基礎講座

SASユーザー総会2011

SAS 言語の強力なデータハンドリング技法

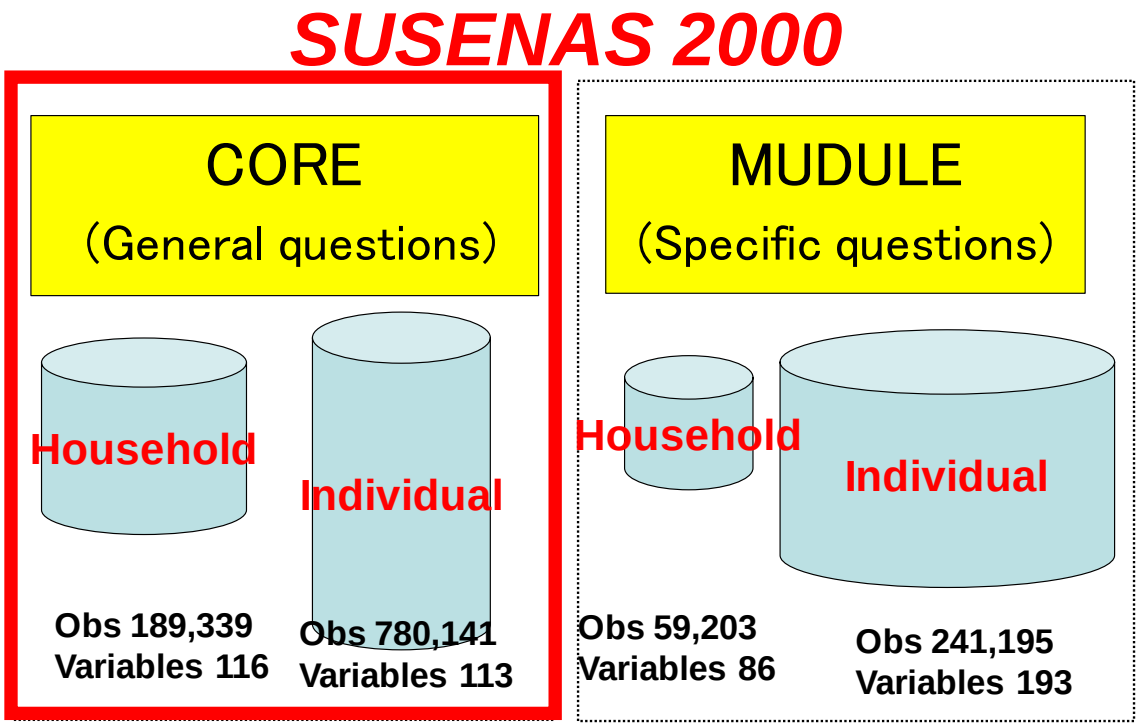
Powerful Data Handling Techniques of SAS Language

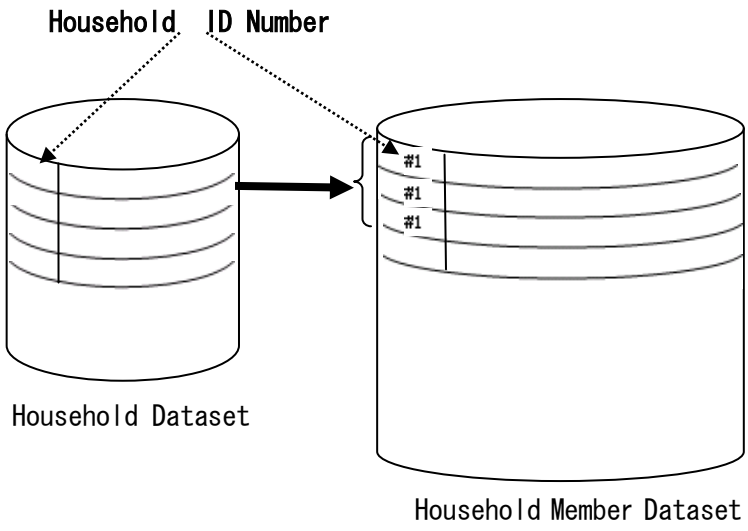
兵庫県立大学　周防 節雄

2011/07/27-28

要旨

他の統計解析ソフトと比較して SAS が抜群に優れている機能の一つに、データハンドリング能力がある。特に、大容量の複雑な構造を持つ複数のデータ間のファイル処理、あるいは、大容量でないファイルの場合でも構造が例えば階層構造になっているような場合には、他の統計解析ソフトでは事実上不可能と言っても過言ではない。このような場合、SAS ユーザーであっても、初期入力データの SAS への取り込みは、他の汎用言語に頼るケースが多い。或いは、企業のユーザーなら、その部分は外注に依存することも多い。SAS では、最初の初期入力データの取り込みから、データセットのマージなどの加工や最後の分析処理まで一貫して行うことができるが、これには熟練を要する。本報告では、インドネシア国家統計局の *National Social Economic Survey 2000 (SUSENAS 2000)*のミクロデータ(個票データ)を使って、多様な SAS のデータハンドリング機能を解説する。





Relation between Household and Household Member Datasets

Today's Focus

1. How to handle large volume of datasets
2. How to merge two datasets by common identification number
3. How to aggregate some attributes of multiple observations into one observation

HOW TO USE MICRODATA FOR ANALYSES

1. SAS PROGRAM TO MAKE COMMON UNIQUE IDENTIFICATION NUMBER

```
/* indonesia_data_with_HHid.sas */ *Original DS + HHid; options nocenter;
libname original "C:¥SAS_Forum¥2011¥basic_course¥sasds¥original";
libname suoh     "C:¥SAS_Forum¥2011¥basic_course¥sasds¥suoh";

data suoh.core_household; /* Add unique household idno to HH dataset */
    length HHid $ 18;
    set original.core_household;
    HHid=K1R1 || K1R2 || K1R3 || K1R4 || K1R5 || K1R8 || K1R9;
run;

proc print data=suoh.core_household(obs=10); title "core_HH + HHid"; run;

data suoh.core_individual;  /* Add unique HH idno to Individual dataset */
    length HHid $ 18;
    set original.core_individual;
    HHid=K1R1 || K1R2 || K1R3 || K1R4 || K1R5 || K1R8 || K1R9;
run;

proc print data=suoh.core_individual(obs=20);
    title "core_individual + HHid";
run;
```

2. SAS PROGRAM TO CREATE DATASET FOR ANALYSIS FROM ORIGINAL MICRODATA

```
/* indonesia_data_extrcted.sas */ *Make dataset with renamed variables;

%include
"C:¥SAS_Forum¥2011¥basic_course¥sasprogram¥format_susenas_hhold.sas";

%include
"C:¥SAS_Forum¥2011¥basic_course¥sasprogram¥format_susenas_individual.sas";

options nocenter nodate nonumber;

libname suoh "C:¥SAS_Forum¥2011¥basic_course¥sasds¥suoh";
```

APPENDIX 3

APPENDIX 4

```
data suoh.core_household_extracted;
    set suoh.core_household;
    keep   HHid
           K8R1
           K8R2
           K8R3
           K8R4
           K8R5
           K8R6A
           K8R7
           K8R8
           K8R9A
           K8R9B
           K8R9C
           K8R10
           K9R26
           K9R27
           K9R30
           K10R4A2
           K10R4A3
           WERT00
           KAB
           K2R2;
```

Household

```
    rename K8R1=house
           K8R2=roof
           K8R3=wall
           K8R4=floor
           K8R5=floor_area
           K8R6A=water_source
           K8R7=water_get
           K8R8=water_facility
           K8R9A=toilet_facility
           K8R9B=toilet_type
           K8R9C=final_dsposal
           K8R10=light
           K9R26=monthly_food
           K9R27=monthly_non_f
           K9R30=celebration
           K10R4A2=own_rice_f
           K10R4A3=own_upland
           WERT00=weight_hhold
           KAB=province
           K2R2=no_hhold_member;
run;

proc print data= suoh.core_household_extracted(obs=10);
    title "core_household_extracted";
        format celebration $celebration.;
run;
```

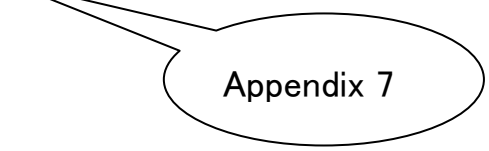
Appendix 6

```
data suoh.core_individual_extracted;
    set suoh.core_individual;
    keep   HHid
           MN0ART
           HB
           JK
           USIA
           ST
           M500i
           K5R15A
           K5R17
           K5R18
           K6R27A
           K6R27B
           K7R28
           K7R29
           K7R30
           K7R31A2
           K7R31A3
           K7R31B2
           K7R31B3
           K7R31C2
           K7R31C3
           WEIND00;
```

Individual

```
    rename MNOART=memberid
           HB=relation
           JK=gender
           USIA=age
           ST=mstatus
           M500i=mother_who  /* specified by "memberid" */
           K5R15A=edu_attend
           K5R17 =edu_own
           K5R18=literacy
           K6R27A=wage_cash
           K6R27B=wage_goods
           K7R28=age_marry1
           K7R29=marry_years
           K7R30=marry_times
           K7R31A2=alive_born_children
           K7R31A3=alive_born_children_Jan_1997
           K7R31B2=still_alive_children
           K7R31B3=still_alive_children_Jan_1997
           K7R31C2=born_die_children
           K7R31C3=die_children_Jan_1997
           WEINDOO=weight_individual;
run;

proc print data= suoh.core_individual_extracted(obs=10);
    title "core_individual_extracted";
    format relation relation.;
    format gender gender.;
    format mstatus mstatus.;
    format edu_attend $edu_attend.;
    format literacy $literacy.;
run;
```



3. SAS PROGRAM FOR MERGING TWO DATASETS FOR FURTHER ANALYSES

```
/* HHstructure_chk.sas */

%let ods_folder=C:¥SAS_Forum¥2011¥basic_course¥ods_output;
*Specify the folder where  the output is saved;

%include
"C:¥SAS_Forum¥2011¥basic_course¥sasprogram¥format_susenas_hhold.sas";

%include
"C:¥SAS_Forum¥2011¥basic_course¥sasprogram¥format_susenas_individual.sas";

options nocenter ls=120;

libname suoh "C:¥SAS_Forum¥2011¥basic_course¥sasds¥suoh";
```

```
data HHstructure; *Observation unit: Household member;
    keep wage_cash wage_goods gender wage_cash wage_goods relation mstatus
         HHid memberid age mother_who age_marry1 marry_years marry_times;
    set suoh.core_individual_extracted;
    memberid=right(memberid);
    if wage_cash =0 then wage_cash=.;
    if wage_goods=0 then wage_goods=.;
run;

proc sort data=HHstructure; by HHid memberid; run;
```

```
ods listing close;
ods html file="&ods_folder¥HHstructure_chk.xls";
```

All results saved here.

```
proc freq; tables gender*relation /missing;
    format relation relation.;
    format gender gender.; title "(1) data=HHstructure (FREQ)";
run;
```

```
proc freq; tables gender*mstatus /missing;
    format mstatus mstatus.;
    format gender gender.;  title "(2) data=HHstructure (FREQ)";
run;

proc tabulate;
    class gender relation mstatus;
    table (relation="" ALL),
          (mstatus="" ALL)*(gender="" ALL);
    format relation relation.;
    format gender gender.;
    format mstatus mstatus.;  title "(3) data=HHstructure (TABULATE)";
run;
```

```
data married_member; *Observation unit: Household;
    *Count the No. of married people in each HouseHold;
    keep HHid no_of_married_member no_of_married_M no_of_married_F;
    set HHstructure; by HHid;
    if first.HHid then do; no_of_married_member=0;
                           no_of_married_M=0;
                           no_of_married_F=0;
                       end;
    if mstatus=2 then do; no_of_married_member+1;
                      if gender=1 then no_of_married_M+1;
                      if gender=2 then no_of_married_F+1;
                  end;
    if last.HHid;
run;

proc print data=married_member(obs=20); title “(4) data=married_member” ; run;
```

```
data HHstructure; *Observation unit: Household member;
    /* Add the No. of married people in each HouseHold. */
    merge HHstructure married_member; by HHid;
run;

proc print data=HHstructure(obs=50);
    title "(5) Add the No. of married people in each HouseHold(1st 50 obs)";
    format relation relation.;
    format gender gender.;
    format mstatus mstatus.;
run;
```

```
data married_in_HH;
    set HHstructure; by HHid;
    if first.HHid;
run;

proc print data=married_in_HH(obs=20); title "(6) married_in_HH(obs=20)"; run;

proc chart data=married_in_HH; hbar no_of_married_member /discrete;
    title "(7) No. of Married Members in each Household";
run;
```

```
data core_HH;
    merge suoh.core_household_extracted married_member; by HHid;
run;

proc print data=core_HH(obs=30); title "(8) data=core_HH"; run;

proc chart data=core_HH; hbar celebration /discrete missing;
    title "(9) celebration(religion) of each HouseHold";
    format celebration $celebration.;
run;

proc freq data=core_HH; tables no_of_married_member*celebration /missing;
    format celebration $celebration.; title "(10) Married members*celeblation";
run;
```

```
data celebration;
    keep HHid no_of_married_member no_of_married_M no_of_married_F celebration;
    set core_HH;
run;

proc print data=celebration(obs=30); title "(11) data=celebration;"; run;

data HHstructure; *Observation unit: Household member;
    *"Celebration"(religion) etc. added. to individual member.;
    merge HHstructure celebration; by HHid;
run;

proc print data=HHstructure (obs=100);
    var HHid memberid relation gender mstatus age celebration marry_times
        no_of_married_member no_of_married_M no_of_married_F;
    title "(12) Add 'celebration' from core_household";
    format relation relation.;
    format gender gender.;
    format mstatus mstatus.;
    format celebration $celebration.;
run;

proc freq data=HHstructure;
    tables marry_times*celebration / norow nocol nopercent;
    title "(13) marry_times and religion";
    format celebration $celebration.;
run;
```

```
data too_many_married;
    set HHstructure;
    if no_of_married_member >= 9;
run;

proc print data=too_many_married(obs=100);
    title "(14) Household no less than 9 married members";
    format celebration $celebration.;
    format relation relation.;
    format gender gender.;
    format mstatus mstatus.;
run;
```

```
data married_people; *Only married people;
    set HHstructure;
    if mstatus=2;
run;

proc print data=married_people(obs=20); title "(15) data=married_people";
run;

proc tabulate data=married_people;
    class relation gender celebration marry_times;
    table relation="" gender="",
          celebration="" marry_times;
    format relation relation.;
    format gender gender.;
    format celebration $celebration.;
    title "(16) Married People Only";
run;
```

```
data marry20;
    set married_people;
    if marry_times=20;
run;

proc print data=marry20;
    var HHid memberid relation gender mstatus age celebration marry_times
        age_marry1 marry_years no_of_married_member no_of_married_M
        no_of_married_F;
    title "(17) data=marry20 (All observations)";
    format relation relation.;
    format gender gender.;
    format mstatus mstatus.;
    format celebration $celebration.;
run;
```

```
data HHstructure;
    *"memberID" and "mother_who" are converted into unique ID in HHstructure;
    set HHstructure;

    length Qmemberid   $ 22;
    length Qmother_who $ 22;

    Qmemberid  =HHid || "X" || right(memberid);   *Unique ID of memberID;
    if mother_who > 0 then Qmother_who=HHid || "X" || right(mother_who);
                           *Unique ID of mother_who;
run;

proc print data=HHstructure(obs=20);
    title "(18) HHstructure with Qmemberid and Qmother_who";
run;
```

```
proc sort data=HHstructure out=QHHstructure; by Qmother_who; run;

data children; * No. of children of a mother;
    keep HHid Qmother_who children;
    rename Qmother_who=Qmemberid;

    set QHHstructure; by Qmother_who;
    if Qmother_who=" " then delete;
    if first.Qmother_who then children=1; *No. of children of this mother ;
                         else children+1;
    if last.Qmother_who;
run;

proc print data=children(obs=20); title "(19) Children"; run;

data xHHstructure; merge HHstructure children; by Qmemberid; run;

proc print data=xHHstructure(obs=20); title "(20) data=xHHstructure"; run;
```

```
ods html close;
ods listing;
```

APPENDIX 1 VARIABLE TABLE (CORE HOUSEHOLD)

SAS Dataset: SSN00K1_HOUSEHOLD
(original file: ssn00k1_household.dbf)

No	Variable	Type	Format	Description	Value code	Value Description
1	K1R1	String	F2	Province code		
2	K1R2	String	F2	Regency/Municipality code		
3	K1R3	String	F3	Subregency code		
4	K1R4	String	F3	Village/"Kelurahan" code		
5	K1R5	String	F1	Village/"Kelurahan" classification	1	Urban
					2	Rural
6	K1R8	String	F5	Sample code number		
7	K1R9	String	F2	Serial No. of sample household		
8	K8R6B	String	F1	Distance to the nearest sewage storage tank	0	Not in universe
					1	<6m
					2	6-10m
					3	11-15m
					4	>=16m
					5	DK
9	FLAG	String	F1	Edit flag		
10	ENTRI	String	F1	Entry number		
11	K2R3	Number	F2	No. of children aged 0-4 years old		
12	K2R4	Number	F2	No. of household members who are being victim of crime/violation during the year 1999		
13	K2R5	Number	F1	No. of household members who passed away since January 1997		
14	K2R6	String	F1	Does this household have a holy book?	1	Yes
					2	No
15	K3R1	Number	F5	Employment identity No. of enumerator		
16	K3R2	String	F1	Enumerator's occupation	1	BPS provincial staff
					2	BPS regency staff
					3	BPS subregency staff
					4	Hired worker
17	K3R5	Number	F5	Employment identity No. of supervisor		
18	K3R6	String	F1	Supervisor's occupation	1	BPS provincial staff
					2	BPS regency staff
					3	BPS subregency staff
					4	Hired worker
19	K8R1	String	F1	Status of house	1	Private, own property
					2	Lease with the agreement on periods
					3	Rental without time limits
					4	Official
					5	Free rent
					6	Others
20	K8R2	String	F1	Type of roof	1	Concrete
					2	Corrugated tile
					3	Shingle roof
					4	Iron sheeting
					5	Asbestos
					6	Sugar palm fiber
					7	Leaves
					8	Others
21	K8R3	String	F1	Type of wall	1	Brick
					2	Wood
					3	Bamboo
					4	Others
22	K8R4	String	F1	Type of floor	1	Marble/ceramic
					2	Floor tile
					3	Cement plaster/bricks
					4	Wood
					5	Bamboo
					6	Earth
					7	Others
23	K8R5	Number	F3	Floor area (m^2)		

65	K9R26	Number	F10	Average of monthly food expenditure (K9R16*30/7)		
66	K9R27	Number	F10	Average of monthly non food expenditure (K9R25B/12)		
67	K9R28	Number	F10	Average of monthly household expenditure (K9R26+K9R27)		
68	K9R29	Number	F10	Main sources of household income		
69	K9R30	String	F1	Celebration on which this household has the largest extra expenditure	1	Idul Fitri(moslems)
					2	New year
					3	Nyepi/Galungan(Hindus)
					4	Christmas/Easter
					5	Waisak(Buddhists)
					6	None
70	K9R31A	Number	F8	Average expenditure of daily household food consumption on common days (Rp.)		
71	K9R31B	Number	F8	Average expenditure of daily household food consumption on holidays (Rp.)		

APPENDIX 2 VARIABLE TABLE (CORE INDIVIDUAL)

SAS Dataset: SSN00K2_INDIVIDU
(original file: ssn00k2_individu.dbf)

No	Variable	Type	Format	Description	Value code	Value Description
1	K1R1	String	F2	Province code		
2	K1R2	String	F2	Regency/Municipality code		
3	K1R3	String	F3	Subregency code		
4	K1R4	String	F3	Village/"Kelurahan" code		
5	K1R5	String	F1	Village/"Kelurahan" classification		
6	K1R8	String	F5	Sample code number		
7	K1R9	String	F2	Serial number of sample household		
8	MNOART	String	F2	Serial number of sample household member		
9	FLAG	String	F1	Edit flag		
10	ENTRI	String	F1	Entry number		
11	HB	Number	F1	Relation to the head of household	1	The head of household
					2	Wife/husband
					3	Children
					4	Son/daughter in-law
					5	Gandchildren
					6	Parent/In-law
					7	Other relative
					8	House maid
					9	Others
12	JK	Number	F1	Sex	1	Male
					2	Female
13	USIA	Number	F2	Age (year)		
14	ST	Number	F1	Marital status (code)	1	Single
					2	Married
					3	Divorced
					4	Widowed
78	K5R15A	String	F2	The highest level and type of education ever or being attended	0	Not in universe
					1	Primary school
					2	"Ibtidaiyah"(islamic)
					3	"A" packet group
					4	Junior high/vocational Jr.school
					5	"Tsanawiyah"(islamic)
					6	"B" packet group
					7	Senior high school
					8	"Aliyah"(islamic)
					9	Vocational high school
					10	Diploma I/II
					11	Diploma III/bachelor"Sarmud"
					12	Diploma IV/graduate
					13	Master
					14	Ph.D.
79	K5R15B	String	F1	Educational organizer	0	Not in universe
					1	Government
					2	Private
					3	Foreign
80	K5R16	Number	F1	The highest level/grade ever or being attended		
81	K5R17	String	F2	The highest certificate that be owned	0	Not in universe
					1	None
					2	Primary school/"Ibtidaiyah"/comparable school
					3	Junior high/"Tsanawiyah"/vocational Jr.school/ comparable school
					4	Senior high school/Aliyah"/comparable school
					5	Vocational high school
					6	Diploma I/II
					7	Diploma III/bachelor"Sarmud"
					8	Diploma IV/graduate
					9	Master
					10	Ph.D.
82	K5R18	String	F1	Can you write and read?	0	Not in universe (USIA = under 5)
					1	Latin
					2	Other alphabets
					3	Can not
83	K6R19A1	Number	F1	Did you work during previous week?	0	Not in universe (USIA = under 10)
					1	Yes
					2	No
98	K7R28	Number	F2	Age first married		
99	K7R29	Number	F2	No. of years of marriage		
100	K7R30	Number	F2	How many times of marriage had been done		
101	K7R31A2	Number	F2	No. of biological children born alive after birth		
102	K7R31A3	Number	F2	No. of biological children born alive since Jan. 1997		
103	K7R31B2	Number	F2	No. of biological children still alive after birth		
104	K7R31B3	Number	F2	No. of biological children still alive since Jan. 1997		
105	K7R31C2	Number	F2	No. of biological children died after birth		
106	K7R31C3	Number	F2	No. of biological children died since Jan. 1997		

APPENDIX 3 DEFINITION OF FORMAT (CORE HOUSEHOLD)

```
* format_susenas_hhold.sas ;

proc format;
value $st_house   /*K8R1*/
"1"="Private, own property"
"2"="Lease with the agreement on periods"
"3"="Rental without time limits"
"4"="Official"
"5"="Free rent"
"6"="Others"
;
value $roof/*K8R2*/
"1"="Concrete"
"2"="Corrugated tile"
"3"="Shingle roof"
"4"="Iron sheeting"
"5"="Asbestos"
"6"="Sugar palm fiber"
"7"="Leaves"
"8"="Others"
;
value $wall /*K8R3*/
"1"="Brick"
"2"="Wood"
"3"="Bamboo"
"4"="Others"
;
value $floor /*K8R4*/
"1"="Marble/ceramic"
"2"="Floor tile"
"3"="Cement plaster/bricks"
"4"="Wood"
"5"="Bamboo"
"6"="Earth"
"7"="Others"
;
value $source_water /*K8R6A*/
"1"="Bottled water"
"2"="Tap water"
"3"="Pump well"
"4"="Protected well"
"5"="Unprotected well"
"6"="Protected spring"
"7"="Unprotected spring"
"8"="River"
"9"="Rain water"
"0"="Others"
```

```
value $get_water   /*K8R7*/
"1"="Purchased"
"2"="Not purchased"
;
value $d_water_facility   /*K8R8*/
"1"="Private"
"2"="Shared"
"3"="Public"
"4"="None"
;
value   $toilet_facility /*K8R9A*/
"1"="Private"
"2"="Shared"
"3"="Public"
"4"="Others"
;
value $toilet /*K8R9B*/
"0"="Not in universe"
"1"="Squat toilet/toilet with U-drain"
"2"="Western style toilet/sit toilet/toilet with straight drain"
"3"="Pit latrine/toilet without drain, directly connecting to container"
"4"="Not use"
;
value $disposal   /*K8R9C*/
"1"="Septic tank"
"2"="Pond/Rice field"
"3"="River/lake/sea"
"4"="Ground hole"
"5"="Shore/open field/gardens"
"6"="Others"
;
value $light   /*K8R10*/
"1"="PLN electricity"
"2"="Electricity non PLN"
"3"="Pump lantern"
"4"="Oil lamp"
"5"="Others"
;
value $celebration /*K9R30*/
"1"="1.moslems"
"2"="2.New year"
"3"="3.Hindus"
"4"="4.Christmas/Easter"
"5"="5.Buddhists"
"6"="6.None"
;
```


Legend; /*K8R1*/ is a variable name.

APPENDIX 4 DEFINITION OF FORMAT (CORE INDIVIDUAL)

```
/* format_susenas_individual.sas */

proc format;
value relation /*HB*/
1="Head of household"
2="Wife/husband"
3="Children"
4="Son/daughter in-law"
5="Grand children"
6="Parent/In-law"
7="Other relative"
8="House maid"
9="Others"
;
value gender /*JK*/
1="Male"
2="Female"
;
value mstatus /*ST*/
1="Single"
2="Married"
3="Divorced"
4="Widowed"
;
value $edu_attend /*K5R15A*/
"0"="Not in universe"
"1"="Primary school"
"2"="'Ibtidaiyah'(islamic)"
"3"="'A' packet group"
"4"="Junior high/vocational Jr.school"
"5"="'Tsanawiyah'(islamic)"
"6"="'B' packet group"
"7"="Senior high school"
"8"="'Aliyah'(islamic)"
"9"="Vocational high school"
"10"="Diploma I/II"
"11"="Diploma III/bachelor'Sarmud'"
"12"="Diploma IV/graduate"
"13"="Master"
"14"="Ph.D."
;
value $edu_own /*K5R17*/
"0"="Not in universe"
"1"="None"
"2"="Primary  school/'Ibtidaiyah'/comparable
school"
"3"="Junior high/'Tsanawiyah'
      /vocational Jr.school
      /comparable school"
"4"="Seniorhigh school/Aliyah'
      /comparable school"
"5"="Vocational high school"
"6"="Diploma I/II"
"7"="Diploma III/bachelor'Sarmud'"
"8"="Diploma IV/graduate"
"9"="Master"
"10"="Ph.D."
;
value $literacy /*K5R18*/
"0"="Not in universe"
"1"="Latin"
"2"="Other alphabets"
"3"="Can not"
;
run;
```

APPENDIX 5 JOB CODE TABLE

code	job
011	Army Force
021	Navy Force
031	Air Force
091	Other Defense Components
111	Legislative member
112	Government Senior Official
113	Non- government Organization Senior Official
121	Prime Director and Executive Leader
122	Agricultural and Mining General Manager
123	Manufacturing, Building and Construction General Manager
124	Services General Manager
131	Resource Manager
132	Technical, Distribution and Processing Manager
133	Marketing and Sales Manager
139	Other Specialists Manager
141	Farmer and Agricultural Manager
211	Chemistry and Physics Expert
212	Architect and Technical Expert
221	Accountant, Auditor, and Company Treasurer
222	Advertising, Marketing and Sales Expert
223	Computing Expert
229	Business and other Information Expert
231	Doctor
232	Nursing Expert
233	Physio- therapist and traditional Therapist Expert
234	Dentist
239	Other Health Experts
241	General School Teacher
242	Handicap School and Vocational Teacher
243	Lecturer
249	Other Education Experts
251	Social Welfare Expert
252	Religion and sects Expert
253	Other Social Expert
254	Artist and related Expert
255	Other Experts
311	Scientific and Medical Technician
312	Technical and Building Expert Assistant
321	Financial Expert Assistant
329	Other Business and Administration Expert Assistant
331	Store Manager
332	Hospitality and Accommodation Manager
339	Other Sales and Service
341	Nurse
342	Welfare Expert Assistant
349	Other Health and Welfare Expert
391	Policeman
399	Unclassified Technician and Expert Assistant

code	job
411	Technical Machine Labor
412	Technical Fabrication Labor Automotive Labor
421	Automotive Labor
431	Electricity and Electric Labor
441	Structural Construction Labor
442	Construction Finishing Touch Labor
443	Water Piping Labor
451	Food Producing Labor
461	Skilled Agricultural Labor
462	Horticulture Producing Labor
491	Printing Labor
492	Wood Man
493	Hair Dresser
494	Textile, Garment and related Goods Labor
498	Handyman and Related work
499	Supported Show labor
511	Secretary and Personal Assistant
591	Advance Financial Administrator
599	Administrator and other Advance Services
611	General Administrator
612	Typist
613	Receptionist
614	Intermediate Financial Administrator and Statistic's staff
615	Recording and Courier Administrator
619	Other Intermediate Administrator
621	Intermediate Sales and its types Staff
631	Nursing and Aid Staff
632	Hotel Service Staff
639	Other Intermediate Service Staff
711	Mobile Manufacture Machine Operator
712	Intermediate Stationary Machine Operator
721	Textile machine, Garment and its types Operator
'729	Other Intermediate Machines Operator
731	Engineer of Locomotive and Driver
791	Intermediate Construction and Mining Labor
799	Unclassified Producing and Transportation Labor
811	Basic Administrator
821	Sales Staff
829	Other Basic Sales Staff
831	Basic Service Staff
911	Cleaning Service Staff
921	Product Processing Labor
922	Product Packaging Labor
991	Mining, Construction, and Related Labor
992	Agriculture and Horticulture Labor
993	Beginning Food Preparing and Related Labor
999	Unskilled and Related Labor

Note: First two digits of code denote large job classification.

APPENDIX 6 SELECTED VARIABLES AFTER RENAMING: HOUSHOLD

core_household_extracted

OBS	HHid	house	roof	wall	floor	floor_ area	water_ source	water_ get
1	XXXXXXXXXXXXXX21	1	7	2	4	66	5	2
2	XXXXXXXXXXXXXX210	1	7	3	3	54	5	2
3	XXXXXXXXXXXXXX211	1	7	2	4	63	5	2
4	XXXXXXXXXXXXXX212	1	7	2	4	70	5	2
5	XXXXXXXXXXXXXX213	1	7	2	4	60	5	2
6	XXXXXXXXXXXXXX214	1	7	2	4	54	5	2
7	XXXXXXXXXXXXXX215	1	7	2	4	63	5	2
8	XXXXXXXXXXXXXX216	1	7	2	4	60	5	2
9	XXXXXXXXXXXXXX22	1	7	2	4	48	5	2
10	XXXXXXXXXXXXXX23	6	7	2	4	40	5	2

OBS	water_ facility	toilet_ facility	toilet_ type	final_ dsposal	light	monthly_ food	monthly_ non_f
1	3	4	0	5	3	274243	101883
2	3	4	0	5	3	381686	122150
3	3	4	0	5	3	474900	117175
4	3	4	0	5	3	407786	101404
5	3	4	0	5	3	527721	165525
6	3	4	0	5	3	276000	92750
7	4	4	0	5	3	356357	101133
8	3	4	0	5	3	377777	97050
9	3	4	0	5	3	673436	131500
10	3	4	0	5	3	463114	123633

OBS	celebration	own_ rice_f	own_ upland	weight_ hhold	province	no_hhold_ member
1	5.Buddhists	0	0	259	1201	4
2	5.Buddhists	0	1	259	1201	6
3	5.Buddhists	0	2	259	1201	5
4	5.Buddhists	0	2	259	1201	4
5	5.Buddhists	0	1	259	1201	7
6	2.New year	0	1	259	1201	3
7	5.Buddhists	0	2	259	1201	4
8	5.Buddhists	0	1	259	1201	4
9	5.Buddhists	0	2	259	1201	7
10	3.Hindus	0	1	259	1201	6

APPENDIX 7 SELECTED VARIABLES AFTER RENAMING: INDIVIDUAL

core_individual_extracted

OBS	HHid	memberid	relation	gender	age	mstatus	mother_who
1	XXXXXXXXXXXXXX21	1	Head of household	Male	35	Married	00
2	XXXXXXXXXXXXXX21	2	Wife/husband	Female	33	Married	00
3	XXXXXXXXXXXXXX21	3	Children	Female	7	Single	2
4	XXXXXXXXXXXXXX21	4	Children	Male	4	Single	2
5	XXXXXXXXXXXXXX210	1	Head of household	Male	47	Married	00
6	XXXXXXXXXXXXXX210	2	Wife/husband	Female	32	Married	00
7	XXXXXXXXXXXXXX210	3	Children	Female	12	Single	2
8	XXXXXXXXXXXXXX210	4	Children	Male	10	Single	2
9	XXXXXXXXXXXXXX210	5	Children	Male	9	Single	2
10	XXXXXXXXXXXXXX210	6	Children	Male	7	Single	2

OBS	edu_attend	edu_own	literacy	wage_cash	Wage_goods	age_marry1	marry_years	marry_times
1	Primary school	2	Latin	0	0	0	0	0
2	Primary school	1	Latin	0	0	25	8	1
3	Primary school	1	Latin	0	0	0	0	0
4	Not in universe	0	Not in universe	0	0	0	0	0
5	Primary school	1	Latin	0	0	0	0	0
6	Primary school	1	Latin	0	0	19	13	1
7	Primary school	1	Latin	0	0	0	0	0
8	Primary school	1	Latin	0	0	0	0	0
9	Primary school	1	Latin	0	0	0	0	0
10	Primary school	1	Latin	0	0	0	0	0

OBS	alive_born_children	alive_born_children_Jan_1997	still_alive_children	still_alive_children_Jan_1997	born_die_children	die_children_Jan_1997	weight_individual
1	0	0	0	0	0	0	245
2	2	0	2	0	0	0	245
3	0	0	0	0	0	0	245
4	0	0	0	0	0	0	245
5	0	0	0	0	0	0	245
6	4	0	4	0	0	0	245
7	0	0	0	0	0	0	245
8	0	0	0	0	0	0	245
9	0	0	0	0	0	0	245
10	0	0	0	0	0	0	245