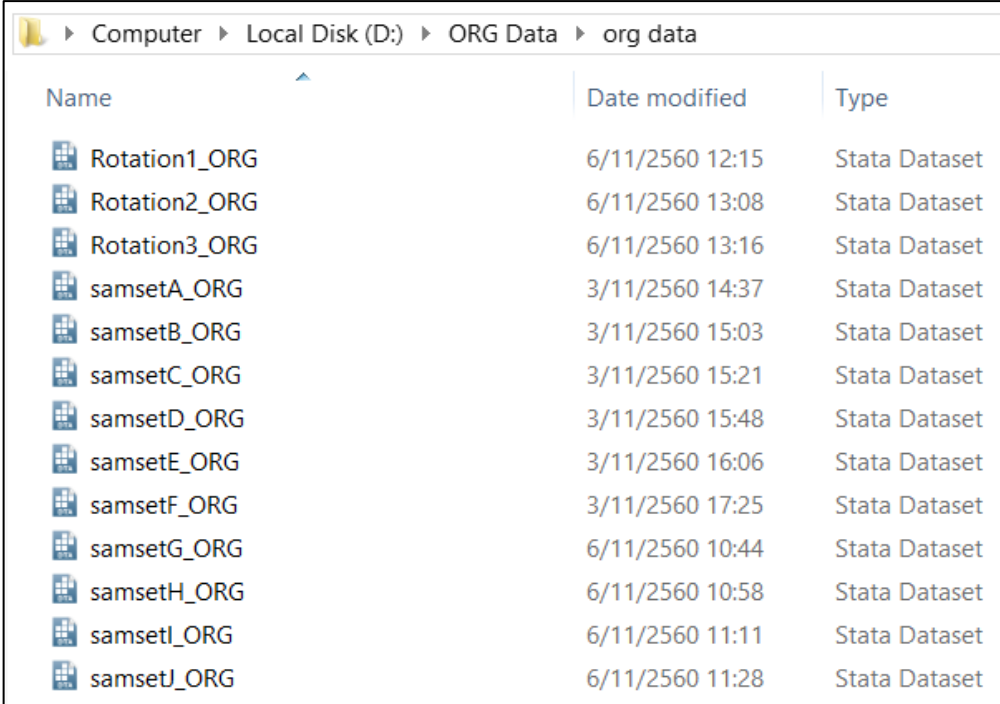
9. ข้อมูล LFS ที่รวมชุดข้อมูลในแต่ละช่วงก่อนที่จะนำไปจัดการให้เป็นข้อมูล ORG จะเก็บอยู่ในโฟลเดอร์ combine data



Computer > Local Disk (D:) > ORG Data > combine data

Name	Date modified	Type
ifs2002-2003	3/11/2560 14:49	Stata Dataset
ifs2004-2006	3/11/2560 15:36	Stata Dataset
ifs2006-2008	3/11/2560 17:12	Stata Dataset
ifs2008-2010	6/11/2560 10:47	Stata Dataset
ifs2010-2011	6/11/2560 11:21	Stata Dataset
ifs2012-2013	6/11/2560 13:11	Stata Dataset

10. ข้อมูล ORG แต่ละชุดจะอยู่ในโฟลเดอร์ org data



Computer > Local Disk (D:) > ORG Data > org data

Name	Date modified	Type
Rotation1_ORG	6/11/2560 12:15	Stata Dataset
Rotation2_ORG	6/11/2560 13:08	Stata Dataset
Rotation3_ORG	6/11/2560 13:16	Stata Dataset
samsetA_ORG	3/11/2560 14:37	Stata Dataset
samsetB_ORG	3/11/2560 15:03	Stata Dataset
samsetC_ORG	3/11/2560 15:21	Stata Dataset
samsetD_ORG	3/11/2560 15:48	Stata Dataset
samsetE_ORG	3/11/2560 16:06	Stata Dataset
samsetF_ORG	3/11/2560 17:25	Stata Dataset
samsetG_ORG	6/11/2560 10:44	Stata Dataset
samsetH_ORG	6/11/2560 10:58	Stata Dataset
samsetI_ORG	6/11/2560 11:11	Stata Dataset
samsetJ_ORG	6/11/2560 11:28	Stata Dataset

ตารางที่ 1: ตัวอย่างตารางข้อมูล ORG สำหรับ samset A

caseid	caseid_new	caseid_new2	wave	sex	age	same_member0	same_member1	same_member2	same_member3	same_member4
11010001A100000010501	11010001A100000010501	11010001A100000010501	3	2	33	same	same	same	same	same
11010001A100000010501	11010001A100000010501	11010001A100000010501	4	2	33	same	same	same	same	same
11010001A100000010502	11010001A100000010502	11010001A100000010502	1	2	21	same	same	same	same	not same
11010001A100000010502	11010001A100000010502	11010001A100000010502	2	2	21	same	same	same	same	not same
11010001A100000010502	11010001A100000010502	11010001A100000010502	3	2	22	same	same	same	same	not same
11010001A100000010502	11010001A100000010502	11010001A100000010502	4	2	22	same	same	same	same	not same
11010001A100000010601	11010001A100000010606	11010001A100000010606	2	1	49	not same	not same	not same	not same	not same
11010001A100000010601	11010001A100000010609	11010001A100000010609	4	1	46	not same	not same	not same	not same	not same
11010001A100000010602	11010001A100000010607	11010001A100000010607	2	2	25	not same	same	same	same	same
11010001A100000010602	11010001A100000010610	11010001A100000010610	4	2	32	not same	not same	not same	not same	not same
11010001A100000010603	11010001A100000010608	11010001A100000010608	2	1	20	not same	same	same	same	not same
11010001A100000010603	11010001A100000010607	11010001A100000010607	4	2	26	not same	same	same	same	not same
11010001A100000010604	11010001A100000010608	11010001A100000010608	4	1	21	not same	same	same	same	same
11010001A100000010605	11010001A100000010611	11010001A100000010611	4	2	1	not same	not same	not same	not same	not same
11010001A100000010701	11010001A100000010701	11010001A100000010701	1	1	54	same	same	same	same	same
11010001A100000010701	11010001A100000010701	11010001A100000010701	2	1	54	same	same	same	same	same
11010001A100000010701	11010001A100000010701	11010001A100000010701	3	1	55	same	same	same	same	same
11010001A100000010701	11010001A100000010701	11010001A100000010701	4	1	56	same	same	not same	not same	not same
11010001A100000010702	11010001A100000010702	11010001A100000010702	1	2	44	same	same	same	same	same
11010001A100000010702	11010001A100000010702	11010001A100000010702	2	2	44	same	same	same	same	same
11010001A100000010702	11010001A100000010702	11010001A100000010702	3	2	45	same	same	same	same	same
11010001A100000010702	11010001A100000010702	11010001A100000010702	4	2	45	same	same	same	same	same
11010001A100000010703	11010001A100000010703	11010001A100000010703	1	2	26	same	same	same	same	same
11010001A100000010703	11010001A100000010703	11010001A100000010703	2	2	26	same	same	same	same	same
11010001A100000010703	11010001A100000010703	11010001A100000010703	3	2	27	same	same	same	same	same

4. เกณฑ์ที่ใช้ในการค้นหาข้อมูลตัวอย่างซ้ำรายบุคคล แบ่งออกเป็น 5 เกณฑ์ ได้แก่

1. กลุ่มตัวอย่างซ้ำคนเดียวกันต้องมีเพศเดียวกัน และอายุไม่เกิน 2 ปี นับตั้งแต่การถูกสัมภาษณ์ ครั้งแรกไปจนถึงครั้งสุดท้ายที่ถูกสัมภาษณ์ และอายุจะต้องเท่าเดิมหรือเพิ่มขึ้นตามครั้งที่ สัมภาษณ์ (ตัวแปรที่ใช้ same_member1)
2. กลุ่มตัวอย่างซ้ำคนเดียวกันต้องมีเพศเดียวกัน และอายุไม่เกิน 1 ปี นับตั้งแต่การถูกสัมภาษณ์ ครั้งแรกไปจนถึงครั้งสุดท้ายที่ถูกสัมภาษณ์ และอายุจะต้องเท่าเดิมหรือเพิ่มขึ้นตามครั้งที่ สัมภาษณ์ (ตัวแปรที่ใช้ same_member2)
3. กลุ่มตัวอย่างซ้ำคนเดียวกันต้องมีเพศเดียวกัน และอายุไม่เกิน 1 ปี นับตั้งแต่การถูกสัมภาษณ์ ครั้งแรกไปจนถึงครั้งสุดท้ายที่ถูกสัมภาษณ์ และอายุจะต้องเท่าเดิมหรือเพิ่มขึ้นตามครั้งที่ สัมภาษณ์ และสถานภาพสมรสต้องเป็นไปตามเงื่อนไขของแต่ละสถานภาพ (ตัวแปรที่ใช้ same_member3)
4. กลุ่มตัวอย่างซ้ำคนเดียวกันต้องมีเพศเดียวกัน และอายุไม่เกิน 1 ปี นับตั้งแต่การถูกสัมภาษณ์ ครั้งแรกไปจนถึงครั้งสุดท้ายที่ถูกสัมภาษณ์ และอายุจะต้องเท่าเดิมหรือเพิ่มขึ้นตามครั้งที่ สัมภาษณ์ และจำนวนปีการศึกษาจะต้องมีจำนวนเท่าเดิมหรือเพิ่มขึ้นไม่เกิน 1 ปีการศึกษา นับตั้งแต่การถูกสัมภาษณ์ครั้งแรกจนถึงครั้งสุดท้าย (ตัวแปรที่ใช้ same_member4)
5. กลุ่มตัวอย่างซ้ำคนเดียวกันต้องมีเพศเดียวกัน และอายุไม่เกิน 1 ปี นับตั้งแต่การถูกสัมภาษณ์ ครั้งแรกไปจนถึงครั้งสุดท้ายที่ถูกสัมภาษณ์ และอายุจะต้องเท่าเดิมหรือเพิ่มขึ้นตามครั้งที่ สัมภาษณ์ และสถานภาพสมรสเป็นไปตามเงื่อนไข และจำนวนปีการศึกษาจะต้องมีจำนวน เท่าเดิมหรือเพิ่มขึ้นไม่เกิน 1 ปีการศึกษานับตั้งแต่การถูกสัมภาษณ์ครั้งแรกจนถึงครั้งสุดท้าย (ตัวแปรที่ใช้ same_member5)

หมายเหตุ : เกณฑ์ที่ 0 เป็นเกณฑ์ที่ใช้แยกระหว่างการเชื่อมโยงโดยใช้ caseid และไม่สามารถใช้ caseid ได้ ซึ่งต้องเชื่อมโยงจากการใช้เพศเดียวกันและอายุที่ใกล้เคียงกัน (ตัวแปรที่ใช้ same_member0)

นอกจากนี้ นักวิจัยได้ทำการตรวจสอบข้อมูลตัวอย่างซ้ำรายบุคคลโดยใช้ caseid ที่สร้างขึ้น จากตัวแปร demographics และ caseid ใหม่ที่สร้างขึ้นจากข้อมูลที่ไม่สามารถใช้ caseid เดิมได้ โดยตั้งชื่อว่า caseid_new เพราะพบว่า อายุและเพศต่างจาก caseid ที่ถูกสัมภาษณ์ในครั้งแรก ตาม ตัวอย่างในตารางที่ 2

ตารางที่ 2: ตัวอย่างตารางข้อมูลที่ต้องมีการสร้าง caseid ใหม่

ครัวเรือนที่	ลำดับที่สมาชิก	ปี	ไตรมาส	ครั้งที่สัมภาษณ์	เพศ	อายุ	caseid
1	1	2545	1	1	ชาย	55	ใช้แบบเดิม
1	1	2545	4	2	ชาย	34	ต้องสร้างใหม่
1	1	2546	1	3	ชาย	34	ต้องสร้างใหม่
1	1	2546	4	4	ชาย	35	ต้องสร้างใหม่
1	2	2545	1	1	หญิง	29	ใช้แบบเดิม
1	2	2545	4	2	หญิง	29	ใช้แบบเดิม
1	2	2546	1	3	ชาย	17	ต้องสร้างใหม่
1	2	2546	4	4	ชาย	18	ต้องสร้างใหม่

จากตารางที่ 2 จะเห็นได้ว่า ครัวเรือนตัวอย่างที่ 1 สมาชิกคนที่ 1 ถูกสัมภาษณ์ครั้งที่ 1 มีอายุ 55 ปี เพศชาย และเมื่อสมาชิกคนที่ 1 ถูกสัมภาษณ์ครั้งที่ 2 มีอายุ 34 ปี เพศชาย และเมื่อถูก สัมภาษณ์ครั้งที่ 3 และ 4 สมาชิกคนที่ 1 มีอายุ 34 ปี เพศชาย และอายุ 35 ปี เพศชาย ตามลำดับ ซึ่งพบว่า สมาชิกคนที่ 1 ที่ถูกสัมภาษณ์ครั้งแรกมีอายุห่างจากการถูกสัมภาษณ์ครั้งที่ 2-4 มาก พอสมควร ทำให้สามารถสรุปได้ว่า สมาชิกคนที่ 1 ที่ถูกสัมภาษณ์ครั้งที่ 1 ไม่ใช่คนเดียวกับสมาชิก คนที่ 1 ที่ถูกสัมภาษณ์ครั้งที่ 2-4 เพราะฉะนั้น จึงจำเป็นต้องสร้าง caseid ใหม่ให้กับสมาชิกที่ถูก สัมภาษณ์ครั้งที่ 2-4

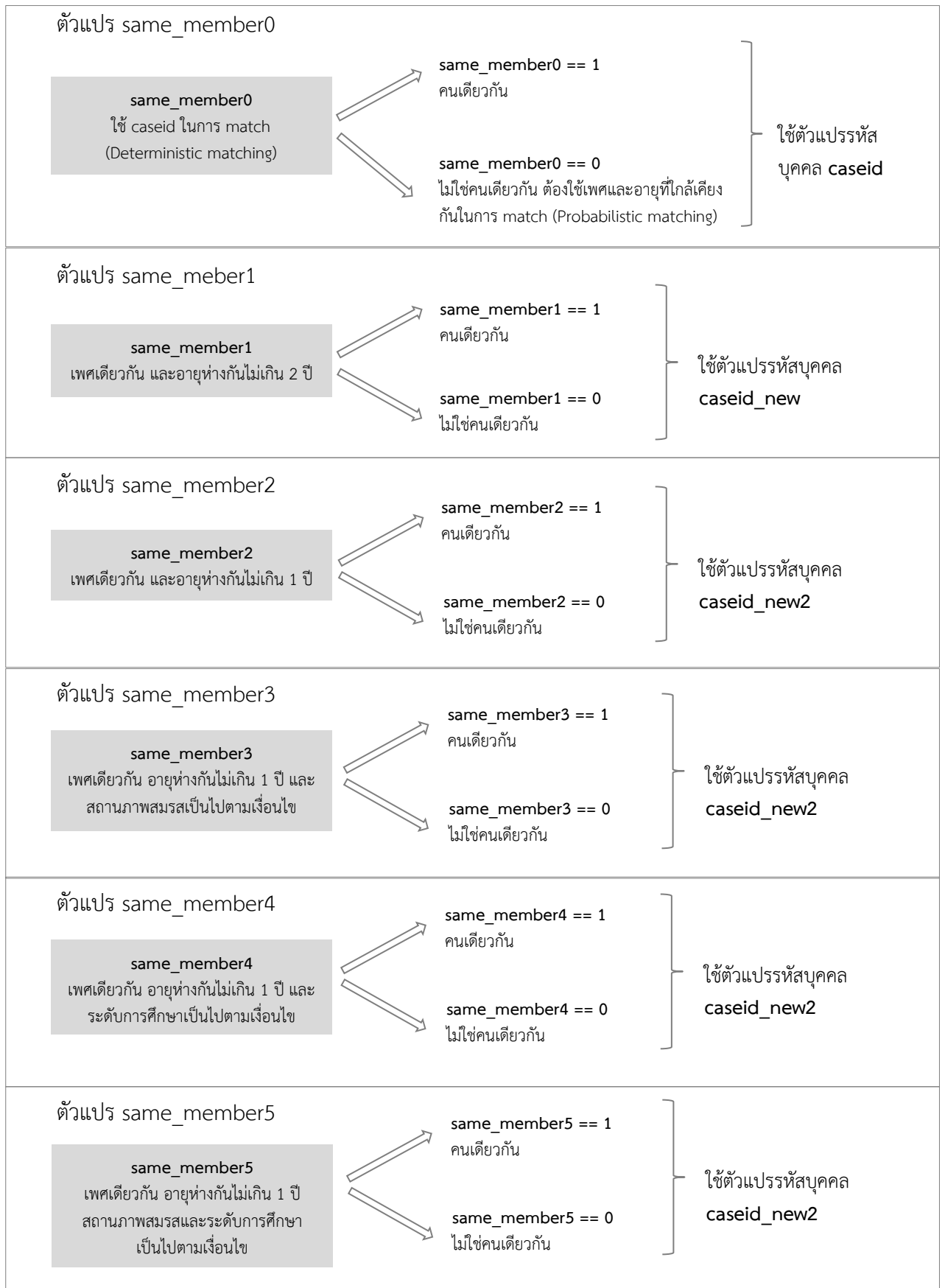
ในการตรวจสอบข้อมูลกลุ่มตัวอย่างซ้ำที่ใช้ caseid ใหม่ ก็ยังคงใช้เกณฑ์ตามเงื่อนไขที่ 1 – 5 เช่นเดียวกับการใช้ caseid แบบเดิม ซึ่งสำหรับการสร้าง caseid ใหม่ นักวิจัยใช้ตัวแปรเพศและ อายุในการเชื่อมโยง โดยได้แบ่งการสร้าง caseid ใหม่จาก 2 เงื่อนไข คือ 1) อายุของคนดังกล่าวที่ ต้องใช้ caseid_new เดียวกันจะต้องมีอายุห่างกันไม่เกิน 2 ปี (กำหนดเป็น caseid_new) และ 2) อายุของคนดังกล่าวที่ต้องใช้ caseid_new เดียวกันจะต้องมีอายุห่างกันไม่เกิน 1 ปี (กำหนดเป็น caseid_new2) นอกจากนี้ นักวิจัยจะใช้เลข 2 หลักสุดท้ายของ caseid_new ซึ่งเป็นลำดับที่ต่อจาก สมาชิกครัวเรือนสุดท้ายของ caseid เดิม ตามตัวอย่างในตารางที่ 3 ดังต่อไปนี้

ตารางที่ 3: ตัวอย่างตารางข้อมูลที่มีการสร้าง caseid_new และ caseid_new2

ครัวเรือนที่	ลำดับที่สมาชิก	ปี	ไตรมาส	ครั้งที่สัมภาษณ์	เพศ	อายุ	caseid	caseid_new	caseid_new2
2	1	2545	1	1	ชาย	42	01	01	01
2	1	2545	4	2	ชาย	42	01	01	01
2	1	2546	1	3	ชาย	35	01	03	03
2	1	2546	4	4	ชาย	37	01	03	04
2	2	2545	1	1	หญิง	29	02	02	02
2	2	2545	4	2	หญิง	29	02	02	02
2	2	2546	1	3	ชาย	18	02	04	05
2	2	2546	4	4	ชาย	18	02	04	05

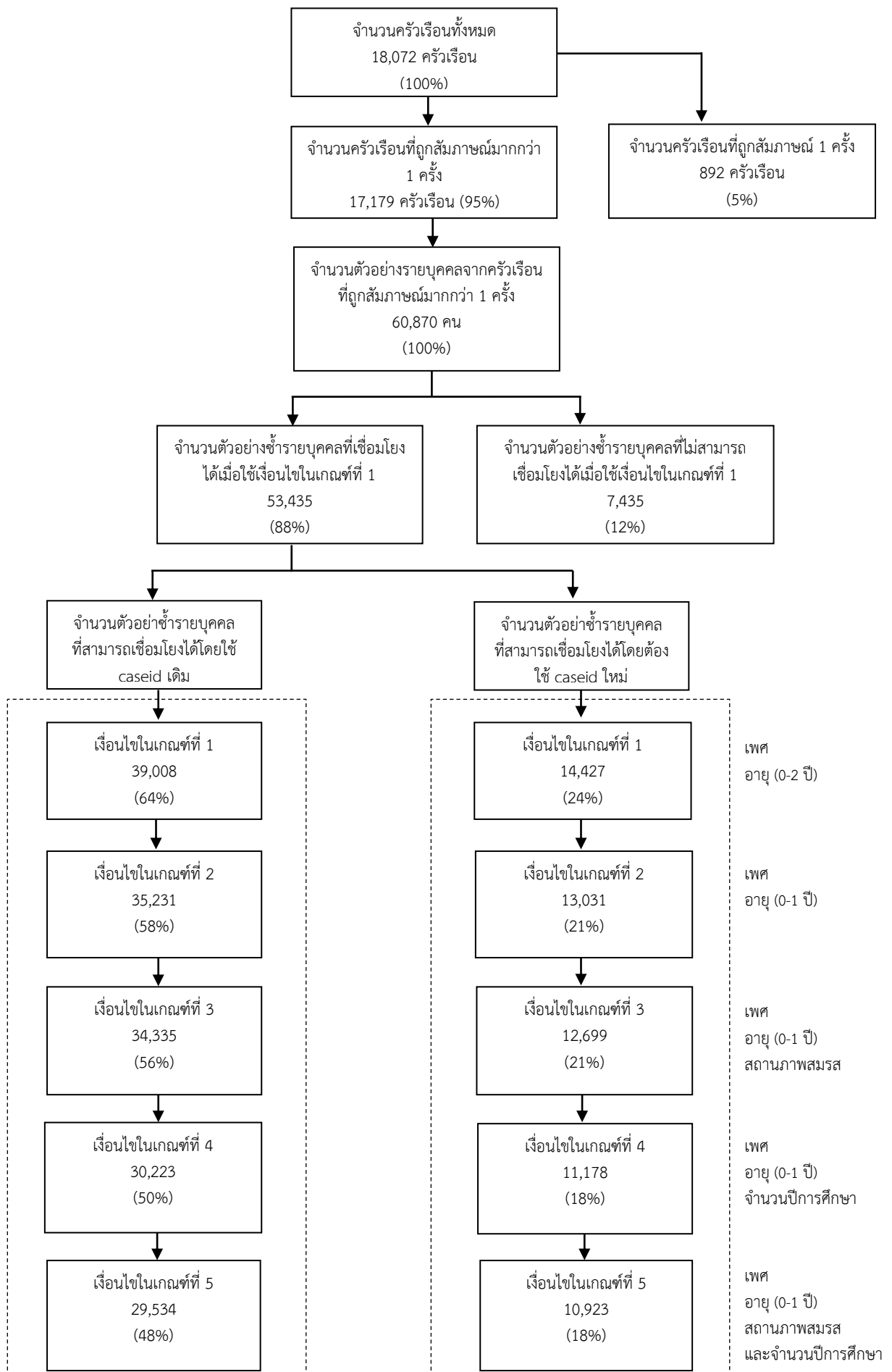
จากตารางที่ 3 พบว่า ครัวเรือนที่ 2 สมาชิกคนที่ 1 ถูกสัมภาษณ์ครั้งที่ 1 มีอายุ 42 ปี เพศชาย เมื่อสมาชิกคนที่ 1 ถูกสัมภาษณ์ครั้งที่ 2 มีอายุ 42 ปี เพศชาย และสมาชิกคนที่ 1 ถูกสัมภาษณ์ครั้งที่ 3 มีอายุ 35 ปี เพศชาย และสมาชิกคนที่ 1 ถูกสัมภาษณ์ครั้งที่ 4 มีอายุ 37 ปี เพศชาย ซึ่งจะเห็นว่า สมาชิกคนที่ 1 ที่ถูกสัมภาษณ์ครั้งที่ 3-4 ไม่ใช่คนเดียวเท่ากับสมาชิกคนที่ 1 ที่ถูกสัมภาษณ์ครั้งที่ 1-2 ดังนั้น เราจึงต้องสร้าง caseid ใหม่ให้กับสมาชิกคนดังกล่าว และเนื่องจากว่า สมาชิกที่ถูกสัมภาษณ์ครั้งที่ 3 มีอายุ 35 ปี และถูกสัมภาษณ์ครั้งที่ 4 มีอายุ 37 ปี เป็นไปตามเงื่อนไขในการสร้าง caseid_new เพราะมีอายุห่างกันไม่เกิน 2 ปี แต่ถ้าต้องสร้าง caseid_new2 มีเงื่อนไขว่าต้องมีอายุห่างกันไม่เกิน 1 ปี ทำให้สมาชิกคนนี้จะเป็นคนเดียวกันเมื่อใช้ caseid_new แต่จะเป็นคนละคนกันถ้าใช้ caseid_new2

ผู้ใช้สามารถเลือกใช้เกณฑ์การจัดทำกลุ่มตัวอย่างซ้ำได้จากตัวแปร same_member ซึ่งมีรายละเอียด ดังนี้



จากเกณฑ์ข้างต้น สามารถสรุปจำนวนข้อมูลกลุ่มตัวอย่างซ้ำรายบุคคลเฉลี่ยต่อ 1 rotation plan โดยพบว่ามีจำนวนครัวเรือนทั้งหมดโดยเฉลี่ย 18,072 ครัวเรือน แบ่งเป็นจำนวนครัวเรือนที่ถูกสัมภาษณ์เพียง 1 ครั้ง เฉลี่ย 892 ครัวเรือน คิดเป็นร้อยละ 5 ของครัวเรือนทั้งหมด และจำนวนครัวเรือนที่ถูกสัมภาษณ์มากกว่า 1 ครั้ง เฉลี่ย 17,179 ครัวเรือน คิดเป็นร้อยละ 95 ของครัวเรือนทั้งหมด โดยที่ครัวเรือนที่ถูกสัมภาษณ์มากกว่า 1 ครั้ง มีจำนวนตัวอย่างรายบุคคลทั้งหมดโดยเฉลี่ย 60,870 คน ซึ่งคิดเป็นตัวอย่างรายบุคคลที่ถูกสัมภาษณ์มากกว่า 1 ครั้งขึ้นไป มีจำนวนเฉลี่ย 53,887 คน คิดเป็นร้อยละ 89 ของจำนวนคนทั้งหมด (ใช้เงื่อนไขอายุห่างกันไม่เกิน 2 ปี) และตัวอย่างรายบุคคลที่ถูกสัมภาษณ์มากกว่า 1 ครั้งขึ้นไป มีจำนวนเฉลี่ย 53,154 คน คิดเป็นร้อยละ 87 ของจำนวนคนทั้งหมด (ใช้เงื่อนไขอายุห่างกันไม่เกิน 1 ปี)

เมื่อใช้เกณฑ์การค้นหาข้อมูลตัวอย่างซ้ำรายบุคคลที่อยู่ในครัวเรือนตัวอย่างซ้ำ (ครัวเรือนที่ถูกสัมภาษณ์มากกว่า 1 ครั้ง) พบว่า หากใช้เกณฑ์ที่ 5 ซึ่งรัดกุมที่สุดจะมีจำนวนตัวอย่างซ้ำรายบุคคลเฉลี่ย 40,457 คน คิดเป็นร้อยละ 66 ของจำนวนคนทั้งหมด และหากใช้เกณฑ์ที่ 3 พบว่า มีจำนวนตัวอย่างซ้ำรายบุคคลเฉลี่ย 41,401 คน คิดเป็นร้อยละ 68 ของจำนวนคนทั้งหมด และหากใช้เกณฑ์ที่ 4 พบว่า มีจำนวนตัวอย่างซ้ำรายบุคคลเฉลี่ย 47,034 คน คิดเป็นร้อยละ 77 ของจำนวนคนทั้งหมด ในขณะที่หากใช้เกณฑ์ที่ 2 จะมีจำนวนตัวอย่างซ้ำรายบุคคลเฉลี่ย 48,262 คน คิดเป็นร้อยละ 79 ของจำนวนคนทั้งหมด และเมื่อใช้เกณฑ์ที่ 1 ซึ่งมีความรัดกมน้อยที่สุด จะพบว่า มีจำนวนตัวอย่างซ้ำรายบุคคลเฉลี่ย 53,435 คน คิดเป็นร้อยละ 88 ของจำนวนคนทั้งหมด ทั้งนี้ขึ้นอยู่กับผู้ใช้ข้อมูลที่ต้องตัดสินใจว่าจะเลือกใช้เกณฑ์ใด โดยจำนวนข้อมูลตัวอย่างซ้ำรายบุคคลสามารถแสดงได้จากรูปดังต่อไปนี้



แผนภาพสรุปจำนวนข้อมูลตัวอย่างซ้ำรายบุคคลเฉลี่ยต่อ 1 rotation plan