(Version 1.1)

Users' Manual for Handling Resampled Micro Data of Sri Lanka Household Income and Expenditure Survey HIES 2009/10

2015
The Institute of Statistical Mathematics (ISM)
and
Statistical Information Institute for Consulting and Analysis (SINFONICA)

History of revision of the manual

- Version 1.1 in March 2015

  - Revised based on the discussion during the Sixth International Workshop on Analysis of Micro Data of Official Statistics in December 2014.
  - Revised errors in the resampled micro data.

- First draft version 1.0 in September 2014

CONTENTS

HIES 2009/10

1. About this Manual

1. This manual was prepared for users to use the next 80% resampled micro data sets of Sri Lanka Household Income and Expenditure Survey 2009/10.

Survey year	Data files	Contents
HIES 2009/10	26 files	R01 to R24: HIES 2009/10 Questionnaire R25, R26: Summary files compiled by the author
	CSV files	
	"R01_80.csv" "R02_80.csv" "R03_80.csv" "R04_80.csv" "R05_80.csv" "R06_80.csv" "R07_80.csv" "R08_80.csv" "R09_80.csv" "R10_80.csv" "R11_80.csv" "R12_80.csv" "R13_80.csv" "R14_80.csv" "R15_80.csv" "R16_80.csv" "R17_80.csv" "R18_80.csv" "R19_80.csv" "R20_80.csv" "R21_80.csv" "R22_80.csv" "R23_80.csv" "R24_80.csv" "R25_80.csv" "R26_80.csv"	
	R data frames	
	"R01.80" "R02.80" "R03.80" "R04.80" "R05.80" "R06.80" "R07.80" "R08.80" "R09.80" "R10.80" "R11.80" "R12.80" "R13.80" "R14.80" "R15.80" "R16.80" "R17.80" "R18.80" "R19.80" "R20.80" "R21.80" "R22.80" "R23.80" "R24.80" "R25.80" "R26.80"	

2. The original micro data sets composed of all the samples were provided by DCS, Sri Lanka based on the Charter for Experimental Laboratory for Research Purpose Statistical Use of Micro Data, and resampled at the rate of 80% by Sinfonica.

3. The above resampled data sets are available through the Institute of Statistical Mathematics (ISM) both in R and CSV format.

4. This manual was first compiled in September 2014 by;

Hiroshige Furuta
Visiting Senior Research Fellow, Sinfonica

Acknowledgements

Special thanks to Mr. K.M.R. Wickramasinghe and Ms. Dilhanie Deepawansa, DCS, Sri Lanka, who assisted my work of compiling the manuals by properly answering to my queries via email.

2. Outline of HIES 2009/10

Objective	To provide information on household income and expenditure in order to measure the levels and changes in the living condition of the people, and compute various other indicators such as poverty etc.
Frequency	HIES had been conducted since 1990 as a separate survey once in every five years until HIES 2006/07. Thereafter, as rapidly changing economic conditions demanded for frequent monitoring of income and expenditure patterns in the country, the DCS decided to conduct the HIES once in every three years.
Topics covered	In general, the survey gathers information on the next three main topics; ● Demographic characteristics ● Household expenditure ● Household income After HIES 2006/07, the following seven topics have been introduced; Education, Health, Durable goods, Access to facilities, Debts, Housing, Agriculture holdings and livestock
Data collection	Direct personal interview The data collection in the field is done in twelve monthly rounds from July 2009 to June 2010.
Reference period	Twelve months to capture the seasonal variation of income and expenditure patterns of households. Reference period of income and expenditure depends on item; a week, a month, 6 months or 12 months.
Coverage and scope	Geographically, whole country excluding Mannar, Kilinochchi and Mullaithivu districts in the Northern province as post liberation settlements of internally displaced persons. Private households, excluding collective (institutional) households. A household may be a one-person household or a multi person household. A one-person household is a unit where a person lives by himself and makes separate provision for his food, either cooking himself or purchasing. A multi person household is a group of two or more persons who lives together and has a common arrangement for cooking and partaking food. Boarders and servants who share the meals and housing facilities with other members of the household are also considered as members of the household.
Sample frame	List of census blocks prepared for Census 2001 List of buildings prepared for Census 2001

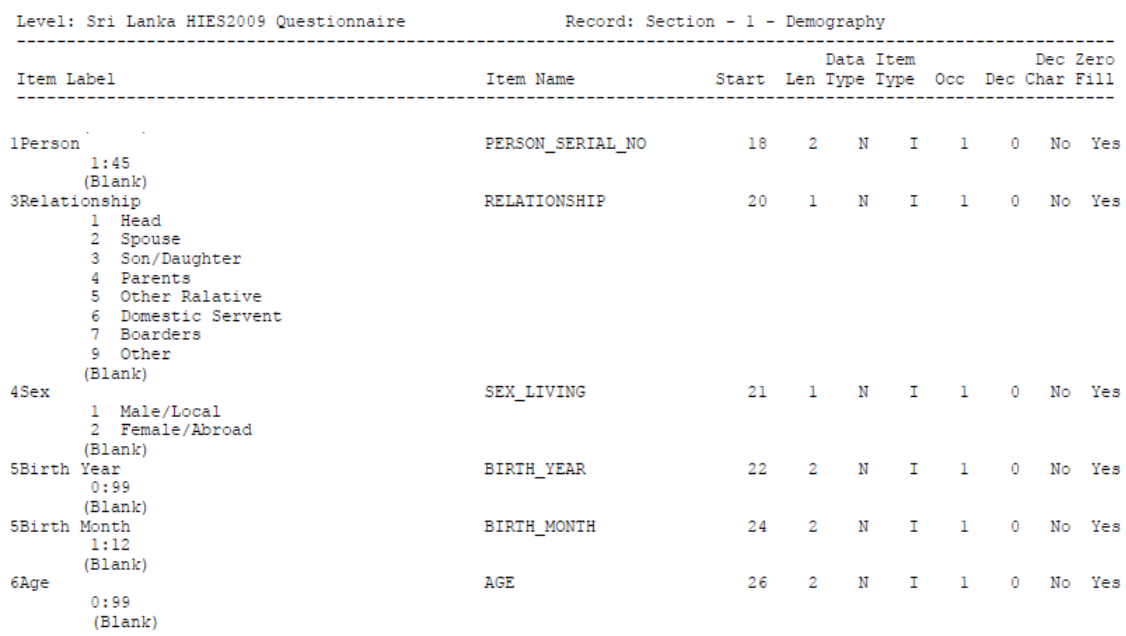
	HIES 2009/10 will be the last HIES sampled from this sampling frame.
Sample design	Two stage stratified random sample design Strata: district (22) and sector (3) Psu: census blocks. 2,500 psu were selected. List of buildings in each selected psu was updated about one month prior to the scheduled interviewing. Ssu: housing units. For each psu, 10 housing units were selected for the survey. Note: Sample households of different year surveys are completely independent.
Sample size and sample allocation	Total sample size is 25,000 housing units. Neymann (optimal) method was employed for allocation of psu for districts and sectors. The district sample was equally distributed among the 12 monthly rounds. In total, 19,958 households responded to the survey.
Data processing	According to the delegates from Sri Lanka for the International Workshop, "Data cleaning is done in districts and head office concurrently with the field work using an online and interactive computer data editing and cleaning program, which reports errors in the data as identified according to the conditions pre-prepared and the corrections are made referring to the hard questionnaire or to the enumerator or the respondent household."
Preliminary report	Based on only the first 3 monthly rounds of the survey to fulfil urgent data needs of the country.
Final report	Based on the 100% data

For more detail, please refer to "1. Introduction" of the survey report.

3. Data and metadata provided by NSO

Two following micro data and metadata were provided to Sinfonica by Mr. S. A. S. Bandulasena, Director of ICT Division, DCS upon request in June 2014.

Type	Filename	Description	Included in this manual
Data			
	HIES-2009-10-100%-Data.txt (66.0M)	Text data file with multi layouts; In total, 1,844,869 lines; 24 record types	
	Folder: HIES-2009-10-100%-Data_CSV Format	The next 24 data files in CSV format	
	[1] "HIES-2009-10-100%-Data-SEC_1_DEMOGRAPHIC.CSV" [2] "HIES-2009-10-100%-Data-SEC_2_SCHOOL_EDUCATION.CSV" [3] "HIES-2009-10-100%-Data-SEC_3_HEALTH.CSV" [4] "HIES-2009-10-100%-Data-SEC_4_1_FOOD_EXP.CSV" [5] "HIES-2009-10-100%-Data-SEC_4_2_NONFOOD.CSV" [6] "HIES-2009-10-100%-Data-SEC_4_3_BOARDERS.CSV" [7] "HIES-2009-10-100%-Data-SEC_4_3_IS_BOADERS.CSV" [8] "HIES-2009-10-100%-Data-SEC_5_1_EMP_INCOME.CSV" [9] "HIES-2009-10-100%-Data-SEC_5_1_IS_EMP_INCOME.CSV" [10] "HIES-2009-10-100%-Data-SEC_5_2_AGRI_INCOME.CSV" [11] "HIES-2009-10-100%-Data-SEC_5_2_IS_AGRI_INCOME.CSV" [12] "HIES-2009-10-100%-Data-SEC_5_3_IS_OTHER_AGRI_INCOME.CSV" [13] "HIES-2009-10-100%-Data-SEC_5_3_OTHER_AGRI_INCOME.CSV" [14] "HIES-2009-10-100%-Data-SEC_5_4_IS_NON_AGRI_INCOME.CSV" [15] "HIES-2009-10-100%-Data-SEC_5_4_NON_AGRI_INCOME.CSV" [16] "HIES-2009-10-100%-Data-SEC_5_5_1_IS_OTHER_INCOME.CSV" [17] "HIES-2009-10-100%-Data-SEC_5_5_1_OTHER_INCOME.CSV" [18] "HIES-2009-10-100%-Data-SEC_5_5_2_IS_WINDFALL_INCOME.CSV" [19] "HIES-2009-10-100%-Data-SEC_5_5_2_WINDFALL_INCOME.CSV" [20] "HIES-2009-10-100%-Data-SEC_6_B_DEBTNESS.CSV" [21] "HIES-2009-10-100%-Data-SEC_6A_DURABLE_GOODS.CSV" [22] "HIES-2009-10-100%-Data-SEC_7_BASIC_FACILITIES.CSV" [23] "HIES-2009-10-100%-Data-SEC_8_HOUSING.CSV"		Yes

	[24] "HIES-2009-10-100%-Data-SEC_9_LAND_ANIMAL.CSV"		
	HIES-2009-10-100%-Factor File.xls	Weight data by district and psu, 2,269 records	Yes
Metadata	HIES-2009_10-Questionnaire.pdf	32 pages	Yes
	HIES-2009-10-Data Layout.pdf	Note: The original file provided was for HIES 2006/07. The revised Data Layout as well as the next Data Dictionary were provided upon request. 25 pages	
	Sri Lanka HIES-2009-10-Data Layout with codes.pdf	Data dictionary/Codebook 65 pages	Yes
	Example: 		
	ICI_OCodesSriLanka.xlsx	4-digit codes of industry and occupation classification	Yes

Level: Sri Lanka HIES2009 Questionnaire Record: Section - 1 - Demography

Item Label	Item Name	Start	Len	Data Type	Item Type	Occ	Dec	Char	Zero Fill
1Person	PERSON_SERIAL_NO	18	2	N	I	1	0	No	Yes
1:45									
(Blank)									
3Relationship	RELATIONSHIP	20	1	N	I	1	0	No	Yes
1 Head									
2 Spouse									
3 Son/Daughter									
4 Parents									
5 Other Ralative									
6 Domestic Servent									
7 Boarders									
9 Other									
(Blank)									
4Sex	SEX_LIVING	21	1	N	I	1	0	No	Yes
1 Male/Local									
2 Female/Abroad									
(Blank)									
5Birth Year	BIRTH_YEAR	22	2	N	I	1	0	No	Yes
0:99									
(Blank)									
5Birth Month	BIRTH_MONTH	24	2	N	I	1	0	No	Yes
1:12									
(Blank)									
6Age	AGE	26	2	N	I	1	0	No	Yes
0:99									
(Blank)									

Remarks:

Enumerator's manual is only given in Sinhala and Tamil, not in English.

☐ Changes from the previous HIES 2006/07

- The questionnaire of HIES 2009/10 is almost the same as HIES 2006/07.
  Changes are shown as below. Most of them are changes of response categories.

Section	Item	Changes from HIES 2006/07
Section 1	Col. 9 & 10	Contents of col. 9 and col. 10 were swapped.
Section 2	Target	"Aged 5-19" from "Aged 5-20"
	Col. 9 & 10	Number of response categories increased.
Section 5.5.2	Col. 6, 8 & 9	Contents were revised.
Section 8	5 Tenure	Response categories
	7A Availability of toilet	Response categories

Remarks: Change of the target age of education section
According to the delegates from Sri Lanka for the International Workshop,
"Since the school education tabulations are made on certain age bands, this has never been an issue."

- Geographical coverage
  The next three districts were covered by HIES 2009 in addition to 19 districts in HIES 2006.

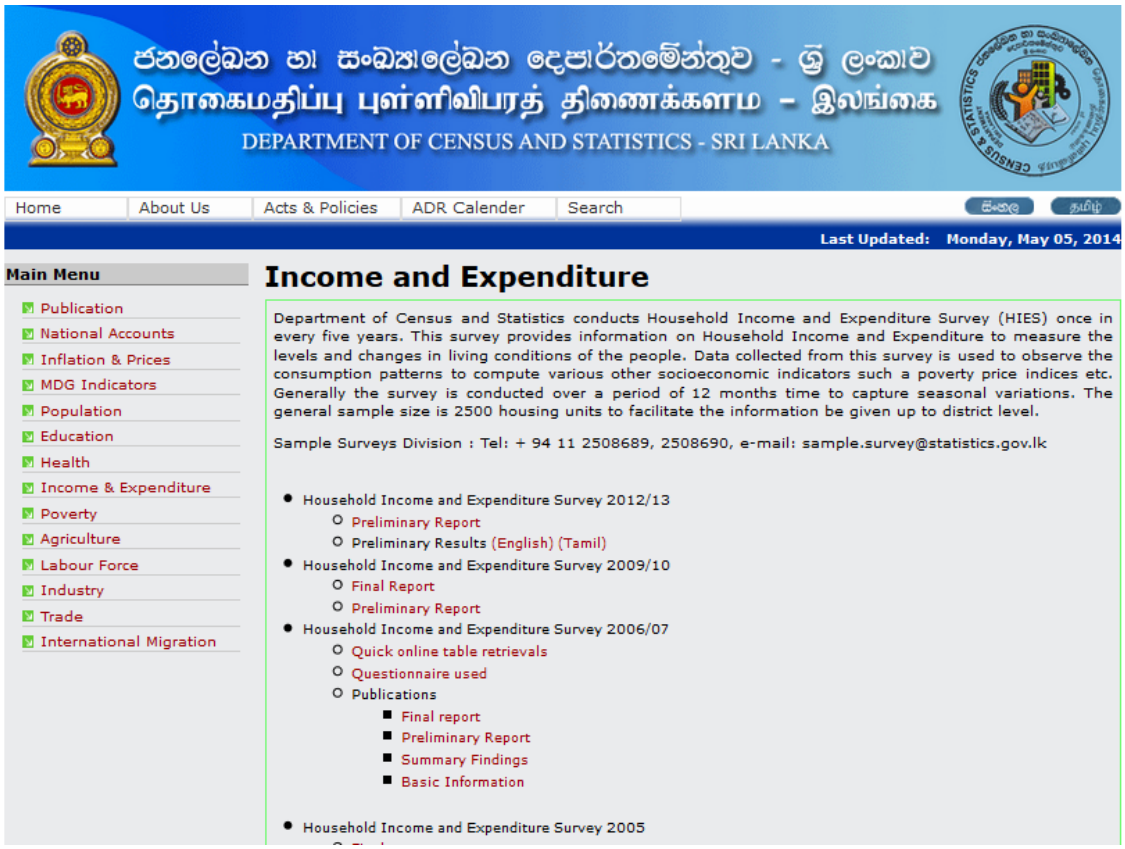
41	Jaffna
43	Vavuniya
53	Trincomalee

References

- **Homepage of DCS**

  http://www.statistics.gov.lk/page.asp?page=Income%20and%20Expenditure

  (Accessed on 9 July 2014)



It provides the preliminary report and the final report of HIES 2009/10.

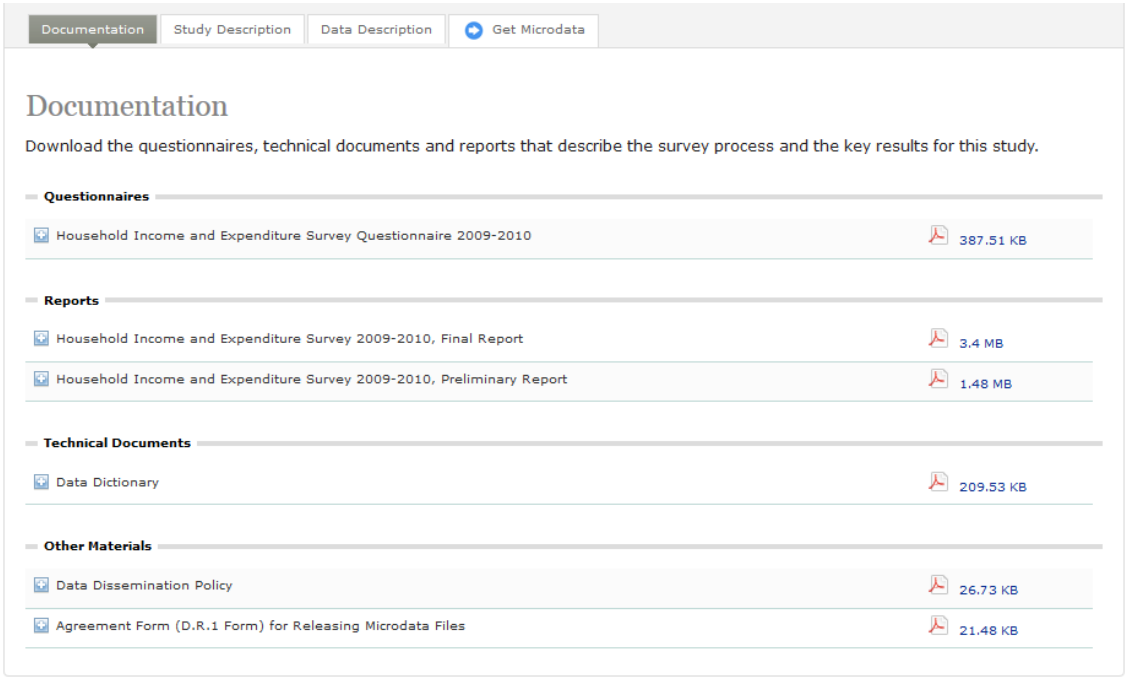
☐ IHSN Survey Catalog:

Sri Lanka - Household Income and Expenditure Survey 2009-2010

http://catalog.ihsn.org/index.php/catalog/3246

(Accessed on 9 July 2014)

The below materials are available on the web.



Note: The file of "Data Dictionary.pdf" was also mistaken. It was for HIES 2006/07.

4. Data Import

Strategy to import micro data into R

Two types of data set were provided. One was a text file with multi layouts. There were 24 types of records, as same as HIES 2006. The other was a set of 24 CSV files, which were already split and given variable names.
Weight data by district and psu was prepared separately.

1) To import CSV data files into R, in order to make use of the variable names prepared by NSO.

2) To combine the top nine variables of each data file and generate the variable ID, household identifier.

3) To generate the variable of PID (person id) as a combination of ID and person number for data files at individual level.

4) To import weight data into R and append weight to all data files.

****

Import CSV files to R

```
# CSV files to be imported
> list.files()
 [1] "HIES-2009-10-100%-Data-SEC_1_DEMOGRAPHIC.CSV"
 [2] "HIES-2009-10-100%-Data-SEC_2_SCHOOL_EDUCATION.CSV"
  :
[23] "HIES-2009-10-100%-Data-SEC_8_HOUSING.CSV"
[24] "HIES-2009-10-100%-Data-SEC_9_LAND_ANIMAL.CSV"

# Names of data frames in R
# using the sequential number in alphabetical order of CSV file names
> (Rnames<-paste("R",formatC(1:24,width=2,flag="0"),sep=""))
 [1] "R01" "R02" "R03" "R04" "R05" "R06" "R07" "R08" "R09" "R10" "R11" "R12"
```

```
[13] "R13" "R14" "R15" "R16" "R17" "R18" "R19" "R20" "R21" "R22" "R23" "R24"

# Imported 24 CSV files to R and stored R data frames in the list hies2009.
> for(j in 1:24){ hies2009<-c(hies2009,list(read.csv(list.files()[j]))) }
> dim(hies2009)

# List of the number of rows and columns of the created data frames
> file.names<-list.files()
> for(j in 1:24){
+ cat(Rnames[j],":",formatC(nrow(hies2009[[j]]),width=7),",",
+ formatC(ncol(hies2009[[j]]),width=3),":  ",
+ sub(".CSV","",sub("HIES-2009-10-100%-Data-","",file.names))[j],"¥n",sep="")
+ }
R01:  85443, 25:  SEC_1_DEMOGRAPHIC
R02:  20853, 20:  SEC_2_SCHOOL_EDUCATION
R03:  80866, 22:  SEC_3_HEALTH
R04: 798046, 14:  SEC_4_1_FOOD_EXP
R05: 568665, 14:  SEC_4_2_NONFOOD
R06:    240, 24:  SEC_4_3_BOARDERS
R07:  19958, 11:  SEC_4_3_IS_BOADERS
R08:  18369, 15:  SEC_5_1_EMP_INCOME
R09:  19958, 11:  SEC_5_1_IS_EMP_INCOME
R10:   4278, 18:  SEC_5_2_AGRI_INCOME
R11:  19958, 11:  SEC_5_2_IS_AGRI_INCOME
R12:  19958, 11:  SEC_5_3_IS_OTHER_AGRI_INCOME
R13:   4196, 17:  SEC_5_3_OTHER_AGRI_INCOME
R14:  19958, 11:  SEC_5_4_IS_NON_AGRI_INCOME
R15:   5528, 14:  SEC_5_4_NON_AGRI_INCOME
R16:  19958, 11:  SEC_5_5_1_IS_OTHER_INCOME
R17:  10735, 19:  SEC_5_5_1_OTHER_INCOME
R18:  19958, 11:  SEC_5_5_2_IS_WINDFALL_INCOME
R19:   8154, 18:  SEC_5_5_2_WINDFALL_INCOME
R20:  19958, 26:  SEC_6_B_DEBTNESS
R21:  19958, 33:  SEC_6A_DURABLE_GOODS
R22:  19958, 45:  SEC_7_BASIC_FACILITIES
R23:  19958, 37:  SEC_8_HOUSING
```

```
R24:  19958, 38:  SEC_9_LAND_ANIMAL

# Saved hies2009
> hies2009.old<-hies2009
```

ID: Household identifier

✓ Generated the variable of ID consisted of the next 9 items for all data files; DISTRICT(2), SECTOR(1), DSD(2), MONTH(2), PSU(3), SAMPLE_N(2), SERIAL_NO(1), NHH(1) and RESULT(1)

```
> for(j in 1:24){
+ d<-hies2009[[j]]
+ d["ID"]<-as.character(d$DISTRICT*10^13+d$SECTOR*10^12+d$DSD*10^10+d$MONTH*10^8+
+ d$PSU*10^5+d$SAMPLE_N*10^3+d$SERIAL_NO*10^2+d$NHH*10+d$RESULT)
+ hies2009[[j]]<-d
+ }
```

PID: Individual identifier

✓ Generated the variable of individual identifier PID consisted of ID and person number for R01, R02, R03, R06, R08, R10, R13, R15, R17 and R19.

```
> ind.files<-c(1,2,3,6,8,10,13,15,17,19)
> for(j in ind.files){
+ d<-hies2009[[j]]
+ d["PID"]<-paste(d$ID,formatC(d[,11],width=2,flag="0"),sep="")
+ hies2009[[j]]<-d
+ }
```

Import weight data by district and psu

✓ Opened the weight file "HIES-2009-10-100%-Factor File.xls", and saved as "wt2009.csv".

✓ Imported "wt2009.csv" to R as data frame "wt".

```
> wt<-read.csv("wt2009.csv")
> dim(wt)
```

```
[1] 2269    3
> colnames(wt)
[1] "District.code" "PSU.Number"    "Weight"
> colnames(wt)<-c("district","psu","WT")
> head(wt)
  district psu        WT
1       11   1  41.42092
2       11   2 224.96960
3       11   3 211.77130
4       11   4 131.90840
5       11   5 160.39880
6       11   6 475.23530
```

Append weight to all data files

```
> for(j in 1:24){
+ d<-hies2009[[j]]
+ d<-merge(d,wt,by.x=c("DISTRICT","PSU"),by.y=c("district","psu"))
+ hies2009[[j]]<-d
+ }


# Number of rows and columns of each data file
> for(j in 1:24){
+ cat(Rnames[j],":",formatC(nrow(hies2009[[j]]),width=7),",",
+ formatC(ncol(hies2009[[j]]),width=3),":  ",
+ sub(".CSV","",sub("HIES-2009-10-100%-Data-","",file.names))[j],"¥n",sep="")
+ }
R01:  85443, 28:  SEC_1_DEMOGRAPHIC
R02:  20853, 23:  SEC_2_SCHOOL_EDUCATION
R03:  80866, 25:  SEC_3_HEALTH
R04: 798046, 16:  SEC_4_1_FOOD_EXP
R05: 568665, 16:  SEC_4_2_NONFOOD
R06:    240, 27:  SEC_4_3_BOARDERS
R07:  19958, 13:  SEC_4_3_IS_BOADERS
R08:  18369, 18:  SEC_5_1_EMP_INCOME
R09:  19958, 13:  SEC_5_1_IS_EMP_INCOME
R10:   4278, 21:  SEC_5_2_AGRI_INCOME
R11:  19958, 13:  SEC_5_2_IS_AGRI_INCOME
R12:  19958, 13:  SEC_5_3_IS_OTHER_AGRI_INCOME
R13:   4196, 20:  SEC_5_3_OTHER_AGRI_INCOME
R14:  19958, 13:  SEC_5_4_IS_NON_AGRI_INCOME
R15:   5528, 17:  SEC_5_4_NON_AGRI_INCOME
R16:  19958, 13:  SEC_5_5_1_IS_OTHER_INCOME
R17:  10735, 22:  SEC_5_5_1_OTHER_INCOME
R18:  19958, 13:  SEC_5_5_2_IS_WINDFALL_INCOME
```

```
R19:   8154, 21:  SEC_5_5_2_WINDFALL_INCOME
R20:  19958, 28:  SEC_6_B_DEBTNESS
R21:  19958, 35:  SEC_6A_DURABLE_GOODS
R22:  19958, 47:  SEC_7_BASIC_FACILITIES
R23:  19958, 39:  SEC_8_HOUSING
R24:  19958, 40:  SEC_9_LAND_ANIMAL


# Example: R01
 > head(R01)
  DISTRICT PSU REC_TYPE SECTOR DSD MONTH SAMPLE_N SERIAL_NO NHH RESULT PERSON_SERIAL_NO
1       11   1        1      1   0     7        1         1   1      1                 1
2       11   1        1      1   0     7        1         1   1      1                 2
3       11   1        1      1   0     7        1         1   1      1                 3
4       11   1        1      1   0     7        1         1   1      1                 4
5       11   1        1      1   0     7        1         1   1      1                 5
6       11   1        1      1   0     7        1         1   1      1                 6
  RELATIONSHIP SEX_LIVING BIRTH_YEAR BIRTH_MONTH AGE ETHNICITY RELIGION CURR_EDUCATION
1            1          2         32           5  77         2        4              7
2            3          1         56          10  52         2        4              7
3            3          1         61          10  47         2        4              7
4            3          2         67          12  41         2        4              7
5            5          1          2          11   6         2        4              2
6            5          2          7           3   2         2        4             NA
  EDUCATION MARITAL_STATUS MAIN_ACTIVITY MAIN_OCCUPATION INDUSTRY EMPLOYMENT_STATUS
1        19              3             5              NA       NA                NA
2         7              2             5              NA       NA                NA
3         8              1             2              NA       NA                NA
4        10              5             1            7433     1810                 5
5         1             NA            NA              NA       NA                NA
6        NA             NA            NA              NA       NA                NA
                ID                PID       WT
1 111000700101111 11100070010111101 41.42092
2 111000700101111 11100070010111102 41.42092
3 111000700101111 11100070010111103 41.42092
4 111000700101111 11100070010111104 41.42092
5 111000700101111 11100070010111105 41.42092
6 111000700101111 11100070010111106 41.42092
```

```
# Saved hies2009
> hies2009.old1<-hies2009
```

Remarks: R01

Out of 85,443, R01 included 4,571 records with PERSON_NO over 40, which were not regarded as household members.

- ✓ **Omitted 4,571 records with PERSON_NO over 40.**
  After omitting such records, the number of records of R01 became 80,872. This is the final number of household members.

```
> R01<-hies2009[[1]]
> dim(R01)
[1] 85443    28
> addmargins(table(R01$PERSON_SERIAL_NO))
    1     2     3     4     5     6     7     8     9    10    11    12    13
19958 19056 16470 12321  7156  3354  1417   582   264   138    71    43    21
   14    15    16    17    41    42    43    44    45    46    47    48    49
   12     7     1     1  3588   746   182    37    11     4     1     1     1
  Sum
85443
> R01.old<-R01
> R01<-subset(R01,PERSON_SERIAL_NO<41)
> dim(R01)
[1] 80872    28
```

****

Remarks: Objectives of recording persons with person number over 40

According to the delegates from Sri Lanka for the International Workshop,

“Sri Lanka is a country with a large number of economic migrators and they bring the highest foreign income to the country. HIES records their migration status and relationship th the household and the questionnaire has a place to record the income the household received from such abroad resources or from other local remittances. This two information is highly helpful and necessary in making a consistent household in between demography and economy.”

☐ Characteristics of persons omitted from R01, for reference

Persons with PERSON_SERIAL_NO > 40 are not regarded as the household members in HIES because their usual residences are not in the household.

```
# Number of persons omitted from R01
> R01.omit<-subset(R01.old,PERSON_SERIAL_NO>40)
> dim(R01.omit)
[1] 4571   28
```

✓ For persons with person number > 40, the variables from BIRTH_YEAR to EMPLOYMENT_STATUS are empty as in the questionnaire, and the variable SEX_LIVING is read as USUAL_RESIDENCE.

```
> R01.omit<-R01.omit[-14:-25]
> colnames(R01.omit)[13]<-"USUAL_RESIDENCE"

# Number of persons omitted by relationship to the household head
> addmargins(table(R01.omit$RELATIONSHIP,useNA="ifany"))
   1    2    3    4    5    9 <NA>  Sum
   3 1346 2336   97  755   10   24 4571
```

Note: Three person are household heads. However, it is confirmed that there is a household head in the remaining household.

```
> R01.omit[!is.na(R01.omit$RELATIONSHIP)&R01.omit$RELATIONSHIP==1,"PID"]
[1] "91200080140111141" "91200080140111142" "91200080140111143"
> R01.old[R01.old$ID=="912000801401111",c("ID","PERSON_SERIAL_NO","RELATIONSHIP",
+ "SEX_LIVING","AGE","MARITAL_STATUS")]
                   ID PERSON_SERIAL_NO RELATIONSHIP SEX_LIVING AGE MARITAL_STATUS
78833 912000801401111                1            1          1  60              2
78834 912000801401111               41            1         NA  NA             NA
78835 912000801401111               42            1         NA  NA             NA
78836 912000801401111               43            1         NA  NA             NA

# Number of persons omitted by usual residence
> t<-addmargins(table(R01.omit$USUAL_RESIDENCE,useNA="ifany"))
```

```
> names(t)<-c("In the country","Abroad","<NA>","Sum")
> t
In the country         Abroad           <NA>            Sum
          3153           1413              5           4571
**************************************************************************
```

Remarks: R03 Health

The number of records of R03 is 80,866, which is 6 less than R01, and the variable of person number in R03 has 3 NA.

✓ Summary:

1) Omitted the records with person number=NA, because all data are NA.

2) 11 records with the next PID are missing as compared with R01.

File	Record	TO-DO
R03	person number = NA (3)	To drop
R03	As compared with PIDs of R01, there are 11 PIDs without a corresponding PID in R03.	The next are PIDs of missing records of R03; "22100040510211107", "51200110330611106" , "53200110270311101", "53200110270311102", "53200110270311103", "53200110270611102", "53200110270611103", "53200110270611105", "53200110270611106", "82200030470211101", "91200060900411103"

```
# Frequency of person number in R03
> R03<-hies2009[[3]]
> dim(R03)
[1] 80866    25
> addmargins(table(R03$R3_PERSON_SERIAL,useNA="ifany"))
    1     2     3     4     5     6     7     8     9    10    11    12    13
19957 19054 16467 12321  7155  3353  1416   582   264   138    71    43    21
   14    15    16    17  <NA>   Sum
   12     7     1     1     3 80866
> R03[is.na(R03$R3_PERSON_SERIAL),c(24,11:22)]
                    PID R3_PERSON_SERIAL DID_ATTEND_HOSPITAL REASON_HOSPITAL
```

```
55805 532001102703111NA              NA                NA             NA
55806 532001102703111NA              NA                NA             NA
55807 532001102703111NA              NA                NA             NA
      IS_STAY_HOSPITAL REASON_STAY IS_ILL_DISABLE WHAT_ILL_DISABLE IS_EMPL_REASON
55805               NA          NA             NA               NA             NA
55806               NA          NA             NA               NA             NA
55807               NA          NA             NA               NA             NA
      DURATION_YEARS DURATION_MONTHS IS_ABSENT_ACT DAYS_ABSENT
55805             NA              NA            NA          NA
55806             NA              NA            NA          NA
55807             NA              NA            NA          NA

# Omitted the above 3 records.
> R03.old<-R03
> R03<-R03[!is.na(R03$R3_PERSON_SERIAL),]
> dim(R03)
[1] 80863    25
```

✓ The difference of PID between R01 and R03 is as follows.

```
> PID.e<-setdiff(R01$PID,R03$PID)
> PID.e
 [1] "22100040510211107" "51200110330611106" "53200110270311101"
 [4] "53200110270311102" "53200110270311103" "53200110270611102"
 [7] "53200110270611103" "53200110270611105" "53200110270611106"
[10] "82200030470211101" "91200060900411103"
```

Un-weighted and weighted number of households and household members

	Un-weighted number	Weighted number
Number of households	19,958	5,079,362
Number of household members	80,872	20,337,761

```
> dim(R01)
```

```
[1] 80872    28
> sum(R01$WT)
[1] 20337761
> R23<-hies2009[[23]]
> dim(R23)
[1] 19958    39
> sum(R23$WT)
[1] 5079362

# Updated R01 and R03, and saved hies2009
> hies2009[[1]]<-R01
> hies2009[[3]]<-R03
> hies2009.old2<-hies2009
```

5 Data Check

5.1 Summary of each variable

```
SUMMARY OF EACH DATA FRAME
> for(j in 1:24){
+ cat("#### ",Rnames[j]," ##################################################¥n")
+ print(summary(hies2009[[j]]))
+ cat("¥n¥n")
+ }

####  R01  ##############################################
    DISTRICT         PSU            REC_TYPE     SECTOR          DSD
 Min.   :11.0   Min.   :  1.00   Min.   :1   Min.   :1.000   Min.   :0
 1st Qu.:13.0   1st Qu.: 28.00   1st Qu.:1   1st Qu.:1.000   1st Qu.:0
 Median :32.0   Median : 55.00   Median :1   Median :2.000   Median :0
 Mean   :40.4   Mean   : 68.77   Mean   :1   Mean   :1.813   Mean   :0
 3rd Qu.:61.0   3rd Qu.: 90.00   3rd Qu.:1   3rd Qu.:2.000   3rd Qu.:0
 Max.   :92.0   Max.   :300.00   Max.   :1   Max.   :3.000   Max.   :0

     MONTH          SAMPLE_N       SERIAL_NO         NHH           RESULT
 Min.   : 1.000   Min.   : 1.000   Min.   :1.000   Min.   :1.000   Min.   :1
 1st Qu.: 3.000   1st Qu.: 3.000   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1
 Median : 6.000   Median : 5.000   Median :1.000   Median :1.000   Median :1
 Mean   : 6.512   Mean   : 5.448   Mean   :1.007   Mean   :1.015   Mean   :1
 3rd Qu.:10.000   3rd Qu.: 8.000   3rd Qu.:1.000   3rd Qu.:1.000   3rd Qu.:1
 Max.   :12.000   Max.   :10.000   Max.   :4.000   Max.   :4.000   Max.   :1

 PERSON_SERIAL_NO  RELATIONSHIP     SEX_LIVING      BIRTH_YEAR
 Min.   : 1.00    Min.   :1.000   Min.   :1.000   Min.   : 1.00
 1st Qu.: 2.00    1st Qu.:2.000   1st Qu.:1.000   1st Qu.:46.00
 Median : 3.00    Median :3.000   Median :2.000   Median :67.00
 Mean   : 2.88    Mean   :2.661   Mean   :1.524   Mean   :60.56
 3rd Qu.: 4.00    3rd Qu.:3.000   3rd Qu.:2.000   3rd Qu.:84.00
 Max.   :17.00    Max.   :9.000   Max.   :2.000   Max.   :99.00
                                                  NA's   :1536
  BIRTH_MONTH          AGE          ETHNICITY        RELIGION
 Min.   : 1.000   Min.   : 1.00   Min.   :1.000   Min.   :1.000
 1st Qu.: 3.000   1st Qu.:15.00   1st Qu.:1.000   1st Qu.:1.000
 Median : 6.000   Median :30.00   Median :1.000   Median :1.000
 Mean   : 6.373   Mean   :32.50   Mean   :1.627   Mean   :1.685
 3rd Qu.: 9.000   3rd Qu.:48.25   3rd Qu.:2.000   3rd Qu.:2.000
 Max.   :12.000   Max.   :99.00   Max.   :9.000   Max.   :9.000
 NA's   :442      NA's   :1316    NA's   :11      NA's   :9
 CURR_EDUCATION    EDUCATION      MARITAL_STATUS  MAIN_ACTIVITY
 Min.   :1.000   Min.   : 1.000   Min.   :1.000   Min.   :1.000
 1st Qu.:2.000   1st Qu.: 6.000   1st Qu.:1.000   1st Qu.:1.000
 Median :7.000   Median :10.000   Median :2.000   Median :3.000
 Mean   :5.648   Mean   : 8.804   Mean   :1.773   Mean   :2.541
 3rd Qu.:7.000   3rd Qu.:11.000   3rd Qu.:2.000   3rd Qu.:4.000
 Max.   :9.000   Max.   :19.000   Max.   :5.000   Max.   :9.000
 NA's   :4209    NA's   :8692     NA's   :14178   NA's   :14197
 MAIN_OCCUPATION    INDUSTRY     EMPLOYMENT_STATUS      ID
```

```
Min.    : 110    Min.    : 111    Min.    :1.00      Length:80872
1st Qu.:4116    1st Qu.: 434    1st Qu.:3.00      Class :character
Median :6153    Median :4521    Median :3.00      Mode  :character
Mean   :6149    Mean   :3987    Mean   :3.44
3rd Qu.:8323    3rd Qu.:6519    3rd Qu.:5.00
Max.   :9411    Max.   :9900    Max.   :9.00
NA's   :53209   NA's   :53218   NA's   :53176
    PID                 WT
Length:80872       Min.    :   3.467
Class :character   1st Qu.:  84.095
Mode  :character   Median : 249.479
                   Mean   : 251.481
                   3rd Qu.: 377.676
                   Max.   :1555.993
```

Remarks: R01

1. The range of the variable AGE is from 1 to 99, and there are 1,316 NA. The reason might be that the age 0 is represented as NA (blank in CSV file) because the question states “Age as at last birthday”. Most of persons with AGE=NA born in the year 2008, 2009 or 2010.

```
> table(subset(R01,is.na(AGE))$BIRTH_YEAR,useNA="ifany")
   8    9   10   28   58   91 <NA>
 164  991  140    1    1    2   17
```

Note: According to the delegates from Sri Lanka for the International Workshop, “Blank represents ‘Not reported’ and 0 is used when a member is less than one year old.” However, blank code of age variable is found in the provided CSV and Text files.

2. The number of missing values of the variables taking into consideration the target age groups

Variable	Target respondent	Number of NA
CURR_EDUCATION	Aged 3 and over	20
EDUCATION	Aged 5 and over	1,538
MARITAL_STATUS	Aged 10 and over	8
MAIN_ACTIVITY	Aged 10 and over	26
MAIN_OCCUPATION	MAIN_ACTIVITY=1 (Employed)	34
INDUSTRY	MAIN_ACTIVITY=1 (Employed)	43
EMPLOYMENT_STATUS	MAIN_ACTIVITY=1 (Employed)	6

```
> nrow(subset(R01,AGE>=3&is.na(CURR_EDUCATION)))
[1] 20
> nrow(subset(R01,AGE>=5&is.na(EDUCATION)))
[1] 1538
> nrow(subset(R01,AGE>=10&is.na(MARITAL_STATUS)))
[1] 8
> nrow(subset(R01,AGE>=10&is.na(MAIN_ACTIVITY)))
[1] 26
> nrow(subset(R01,MAIN_ACTIVITY==1&is.na(MAIN_OCCUPATION)))
[1] 34
> nrow(subset(R01,MAIN_ACTIVITY==1&is.na(INDUSTRY)))
[1] 43
> nrow(subset(R01,MAIN_ACTIVITY==1&is.na(EMPLOYMENT_STATUS)))
[1] 6
```

```
####  R02  ##############################################
    DISTRICT         PSU            REC_TYPE     SECTOR          DSD
 Min.   :11.0   Min.   :  1.00   Min.   :2   Min.   :1.00   Min.   :0
 1st Qu.:13.0   1st Qu.: 27.00   1st Qu.:2   1st Qu.:1.00   1st Qu.:0
 Median :33.0   Median : 53.00   Median :2   Median :2.00   Median :0
 Mean   :41.7   Mean   : 65.95   Mean   :2   Mean   :1.82   Mean   :0
 3rd Qu.:61.0   3rd Qu.: 85.00   3rd Qu.:2   3rd Qu.:2.00   3rd Qu.:0
 Max.   :92.0   Max.   :300.00   Max.   :2   Max.   :3.00   Max.   :0

     MONTH           SAMPLE_N        SERIAL_NO          NHH            RESULT
 Min.   : 1.000   Min.   : 1.000   Min.   :1.000   Min.   :1.000   Min.   :1
 1st Qu.: 3.000   1st Qu.: 3.000   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1
 Median : 6.000   Median : 5.000   Median :1.000   Median :1.000   Median :1
 Mean   : 6.509   Mean   : 5.442   Mean   :1.008   Mean   :1.014   Mean   :1
 3rd Qu.:10.000   3rd Qu.: 8.000   3rd Qu.:1.000   3rd Qu.:1.000   3rd Qu.:1
 Max.   :12.000   Max.   :10.000   Max.   :4.000   Max.   :4.000   Max.   :1

 R2_PERSON_SERIAL R2_SCHOOL_EDUCATION GRADE_THIS_YEAR  GRADE_LAST_YEAR
 Min.   : 1.000   Min.   :1.000       Min.   : 1.000   Min.   : 1.000
 1st Qu.: 3.000   1st Qu.:1.000       1st Qu.: 4.000   1st Qu.: 3.000
 Median : 4.000   Median :1.000       Median : 7.000   Median : 6.000
 Mean   : 3.941   Mean   :1.307       Mean   : 6.689   Mean   : 7.013
 3rd Qu.: 5.000   3rd Qu.:1.000       3rd Qu.:10.000   3rd Qu.:10.000
 Max.   :15.000   Max.   :9.000       Max.   :19.000   Max.   :19.000
                  NA's   :6           NA's   :3574     NA's   :3636
    DISTANCE      TRANSPORT_MEDIUM TIME_TO_SCHOOL   NOSCHOOLING_REASON
 Min.   : 1.000   Min.   :1.000    Min.   :  1.00   Min.   :1.000
 1st Qu.: 1.000   1st Qu.:1.000    1st Qu.: 10.00   1st Qu.:7.000
 Median : 2.000   Median :2.000    Median : 15.00   Median :7.000
 Mean   : 4.593   Mean   :2.465    Mean   : 20.45   Mean   :6.611
 3rd Qu.: 5.000   3rd Qu.:4.000    3rd Qu.: 30.00   3rd Qu.:7.000
```

```
 Max.    :80.000   Max.    :9.000    Max.    :150.00   Max.    :9.000
 NA's    :9026     NA's    :3598     NA's    :3614     NA's    :20111
 REASON_NOT_GOING WHEN_STOP_SCHOOLING      ID                PID
 Min.    :1.000    Min.    :   1        Length:20853       Length:20853
 1st Qu.:3.000     1st Qu.:2006         Class :character   Class :character
 Median :6.000     Median :2007         Mode  :character   Mode  :character
 Mean    :5.537    Mean    :1997
 3rd Qu.:7.000     3rd Qu.:2008
 Max.    :9.000    Max.    :2078
 NA's    :18043    NA's    :18292
       WT
 Min.    :   3.467
 1st Qu.:  82.206
 Median : 241.497
 Mean    : 247.548
 3rd Qu.: 373.299
 Max.    :1555.993
```

Remarks: R02

The target of R02 is persons aged 5-19 years. The number of persons aged 5-19 of R02 who matched with the corresponding record in R01 is 20,847. The records with the following PID are missing.

```
> nrow(merge(R02,subset(R01,AGE>=5&AGE<=19),by="PID"))
[1] 20847

# Household members aged 5-19 in R01 who have no corresponding records in R02;
> setdiff(subset(R01,AGE>=5&AGE<=19)$PID,R02$PID)
 [1] "11200111110211104" "12200021280211104" "13200080260511103"
 [4] "13200090360411103" "13200090410811106" "13200090410811107"
 [7] "13200100540811103" "13200100540911103" "13300100580911103"
[10] "13200120760411103" "13200120770111103" "13200120770611104"
[13] "22100040510211107" "22200050541011105" "31200070070311105"
[16] "52200110390411103" "53200110270111102" "53200110270411103"
[19] "53200110270511103" "53100120290711107" "71200120430311103"
[22] "82200030470211102" "91200080160611102" "91300080170611103"
[25] "91200090200911103" "91200090201011104" "91300070040911105"
[28] "91200060900111104" "91200060900311103"
```

```
####  R03  ##############################################
   DISTRICT         PSU             REC_TYPE     SECTOR          DSD
 Min.   :11.0   Min.   :  1.00   Min.   :3   Min.   :1.000   Min.   :0
 1st Qu.:13.0   1st Qu.: 28.00   1st Qu.:3   1st Qu.:1.000   1st Qu.:0
 Median :32.0   Median : 55.00   Median :3   Median :2.000   Median :0
 Mean   :40.4   Mean   : 68.77   Mean   :3   Mean   :1.813   Mean   :0
 3rd Qu.:61.0   3rd Qu.: 90.00   3rd Qu.:3   3rd Qu.:2.000   3rd Qu.:0
 Max.   :92.0   Max.   :300.00   Max.   :3   Max.   :3.000   Max.   :0

     MONTH           SAMPLE_N        SERIAL_NO          NHH            RESULT
 Min.   : 1.000   Min.   : 1.000   Min.   :1.000   Min.   :1.000   Min.   :1
 1st Qu.: 3.000   1st Qu.: 3.000   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1
 Median : 6.000   Median : 5.000   Median :1.000   Median :1.000   Median :1
 Mean   : 6.512   Mean   : 5.448   Mean   :1.007   Mean   :1.015   Mean   :1
 3rd Qu.:10.000   3rd Qu.: 8.000   3rd Qu.:1.000   3rd Qu.:1.000   3rd Qu.:1
 Max.   :12.000   Max.   :10.000   Max.   :4.000   Max.   :4.000   Max.   :1

 R3_PERSON_SERIAL DID_ATTEND_HOSPITAL REASON_HOSPITAL IS_STAY_HOSPITAL
 Min.   : 1.000   Min.   :1.000       Min.   :1.00    Min.   :1.000
 1st Qu.: 2.000   1st Qu.:1.000       1st Qu.:1.00    1st Qu.:2.000
 Median : 3.000   Median :2.000       Median :1.00    Median :2.000
 Mean   : 2.879   Mean   :1.703       Mean   :1.24    Mean   :1.896
 3rd Qu.: 4.000   3rd Qu.:2.000       3rd Qu.:1.00    3rd Qu.:2.000
 Max.   :17.000   Max.   :4.000       Max.   :9.00    Max.   :9.000
                  NA's   :38          NA's   :56848   NA's   :51
  REASON_STAY    IS_ILL_DISABLE  WHAT_ILL_DISABLE IS_EMPL_REASON
 Min.   :1.00    Min.   :1.000   Min.   : 1.0     Min.   :1.00
 1st Qu.:1.00    1st Qu.:2.000   1st Qu.: 2.0     1st Qu.:2.00
 Median :1.00    Median :2.000   Median : 4.0     Median :2.00
 Mean   :2.11    Mean   :1.856   Mean   : 6.8     Mean   :1.93
 3rd Qu.:3.00    3rd Qu.:2.000   3rd Qu.:11.0     3rd Qu.:2.00
 Max.   :9.00    Max.   :2.000   Max.   :51.0     Max.   :4.00
 NA's   :72476   NA's   :57      NA's   :69233    NA's   :69309
 DURATION_YEARS  DURATION_MONTHS IS_ABSENT_ACT    DAYS_ABSENT
 Min.   : 1.00   Min.   : 1.00   Min.   :1.00    Min.   : 1.0
 1st Qu.: 3.00   1st Qu.: 2.00   1st Qu.:2.00    1st Qu.: 7.0
 Median : 5.00   Median : 5.00   Median :2.00    Median :25.0
 Mean   : 8.15   Mean   : 4.69   Mean   :1.81    Mean   :19.5
 3rd Qu.:10.00   3rd Qu.: 6.00   3rd Qu.:2.00    3rd Qu.:30.0
 Max.   :75.00   Max.   :26.00   Max.   :2.00    Max.   :90.0
 NA's   :70163   NA's   :76785   NA's   :69404   NA's   :78755
      ID                PID                  WT
 Length:80863       Length:80863       Min.   :   3.467
 Class :character   Class :character   1st Qu.:  84.095
 Mode  :character   Mode  :character   Median : 249.479
                                       Mean   : 251.483
                                       3rd Qu.: 377.676
                                       Max.   :1555.993


####  R04  ##############################################
   DISTRICT         PSU             REC_TYPE     SECTOR          DSD
 Min.   :11.00   Min.   :  1.00   Min.   :4   Min.   :1.000   Min.   :0
 1st Qu.:13.00   1st Qu.: 27.00   1st Qu.:4   1st Qu.:1.000   1st Qu.:0
 Median :32.00   Median : 55.00   Median :4   Median :2.000   Median :0
```

```
Mean   :40.12   Mean   : 69.01   Mean   :4   Mean   :1.803   Mean   :0
3rd Qu.:61.00   3rd Qu.: 90.00   3rd Qu.:4   3rd Qu.:2.000   3rd Qu.:0
Max.   :92.00   Max.   :300.00   Max.   :4   Max.   :3.000   Max.   :0

    MONTH          SAMPLE_N       SERIAL_NO         NHH           RESULT
Min.   : 1.000   Min.   : 1.000   Min.   :1.000   Min.   :1.000   Min.   :1
1st Qu.: 3.000   1st Qu.: 3.000   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1
Median : 7.000   Median : 5.000   Median :1.000   Median :1.000   Median :1
Mean   : 6.514   Mean   : 5.441   Mean   :1.008   Mean   :1.016   Mean   :1
3rd Qu.:10.000   3rd Qu.: 8.000   3rd Qu.:1.000   3rd Qu.:1.000   3rd Qu.:1
Max.   :12.000   Max.   :10.000   Max.   :4.000   Max.   :4.000   Max.   :1

     CODE          QUANTITY          VALUE          INKIND_VALUE
Min.   : 101.0   Min.   :    1.0   Min.   :    1.00   Min.   :   1.0
1st Qu.: 415.0   1st Qu.:   50.0   1st Qu.:   20.00   1st Qu.:  10.0
Median :1102.0   Median :  200.0   Median :   40.00   Median :  30.0
Mean   : 904.4   Mean   :  596.3   Mean   :   83.87   Mean   :  94.9
3rd Qu.:1120.0   3rd Qu.:  500.0   3rd Qu.:   90.00   3rd Qu.:  80.0
Max.   :1924.0   Max.   :50000.0   Max.   :30000.00   Max.   :6000.0
                 NA's   :145525                       NA's   :728119
     ID                WT
Length:798046      Min.   :   3.467
Class :character   1st Qu.:  87.500
Mode  :character   Median : 258.978
                   Mean   : 255.198
                   3rd Qu.: 381.289
                   Max.   :1555.993



####  R05  ##############################################
   DISTRICT         PSU            REC_TYPE     SECTOR         DSD
Min.   :11.00   Min.   :  1.00   Min.   :5   Min.   :1.0   Min.   :0
1st Qu.:13.00   1st Qu.: 26.00   1st Qu.:5   1st Qu.:1.0   1st Qu.:0
Median :32.00   Median : 54.00   Median :5   Median :2.0   Median :0
Mean   :40.27   Mean   : 67.32   Mean   :5   Mean   :1.8   Mean   :0
3rd Qu.:61.00   3rd Qu.: 89.00   3rd Qu.:5   3rd Qu.:2.0   3rd Qu.:0
Max.   :92.00   Max.   :300.00   Max.   :5   Max.   :3.0   Max.   :0

    MONTH          SAMPLE_N       SERIAL_NO         NHH           RESULT
Min.   : 1.000   Min.   : 1.000   Min.   :1.000   Min.   :1.000   Min.   :1
1st Qu.: 4.000   1st Qu.: 3.000   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1
Median : 7.000   Median : 5.000   Median :1.000   Median :1.000   Median :1
Mean   : 6.528   Mean   : 5.427   Mean   :1.007   Mean   :1.015   Mean   :1
3rd Qu.:10.000   3rd Qu.: 8.000   3rd Qu.:1.000   3rd Qu.:1.000   3rd Qu.:1
Max.   :12.000   Max.   :10.000   Max.   :4.000   Max.   :4.000   Max.   :1

   NF_CODE       NF_QUANTITY         NF_VALUE        NF_INKIND_VALUE
Min.   :2001   Min.   :    1.0   Min.   :      1   Min.   :      1
1st Qu.:2202   1st Qu.:    1.0   1st Qu.:    100   1st Qu.:    300
Median :2605   Median :    2.0   Median :    276   Median :    650
Mean   :2647   Mean   :  323.5   Mean   :   1736   Mean   :   2694
3rd Qu.:3013   3rd Qu.:    4.0   3rd Qu.:    800   3rd Qu.:   2000
Max.   :3509   Max.   :75000.0   Max.   :7000000   Max.   :4000000
               NA's   :354022    NA's   :1         NA's   :525297
     ID                WT
```

```
 Length:568665      Min.    :   3.467
 Class :character   1st Qu.:  84.460
 Mode  :character   Median : 253.296
                    Mean   : 252.226
                    3rd Qu.: 377.893
                    Max.   :1555.993



####  R06  ############################################
   DISTRICT          PSU           REC_TYPE     SECTOR          DSD
 Min.   :11.00   Min.   :  1.0   Min.   :7   Min.   :1.00   Min.   :0
 1st Qu.:11.00   1st Qu.: 29.0   1st Qu.:7   1st Qu.:1.00   1st Qu.:0
 Median :13.00   Median : 67.0   Median :7   Median :1.00   Median :0
 Mean   :28.56   Mean   : 83.8   Mean   :7   Mean   :1.45   Mean   :0
 3rd Qu.:33.00   3rd Qu.:130.2   3rd Qu.:7   3rd Qu.:2.00   3rd Qu.:0
 Max.   :92.00   Max.   :271.0   Max.   :7   Max.   :3.00   Max.   :0

    MONTH           SAMPLE_N        SERIAL_NO      NHH        RESULT
 Min.   : 1.000   Min.   : 1.000   Min.   :1   Min.   :1   Min.   :1
 1st Qu.: 3.000   1st Qu.: 2.000   1st Qu.:1   1st Qu.:1   1st Qu.:1
 Median : 7.000   Median : 5.000   Median :1   Median :1   Median :1
 Mean   : 6.287   Mean   : 5.021   Mean   :1   Mean   :1   Mean   :1
 3rd Qu.: 9.000   3rd Qu.: 7.000   3rd Qu.:1   3rd Qu.:1   3rd Qu.:1
 Max.   :12.000   Max.   :10.000   Max.   :1   Max.   :1   Max.   :1

    COL_2            COL_3            COL_4            COL_5
 Min.   : 2.000   Min.   : 100.0   Min.   :  50.0   Min.   :  240.0
 1st Qu.: 3.000   1st Qu.: 200.0   1st Qu.:  53.0   1st Qu.:  812.5
 Median : 4.000   Median : 500.0   Median : 150.0   Median : 1500.0
 Mean   : 4.642   Mean   : 774.7   Mean   : 330.8   Mean   : 2127.0
 3rd Qu.: 6.000   3rd Qu.:1000.0   3rd Qu.: 500.0   3rd Qu.: 3000.0
 Max.   :12.000   Max.   :4320.0   Max.   :1500.0   Max.   :14500.0
                  NA's   :167      NA's   :228      NA's   :134
    COL_6             COL_7           COL_8            COL_9
 Min.   :   50.0   Min.   : 200   Min.   :  38.0   Min.   :  50
 1st Qu.:  200.0   1st Qu.: 625   1st Qu.: 200.0   1st Qu.: 200
 Median :  350.0   Median :3500   Median : 300.0   Median : 500
 Mean   :  788.5   Mean   :3400   Mean   : 538.4   Mean   : 750
 3rd Qu.:  500.0   3rd Qu.:6000   3rd Qu.: 600.0   3rd Qu.:1050
 Max.   :14000.0   Max.   :7000   Max.   :4500.0   Max.   :3500
 NA's   :201       NA's   :232    NA's   :131      NA's   :133
    COL_10          COL_11          COL_12          COL_13          COL_14
 Min.   : 100   Min.   :  250   Min.   : 500   Min.   :  500   Min.   :   75
 1st Qu.: 200   1st Qu.: 1625   1st Qu.:1500   1st Qu.: 2000   1st Qu.: 1000
 Median : 720   Median : 4750   Median :4000   Median : 3000   Median : 2000
 Mean   :1209   Mean   : 8657   Mean   :3314   Mean   : 3635   Mean   : 3003
 3rd Qu.:1625   3rd Qu.: 9410   3rd Qu.:4500   3rd Qu.: 5000   3rd Qu.: 4450
 Max.   :5000   Max.   :55000   Max.   :8000   Max.   :12000   Max.   :12837
 NA's   :204    NA's   :221     NA's   :165    NA's   :126     NA's   :182
    COL_15            ID                PID                WT
 Min.   : 100.0   Length:240         Length:240         Min.   :   7.483
 1st Qu.: 200.0   Class :character   Class :character   1st Qu.:  74.156
 Median : 500.0   Mode  :character   Mode  :character   Median : 228.936
 Mean   : 918.6                                        Mean   : 242.918
 3rd Qu.:1000.0                                        3rd Qu.: 366.536
```

```
 Max.    :4500.0                                        Max.    :1488.380
 NA's    :208

####  R07  ##############################################
    DISTRICT          PSU             REC_TYPE     SECTOR          DSD
 Min.    :11.00   Min.    :  1.00   Min.    :6   Min.    :1.000   Min.    :0
 1st Qu.:13.00   1st Qu.: 28.00   1st Qu.:6   1st Qu.:1.000   1st Qu.:0
 Median :32.00   Median : 55.00   Median :6   Median :2.000   Median :0
 Mean    :40.82   Mean    : 68.22   Mean    :6   Mean    :1.823   Mean    :0
 3rd Qu.:61.00   3rd Qu.: 89.00   3rd Qu.:6   3rd Qu.:2.000   3rd Qu.:0
 Max.    :92.00   Max.    :300.00   Max.    :6   Max.    :3.000   Max.    :0

     MONTH           SAMPLE_N        SERIAL_NO          NHH            RESULT
 Min.    : 1.000   Min.    : 1.000   Min.    :1.000   Min.    :1.000   Min.    :1
 1st Qu.: 3.000   1st Qu.: 3.000   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1
 Median : 6.000   Median : 5.000   Median :1.000   Median :1.000   Median :1
 Mean    : 6.506   Mean    : 5.451   Mean    :1.008   Mean    :1.017   Mean    :1
 3rd Qu.:10.000   3rd Qu.: 8.000   3rd Qu.:1.000   3rd Qu.:1.000   3rd Qu.:1
 Max.    :12.000   Max.    :10.000   Max.    :4.000   Max.    :4.000   Max.    :1

 IS_BOARDERS_SERVENTS      ID                  WT
 Min.    :1.000        Length:19958       Min.    :   3.467
 1st Qu.:2.000        Class :character   1st Qu.:  85.574
 Median :2.000        Mode  :character   Median : 254.568
 Mean    :1.991                           Mean    : 254.503
 3rd Qu.:2.000                           3rd Qu.: 381.873
 Max.    :2.000                           Max.    :1555.993
 NA's    :4

####  R08  ##############################################
    DISTRICT          PSU             REC_TYPE     SECTOR          DSD
 Min.    :11.00   Min.    :  1.00   Min.    :9   Min.    :1.000   Min.    :0
 1st Qu.:13.00   1st Qu.: 29.00   1st Qu.:9   1st Qu.:1.000   1st Qu.:0
 Median :31.00   Median : 56.00   Median :9   Median :2.000   Median :0
 Mean    :38.96   Mean    : 71.11   Mean    :9   Mean    :1.887   Mean    :0
 3rd Qu.:61.00   3rd Qu.: 93.00   3rd Qu.:9   3rd Qu.:2.000   3rd Qu.:0
 Max.    :92.00   Max.    :300.00   Max.    :9   Max.    :3.000   Max.    :0

     MONTH           SAMPLE_N        SERIAL_NO          NHH            RESULT
 Min.    : 1.000   Min.    : 1.000   Min.    :1.000   Min.    :1.000   Min.    :1
 1st Qu.: 3.000   1st Qu.: 3.000   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1
 Median : 7.000   Median : 5.000   Median :1.000   Median :1.000   Median :1
 Mean    : 6.536   Mean    : 5.468   Mean    :1.008   Mean    :1.016   Mean    :1
 3rd Qu.:10.000   3rd Qu.: 8.000   3rd Qu.:1.000   3rd Qu.:1.000   3rd Qu.:1
 Max.    :12.000   Max.    :10.000   Max.    :4.000   Max.    :4.000   Max.    :1

 SERIAL_NO_SEC_1     PRI_SEC       WAGES_SALARIES     ALLOWENCES
 Min.    : 1.000   Min.    :1.000   Min.    :     1   Min.    :   30
 1st Qu.: 1.000   1st Qu.:1.000   1st Qu.:  6300   1st Qu.: 2000
 Median : 2.000   Median :1.000   Median : 10000   Median : 4000
 Mean    : 2.096   Mean    :1.011   Mean    : 12877   Mean    : 5021
 3rd Qu.: 3.000   3rd Qu.:1.000   3rd Qu.: 16000   3rd Qu.: 5500
 Max.    :14.000   Max.    :2.000   Max.    :600000   Max.    :95000
                  NA's    :9       NA's    :128      NA's    :15598
```

```
     BONUS             ID                PID                  WT
 Min.    :     6   Length:18369       Length:18369       Min.    :   3.467
 1st Qu.:  2500   Class :character   Class :character   1st Qu.:  77.523
 Median :  5000   Mode  :character   Mode  :character   Median : 236.602
 Mean   : 17393                                        Mean   : 244.958
 3rd Qu.: 15000                                        3rd Qu.: 373.210
 Max.   :600000                                        Max.   :1555.993
 NA's   :17179


####  R09  ##############################################
    DISTRICT          PSU            REC_TYPE     SECTOR           DSD
 Min.   :11.00   Min.   :  1.00   Min.   :8   Min.   :1.000   Min.   :0
 1st Qu.:13.00   1st Qu.: 28.00   1st Qu.:8   1st Qu.:1.000   1st Qu.:0
 Median :32.00   Median : 55.00   Median :8   Median :2.000   Median :0
 Mean   :40.82   Mean   : 68.22   Mean   :8   Mean   :1.823   Mean   :0
 3rd Qu.:61.00   3rd Qu.: 89.00   3rd Qu.:8   3rd Qu.:2.000   3rd Qu.:0
 Max.   :92.00   Max.   :300.00   Max.   :8   Max.   :3.000   Max.   :0

     MONTH           SAMPLE_N        SERIAL_NO          NHH            RESULT
 Min.   : 1.000   Min.   : 1.000   Min.   :1.000   Min.   :1.000   Min.   :1
 1st Qu.: 3.000   1st Qu.: 3.000   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1
 Median : 6.000   Median : 5.000   Median :1.000   Median :1.000   Median :1
 Mean   : 6.506   Mean   : 5.451   Mean   :1.008   Mean   :1.017   Mean   :1
 3rd Qu.:10.000   3rd Qu.: 8.000   3rd Qu.:1.000   3rd Qu.:1.000   3rd Qu.:1
 Max.   :12.000   Max.   :10.000   Max.   :4.000   Max.   :4.000   Max.   :1

 IS_EMPLOYMENT_INCOME      ID                 WT
 Min.   :1.000        Length:19958       Min.   :   3.467
 1st Qu.:1.000        Class :character   1st Qu.:  85.574
 Median :1.000        Mode  :character   Median : 254.568
 Mean   :1.362                           Mean   : 254.503
 3rd Qu.:2.000                           3rd Qu.: 381.873
 Max.   :2.000                           Max.   :1555.993
 NA's   :1


####  R10  ##############################################
    DISTRICT          PSU            REC_TYPE     SECTOR           DSD
 Min.   :11.00   Min.   :  1.00   Min.   :11   Min.   :1.000   Min.   :0
 1st Qu.:33.00   1st Qu.: 20.00   1st Qu.:11   1st Qu.:2.000   1st Qu.:0
 Median :61.00   Median : 41.00   Median :11   Median :2.000   Median :0
 Mean   :57.02   Mean   : 46.74   Mean   :11   Mean   :2.001   Mean   :0
 3rd Qu.:81.00   3rd Qu.: 65.00   3rd Qu.:11   3rd Qu.:2.000   3rd Qu.:0
 Max.   :92.00   Max.   :288.00   Max.   :11   Max.   :3.000   Max.   :0

     MONTH           SAMPLE_N       SERIAL_NO          NHH            RESULT
 Min.   : 1.000   Min.   : 1.00   Min.   :1.000   Min.   :1.000   Min.   :1
 1st Qu.: 4.000   1st Qu.: 3.00   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1
 Median : 7.000   Median : 5.00   Median :1.000   Median :1.000   Median :1
 Mean   : 6.668   Mean   : 5.46   Mean   :1.002   Mean   :1.005   Mean   :1
 3rd Qu.:10.000   3rd Qu.: 8.00   3rd Qu.:1.000   3rd Qu.:1.000   3rd Qu.:1
 Max.   :12.000   Max.   :10.00   Max.   :2.000   Max.   :2.000   Max.   :1

 SER_NO_SEC_5_2  SEAS_CROPS_CODE    ACR_5_2          RT_5_2
 Min.   :1.000   Min.   :1.000   Min.   : 1.000   Min.   :1.00
```

```
 1st Qu.:1.000   1st Qu.:1.000   1st Qu.: 1.000   1st Qu.:1.00
 Median :1.000   Median :1.000   Median : 1.000   Median :2.00
 Mean   :1.283   Mean   :2.211   Mean   : 2.057   Mean   :1.81
 3rd Qu.:1.000   3rd Qu.:4.000   3rd Qu.: 2.000   3rd Qu.:2.00
 Max.   :8.000   Max.   :9.000   Max.   :40.000   Max.   :8.00
                 NA's   :11      NA's   :2027     NA's   :2125
       P           OUTPUT_5_2      HH_CONSUMPTION     INPUT_5_2
 Min.   : 1.00   Min.   :     50   Min.   :    10   Min.   :    10
 1st Qu.:10.00   1st Qu.:  17000   1st Qu.:  3000   1st Qu.:  6000
 Median :20.00   Median :  40000   Median : 12000   Median : 14850
 Mean   :17.06   Mean   :  77518   Mean   : 15430   Mean   : 29865
 3rd Qu.:20.00   3rd Qu.:  92625   3rd Qu.: 22580   3rd Qu.: 34200
 Max.   :39.00   Max.   :1512000   Max.   :250000   Max.   :974000
 NA's   :3844    NA's   :134       NA's   :735      NA's   :361
      ID                PID                  WT
 Length:4278        Length:4278        Min.   :   3.732
 Class :character   Class :character   1st Qu.: 249.934
 Mode  :character   Mode  :character   Median : 322.222
                                       Mean   : 322.967
                                       3rd Qu.: 395.029
                                       Max.   :1555.993


####  R11  ############################################
    DISTRICT          PSU            REC_TYPE      SECTOR           DSD
 Min.   :11.00   Min.   :  1.00   Min.   :10   Min.   :1.000   Min.   :0
 1st Qu.:13.00   1st Qu.: 28.00   1st Qu.:10   1st Qu.:1.000   1st Qu.:0
 Median :32.00   Median : 55.00   Median :10   Median :2.000   Median :0
 Mean   :40.82   Mean   : 68.22   Mean   :10   Mean   :1.823   Mean   :0
 3rd Qu.:61.00   3rd Qu.: 89.00   3rd Qu.:10   3rd Qu.:2.000   3rd Qu.:0
 Max.   :92.00   Max.   :300.00   Max.   :10   Max.   :3.000   Max.   :0

     MONTH           SAMPLE_N        SERIAL_NO          NHH            RESULT
 Min.   : 1.000   Min.   : 1.000   Min.   :1.000   Min.   :1.000   Min.   :1
 1st Qu.: 3.000   1st Qu.: 3.000   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1
 Median : 6.000   Median : 5.000   Median :1.000   Median :1.000   Median :1
 Mean   : 6.506   Mean   : 5.451   Mean   :1.008   Mean   :1.017   Mean   :1
 3rd Qu.:10.000   3rd Qu.: 8.000   3rd Qu.:1.000   3rd Qu.:1.000   3rd Qu.:1
 Max.   :12.000   Max.   :10.000   Max.   :4.000   Max.   :4.000   Max.   :1

 IS_AGRICULTURAL_INCOME      ID                  WT
 Min.   :1.000          Length:19958       Min.   :   3.467
 1st Qu.:2.000          Class :character   1st Qu.:  85.574
 Median :2.000          Mode  :character   Median : 254.568
 Mean   :1.828                             Mean   : 254.503
 3rd Qu.:2.000                             3rd Qu.: 381.873
 Max.   :2.000                             Max.   :1555.993
 NA's   :4


####  R12  ############################################
    DISTRICT          PSU            REC_TYPE      SECTOR           DSD
 Min.   :11.00   Min.   :  1.00   Min.   :12   Min.   :1.000   Min.   :0
 1st Qu.:13.00   1st Qu.: 28.00   1st Qu.:12   1st Qu.:1.000   1st Qu.:0
 Median :32.00   Median : 55.00   Median :12   Median :2.000   Median :0
```

```
Mean   :40.82   Mean   : 68.22   Mean   :12   Mean   :1.823   Mean   :0
3rd Qu.:61.00   3rd Qu.: 89.00   3rd Qu.:12   3rd Qu.:2.000   3rd Qu.:0
Max.   :92.00   Max.   :300.00   Max.   :12   Max.   :3.000   Max.   :0

     MONTH           SAMPLE_N        SERIAL_NO          NHH            RESULT
Min.   : 1.000   Min.   : 1.000   Min.   :1.000   Min.   :1.000   Min.   :1
1st Qu.: 3.000   1st Qu.: 3.000   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1
Median : 6.000   Median : 5.000   Median :1.000   Median :1.000   Median :1
Mean   : 6.506   Mean   : 5.451   Mean   :1.008   Mean   :1.017   Mean   :1
3rd Qu.:10.000   3rd Qu.: 8.000   3rd Qu.:1.000   3rd Qu.:1.000   3rd Qu.:1
Max.   :12.000   Max.   :10.000   Max.   :4.000   Max.   :4.000   Max.   :1

IS_OTHER_AGRRI_INCOME      ID                  WT
Min.   :1.000         Length:19958       Min.   :   3.467
1st Qu.:2.000         Class :character   1st Qu.:  85.574
Median :2.000         Mode  :character   Median : 254.568
Mean   :1.827                            Mean   : 254.503
3rd Qu.:2.000                            3rd Qu.: 381.873
Max.   :2.000                            Max.   :1555.993
NA's   :3


####  R13  ##############################################
   DISTRICT          PSU            REC_TYPE      SECTOR           DSD
Min.   :11.00   Min.   :  1.00   Min.   :13   Min.   :1.000   Min.   :0
1st Qu.:22.00   1st Qu.: 24.00   1st Qu.:13   1st Qu.:2.000   1st Qu.:0
Median :33.00   Median : 50.00   Median :13   Median :2.000   Median :0
Mean   :48.68   Mean   : 57.46   Mean   :13   Mean   :1.956   Mean   :0
3rd Qu.:81.00   3rd Qu.: 79.00   3rd Qu.:13   3rd Qu.:2.000   3rd Qu.:0
Max.   :92.00   Max.   :288.00   Max.   :13   Max.   :3.000   Max.   :0

     MONTH           SAMPLE_N        SERIAL_NO          NHH            RESULT
Min.   : 1.000   Min.   : 1.000   Min.   :1.000   Min.   :1.000   Min.   :1
1st Qu.: 3.000   1st Qu.: 3.000   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1
Median : 7.000   Median : 5.000   Median :1.000   Median :1.000   Median :1
Mean   : 6.482   Mean   : 5.437   Mean   :1.002   Mean   :1.008   Mean   :1
3rd Qu.: 9.000   3rd Qu.: 8.000   3rd Qu.:1.000   3rd Qu.:1.000   3rd Qu.:1
Max.   :12.000   Max.   :10.000   Max.   :2.000   Max.   :2.000   Max.   :1

SER_NO_SEC_5_3   SEASONAL_CROP      ACRES_5_3        ROOTS_5_3
Min.   : 1.000   Min.   : 1.000   Min.   : 1.000   Min.   :1.000
1st Qu.: 1.000   1st Qu.: 2.000   1st Qu.: 1.000   1st Qu.:1.000
Median : 1.000   Median : 2.000   Median : 1.000   Median :2.000
Mean   : 1.316   Mean   : 3.898   Mean   : 2.347   Mean   :1.741
3rd Qu.: 1.000   3rd Qu.: 4.000   3rd Qu.: 2.000   3rd Qu.:2.000
Max.   :13.000   Max.   :19.000   Max.   :52.000   Max.   :5.000
                 NA's   :4        NA's   :3183     NA's   :2486
  PERCHS_5_3     OUTPUT_5_3        INPUT_5_3            ID
Min.   : 1.0   Min.   :     50   Min.   :     12   Length:4196
1st Qu.:10.0   1st Qu.:   1200   1st Qu.:    400   Class :character
Median :20.0   Median :   4500   Median :   1500   Mode  :character
Mean   :17.1   Mean   :  24679   Mean   :  12608
3rd Qu.:20.0   3rd Qu.:  15000   3rd Qu.:   5662
Max.   :38.0   Max.   :2250000   Max.   :1400000
NA's   :3637   NA's   :111       NA's   :1358
    PID                 WT
```

```
 Length:4196        Min.    :   3.502
 Class :character   1st Qu.: 239.957
 Mode  :character   Median : 336.859
                    Mean   : 314.812
                    3rd Qu.: 413.665
                    Max.   :1467.255


####  R14  ############################################
   DISTRICT          PSU            REC_TYPE      SECTOR           DSD
 Min.   :11.00   Min.   :  1.00   Min.   :14   Min.   :1.000   Min.   :0
 1st Qu.:13.00   1st Qu.: 28.00   1st Qu.:14   1st Qu.:1.000   1st Qu.:0
 Median :32.00   Median : 55.00   Median :14   Median :2.000   Median :0
 Mean   :40.82   Mean   : 68.22   Mean   :14   Mean   :1.823   Mean   :0
 3rd Qu.:61.00   3rd Qu.: 89.00   3rd Qu.:14   3rd Qu.:2.000   3rd Qu.:0
 Max.   :92.00   Max.   :300.00   Max.   :14   Max.   :3.000   Max.   :0

     MONTH           SAMPLE_N        SERIAL_NO          NHH            RESULT
 Min.   : 1.000   Min.   : 1.000   Min.   :1.000   Min.   :1.000   Min.   :1
 1st Qu.: 3.000   1st Qu.: 3.000   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1
 Median : 6.000   Median : 5.000   Median :1.000   Median :1.000   Median :1
 Mean   : 6.506   Mean   : 5.451   Mean   :1.008   Mean   :1.017   Mean   :1
 3rd Qu.:10.000   3rd Qu.: 8.000   3rd Qu.:1.000   3rd Qu.:1.000   3rd Qu.:1
 Max.   :12.000   Max.   :10.000   Max.   :4.000   Max.   :4.000   Max.   :1

 IS_NON_AGRI_INCOME      ID                 WT
 Min.   :1.000      Length:19958       Min.   :   3.467
 1st Qu.:2.000      Class :character   1st Qu.:  85.574
 Median :2.000      Mode  :character   Median : 254.568
 Mean   :1.755                         Mean   : 254.503
 3rd Qu.:2.000                         3rd Qu.: 381.873
 Max.   :2.000                         Max.   :1555.993
 NA's   :2


####  R15  ############################################
   DISTRICT          PSU            REC_TYPE      SECTOR           DSD
 Min.   :11.00   Min.   :  1.00   Min.   :15   Min.   :1.000   Min.   :0
 1st Qu.:12.00   1st Qu.: 29.00   1st Qu.:15   1st Qu.:1.000   1st Qu.:0
 Median :32.00   Median : 58.00   Median :15   Median :2.000   Median :0
 Mean   :38.27   Mean   : 73.15   Mean   :15   Mean   :1.709   Mean   :0
 3rd Qu.:61.00   3rd Qu.:100.00   3rd Qu.:15   3rd Qu.:2.000   3rd Qu.:0
 Max.   :92.00   Max.   :300.00   Max.   :15   Max.   :3.000   Max.   :0

     MONTH           SAMPLE_N        SERIAL_NO          NHH            RESULT
 Min.   : 1.000   Min.   : 1.000   Min.   :1.000   Min.   :1.000   Min.   :1
 1st Qu.: 3.000   1st Qu.: 3.000   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1
 Median : 6.000   Median : 5.000   Median :1.000   Median :1.000   Median :1
 Mean   : 6.484   Mean   : 5.393   Mean   :1.005   Mean   :1.014   Mean   :1
 3rd Qu.:10.000   3rd Qu.: 8.000   3rd Qu.:1.000   3rd Qu.:1.000   3rd Qu.:1
 Max.   :12.000   Max.   :10.000   Max.   :3.000   Max.   :4.000   Max.   :1

   SERIAL_5_4        NON_AGRI       OUTPUT_5_4          INPUT_5_4
 Min.   : 1.000   Min.   :1.000   Min.   :        9   Min.   :       25
 1st Qu.: 1.000   1st Qu.:4.000   1st Qu.:    10000   1st Qu.:     4600
```

```
 Median : 1.000   Median :4.000   Median :    22800   Median :    13500
 Mean   : 1.721   Mean   :4.317   Mean   :   148736   Mean   :   135603
 3rd Qu.: 2.000   3rd Qu.:5.000   3rd Qu.:    56000   3rd Qu.:    45000
 Max.   :13.000   Max.   :7.000   Max.   :180000000   Max.   :143680000
                  NA's   :6       NA's   :91          NA's   :792
     ID                PID                 WT
 Length:5528        Length:5528        Min.   :   3.467
 Class :character   Class :character   1st Qu.:  78.632
 Mode  :character   Mode  :character   Median : 244.886
                                       Mean   : 249.758
                                       3rd Qu.: 382.143
                                       Max.   :1555.993


####  R16  ############################################
   DISTRICT          PSU             REC_TYPE      SECTOR           DSD
 Min.   :11.00   Min.   :  1.00   Min.   :16   Min.   :1.000   Min.   :0
 1st Qu.:13.00   1st Qu.: 28.00   1st Qu.:16   1st Qu.:1.000   1st Qu.:0
 Median :32.00   Median : 55.00   Median :16   Median :2.000   Median :0
 Mean   :40.82   Mean   : 68.22   Mean   :16   Mean   :1.823   Mean   :0
 3rd Qu.:61.00   3rd Qu.: 89.00   3rd Qu.:16   3rd Qu.:2.000   3rd Qu.:0
 Max.   :92.00   Max.   :300.00   Max.   :16   Max.   :3.000   Max.   :0
    MONTH           SAMPLE_N        SERIAL_NO          NHH            RESULT
 Min.   : 1.000   Min.   : 1.000   Min.   :1.000   Min.   :1.000   Min.   :1
 1st Qu.: 3.000   1st Qu.: 3.000   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1
 Median : 6.000   Median : 5.000   Median :1.000   Median :1.000   Median :1
 Mean   : 6.506   Mean   : 5.451   Mean   :1.008   Mean   :1.017   Mean   :1
 3rd Qu.:10.000   3rd Qu.: 8.000   3rd Qu.:1.000   3rd Qu.:1.000   3rd Qu.:1
 Max.   :12.000   Max.   :10.000   Max.   :4.000   Max.   :4.000   Max.   :1
 IS_OTHER_INCOME      ID                 WT
 Min.   :1.000   Length:19958       Min.   :   3.467
 1st Qu.:1.000   Class :character   1st Qu.:  85.574
 Median :2.000   Mode  :character   Median : 254.568
 Mean   :1.508                      Mean   : 254.503
 3rd Qu.:2.000                      3rd Qu.: 381.873
 Max.   :2.000                      Max.   :1555.993


####  R17  ############################################
   DISTRICT          PSU             REC_TYPE      SECTOR          DSD
 Min.   :11.00   Min.   :  1.00   Min.   :17   Min.   :1.00   Min.   :0
 1st Qu.:21.00   1st Qu.: 26.00   1st Qu.:17   1st Qu.:1.00   1st Qu.:0
 Median :32.00   Median : 53.00   Median :17   Median :2.00   Median :0
 Mean   :40.86   Mean   : 64.33   Mean   :17   Mean   :1.78   Mean   :0
 3rd Qu.:61.00   3rd Qu.: 85.00   3rd Qu.:17   3rd Qu.:2.00   3rd Qu.:0
 Max.   :92.00   Max.   :300.00   Max.   :17   Max.   :3.00   Max.   :0

    MONTH           SAMPLE_N        SERIAL_NO          NHH            RESULT
 Min.   : 1.000   Min.   : 1.000   Min.   :1.000   Min.   :1.000   Min.   :1
 1st Qu.: 3.000   1st Qu.: 3.000   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1
 Median : 7.000   Median : 5.000   Median :1.000   Median :1.000   Median :1
 Mean   : 6.546   Mean   : 5.422   Mean   :1.007   Mean   :1.017   Mean   :1
 3rd Qu.:10.000   3rd Qu.: 8.000   3rd Qu.:1.000   3rd Qu.:1.000   3rd Qu.:1
 Max.   :12.000   Max.   :10.000   Max.   :3.000   Max.   :4.000   Max.   :1
```

```
  SERIAL_5_5_1        PENSION        DISABILITY_AND_RELIEF PROPERTY_RENTS
 Min.    : 1.000   Min.    :   100   Min.    :  100         Min.    :   150
 1st Qu.: 1.000   1st Qu.:  9031   1st Qu.:  100         1st Qu.:  3000
 Median : 1.000   Median : 12000   Median :  200         Median :  6000
 Mean    : 1.415   Mean    : 13471   Mean    : 1046         Mean    : 14048
 3rd Qu.: 1.000   3rd Qu.: 17000   3rd Qu.:  500         3rd Qu.: 13575
 Max.    :16.000   Max.    :150000   Max.    :35000         Max.    :300000
                  NA's    :9120     NA's    :10185         NA's    :9957
    SAMURDHI         DIVIDENDS          OTHER             ABROAD
 Min.    :   61.0   Min.    :     8   Min.    :    100   Min.    :    608
 1st Qu.:  340.0   1st Qu.:   750   1st Qu.:   2200   1st Qu.:  20000
 Median :  360.0   Median :  4500   Median :   5000   Median :  60000
 Mean    :  513.1   Mean    : 11127   Mean    :  12688   Mean    : 129946
 3rd Qu.:  600.0   3rd Qu.: 11750   3rd Qu.:  12000   3rd Qu.: 156000
 Max.    :14400.0   Max.    :200000   Max.    :2100000   Max.    :2400000
 NA's    :6138      NA's    :10437    NA's    :8798      NA's    :9235
     LOCAL              ID                PID                 WT
 Min.    :    200   Length:10735       Length:10735       Min.    :   3.467
 1st Qu.:   8000   Class :character   Class :character   1st Qu.:  85.321
 Median :  24000   Mode  :character   Mode  :character   Median : 252.779
 Mean    :  62412                                        Mean    : 252.606
 3rd Qu.:  72000                                        3rd Qu.: 378.331
 Max.    :2400000                                       Max.    :1555.993
 NA's    :9195


####  R18  ##############################################
    DISTRICT          PSU            REC_TYPE      SECTOR           DSD
 Min.    :11.00   Min.    :  1.00   Min.    :18   Min.    :1.000   Min.    :0
 1st Qu.:13.00   1st Qu.: 28.00   1st Qu.:18   1st Qu.:1.000   1st Qu.:0
 Median :32.00   Median : 55.00   Median :18   Median :2.000   Median :0
 Mean    :40.82   Mean    : 68.22   Mean    :18   Mean    :1.823   Mean    :0
 3rd Qu.:61.00   3rd Qu.: 89.00   3rd Qu.:18   3rd Qu.:2.000   3rd Qu.:0
 Max.    :92.00   Max.    :300.00   Max.    :18   Max.    :3.000   Max.    :0

     MONTH           SAMPLE_N        SERIAL_NO          NHH            RESULT
 Min.    : 1.000   Min.    : 1.000   Min.    :1.000   Min.    :1.000   Min.    :1
 1st Qu.: 3.000   1st Qu.: 3.000   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1
 Median : 6.000   Median : 5.000   Median :1.000   Median :1.000   Median :1
 Mean    : 6.506   Mean    : 5.451   Mean    :1.008   Mean    :1.017   Mean    :1
 3rd Qu.:10.000   3rd Qu.: 8.000   3rd Qu.:1.000   3rd Qu.:1.000   3rd Qu.:1
 Max.    :12.000   Max.    :10.000   Max.    :4.000   Max.    :4.000   Max.    :1

 IS_WINDFALL_INCOME       ID                 WT
 Min.    :1.000      Length:19958       Min.    :   3.467
 1st Qu.:1.000      Class :character   1st Qu.:  85.574
 Median :2.000      Mode  :character   Median : 254.568
 Mean    :1.619                         Mean    : 254.503
 3rd Qu.:2.000                         3rd Qu.: 381.873
 Max.    :2.000                         Max.    :1555.993
 NA's    :4


####  R19  ##############################################
    DISTRICT          PSU            REC_TYPE      SECTOR           DSD
 Min.    :11.00   Min.    :  1.0   Min.    :19   Min.    :1.000   Min.    :0
```

```
1st Qu.:13.00   1st Qu.: 25.0   1st Qu.:19   1st Qu.:1.000   1st Qu.:0
Median :32.00   Median : 50.0   Median :19   Median :2.000   Median :0
Mean   :41.19   Mean   : 63.7   Mean   :19   Mean   :1.856   Mean   :0
3rd Qu.:62.00   3rd Qu.: 84.0   3rd Qu.:19   3rd Qu.:2.000   3rd Qu.:0
Max.   :92.00   Max.   :299.0   Max.   :19   Max.   :3.000   Max.   :0

     MONTH           SAMPLE_N        SERIAL_NO          NHH            RESULT
Min.   : 1.000   Min.   : 1.000   Min.   :1.000   Min.   :1.000   Min.   :1
1st Qu.: 4.000   1st Qu.: 3.000   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1
Median : 7.000   Median : 5.000   Median :1.000   Median :1.000   Median :1
Mean   : 6.618   Mean   : 5.479   Mean   :1.008   Mean   :1.014   Mean   :1
3rd Qu.:10.000   3rd Qu.: 8.000   3rd Qu.:1.000   3rd Qu.:1.000   3rd Qu.:1
Max.   :12.000   Max.   :10.000   Max.   :3.000   Max.   :3.000   Max.   :1

 PERSON_5_5_2        LOANS         PAWNING_SELLING   DEPOSITS_PENSIONS_EPF
Min.   : 1.000   Min.   :    500   Min.   :    500   Min.   :    100
1st Qu.: 1.000   1st Qu.:  15000   1st Qu.:  12000   1st Qu.:  10000
Median : 1.000   Median :  40000   Median :  30000   Median :  25000
Mean   : 1.561   Mean   :  91659   Mean   :  59650   Mean   : 107267
3rd Qu.: 2.000   3rd Qu.: 100000   3rd Qu.:  60000   3rd Qu.: 100000
Max.   :41.000   Max.   :3200000   Max.   :7500000   Max.   :2400000
NA's   :1        NA's   :5704      NA's   :2341      NA's   :7662
   LOTTERY        SITTU_DEBTS      COMPENSATION     OTHER_WINDFALL
Min.   :   500   Min.   :   100   Min.   :    700   Min.   :    20
1st Qu.:  7000   1st Qu.: 10000   1st Qu.:   3500   1st Qu.:   475
Median : 12000   Median : 20000   Median :  15000   Median :  1000
Mean   : 24199   Mean   : 31181   Mean   :  90120   Mean   : 18255
3rd Qu.: 30000   3rd Qu.: 36000   3rd Qu.:  37400   3rd Qu.:  5000
Max.   :435000   Max.   :500000   Max.   :3696000   Max.   :300000
NA's   :7953     NA's   :6866     NA's   :8067      NA's   :7986
     ID                PID                 WT
Length:8154        Length:8154        Min.   :   3.467
Class :character   Class :character   1st Qu.:  98.296
Mode  :character   Mode  :character   Median : 257.853
                                      Mean   : 253.494
                                      3rd Qu.: 376.335
                                      Max.   :1555.993


####  R20  ##############################################
   DISTRICT          PSU            REC_TYPE      SECTOR           DSD
Min.   :11.00   Min.   :  1.00   Min.   :21   Min.   :1.000   Min.   :0
1st Qu.:13.00   1st Qu.: 28.00   1st Qu.:21   1st Qu.:1.000   1st Qu.:0
Median :32.00   Median : 55.00   Median :21   Median :2.000   Median :0
Mean   :40.82   Mean   : 68.22   Mean   :21   Mean   :1.823   Mean   :0
3rd Qu.:61.00   3rd Qu.: 89.00   3rd Qu.:21   3rd Qu.:2.000   3rd Qu.:0
Max.   :92.00   Max.   :300.00   Max.   :21   Max.   :3.000   Max.   :0

     MONTH           SAMPLE_N        SERIAL_NO          NHH            RESULT
Min.   : 1.000   Min.   : 1.000   Min.   :1.000   Min.   :1.000   Min.   :1
1st Qu.: 3.000   1st Qu.: 3.000   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1
Median : 6.000   Median : 5.000   Median :1.000   Median :1.000   Median :1
Mean   : 6.506   Mean   : 5.451   Mean   :1.008   Mean   :1.017   Mean   :1
3rd Qu.:10.000   3rd Qu.: 8.000   3rd Qu.:1.000   3rd Qu.:1.000   3rd Qu.:1
Max.   :12.000   Max.   :10.000   Max.   :4.000   Max.   :4.000   Max.   :1
```

```
     BANKS         BANK_AMOUNT          FINANCE      FINANCE_AMOUNT
 Min.    :1.000   Min.    :    300   Min.    :1.000   Min.    :      2
 1st Qu.:2.000   1st Qu.:  16000   1st Qu.:2.000   1st Qu.:  35000
 Median :2.000   Median :  40000   Median :2.000   Median : 105000
 Mean   :1.755   Mean   : 145711   Mean   :1.955   Mean   : 270153
 3rd Qu.:2.000   3rd Qu.: 100000   3rd Qu.:2.000   3rd Qu.: 300000
 Max.    :2.000   Max.    :9999999   Max.    :2.000   Max.    :4000000
 NA's    :3       NA's    :15066     NA's    :5       NA's    :19062
    EMPLOYER     EMPLOYER_AMOUNT       LENDER      LENDER_AMOUNT
 Min.    :1.000   Min.    :    150   Min.    :1.000   Min.    :    100
 1st Qu.:2.000   1st Qu.:  10000   1st Qu.:2.000   1st Qu.:   6638
 Median :2.000   Median :  41000   Median :2.000   Median :  20000
 Mean   :1.922   Mean   : 114723   Mean   :1.914   Mean   :  66158
 3rd Qu.:2.000   3rd Qu.: 113695   3rd Qu.:2.000   3rd Qu.:  50000
 Max.    :2.000   Max.    :4000000   Max.    :2.000   Max.    :3000000
 NA's    :4       NA's    :18409     NA's    :4       NA's    :18248
  RETAIL_SHOPS   RETAIL_SHOP_AMOUNT    PAWNING      PAWNING_AMOUNT
 Min.    :1.000   Min.    :     50     Min.    :1.000   Min.    :    160
 1st Qu.:2.000   1st Qu.:   1500     1st Qu.:1.000   1st Qu.:  15000
 Median :2.000   Median :   3000     Median :2.000   Median :  33800
 Mean   :1.837   Mean   :   4789     Mean   :1.663   Mean   :  68584
 3rd Qu.:2.000   3rd Qu.:   6000     3rd Qu.:2.000   3rd Qu.:  75000
 Max.    :2.000   Max.    :150000     Max.    :2.000   Max.    :3500000
 NA's    :5       NA's    :16698     NA's    :3       NA's    :13234
 INSTALMENT_GOODS INSTALEMENT_AMOUNT  OTHER_DEBTS     OTHER_AMOUNT
 Min.    :1.000    Min.    :    250    Min.    :1.000   Min.    :   250
 1st Qu.:2.000    1st Qu.:   4000    1st Qu.:2.000   1st Qu.:  6000
 Median :2.000    Median :  10000    Median :2.000   Median : 15000
 Mean   :1.957    Mean   :  21225    Mean   :1.973   Mean   : 38442
 3rd Qu.:2.000    3rd Qu.:  25000    3rd Qu.:2.000   3rd Qu.: 35500
 Max.    :2.000    Max.    :1000000    Max.    :2.000   Max.    :800000
 NA's    :4       NA's    :19096     NA's    :4       NA's    :19427
      ID                 WT
 Length:19958       Min.    :   3.467
 Class :character   1st Qu.:  85.574
 Mode  :character   Median : 254.568
                    Mean   : 254.503
                    3rd Qu.: 381.873
                    Max.    :1555.993


#### R21 ##############################################
    DISTRICT          PSU            REC_TYPE      SECTOR           DSD
 Min.    :11.00   Min.    :  1.00   Min.    :20   Min.    :1.000   Min.    :0
 1st Qu.:13.00   1st Qu.: 28.00   1st Qu.:20   1st Qu.:1.000   1st Qu.:0
 Median :32.00   Median : 55.00   Median :20   Median :2.000   Median :0
 Mean   :40.82   Mean   : 68.22   Mean   :20   Mean   :1.823   Mean   :0
 3rd Qu.:61.00   3rd Qu.: 89.00   3rd Qu.:20   3rd Qu.:2.000   3rd Qu.:0
 Max.    :92.00   Max.    :300.00   Max.    :20   Max.    :3.000   Max.    :0

     MONTH           SAMPLE_N        SERIAL_NO          NHH            RESULT
 Min.    : 1.000   Min.    : 1.000   Min.    :1.000   Min.    :1.000   Min.    :1
 1st Qu.: 3.000   1st Qu.: 3.000   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1
 Median : 6.000   Median : 5.000   Median :1.000   Median :1.000   Median :1
```

```
Mean   : 6.506   Mean   : 5.451   Mean   :1.008   Mean   :1.017   Mean   :1
3rd Qu.:10.000   3rd Qu.: 8.000   3rd Qu.:1.000   3rd Qu.:1.000   3rd Qu.:1
Max.   :12.000   Max.   :10.000   Max.   :4.000   Max.   :4.000   Max.   :1

    RADIO             TV             VCD         SEWING_MECHINE
Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000
1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1.000
Median :1.000   Median :1.000   Median :2.000   Median :2.000
Mean   :1.263   Mean   :1.212   Mean   :1.638   Mean   :1.592
3rd Qu.:2.000   3rd Qu.:1.000   3rd Qu.:2.000   3rd Qu.:2.000
Max.   :2.000   Max.   :2.000   Max.   :2.000   Max.   :2.000
NA's   :1       NA's   :2       NA's   :2       NA's   :1
WASHING_MECHINE     FRIDGE         COOKERT     ELECTRIC_FANS
Min.   :1.000   Min.   :1.000   Min.   :1.00   Min.   :1.000
1st Qu.:2.000   1st Qu.:1.000   1st Qu.:1.00   1st Qu.:1.000
Median :2.000   Median :2.000   Median :2.00   Median :1.000
Mean   :1.865   Mean   :1.604   Mean   :1.59   Mean   :1.463
3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2.00   3rd Qu.:2.000
Max.   :2.000   Max.   :2.000   Max.   :2.00   Max.   :2.000
NA's   :1       NA's   :1       NA's   :1      NA's   :1
  TELEPHONE     TELEPHONE_MOBILE   COMPUTERS       BICYCLE
Min.   :1.000   Min.   :1.000    Min.   :1.000   Min.   :1.000
1st Qu.:1.000   1st Qu.:1.000    1st Qu.:2.000   1st Qu.:1.000
Median :2.000   Median :1.000    Median :2.000   Median :2.000
Mean   :1.545   Mean   :1.396    Mean   :1.876   Mean   :1.586
3rd Qu.:2.000   3rd Qu.:2.000    3rd Qu.:2.000   3rd Qu.:2.000
Max.   :2.000   Max.   :2.000    Max.   :2.000   Max.   :2.000
NA's   :4       NA's   :1        NA's   :2       NA's   :1
MOTOR_BICYCLE   THREE_WHEELER   MOTOR_CAR_VAN     BUS_LORRY
Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000
1st Qu.:2.000   1st Qu.:2.000   1st Qu.:2.000   1st Qu.:2.000
Median :2.000   Median :2.000   Median :2.000   Median :2.000
Mean   :1.757   Mean   :1.941   Mean   :1.945   Mean   :1.985
3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2.000
Max.   :2.000   Max.   :2.000   Max.   :2.000   Max.   :2.000
NA's   :1       NA's   :1       NA's   :1       NA's   :1
TRACTOR_2_WHEEL TRACTOR_4_WHEEL   PESTICIDER     PADDY_BLOWER
Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000
1st Qu.:2.000   1st Qu.:2.000   1st Qu.:2.000   1st Qu.:2.000
Median :2.000   Median :2.000   Median :2.000   Median :2.000
Mean   :1.974   Mean   :1.994   Mean   :1.968   Mean   :1.997
3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2.000
Max.   :2.000   Max.   :2.000   Max.   :2.000   Max.   :2.000
NA's   :2       NA's   :2       NA's   :2       NA's   :2
 WATER_PUMPS        BOATS        FISHING_NETS       ID
Min.   :1.000   Min.   :1.000   Min.   :1.00   Length:19958
1st Qu.:2.000   1st Qu.:2.000   1st Qu.:2.00   Class :character
Median :2.000   Median :2.000   Median :2.00   Mode  :character
Mean   :1.979   Mean   :1.986   Mean   :1.98
3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2.00
Max.   :2.000   Max.   :2.000   Max.   :2.00
NA's   :2       NA's   :4       NA's   :4
      WT
Min.   :   3.467
1st Qu.:  85.574
Median : 254.568
```

```
 Mean   : 254.503
 3rd Qu.: 381.873
 Max.   :1555.993


####  R22  ############################################
   DISTRICT          PSU             REC_TYPE      SECTOR           DSD
 Min.   :11.00   Min.   :  1.00   Min.   :22   Min.   :1.000   Min.   :0
 1st Qu.:13.00   1st Qu.: 28.00   1st Qu.:22   1st Qu.:1.000   1st Qu.:0
 Median :32.00   Median : 55.00   Median :22   Median :2.000   Median :0
 Mean   :40.82   Mean   : 68.22   Mean   :22   Mean   :1.823   Mean   :0
 3rd Qu.:61.00   3rd Qu.: 89.00   3rd Qu.:22   3rd Qu.:2.000   3rd Qu.:0
 Max.   :92.00   Max.   :300.00   Max.   :22   Max.   :3.000   Max.   :0

     MONTH           SAMPLE_N        SERIAL_NO          NHH            RESULT
 Min.   : 1.000   Min.   : 1.000   Min.   :1.000   Min.   :1.000   Min.   :1
 1st Qu.: 3.000   1st Qu.: 3.000   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1
 Median : 6.000   Median : 5.000   Median :1.000   Median :1.000   Median :1
 Mean   : 6.506   Mean   : 5.451   Mean   :1.008   Mean   :1.017   Mean   :1
 3rd Qu.:10.000   3rd Qu.: 8.000   3rd Qu.:1.000   3rd Qu.:1.000   3rd Qu.:1
 Max.   :12.000   Max.   :10.000   Max.   :4.000   Max.   :4.000   Max.   :1

   BUS_HALT      BUS_HALT_TIME      PRE_SCHOOL     PRE_SCHOOL_TIME
 Min.   : 1.000   Min.   :  1.00   Min.   : 1.000   Min.   :  1.00
 1st Qu.: 1.000   1st Qu.:  5.00   1st Qu.: 1.000   1st Qu.:  5.00
 Median : 1.000   Median : 10.00   Median : 1.000   Median : 10.00
 Mean   : 1.764   Mean   : 10.81   Mean   : 1.613   Mean   : 11.48
 3rd Qu.: 2.000   3rd Qu.: 15.00   3rd Qu.: 2.000   3rd Qu.: 15.00
 Max.   :60.000   Max.   :150.00   Max.   :15.000   Max.   :150.00
 NA's   :14350    NA's   :52       NA's   :14063    NA's   :56
 PRIMERY_SCHOOL   PRIMERY_SCHOOL_TIME SECONDERY_SCHOOL SEC_SCHOOL_TIME
 Min.   : 1.000   Min.   :  1.00      Min.   : 1.000   Min.   :  1.00
 1st Qu.: 1.000   1st Qu.: 10.00      1st Qu.: 1.000   1st Qu.: 10.00
 Median : 1.000   Median : 10.00      Median : 2.000   Median : 15.00
 Mean   : 1.899   Mean   : 14.48      Mean   : 3.491   Mean   : 20.84
 3rd Qu.: 2.000   3rd Qu.: 20.00      3rd Qu.: 4.000   3rd Qu.: 30.00
 Max.   :20.000   Max.   :120.00      Max.   :45.000   Max.   :180.00
 NA's   :10456    NA's   :525         NA's   :6286     NA's   :38
   HOSPITAL      HOSPITAL_TIME    MATRENITY_HOME   MATERNITY_HOME_TIME
 Min.   : 1.000   Min.   :  1.00   Min.   : 1.000   Min.   :  1.0
 1st Qu.: 3.000   1st Qu.: 20.00   1st Qu.: 2.000   1st Qu.: 15.0
 Median : 5.000   Median : 30.00   Median : 4.000   Median : 25.0
 Mean   : 6.701   Mean   : 32.98   Mean   : 5.609   Mean   : 29.1
 3rd Qu.: 9.000   3rd Qu.: 40.00   3rd Qu.: 7.000   3rd Qu.: 40.0
 Max.   :63.000   Max.   :190.00   Max.   :80.000   Max.   :190.0
 NA's   :1722     NA's   :18       NA's   :3565     NA's   :915
 GOV_DISPENSARZ   GOV_DISPENSARY_TIME PRIVATE_DISPENSARY
 Min.   : 1.000   Min.   :  1.00      Min.   : 1.000
 1st Qu.: 2.000   1st Qu.: 15.00      1st Qu.: 1.000
 Median : 3.000   Median : 20.00      Median : 2.000
 Mean   : 4.317   Mean   : 25.36      Mean   : 3.799
 3rd Qu.: 5.000   3rd Qu.: 30.00      3rd Qu.: 5.000
 Max.   :45.000   Max.   :190.00      Max.   :42.000
 NA's   :4123     NA's   :858         NA's   :8204
 PRIVATE_DISPENSARY_TIME MATERNITY_CLINIC MATERNITY_CLINIC_TIME
```

```
 Min.    :  1.00          Min.    : 1.000   Min.    :  1.00
 1st Qu.: 10.00          1st Qu.: 1.000   1st Qu.: 10.00
 Median : 15.00          Median : 2.000   Median : 15.00
 Mean   : 19.07          Mean   : 3.476   Mean   : 20.88
 3rd Qu.: 25.00          3rd Qu.: 4.000   3rd Qu.: 30.00
 Max.   :190.00          Max.   :40.000   Max.   :240.00
 NA's   :199             NA's   :6083     NA's   :153
      DMO            DMO_TIME          MCUCPC        MCUCPC_TIME
 Min.   : 1.000   Min.   :  1.00   Min.   : 1.000   Min.   :  1.0
 1st Qu.: 2.000   1st Qu.: 20.00   1st Qu.: 3.000   1st Qu.: 18.0
 Median : 5.000   Median : 30.00   Median : 5.000   Median : 30.0
 Mean   : 6.649   Mean   : 32.51   Mean   : 7.362   Mean   : 33.9
 3rd Qu.: 9.000   3rd Qu.: 40.00   3rd Qu.:10.000   3rd Qu.: 45.0
 Max.   :50.000   Max.   :240.00   Max.   :90.000   Max.   :240.0
 NA's   :2093     NA's   :151      NA's   :1852     NA's   :11
   DS_OFFICE      DS_OFFICE_TIME     GN_OFFICE      GN_OFFICE_TIME
 Min.   : 1.000   Min.   :  1.00   Min.   : 1.000   Min.   :  1.00
 1st Qu.: 3.000   1st Qu.: 20.00   1st Qu.: 1.000   1st Qu.:  5.00
 Median : 5.000   Median : 30.00   Median : 1.000   Median : 10.00
 Mean   : 7.828   Mean   : 35.53   Mean   : 1.869   Mean   : 12.43
 3rd Qu.:10.000   3rd Qu.: 45.00   3rd Qu.: 2.000   3rd Qu.: 15.00
 Max.   :90.000   Max.   :300.00   Max.   :30.000   Max.   :150.00
 NA's   :1706     NA's   :14       NA's   :12911    NA's   :81
  POST_OFFICE     POST_OFFICE_TIME      BANK         BANK_TIME
 Min.   : 1.000   Min.   :  1.00   Min.   : 1.0   Min.   :  1.00
 1st Qu.: 1.000   1st Qu.: 10.00   1st Qu.: 2.0   1st Qu.: 10.00
 Median : 2.000   Median : 15.00   Median : 3.0   Median : 20.00
 Mean   : 2.502   Mean   : 17.14   Mean   : 4.6   Mean   : 23.56
 3rd Qu.: 3.000   3rd Qu.: 20.00   3rd Qu.: 6.0   3rd Qu.: 30.00
 Max.   :30.000   Max.   :180.00   Max.   :40.0   Max.   :200.00
 NA's   :7723     NA's   :33       NA's   :5070   NA's   :35
  AGRI_OFFICE     AGRI_OFFICE_TIME IS_POWER_LINES_NEAR IS_TEL_LINES_NEAR
 Min.   : 1.000   Min.   :  1.00   Min.   :1.000       Min.   :1.00
 1st Qu.: 3.000   1st Qu.: 20.00   1st Qu.:1.000       1st Qu.:1.00
 Median : 5.000   Median : 30.00   Median :1.000       Median :1.00
 Mean   : 7.241   Mean   : 35.78   Mean   :1.055       Mean   :1.19
 3rd Qu.:10.000   3rd Qu.: 45.00   3rd Qu.:1.000       3rd Qu.:1.00
 Max.   :52.000   Max.   :200.00   Max.   :2.000       Max.   :2.00
 NA's   :1310     NA's   :201
 IS_WATER_SERVICE_NEAR      ID                  WT
 Min.   :1.000         Length:19958       Min.   :   3.467
 1st Qu.:1.000         Class :character   1st Qu.:  85.574
 Median :1.000         Mode  :character   Median : 254.568
 Mean   :1.342                            Mean   : 254.503
 3rd Qu.:2.000                            3rd Qu.: 381.873
 Max.   :2.000                            Max.   :1555.993



####  R23  ##############################################
    DISTRICT          PSU            REC_TYPE      SECTOR          DSD
 Min.   :11.00   Min.   :  1.00   Min.   :23   Min.   :1.000   Min.   :0
 1st Qu.:13.00   1st Qu.: 28.00   1st Qu.:23   1st Qu.:1.000   1st Qu.:0
 Median :32.00   Median : 55.00   Median :23   Median :2.000   Median :0
 Mean   :40.82   Mean   : 68.22   Mean   :23   Mean   :1.823   Mean   :0
 3rd Qu.:61.00   3rd Qu.: 89.00   3rd Qu.:23   3rd Qu.:2.000   3rd Qu.:0
```

```
Max.    :92.00   Max.    :300.00   Max.    :23   Max.    :3.000   Max.    :0

     MONTH            SAMPLE_N         SERIAL_NO           NHH             RESULT
Min.    : 1.000   Min.    : 1.000   Min.    :1.000   Min.    :1.000   Min.    :1
1st Qu.: 3.000   1st Qu.: 3.000   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1
Median : 6.000   Median : 5.000   Median :1.000   Median :1.000   Median :1
Mean    : 6.506   Mean    : 5.451   Mean    :1.008   Mean    :1.017   Mean    :1
3rd Qu.:10.000   3rd Qu.: 8.000   3rd Qu.:1.000   3rd Qu.:1.000   3rd Qu.:1
Max.    :12.000   Max.    :10.000   Max.    :4.000   Max.    :4.000   Max.    :1

  STRUCTURE        BED_ROOMS          AREA            WALLS
Min.    :1.000   Min.    :1.00   Min.    :1.000   Min.    :1.000
1st Qu.:1.000   1st Qu.:2.00   1st Qu.:3.000   1st Qu.:1.000
Median :1.000   Median :2.00   Median :4.000   Median :1.000
Mean    :1.358   Mean    :2.32   Mean    :3.669   Mean    :2.073
3rd Qu.:1.000   3rd Qu.:3.00   3rd Qu.:5.000   3rd Qu.:3.000
Max.    :9.000   Max.    :9.00   Max.    :5.000   Max.    :9.000
NA's    :3       NA's    :480    NA's    :4       NA's    :4
     FLOOR            ROOF          OWNERSHIP     DRINKING_WATER
Min.    :1.000   Min.    :1.000   Min.    :1.000   Min.    :1.000
1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1.000
Median :1.000   Median :2.000   Median :1.000   Median :5.000
Mean    :1.438   Mean    :1.881   Mean    :2.176   Mean    :3.636
3rd Qu.:1.000   3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:5.000
Max.    :9.000   Max.    :9.000   Max.    :9.000   Max.    :9.000
NA's    :3       NA's    :4       NA's    :5       NA's    :6
  OWN_WATER     WATER_DISTANCE   WATER_SUFFICENCY OTHER_WATER_SUFFICENCY
Min.    :1.000   Min.    :   1.0   Min.    :1.00     Min.    :1.000
1st Qu.:1.000   1st Qu.:  25.0   1st Qu.:1.00     1st Qu.:1.000
Median :1.000   Median : 100.0   Median :1.00     Median :1.000
Mean    :1.232   Mean    : 185.4   Mean    :1.07     Mean    :1.106
3rd Qu.:1.000   3rd Qu.: 200.0   3rd Qu.:1.00     3rd Qu.:1.000
Max.    :2.000   Max.    :4500.0   Max.    :2.00     Max.    :2.000
NA's    :36      NA's    :15441   NA's    :26       NA's    :35
 TIOILET_USE     TOILET_TYPE    GARBAGE_DUMPING  LITE_SOURCE
Min.    :1.000   Min.    :1.000   Min.    :1.000   Min.    :1.000
1st Qu.:1.000   1st Qu.:1.000   1st Qu.:2.000   1st Qu.:2.000
Median :1.000   Median :1.000   Median :2.000   Median :2.000
Mean    :1.186   Mean    :1.272   Mean    :2.348   Mean    :1.889
3rd Qu.:1.000   3rd Qu.:1.000   3rd Qu.:3.000   3rd Qu.:2.000
Max.    :4.000   Max.    :9.000   Max.    :9.000   Max.    :9.000
NA's    :7       NA's    :514     NA's    :6       NA's    :3
 COOKING_FUEL   IS_COLLECT_FIREWOOD FIRE_WOOD_OWN   OTHER_DISTANCE
Min.    :1.000   Min.    :1.000       Min.    :1.000   Min.    :   1.0
1st Qu.:1.000   1st Qu.:1.000       1st Qu.:1.000   1st Qu.: 100.0
Median :1.000   Median :1.000       Median :1.000   Median : 200.0
Mean    :1.327   Mean    :1.364       Mean    :1.703   Mean    : 487.1
3rd Qu.:1.000   3rd Qu.:2.000       3rd Qu.:2.000   3rd Qu.: 500.0
Max.    :9.000   Max.    :2.000       Max.    :3.000   Max.    :9000.0
NA's    :9       NA's    :8           NA's    :7271    NA's    :14273
NATURAL_CALAMITY    FLOODING        DROUGHT       LAND_SLIDES
Min.    :1.000    Min.    :1.000   Min.    :1.00    Min.    :1.000
1st Qu.:2.000    1st Qu.:1.000   1st Qu.:1.00    1st Qu.:2.000
Median :2.000    Median :2.000   Median :2.00    Median :2.000
Mean    :1.937    Mean    :1.555   Mean    :1.74    Mean    :1.956
3rd Qu.:2.000    3rd Qu.:2.000   3rd Qu.:2.00    3rd Qu.:2.000
```

```
 Max.    :2.000     Max.    :2.000    Max.    :2.00     Max.    :2.000
 NA's    :13        NA's    :18696    NA's    :18706    NA's    :18706
  CIVIL_UNREST     WILD_ANIMAL     OTHER_CALAMITY        ID
 Min.    :1.000    Min.    :1.000    Min.    :1.000    Length:19958
 1st Qu.:2.000    1st Qu.:2.000    1st Qu.:2.000    Class :character
 Median :2.000    Median :2.000    Median :2.000    Mode  :character
 Mean    :1.992    Mean    :1.765    Mean    :1.881
 3rd Qu.:2.000    3rd Qu.:2.000    3rd Qu.:2.000
 Max.    :2.000    Max.    :2.000    Max.    :2.000
 NA's    :18706    NA's    :18706    NA's    :18706
       WT
 Min.    :   3.467
 1st Qu.:  85.574
 Median : 254.568
 Mean    : 254.503
 3rd Qu.: 381.873
 Max.    :1555.993



####  R24  ############################################
    DISTRICT          PSU             REC_TYPE      SECTOR          DSD
 Min.    :11.00    Min.    :  1.00    Min.    :24    Min.    :1.000    Min.    :0
 1st Qu.:13.00    1st Qu.: 28.00    1st Qu.:24    1st Qu.:1.000    1st Qu.:0
 Median :32.00    Median : 55.00    Median :24    Median :2.000    Median :0
 Mean    :40.82    Mean    : 68.22    Mean    :24    Mean    :1.823    Mean    :0
 3rd Qu.:61.00    3rd Qu.: 89.00    3rd Qu.:24    3rd Qu.:2.000    3rd Qu.:0
 Max.    :92.00    Max.    :300.00    Max.    :24    Max.    :3.000    Max.    :0

     MONTH           SAMPLE_N        SERIAL_NO          NHH            RESULT
 Min.    : 1.000    Min.    : 1.000    Min.    :1.000    Min.    :1.000    Min.    :1
 1st Qu.: 3.000    1st Qu.: 3.000    1st Qu.:1.000    1st Qu.:1.000    1st Qu.:1
 Median : 6.000    Median : 5.000    Median :1.000    Median :1.000    Median :1
 Mean    : 6.506    Mean    : 5.451    Mean    :1.008    Mean    :1.017    Mean    :1
 3rd Qu.:10.000    3rd Qu.: 8.000    3rd Qu.:1.000    3rd Qu.:1.000    3rd Qu.:1
 Max.    :12.000    Max.    :10.000    Max.    :4.000    Max.    :4.000    Max.    :1

 IS_AGRILAND_OWNER PADDY_OWN_ACR      PADDY_OWN_RT    PADDY_OWN_PERCH
 Min.    :1.000     Min.    :  1.000    Min.    :1.000    Min.    : 1
 1st Qu.:1.000     1st Qu.:  1.000    1st Qu.:1.000    1st Qu.:15
 Median :1.000     Median :  2.000    Median :2.000    Median :20
 Mean    :1.172     Mean    :  2.642    Mean    :1.837    Mean    :19
 3rd Qu.:1.000     3rd Qu.:  2.000    3rd Qu.:2.000    3rd Qu.:24
 Max.    :3.000     Max.    :802.000    Max.    :3.000    Max.    :38
 NA's    :1         NA's    :18443     NA's    :18520    NA's    :19750
 PADDY_OTHER_ACR PADDY_OTHER_RT  PADDY_OTHER_PERCH  LAND_OWN_ACR
 Min.    : 1.00   Min.    :1.000   Min.    : 1.00     Min.    :  1.000
 1st Qu.: 1.00   1st Qu.:1.000   1st Qu.:14.50     1st Qu.:  1.000
 Median : 2.00   Median :2.000   Median :20.00     Median :  1.000
 Mean    : 2.17   Mean    :1.849   Mean    :19.56     Mean    :  2.599
 3rd Qu.: 2.00   3rd Qu.:2.000   3rd Qu.:26.00     3rd Qu.:  2.000
 Max.    :80.00   Max.    :3.000   Max.    :35.00     Max.    :111.000
 NA's    :18853   NA's    :18894   NA's    :19826     NA's    :18848
  LAND_OWN_RT    LAND_OWN_PERCH  LAND_OTHER_ACR   LAND_OTHER_RT
 Min.    :1.000   Min.    : 1.00   Min.    : 1.000   Min.    :1.000
 1st Qu.:1.000   1st Qu.:10.00   1st Qu.: 1.000   1st Qu.:1.000
```

```
Median :2.000   Median :15.00   Median : 1.000   Median :2.000
Mean   :1.783   Mean   :16.16   Mean   : 2.445   Mean   :1.815
3rd Qu.:2.000   3rd Qu.:20.00   3rd Qu.: 2.000   3rd Qu.:2.000
Max.   :5.000   Max.   :40.00   Max.   :50.000   Max.   :3.000
NA's   :18858   NA's   :19130   NA's   :19347    NA's   :19374
LAND_OTHER_PERCH  HOME_OWN_ACR     HOME_OWN_RT    HOME_OWN_PERCH
Min.   : 1.00    Min.   : 1.000   Min.   :1.000   Min.   : 1.0
1st Qu.:10.00    1st Qu.: 1.000   1st Qu.:1.000   1st Qu.: 7.0
Median :18.00    Median : 1.000   Median :2.000   Median :11.0
Mean   :16.43    Mean   : 1.644   Mean   :1.628   Mean   :13.5
3rd Qu.:20.00    3rd Qu.: 2.000   3rd Qu.:2.000   3rd Qu.:20.0
Max.   :38.00    Max.   :25.000   Max.   :4.000   Max.   :50.0
NA's   :19768    NA's   :18011    NA's   :15510   NA's   :10037
HOME_OTHER_ACR   HOME_OTHER_RT   HOME_OTHER_PERCH COWS_BUFFALOWS
Min.   : 1.000   Min.   :1.000   Min.   : 1.0     Min.   :1.000
1st Qu.: 1.000   1st Qu.:1.000   1st Qu.:10.0     1st Qu.:2.000
Median : 1.000   Median :2.000   Median :15.0     Median :2.000
Mean   : 1.788   Mean   :1.692   Mean   :15.8     Mean   :1.954
3rd Qu.: 2.000   3rd Qu.:2.000   3rd Qu.:20.0     3rd Qu.:2.000
Max.   :35.000   Max.   :3.000   Max.   :50.0     Max.   :2.000
NA's   :19373    NA's   :18738   NA's   :18955    NA's   :9
  COWS_COUNT     GOATS_SHEEPS     GOAT_COUNT         PIGS
Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000
1st Qu.:1.000   1st Qu.:2.000   1st Qu.:1.000   1st Qu.:2.000
Median :1.000   Median :2.000   Median :1.000   Median :2.000
Mean   :1.331   Mean   :1.983   Mean   :1.409   Mean   :1.998
3rd Qu.:1.000   3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2.000
Max.   :8.000   Max.   :2.000   Max.   :8.000   Max.   :2.000
NA's   :19045   NA's   :9       NA's   :19621   NA's   :9
  PIGS_COUNT       CHICKENS     CHICKEN_COUNT   OTHER_ANIMALS
Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.   :1.000
1st Qu.:1.000   1st Qu.:2.000   1st Qu.:1.000   1st Qu.:2.000
Median :1.000   Median :2.000   Median :2.000   Median :2.000
Mean   :1.386   Mean   :1.921   Mean   :1.806   Mean   :1.996
3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2.000   3rd Qu.:2.000
Max.   :4.000   Max.   :2.000   Max.   :8.000   Max.   :2.000
NA's   :19914   NA's   :9       NA's   :18381   NA's   :8
     ID                 WT
Length:19958       Min.   :   3.467
Class :character   1st Qu.:  85.574
Mode  :character   Median : 254.568
                   Mean   : 254.503
                   3rd Qu.: 381.873
                   Max.   :1555.993
```

5.2 Frequency table of categorical variables

Data check of categorical variables

```
> # file.names: Rnames[j]
> # file.list: hies2009[[j]]
> # list of column numbers of categorical variables
> check.list<-list()
> check.list[[1]]<-c(11:13,17:22,25)
> check.list[[2]]<-c(11:14,16:19)
> check.list[[3]]<-c(11:18,21)
> check.list[[4]]<-c()
> check.list[[5]]<-c()
> check.list[[6]]<-c(11)
> check.list[[7]]<-c(11)
> check.list[[8]]<-c(11,12)
> check.list[[9]]<-c(11)
> check.list[[10]]<-c(11,12)
> check.list[[11]]<-c(11)
> check.list[[12]]<-c(11)
> check.list[[13]]<-c(11,12)
> check.list[[14]]<-c(11)
> check.list[[15]]<-c(11,12)
> check.list[[16]]<-c(11)
> check.list[[17]]<-c(11)
> check.list[[18]]<-c(11)
> check.list[[19]]<-c(11)
> check.list[[20]]<-c(seq(11,25,by=2))
> check.list[[21]]<-c(11:33)
> check.list[[22]]<-c(43:45)
> check.list[[23]]<-c(11:19,21:29,31:37)
> check.list[[24]]<-c(11,30,32,34,36,38)

> for(j in 1:24){
+ if(length(check.list[[j]])==0){ next }
+ cat(c("¥n¥n","#### FREQUENCY OF VARIABLES IN", Rnames[j],
+ "################################################"),"¥n¥n")
+ for(k in check.list[[j]]){
+ variable.name<-colnames(hies2009[[j]])[k]
+ cat(c("-----",variable.name,"-----------------------------------------------------"))
+ print(table(hies2009[[j]][k],useNA="ifany"))
+ }}



 #### FREQUENCY OF VARIABLES IN R01 ################################################

----- PERSON_SERIAL_NO -----------------------------------------------------
    1     2     3     4     5     6     7     8     9    10    11    12    13
19958 19056 16470 12321  7156  3354  1417   582   264   138    71    43    21
   14    15    16    17
   12     7     1     1
----- RELATIONSHIP ---------------------------------------------------------
    1     2     3     4     5     6     7     9
19959 14223 32929  1555 11849   148    95   114
```

```
----- SEX_LIVING ----------------------------------------------------
    1     2
38485 42387
----- ETHNICITY -----------------------------------------------------
    1     2     3     4     5     6     9  <NA>
55037 11781  3953  9713   231    83    63    11
----- RELIGION ------------------------------------------------------
    1     2     3     4     9  <NA>
50341 13146  9946  7424     6     9
----- CURR_EDUCATION ------------------------------------------------
    1     2     3     4     5     6     7     9  <NA>
 2102 17342   290   681   433   250 55564     1  4209
----- EDUCATION -----------------------------------------------------
    1     2     3     4     5     6     7     8     9    10    11    12    13
 2144  3117  3192  4019  5169  3614  4093  5553  4061 14967  7029  3766  6571
   14    15    16    17    19  <NA>
  184  1051   279    67  3304  8692
----- MARITAL_STATUS ------------------------------------------------
    1     2     3     4     5  <NA>
22784 37794  5275   157   684 14178
----- MAIN_ACTIVITY -------------------------------------------------
    1     2     3     4     5     6     9  <NA>
27693  5157 12355 14627  6433     1   409 14197
----- EMPLOYMENT_STATUS ---------------------------------------------
    1     2     3     4     5     6     9  <NA>
 3414  1191 13437   542  7562  1549     1 53176
```

Remarks: R01
The number of RELATIONSHIP=1 is 19,959, one larger than the number of households.
As for PID= “91200080140511102”, RELATIONSHIP should be read as 2 and SEX_LIVING should be read as 2.

```
> table(subset(R01,RELATIONSHIP==1)$PERSON_SERIAL_NO)
    1     2
19958     1
> R01[R01$RELATIONSHIP==1&R01$PERSON_SERIAL_NO==2,"ID"]
[1] "912000801405111"
> R01[R01$ID=="912000801405111",c(11:13,16,21,22,26,27)]
      PERSON_SERIAL_NO RELATIONSHIP SEX_LIVING AGE MARITAL_STATUS
78847                1            1          1  33              2
78848                2            1          1  28              2
78849                3            3          1  11              1
78850                4            3          2   5             NA
      MAIN_ACTIVITY             ID               PID
78847             1 912000801405111 91200080140511101
78848             4 912000801405111 91200080140511102
78849             3 912000801405111 91200080140511103
78850            NA 912000801405111 91200080140511104
```

✓ **Revised the record with PID="91200080140511102"**

```
> R01[R01$PID=="91200080140511102","RELATIONSHIP"]<-2
> R01[R01$PID=="91200080140511102","SEX_LIVING"]<-2
> R01[R01$PID=="91200080140511102",]
      DISTRICT PSU REC_TYPE SECTOR DSD MONTH SAMPLE_N SERIAL_NO NHH RESULT
```

```
78848       91  14        1      2   0     8        5         1   1      1
      PERSON_SERIAL_NO RELATIONSHIP SEX_LIVING BIRTH_YEAR BIRTH_MONTH AGE
78848                2            2          2         80          11  28
      ETHNICITY RELIGION CURR_EDUCATION EDUCATION MARITAL_STATUS MAIN_ACTIVITY
78848         1        1              7        10              2             4
      MAIN_OCCUPATION INDUSTRY EMPLOYMENT_STATUS             ID               PID
78848              NA       NA                NA 912000801405111 91200080140511102
           WT
78848 440.1271

# Updated R01 in hies2009
> hies2009[[1]]<-R01



 #### FREQUENCY OF VARIABLES IN R02 ##################################################

----- R2_PERSON_SERIAL ------------------------------------------------------
   1    2    3    4    5    6    7    8    9   10   11   12   13   14   15
  30 1451 7590 6279 3305 1307  496  207   98   47   19   10    6    6    2
----- R2_SCHOOL_EDUCATION ---------------------------------------------------
    1     2     3     6     9  <NA>
17277   742  2826     1     1     6
----- GRADE_THIS_YEAR -------------------------------------------------------
   1    2    3    4    5    6    7    8    9   10   11   12   13   14   19
1276 1509 1449 1465 1413 1422 1359 1371 1316 1398 1537  930  755   43   36
<NA>
3574
----- GRADE_LAST_YEAR -------------------------------------------------------
   1    2    3    4    5    6    7    8    9   10   11   12   13   14   19
1553 1454 1464 1411 1416 1369 1357 1324 1409 1507  955  732   22   21 1223
<NA>
3636
----- TRANSPORT_MEDIUM ------------------------------------------------------
   1    2    3    4    5    6    9 <NA>
8445 1696 1553 2003 3448   67   43 3598
----- TIME_TO_SCHOOL --------------------------------------------------------
   1    2    3    4    5    6    7    8    9   10   12   13   14   15   16
 106  112  162   83 1488   70   87  155   10 3830  124   14    5 3098    1
  17   18   20   21   22   23   25   26   27   28   29   30   33   35   36
  12   17 2663    2    7    4  612    4    1    8    2 2329    1  256    3
  39   40   44   45   50   55   60   65   70   75   80   85   90   95  100
   2  478    2  633   75   16  488    9   27   48   30    2  112    2    8
 105  110  120  150 <NA>
   3    2   35    1 3614
----- NOSCHOOLING_REASON ----------------------------------------------------
    1     2     3     4     6     7     9  <NA>
    8    32     5    57    10   570    60 20111
----- REASON_NOT_GOING ------------------------------------------------------
    1     2     3     4     5     6     7     8     9  <NA>
   11   482   311    55    25  1065   327   180   354 18043
```

```
 #### FREQUENCY OF VARIABLES IN RO3 ################################################

----- R3_PERSON_SERIAL ------------------------------------------------------
    1     2     3     4     5     6     7     8     9    10    11    12    13
19957 19054 16467 12321  7155  3353  1416   582   264   138    71    43    21
   14    15    16    17
   12     7     1     1
----- DID_ATTEND_HOSPITAL ---------------------------------------------------
    1     2     4  <NA>
24039 56785     1    38
----- REASON_HOSPITAL -------------------------------------------------------
    1     2     3     4     5     9  <NA>
21897   655   924   127   123   289 56848
----- IS_STAY_HOSPITAL ------------------------------------------------------
    1     2     3     9  <NA>
 8397 72413     1     1    51
----- REASON_STAY -----------------------------------------------------------
    1     2     3     4     5     6     7     9  <NA>
 5354   592   873   861   130   240     1   336 72476
----- IS_ILL_DISABLE --------------------------------------------------------
    1     2  <NA>
11644 69162    57
----- WHAT_ILL_DISABLE ------------------------------------------------------
    1     2     3     4     5     6     7     8     9    10    11    12    13
 1041  2500  1899  1389   237   142   295   255   631   332   150   486   174
   14    15    19    20    51  <NA>
  372   296  1429     1     1 69233
----- IS_EMPL_REASON --------------------------------------------------------
    1     2     4  <NA>
  852 10701     1 69309
----- IS_ABSENT_ACT ---------------------------------------------------------
    1     2  <NA>
 2136  9323 69404


 #### FREQUENCY OF VARIABLES IN RO6 ################################################

----- COL_2 -----------------------------------------------------------------
 2  3  4  5  6  7  8  9 10 12
23 40 58 45 45 21  3  2  2  1


 #### FREQUENCY OF VARIABLES IN RO7 ################################################

----- IS_BOARDERS_SERVENTS --------------------------------------------------
    1     2  <NA>
  187 19767     4
```

```
 #### FREQUENCY OF VARIABLES IN R08 ################################################

----- SERIAL_NO_SEC_1 ----------------------------------------------------
   1    2    3    4    5    6    7    8    9   10   11   12   13   14
8674 4050 3062 1446  559  310  150   61   31   10    8    3    4    1
----- PRI_SEC ------------------------------------------------------------
    1     2  <NA>
18155   205     9


 #### FREQUENCY OF VARIABLES IN R09 ################################################

----- IS_EMPLOYMENT_INCOME -----------------------------------------------
    1     2  <NA>
12739  7218     1


 #### FREQUENCY OF VARIABLES IN R10 ################################################

----- SER_NO_SEC_5_2 -----------------------------------------------------
   1    2    3    4    5    6    7    8
3480  515  198   53   23    5    3    1
----- SEAS_CROPS_CODE ----------------------------------------------------
   1    2    3    4    5    6    7    9 <NA>
2782  122   76  695  410   81   23   78   11


 #### FREQUENCY OF VARIABLES IN R11 ################################################

----- IS_AGRICULTURAL_INCOME ---------------------------------------------
    1     2  <NA>
 3425 16529     4


 #### FREQUENCY OF VARIABLES IN R12 ################################################

----- IS_OTHER_AGRRI_INCOME ----------------------------------------------
    1     2  <NA>
 3449 16506     3


 #### FREQUENCY OF VARIABLES IN R13 ################################################

----- SER_NO_SEC_5_3 -----------------------------------------------------
   1    2    3    4    5    6    7    8   10   13
3343  581  164   54   35   11    4    2    1    1
----- SEASONAL_CROP ------------------------------------------------------
   1    2    3    4    5    6    7    8    9   10   19 <NA>
 989 1355  424  384   68  329  141  216   53   11  222    4
```

```
 #### FREQUENCY OF VARIABLES IN R14 ##################################################

----- IS_NON_AGRI_INCOME -------------------------------------------------------
    1     2  <NA>
 4889 15067     2


 #### FREQUENCY OF VARIABLES IN R15 ##################################################

----- SERIAL_5_4 ---------------------------------------------------------
   1    2    3    4    5    6    7    8    9   10   11   12   13
3375 1129  578  251  103   54   21    7    5    2    1    1    1
----- NON_AGRI -----------------------------------------------------------
   1    2    3    4    5    6    7 <NA>
 126  980  152 2282  797  114 1071    6


 #### FREQUENCY OF VARIABLES IN R16 ##################################################

----- IS_OTHER_INCOME ----------------------------------------------------------
    1     2
 9826 10132


 #### FREQUENCY OF VARIABLES IN R17 ##################################################

----- SERIAL_5_5_1 -------------------------------------------------------
   1    2    3    4    5    6    7    8    9   10   11   12   13   16
8408 1343  431  232  181   81   31   12    5    5    3    1    1    1


 #### FREQUENCY OF VARIABLES IN R18 ##################################################

----- IS_WINDFALL_INCOME -------------------------------------------------------
    1     2  <NA>
 7611 12343     4


 #### FREQUENCY OF VARIABLES IN R19 ##################################################

----- PERSON_5_5_2 -------------------------------------------------------
   1    2    3    4    5    6    7    8    9   11   41 <NA>
5073 2206  517  228   69   37   11    9    1    1    1    1


 #### FREQUENCY OF VARIABLES IN R20 ##################################################

----- BANKS --------------------------------------------------------------
```

```
    1     2  <NA>
 4897 15058     3
----- FINANCE -------------------------------------------------------
    1     2  <NA>
  896 19057     5
----- EMPLOYER ------------------------------------------------------
    1     2  <NA>
 1552 18402     4
----- LENDER --------------------------------------------------------
    1     2  <NA>
 1710 18244     4
----- RETAIL_SHOPS --------------------------------------------------
    1     2  <NA>
 3262 16691     5
----- PAWNING -------------------------------------------------------
    1     2  <NA>
 6731 13224     3
----- INSTALMENT_GOODS ----------------------------------------------
    1     2  <NA>
  862 19092     4
----- OTHER_DEBTS ---------------------------------------------------
    1     2  <NA>
  533 19421     4


 #### FREQUENCY OF VARIABLES IN R21 ##################################################

----- RADIO ---------------------------------------------------------
    1     2  <NA>
14705  5252     1
----- TV ------------------------------------------------------------
    1     2  <NA>
15735  4221     2
----- VCD -----------------------------------------------------------
    1     2  <NA>
 7225 12731     2
----- SEWING_MECHINE ------------------------------------------------
    1     2  <NA>
 8143 11814     1
----- WASHING_MECHINE -----------------------------------------------
    1     2  <NA>
 2685 17272     1
----- FRIDGE --------------------------------------------------------
    1     2  <NA>
 7909 12048     1
----- COOKERT -------------------------------------------------------
    1     2  <NA>
 8180 11777     1
----- ELECTRIC_FANS -------------------------------------------------
    1     2  <NA>
```

```
10717  9240     1
----- TELEPHONE ----------------------------------------------------
    1     2  <NA>
 9073 10881     4
----- TELEPHONE_MOBILE ---------------------------------------------------------
    1     2  <NA>
12060  7897     1
----- COMPUTERS ----------------------------------------------------
    1     2  <NA>
 2465 17491     2
----- BICYCLE ------------------------------------------------------
    1     2  <NA>
 8272 11685     1
----- MOTOR_BICYCLE ------------------------------------------------------
    1     2  <NA>
 4846 15111     1
----- THREE_WHEELER ------------------------------------------------------
    1     2  <NA>
 1179 18778     1
----- MOTOR_CAR_VAN ------------------------------------------------------
    1     2  <NA>
 1092 18865     1
----- BUS_LORRY ----------------------------------------------------
    1     2  <NA>
  306 19651     1
----- TRACTOR_2_WHEEL ----------------------------------------------------
    1     2  <NA>
  517 19439     2
----- TRACTOR_4_WHEEL ----------------------------------------------------
    1     2  <NA>
  123 19833     2
----- PESTICIDER ---------------------------------------------------
    1     2  <NA>
  648 19308     2
----- PADDY_BLOWER -------------------------------------------------------
    1     2  <NA>
   59 19897     2
----- WATER_PUMPS --------------------------------------------------
    1     2  <NA>
  410 19546     2
----- BOATS --------------------------------------------------------
    1     2  <NA>
  287 19667     4
----- FISHING_NETS -------------------------------------------------------
    1     2  <NA>
  406 19548     4


 #### FREQUENCY OF VARIABLES IN R22 ################################################
```

```
----- IS_POWER_LINES_NEAR -------------------------------------------------------
    1     2
18857  1101
----- IS_TEL_LINES_NEAR ---------------------------------------------------------
    1     2
16172  3786
----- IS_WATER_SERVICE_NEAR -----------------------------------------------------
    1     2
13134  6824


 #### FREQUENCY OF VARIABLES IN R23 ##################################################

----- STRUCTURE -----------------------------------------------------------
    1     2     3     4     5     9  <NA>
17078   519   788  1401   126    43     3
----- BED_ROOMS -----------------------------------------------------------
   1    2    3    4    5    6    7    8    9 <NA>
4702 7033 5293 1907  412   97   17   10    7  480
----- AREA ----------------------------------------------------
   1    2    3    4    5 <NA>
 968 2979 4411 4934 6662    4
----- WALLS ---------------------------------------------------------
    1     2     3     4     5     6     7     9  <NA>
11609   829  5226   673   795   486    94   242     4
----- FLOOR ---------------------------------------------------------
    1     2     3     4     9  <NA>
16132  1832  1503     1   487     3
----- ROOF ----------------------------------------------------
   1    2    3    4    5    9 <NA>
9778 6440  711 2555  426   44    4
----- OWNERSHIP -----------------------------------------------------------
    1     2     3     4     5     6     7     8     9  <NA>
11539  3812  1096   294  1650   225   833   182   322     5
----- DRINKING_WATER ------------------------------------------------------
   1    2    3    4    5    6    7    8    9 <NA>
5783 2476  647  693 7448 1247 1183  178  297    6
----- OWN_WATER -----------------------------------------------------
    1     2  <NA>
15308  4614    36
----- WATER_SUFFICENCY ----------------------------------------------------
    1     2  <NA>
18540  1392    26
----- OTHER_WATER_SUFFICENCY -----------------------------------------------------
    1     2  <NA>
17814  2109    35
----- TIOILET_USE ---------------------------------------------------
    1     2     3     4  <NA>
17641  1418   392   500     7
----- TOILET_TYPE ----------------------------------------------------
```

```
    1     2     3     9  <NA>
14941  3812   676    15   514
----- GARBAGE_DUMPING ----------------------------------------------------
    1     2     3     4     5     9  <NA>
 4631 10062  1111  2602  1389   157     6
----- LITE_SOURCE --------------------------------------------------------
    1     2     3     4     9  <NA>
 2516 17148   286     3     2     3
----- COOKING_FUEL -------------------------------------------------------
    1     2     3     4     5     9  <NA>
15310  3712   692    57    41   137     9
----- IS_COLLECT_FIREWOOD ------------------------------------------------
    1     2  <NA>
12692  7258     8
----- FIRE_WOOD_OWN ------------------------------------------------------
   1    2    3 <NA>
6924 2610 3153 7271
----- NATURAL_CALAMITY ---------------------------------------------------
    1     2  <NA>
 1260 18685    13
----- FLOODING -----------------------------------------------------------
    1     2  <NA>
  562   700 18696
----- DROUGHT ------------------------------------------------------------
    1     2  <NA>
  326   926 18706
----- LAND_SLIDES --------------------------------------------------------
    1     2  <NA>
   55  1197 18706
----- CIVIL_UNREST -------------------------------------------------------
    1     2  <NA>
   10  1242 18706
----- WILD_ANIMAL --------------------------------------------------------
    1     2  <NA>
  294   958 18706
----- OTHER_CALAMITY -----------------------------------------------------
    1     2  <NA>
  149  1103 18706


 #### FREQUENCY OF VARIABLES IN R24 ##################################################

----- IS_AGRILAND_OWNER --------------------------------------------------
    1     2     3  <NA>
16524  3432     1     1
----- COWS_BUFFALOWS -----------------------------------------------------
    1     2  <NA>
  913 19036     9
----- GOATS_SHEEPS -------------------------------------------------------
    1     2  <NA>
```

```
  336 19613     9
----- PIGS ---------------------------------------------------
    1     2  <NA>
   43 19906     9
----- CHICKENS -----------------------------------------------
    1     2  <NA>
 1577 18372     9
----- OTHER_ANIMALS ------------------------------------------
    1     2  <NA>
   72 19878     8
```

✓ **Summary:**

Problems found in Chapter 4, Chapter 5.1 and Chapter 5.2

File	Problems	TO-DO
R01	Person number >40 (4,571)	Dropped, as discussed in Chapter 4
R01	Number of RELATIONSHIP=1 (Head) is 19,959, larger than that of households.	Already discussed in Chapter 5.1
R03	Person number = NA (3)	Dropped, as discussed in Chapter 4.
R08	PRI_SEC = NA (9)	● To omit the 5 records with the next PID, because all income variables were NA. 12200031460311102, 22300030450911101, 51100110340711104, 62200070080111107, 92200050720911101 ● To read NA as 1 (Primary) in the 4 records with the next PID, because wages were non-zero and no primary occupation was found. 53200110270211106, 53200110270611101, 53200110270811101, 53200110270911101
R10	SENONAS_CROPS_CODE = NA (11)	● To omit the record with the next PID, because all income variables were NA. 33200120420811101 ● To read NA as 9 (Other) in the 10 records with the next PID , because outputs were non-zero. 13200110720411101, 53200100210111101, 61200061140611101, 61200061140611102, 61200120570111101, 91200080100511101, 91200080160511103, 91200080160911101, 91200090230611101, 91200110380711101
R13	SEASONAL_CROP = NA (4)	● To omit the 2 records with the next PID, because all income variables were NA. 91200030681011102, 91200030681011103 ● To read NA as 19 (Other) in the 2 records with the next PID , because outputs were non-zero. 31100080270311102, 91200080140511102

R15	NON_AGRI = NA (6)	● To omit the 3 records with the next PID, because all income variables were NA. 32100031030411101, 61200100340711101, 61200100340711101 ● To read NA as 19 (Other) , in the 3 records with the next PID because outputs were non-zero. 12200021280211102, 31100070130511101, 43100010270411101
R19	Person number > 40 (1)	To drop
R19	Person number = NA (1)	To drop

```
== R08 ==
> R08<-hies2009.old2[[8]]
> table(R08$PRI_SEC,useNA="ifany")
    1     2  <NA>
18155   205     9
> R08[is.na(R08$PRI_SEC),c("PID","PRI_SEC","WAGES_SALARIES","ALLOWENCES","BONUS")]
                   PID PRI_SEC WAGES_SALARIES ALLOWENCES BONUS
3032  12200031460311102      NA             NA         NA    NA
7075  22300030450911101      NA             NA         NA    NA
11814 51100110340711104      NA             NA         NA    NA
12981 53200110270211106      NA           5500         NA    NA
12982 53200110270611101      NA          10000         NA    NA
12983 53200110270811101      NA           6500         NA    NA
12984 53200110270911101      NA           6000         NA    NA
14702 62200070080111107      NA             NA         NA    NA
18216 92200050720911101      NA             NA         NA    NA
> pid1<-c("12200031460311102", "22300030450911101", "51100110340711104",
+ "62200070080111107", "92200050720911101")
> R08<-subset(R08,!is.na(R08$PRI_SEC)|!is.element(R08$PID,pid1))
> dim(R08)
[1] 18364    18
> pid2<-c("53200110270211106", "53200110270611101", "53200110270811101", "53200110270911101")
> R08["PRI_SEC"]<-ifelse(is.element(R08$PID,pid2)&is.na(R08$PRI_SEC),1,R08$PRI_SEC)
```

```
> table(R08$PRI_SEC,useNA="ifany")
    1     2
18159   205
```

== R10 ==

```
> R10<-hies2009.old2[[10]]
> dim(R10)
[1] 4278   21
> table(R10$SEAS_CROPS_CODE,useNA="ifany")
   1    2    3    4    5    6    7    9 <NA>
2782  122   76  695  410   81   23   78   11
> R10[is.na(R10$SEAS_CROPS_CODE),c(20,12:18)]
                  PID SEAS_CROPS_CODE ACR_5_2 RT_5_2  P OUTPUT_5_2 HH_CONSUMPTION INPUT_5_2
257  13200110720411101              NA       2      NA NA       55000           18000      49000
1121 33200120420811101              NA       1      NA NA          NA              NA         NA
1566 53200100210111101              NA       2       2 NA       80000           15000      37000
1711 61200061140611101              NA      NA       2 NA       13200            3900       9300
1712 61200061140611102              NA      NA       2 NA       30000             150       8830
1989 61200120570111101              NA      NA       2 NA       21000           21000       7000
3974 91200080100511101              NA      NA       1 NA       16000           16000       4000
3991 91200080160511103              NA      NA       2 NA        9225            9225       6350
3993 91200080160911101              NA       1      NA NA      280000            1000     159700
3995 91200090230611101              NA      NA       2 NA        7000              NA       5000
4041 91200110380711101              NA      NA       1 NA        1800            1800        175
> pid1<-c("33200120420811101")
> R10<-subset(R10,!is.element(R10$PID,pid1)|!is.na(R10$SEAS_CROPS_CODE))
> dim(R10)
[1] 4277   21
> pid2<-c("13200110720411101", "53200100210111101", "61200061140611101", "61200061140611102",
+ "61200120570111101", "91200080100511101", "91200080160511103", "91200080160911101",
+ "91200090230611101", "91200110380711101")
> R10["SEAS_CROPS_CODE"]<-ifelse(is.element(R10$PID,pid2)&is.na(R10$SEAS_CROPS_CODE),
+ 9,R10$SEAS_CROPS_CODE)
> table(R10$SEAS_CROPS_CODE,useNA="ifany")
   1    2    3    4    5    6    7    9
```

```
2782  122   76  695  410   81   23   88


== R13 ==
> R13<-hies2009.old2[[13]]
> dim(R13)
[1] 4196   20
> table(R13$SEASONAL_CROP,useNA="ifany")
   1    2    3    4    5    6    7    8    9   10   19 <NA>
 989 1355  424  384   68  329  141  216   53   11  222    4
> R13[is.na(R13$SEASONAL_CROP),c(19,12:17)]
                  PID SEASONAL_CROP ACRES_5_3 ROOTS_5_3 PERCHS_5_3 OUTPUT_5_3 INPUT_5_3
1269 31100080270311102            NA         1        NA         NA       6000       1000
3510 91200080140511102            NA        NA        NA         30       2000         NA
3716 91200030681011102            NA        NA        NA         NA         NA         NA
3717 91200030681011103            NA        NA        NA         NA         NA         NA
> R13<-R13[R13$PID!="91200030681011102",]
> R13<-R13[R13$PID!="91200030681011103",]
> dim(R13)
[1] 4194   20
> R13[R13$PID=="31100080270311102","SEASONAL_CROP"]<-19
> R13[R13$PID=="91200080140511102","SEASONAL_CROP"]<-19


== R15==
> R15<-hies2009.old2[[15]]
> dim(R15)
[1] 5528   17
> table(R15$NON_AGRI,useNA="ifany")
   1    2    3    4    5    6    7 <NA>
 126  980  152 2282  797  114 1071    6
> R15[is.na(R15$NON_AGRI),c(16,12:14)]
                  PID NON_AGRI OUTPUT_5_4 INPUT_5_4
913  12200021280211102       NA      10300       2300
2433 31100070130511101       NA     300000     207000
2763 32100031030411101       NA         NA         NA
```

```
3486 43100010270411101        NA     2628500    1578500
4076 61200100340711101        NA          NA         NA
4078 61200100340711101        NA          NA         NA
> pid1<-c("32100031030411101", "61200100340711101", "61200100340711101")
> R15<-subset(R15,!is.element(R15$PID,pid1)|!is.na(R15$NON_AGRI))
> dim(R15)
[1] 5525   17
> pid2<-c("12200021280211102", "31100070130511101", "43100010270411101")
> R15["NON_AGRI"]<-ifelse(is.element(R15$PID,pid2)&is.na(R15$NON_AGRI),7,R15$NON_AGRI)


== R19 ==
> R19<-hies2009.old2[[19]]
> dim(R19)
[1] 8154   21
> table(R19$PERSON_5_5_2,useNA="ifany")
   1    2    3    4    5    6    7    8    9   11   41 <NA>
5073 2206  517  228   69   37   11    9    1    1    1    1
> R19[is.na(R19$PERSON_5_5_2),c(20,11:18)]
                  PID PERSON_5_5_2 LOANS PAWNING_SELLING DEPOSITS_PENSIONS_EPF LOTTERY
5393 531001102508111NA           NA    NA              NA                     NA      NA
     SITTU_DEBTS COMPENSATION OTHER_WINDFALL
5393          NA           NA             NA
> R19<-R19[R19$PID!="531001102508111NA",]
> R19[R19$PERSON_5_5_2>40,c(20,11:18)]
                  PID PERSON_5_5_2  LOANS PAWNING_SELLING DEPOSITS_PENSIONS_EPF LOTTERY
6764 72200110270711141           41 150000              NA                     NA      NA
     SITTU_DEBTS COMPENSATION OTHER_WINDFALL
6764          NA           NA             NA
> R19<-R19[R19$PERSON_5_5_2<=40,]
> dim(R19)
[1] 8152   21

# Updated R03, R08, R10, R13, R15 and R19 in hies2009
> hies2009[[3]]<-R03
> hies2009[[8]]<-R08
```

```
> hies2009[[10]]<-R10
> hies2009[[13]]<-R13
> hies2009[[15]]<-R15
> hies2009[[19]]<-R19

# Saved hies2009
> hies2009.old3<-hies2009
```

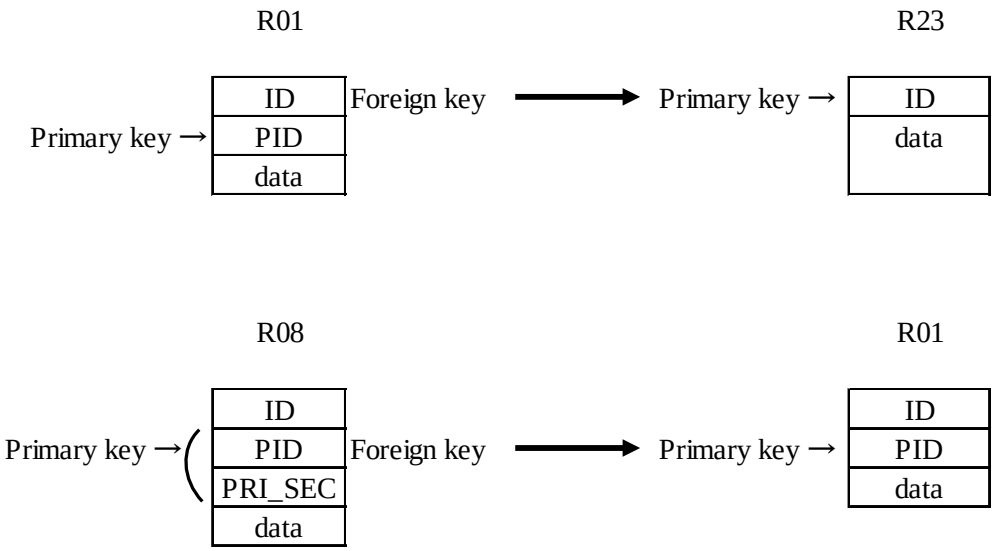
5.3 Primary key and foreign key of each data set

Background: Relationship between data sets

The questionnaire has 24 sections and the micro data is split into 24 data files. Among 24 data sets, R01 and R23 are the most basic. R23 provide the list of all sample households and R01 provides the list of all household members. The next table shows the primary key and the foreign key of each data set.

File	Primary key	No. of duplication	Foreign key	No. of isolation
R01	PID		PID	
R02	PID		PID	
R03	PID		PID	1
R04	ID + CODE (101-1924)		ID	
R05	ID + NF_CODE (2001-3509)		ID	
R06	PID		PID	
R07	ID		ID	
R08	PID + PRI_SEC (1-2)		PID	
R09	ID		ID	
R10	PID + SEAS_CROPS_CODE (1-9)		PID	
R11	ID		ID	
R12	ID		ID	
R13	PID + SEASONAL_CROP (1-19)		PID	
R14	ID		ID	
R15	PID + NON_AGRI (1-7)		PID	1
R16	ID		ID	
R17	PID		PID	1
R18	ID		ID	
R19	PID		PID	
R20	ID		ID	
R21	ID		ID	
R22	ID		ID	
R23	ID		ID	
R24	ID		ID	

Figure: Example of relationship between data sets



☐ **Primary key: Uniqueness**

In case of the primary key = ID

Primary key should be unique. There should be no duplicated key.

✓ **The uniqueness of the primary key was satisfied for all data set.**

```
> file.no<-c(7,9,11,12,14,16,18,20:24)
> for(j in file.no){
+ n<-sum(duplicated(hies2009[[j]]$ID))
+ cat("Number of duplicated IDs in",Rnames[j],"= ",n,"¥n")
+ }
Number of duplicated IDs in R07 =  0
Number of duplicated IDs in R09 =  0
Number of duplicated IDs in R11 =  0
Number of duplicated IDs in R12 =  0
Number of duplicated IDs in R14 =  0
Number of duplicated IDs in R16 =  0
Number of duplicated IDs in R18 =  0
```

```
Number of duplicated IDs in R20 =  0
Number of duplicated IDs in R21 =  0
Number of duplicated IDs in R22 =  0
Number of duplicated IDs in R23 =  0
Number of duplicated IDs in R24 =  0
```

In case of the primary key = PID

Primary key should be unique. There should be no duplicated key.

✓ **The uniqueness of the primary key is satisfied for all data sets.**

```
> file.no<-c(1,2,3,6,17,19)
> for(j in file.no){
+ n<-sum(duplicated(hies2009[[j]]$PID))
+ cat("Number of duplicated PIDs of",Rnames[j],"= ",n,"¥n")
+ }
Number of duplicated PIDs of R01 =  0
Number of duplicated PIDs of R02 =  0
Number of duplicated PIDs of R03 =  0
Number of duplicated PIDs of R06 =  0
Number of duplicated PIDs of R17 =  0
Number of duplicated PIDs of R19 =  0
```

In case that the primary key is a combination of variables

✓ Generated the function to verify the uniqueness of the primary key

```
> verify.unique<-function(j,k1,k2){
+ # j: file no.
+ # k1: ID or PID
+ # k2: unit variable (maximun nchar=4)
+ key1<-hies2009[[j]][,k1]
+ key2<-hies2009[[j]][,k2]
+ pkey<-paste(key1,formatC(key2,width=4,flag="0"),sep="")
+ n<-sum(duplicated(hies2009[[j]]$pkey))
```

```
+ cat("Number of duplicated combination",k1,"+",k2," in",Rnames[j],"= ",n,"¥n")
+ }
```

✓ **The uniqueness of the primary key was satisfied for all data set.**

```
> verify.unique(4,"ID","CODE")
Number of duplicated combination ID + CODE  in R04 =  0
> verify.unique(5,"ID","NF_CODE")
Number of duplicated combination ID + NF_CODE  in R05 =  0
> verify.unique(8,"PID","PRI_SEC")
Number of duplicated combination PID + PRI_SEC  in R08 =  0
> verify.unique(10,"PID","SEAS_CROPS_CODE")
Number of duplicated combination PID + SEAS_CROPS_CODE  in R10 =  0
> verify.unique(13,"PID","SEASONAL_CROP")
Number of duplicated combination PID + SEASONAL_CROP  in R13 =  0
> verify.unique(15,"PID","NON_AGRI")
Number of duplicated combination PID + NON_AGRI  in R15 =  0
```

☐ **Foreign key: Referential integrity**

In case of the foreign key = ID

A corresponding ID should exist in R23.

✓ **The referential integrity was satisfied for all data sets.**

```
> R23<-hies2009[[23]]
> id.all<-R23$ID
> file.no<-c(4,5,7,9,11,12,14,16,18,20:24)
> for(j in file.no){
+ n<-sum(!is.element(hies2009[[j]]$ID,id.all))
+ cat("Number of isolated IDs of",Rnames[j],"= ",n,"¥n")
+ }
Number of isolated IDs of R04 =  0
Number of isolated IDs of R05 =  0
Number of isolated IDs of R07 =  0
```

```
Number of isolated IDs of R09 =  0
Number of isolated IDs of R11 =  0
Number of isolated IDs of R12 =  0
Number of isolated IDs of R14 =  0
Number of isolated IDs of R16 =  0
Number of isolated IDs of R18 =  0
Number of isolated IDs of R20 =  0
Number of isolated IDs of R21 =  0
Number of isolated IDs of R22 =  0
Number of isolated IDs of R23 =  0
Number of isolated IDs of R24 =  0
```

In case of the foreign key = PID

A corresponding PID should exist in R01.

If not, all records within the household with the isolated foreign key PID as well as all household members within the household should be displayed, in order to make it easy to identify the causes of the errors and revise the errors.

- ✓ **The referential integrity was satisfied for R02, R06, R08, R10 and R13. However, there is one isolated foreign key PID in R03, one in R15, one in R17 and two in R19, as follows.**

```
> R01<-hies2009[[1]]
> pid.all<-R01$PID
> file.no<-c(1:3,6,8,10,13,15,17,19)
> for(j in file.no){
+ n<-sum(!is.element(hies2009[[j]]$PID,pid.all))
+ cat("Number of isolated PIDs of",Rnames[j],"= ",n,"¥n")
+ if(n>0){
+ d<-hies2009[[j]]
+ pid.e<-d[!is.element(d$PID,pid.all),"PID"]
+ id.e<-substr(pid.e,1,15)
+ for(k in 1:n){
+ cat("Isolated foreign key: PID=",pid.e[k],"¥n")
+ cat("Records of the household with ID=",id.e[k],"¥n")
+ print(d[d$ID==id.e[k],])
```

```
+ cat("Records of household members with ID=",id.e[k],"¥n")
+ print(R01[R01$ID==id.e[k],])
+ }
+ }
+ }
Number of isolated PIDs of R01 =  0
Number of isolated PIDs of R02 =  0
Number of isolated PIDs of R03 =  1
Isolated foreign key: PID= 82200030470211106
Records of the household with ID= 822000304702111
      DISTRICT PSU REC_TYPE SECTOR DSD MONTH SAMPLE_N SERIAL_NO NHH RESULT R3_PERSON_SERIAL
73770       82  47        3      2   0     3        2         1   1      1                2
73771       82  47        3      2   0     3        2         1   1      1                3
73772       82  47        3      2   0     3        2         1   1      1                4
73773       82  47        3      2   0     3        2         1   1      1                5
73774       82  47        3      2   0     3        2         1   1      1                6
      DID_ATTEND_HOSPITAL REASON_HOSPITAL IS_STAY_HOSPITAL REASON_STAY IS_ILL_DISABLE
73770                   2              NA                2          NA              2
73771                   2              NA                2          NA              2
73772                   2              NA                2          NA              2
73773                   2              NA                2          NA              2
73774                   1               1                1           5              2
      WHAT_ILL_DISABLE IS_EMPL_REASON DURATION_YEARS DURATION_MONTHS IS_ABSENT_ACT DAYS_ABSENT
73770               NA              NA             NA              NA            NA          NA
73771               NA              NA             NA              NA            NA          NA
73772               NA              NA             NA              NA            NA          NA
73773               NA              NA             NA              NA            NA          NA
73774               NA              NA             NA              NA            NA          NA
                   ID               PID       WT
73770 822000304702111 82200030470211102 174.0506
73771 822000304702111 82200030470211103 174.0506
73772 822000304702111 82200030470211104 174.0506
73773 822000304702111 82200030470211105 174.0506
73774 822000304702111 82200030470211106 174.0506
Records of household members with ID= 822000304702111
      DISTRICT PSU REC_TYPE SECTOR DSD MONTH SAMPLE_N SERIAL_NO NHH RESULT PERSON_SERIAL_NO
```

```
77979       82  47        1      2   0     3        2         1   1      1                1
77980       82  47        1      2   0     3        2         1   1      1                2
77981       82  47        1      2   0     3        2         1   1      1                3
77982       82  47        1      2   0     3        2         1   1      1                4
77983       82  47        1      2   0     3        2         1   1      1                5
      RELATIONSHIP SEX_LIVING BIRTH_YEAR BIRTH_MONTH AGE ETHNICITY RELIGION CURR_EDUCATION
77979            1          2         72          10  37         1        1              7
77980            3          1         95          12  14         1        1              2
77981            3          2         98           2  12         1        1              2
77982            3          1         NA           4   9         1        1              2
77983            3          2          8           9   1         1        1             NA
      EDUCATION MARITAL_STATUS MAIN_ACTIVITY MAIN_OCCUPATION INDUSTRY EMPLOYMENT_STATUS
77979         5              2             1            6111      111                5
77980         9              1            NA              NA       NA               NA
77981         6              1            NA              NA       NA               NA
77982         4             NA            NA              NA       NA               NA
77983        NA             NA            NA              NA       NA               NA
                   ID               PID       WT
77979 822000304702111 82200030470211101 174.0506
77980 822000304702111 82200030470211102 174.0506
77981 822000304702111 82200030470211103 174.0506
77982 822000304702111 82200030470211104 174.0506
77983 822000304702111 82200030470211105 174.0506
Number of isolated PIDs of R06 =  0
Number of isolated PIDs of R08 =  0
Number of isolated PIDs of R10 =  0
Number of isolated PIDs of R13 =  0
Number of isolated PIDs of R15 =  1
Isolated foreign key: PID= 53100110250111104
Records of the household with ID= 531001102501111
     DISTRICT PSU REC_TYPE SECTOR DSD MONTH SAMPLE_N SERIAL_NO NHH RESULT SERIAL_5_4 NON_AGRI
3867       53  25       15      1   0    11        1         1   1      1          4        4
     OUTPUT_5_4 INPUT_5_4              ID              PID       WT
3867      10000      4000 531001102501111 53100110250111104 49.91331
Records of household members with ID= 531001102501111
      DISTRICT PSU REC_TYPE SECTOR DSD MONTH SAMPLE_N SERIAL_NO NHH RESULT PERSON_SERIAL_NO
```

```
58913       53  25        1      1   0    11        1         1   1      1                1
58914       53  25        1      1   0    11        1         1   1      1                2
58915       53  25        1      1   0    11        1         1   1      1                3
      RELATIONSHIP SEX_LIVING BIRTH_YEAR BIRTH_MONTH AGE ETHNICITY RELIGION CURR_EDUCATION
58913            1          1         53           4  56         1        1              7
58914            2          2         61           6  48         1        1              7
58915            3          2         87           2  22         1        1              7
      EDUCATION MARITAL_STATUS MAIN_ACTIVITY MAIN_OCCUPATION INDUSTRY EMPLOYMENT_STATUS
58913         8              2             1            6152      601                 3
58914         8              2             1            1412     5211                 5
58915        11              1             2              NA       NA                NA
                  ID               PID       WT
58913 531001102501111 53100110250111101 49.91331
58914 531001102501111 53100110250111102 49.91331
58915 531001102501111 53100110250111103 49.91331
Number of isolated PIDs of R17 =  1
Isolated foreign key: PID= 52100100190711105
Records of the household with ID= 521001001907111
     DISTRICT PSU REC_TYPE SECTOR DSD MONTH SAMPLE_N SERIAL_NO NHH RESULT SERIAL_5_5_1 PENSION
7017       52  19       17      1   0    10        7         1   1      1             5      NA
    DISABILITY_AND_RELIEF PROPERTY_RENTS SAMURDHI DIVIDENDS OTHER ABROAD LOCAL             ID
7017                    NA            NA       NA        NA    NA 120000    NA 521001001907111
                   PID       WT
7017 52100100190711105 41.44451
Records of household members with ID= 521001001907111
      DISTRICT PSU REC_TYPE SECTOR DSD MONTH SAMPLE_N SERIAL_NO NHH RESULT PERSON_SERIAL_NO
55279       52  19        1      1   0    10        7         1   1      1                1
55280       52  19        1      1   0    10        7         1   1      1                2
55281       52  19        1      1   0    10        7         1   1      1                3
55282       52  19        1      1   0    10        7         1   1      1                4
      RELATIONSHIP SEX_LIVING BIRTH_YEAR BIRTH_MONTH AGE ETHNICITY RELIGION CURR_EDUCATION
55279            1          2         69           5  40         2        2              7
55280            3          2         95           8  13         2        2              2
55281            3          2         98          11  10         2        2              2
55282            3          1         NA          12   8         2        2              2
      EDUCATION MARITAL_STATUS MAIN_ACTIVITY MAIN_OCCUPATION INDUSTRY EMPLOYMENT_STATUS
```

```
55279        12              2              4              NA        NA              NA
55280         8              1              3              NA        NA              NA
55281         5              1              3              NA        NA              NA
55282         3             NA             NA              NA        NA              NA
                    ID                PID       WT
55279 521001001907111 52100100190711101 41.44451
55280 521001001907111 52100100190711102 41.44451
55281 521001001907111 52100100190711103 41.44451
55282 521001001907111 52100100190711104 41.44451
Number of isolated PIDs of R19 =  0
```

✓ **Summary:**

Revised the next records.

File	Record	To be read as;	Justification
R03	PID=82200030470211106, R3_PERSON_SERIAL=6	PID=82200030470211101, R3_PERSON_SERIAL=1	Among 1 to 5 of person number in R01, person number=2 to 5 have already counterparts.
R15	PID=82200030470211106 & NON_AGRI=4, SERIAL_5_4=4	PID=82200030470211102 & NON_AGRI=4, SERIAL_5_4=2	Income is non-agri.inc. While the husband (person number=1) is a private sector employee, the wife (person number=2) is an own account worker.
R17	PID=52100100190711105, SERIAL_5_5_1=5	PID=52100100190711101, SERIAL_5_5_1=1	Married female head has person number=1, and her children have person number=2 to 4. Income may be remittance from abroad by her husband.

```
# Revised the three records.
> R03<-hies2009.old3[[3]]
> R15<-hies2009.old3[[15]]
> R17<-hies2009.old3[[17]]

> R03.old<-R03
```

```
> R03[R03$PID=="82200030470211106","R3_PERSON_SERIAL"]<-1
> R03[R03$PID=="82200030470211106","PID"]<-"82200030470211101"
> hies2009[[3]]<-R03

> R15.old<-R15
> R15[R15$PID=="53100110250111104","SERIAL_5_4"]<-2
> R15[R15$PID=="53100110250111104","PID"]<-"53100110250111102"
> hies2009[[15]]<-R15

> R17.old<-R17
> R17[R17$PID=="52100100190711105","SERIAL_5_5_1"]<-2
> R17[R17$PID=="52100100190711105","PID"]<-"52100100190711101"
> hies2009[[17]]<-R17


# Saved hies2009
> hies2009.old4<-hies2009


# Number of records and variables of data files
> file.names<-list.files()
> for(j in 1:24){
+ cat(Rnames[j],":",formatC(nrow(hies2009[[j]]),width=7),",",
+ formatC(ncol(hies2009[[j]]),width=3),":  ",
+ sub(".CSV","",sub("HIES-2009-10-100%-Data-","",file.names))[j],"¥n",sep="")
+ }
R01:  80872, 28:  SEC_1_DEMOGRAPHIC
R02:  20853, 23:  SEC_2_SCHOOL_EDUCATION
R03:  80862, 25:  SEC_3_HEALTH
R04: 798046, 16:  SEC_4_1_FOOD_EXP
R05: 568665, 16:  SEC_4_2_NONFOOD
R06:    240, 27:  SEC_4_3_BOARDERS
R07:  19958, 13:  SEC_4_3_IS_BOADERS
R08:  18364, 18:  SEC_5_1_EMP_INCOME
R09:  19958, 13:  SEC_5_1_IS_EMP_INCOME
R10:   4277, 21:  SEC_5_2_AGRI_INCOME
```

R11: 19958, 13: SEC_5_2_IS_AGRI_INCOME

R12: 19958, 13: SEC_5_3_IS_OTHER_AGRI_INCOME

R13: 4194, 20: SEC_5_3_OTHER_AGRI_INCOME

R14: 19958, 13: SEC_5_4_IS_NON_AGRI_INCOME

R15: 5525, 17: SEC_5_4_NON_AGRI_INCOME

R16: 19958, 13: SEC_5_5_1_IS_OTHER_INCOME

R17: 10735, 22: SEC_5_5_1_OTHER_INCOME

R18: 19958, 13: SEC_5_5_2_IS_WINDFALL_INCOME

R19: 8152, 21: SEC_5_5_2_WINDFALL_INCOME

R20: 19958, 28: SEC_6_B_DEBTNESS

R21: 19958, 35: SEC_6A_DURABLE_GOODS

R22: 19958, 47: SEC_7_BASIC_FACILITIES

R23: 19958, 39: SEC_8_HOUSING

R24: 19958, 40: SEC_9_LAND_ANIMAL

Relationship among R01, R02 and R03

- R01 excluding person number > 40 is the data of all household members.

- The target of R02 is persons aged 5-19 years old. There are 29 missing records in R02, as discussed in Chapter 4.
  PIDs of R02 should have corresponding records in subset of R01 with age 5-19.

```
# Household members in R02 who have no corresponding records in R01;
> R01s<-subset(R01,AGE>=5&AGE<=19)
> setdiff(R02$PID,R01s$PID)
[1] "53200110270211105" "53200110270411101" "53200110270511101"
[4] "71200120430311101" "82200030470211105" "91200060900311101"
```

✓ PID="53200110270211105" should be dropped, because age is 4.

```
> R02[R02$ID=="532001102702111",c("PID","R2_SCHOOL_EDUCATION","GRADE_THIS_YEAR")]
                   PID R2_SCHOOL_EDUCATION GRADE_THIS_YEAR
14189 53200110270211103                  1              6
14190 53200110270211104                  1              4
14191 53200110270211105                  1             14
> R01[R01$ID=="532001102702111",c("PID","RELATIONSHIP","SEX_LIVING","AGE","EDUCATION")]
                   PID RELATIONSHIP SEX_LIVING AGE EDUCATION
59001 53200110270211101            1          1  48        NA
59002 53200110270211102            2          2  39        10
59003 53200110270211103            3          1  10         5
59004 53200110270211104            3          2   9         3
59005 53200110270211105            3          2   4        19
59006 53200110270211106            6          2  44         2
```

✓ For PID="53200110270211105", person number should be read as 3, because only it belongs to the target.

```
> R02[R02$ID=="532001102704111",c("PID","R2_SCHOOL_EDUCATION","GRADE_THIS_YEAR")]
                   PID R2_SCHOOL_EDUCATION GRADE_THIS_YEAR
14192 53200110270411101                  2             NA
> R01[R01$ID=="532001102704111",c("PID","RELATIONSHIP","SEX_LIVING","AGE","EDUCATION")]
                   PID RELATIONSHIP SEX_LIVING AGE EDUCATION
59010 53200110270411101            1          2  33        12
```

```
59011 53200110270411102              2          1  38         12
59012 53200110270411103              3          1   5         19
59013 53200110270411104              3          1   2         NA
```

✓ For PID="53200110270511101", person number should be read as 3, because only it belongs to the target.

```
> R02[R02$ID=="532001102705111",c("PID","R2_SCHOOL_EDUCATION","GRADE_THIS_YEAR")]
                    PID R2_SCHOOL_EDUCATION GRADE_THIS_YEAR
14193 53200110270511101                   1              11
> R01[R01$ID=="532001102705111",c("PID","RELATIONSHIP","SEX_LIVING","AGE","EDUCATION")]
                    PID RELATIONSHIP SEX_LIVING AGE EDUCATION
59014 53200110270511101              1          1  42         10
59015 53200110270511102              2          2  47          6
59016 53200110270511103              3          1  16         10
59017 53200110270511104              3          1   1          6
```

✓ For PID="71200120430311101", person number should be read as 3, because only it belongs to the target.

```
> R02[R02$ID=="712001204303111",c("PID","R2_SCHOOL_EDUCATION","GRADE_THIS_YEAR")]
                    PID R2_SCHOOL_EDUCATION GRADE_THIS_YEAR
16917 71200120430311101                   1               8
> R01[R01$ID=="712001204303111",c("PID","RELATIONSHIP","SEX_LIVING","AGE","EDUCATION")]
                    PID RELATIONSHIP SEX_LIVING AGE EDUCATION
69536 71200120430311101              1          1  45          2
69537 71200120430311102              2          2  46         10
69538 71200120430311103              3          2  13          8
```

✓ As for PID="82200030470211105", person numbers in R02 are shifted one by one.

```
> R02[R02$ID=="822000304702111",c("PID","R2_SCHOOL_EDUCATION","GRADE_THIS_YEAR")]
                    PID R2_SCHOOL_EDUCATION GRADE_THIS_YEAR
19076 82200030470211103                   1              10
19077 82200030470211104                   1               7
19078 82200030470211105                   1               5
> R01[R01$ID=="822000304702111",c("PID","RELATIONSHIP","SEX_LIVING","AGE","EDUCATION")]
                    PID RELATIONSHIP SEX_LIVING AGE EDUCATION
```

```
77979 82200030470211101              1          2  37         5
77980 82200030470211102              3          1  14         9
77981 82200030470211103              3          2  12         6
77982 82200030470211104              3          1   9         4
77983 82200030470211105              3          2   1        NA
```

✓ For PID="91200060900311101", person number should be read as 3, because only it belongs to the target.

```
> R02[R02$ID=="912000609003111",c("PID","R2_SCHOOL_EDUCATION","GRADE_THIS_YEAR")]
                     PID R2_SCHOOL_EDUCATION GRADE_THIS_YEAR
20105 91200060900311101                    1              13
> R01[R01$ID=="912000609003111",c("PID","RELATIONSHIP","SEX_LIVING","AGE","EDUCATION")]
                     PID RELATIONSHIP SEX_LIVING AGE EDUCATION
82095 91200060900311101            1          1  52        19
82096 91200060900311102            2          2  50        10
82097 91200060900311103            3          1  17        12
```

✓ **Summary:**

The next records in R02 should be omitted or revised.

PID in R02	TO-DO
"53200110270211105"	The record should be omitted.
"53200110270411101"	R2_PERSON_SERIAL should be read as 3.
"53200110270511101"	R2_PERSON_SERIAL should be read as 3.
"71200120430311101"	R2_PERSON_SERIAL should be read as 3.
"82200030470211105"	In ID="822000304702111", R2_PERSON_SERIAL =3 should be read as 2. R2_PERSON_SERIAL =4 should be read as 3. R2_PERSON_SERIAL =5 should be read as 4.
"91200060900311101"	R2_PERSON_SERIAL should be read as 3.

```
# Deleted
> R02.old<-R02
> R02<-R02[R02$PID!="53200110270211105",]
> dim(R02)
[1] 20852    23
```

```
# Revised
> R02[R02$PID=="53200110270411101","R2_PERSON_SERIAL"]<-3
> R02[R02$PID=="53200110270511101","R2_PERSON_SERIAL"]<-3
> R02[R02$PID=="71200120430311101","R2_PERSON_SERIAL"]<-3
> R02[R02$PID=="82200030470211103","R2_PERSON_SERIAL"]<-2
> R02[R02$PID=="82200030470211104","R2_PERSON_SERIAL"]<-3
> R02[R02$PID=="82200030470211105","R2_PERSON_SERIAL"]<-4
> R02[R02$PID=="91200060900311101","R2_PERSON_SERIAL"]<-3

# Updated PID because the variable of R2_PERSON_SERIAL was revised.
> R02.old2<-R02
> R02["PID"]<-paste(R02$ID,formatC(R02$ R2_PERSON_SERIAL,width=2,flag="0"),sep="")

# Confirmed again
> setdiff(R02$PID,R01s$PID)
character(0)

# Updated R02 in hies2009 and saved hies2009.
> hies2009[[2]]<-R02
> hies2009.old5<-hies2009
```

- The target of R03 is all household members. As discussed in Chapter 4, there are 11 missing records in R03.
  PIDs of R03 should have corresponding records in R01.

✓ This referential integrity was satisfied.

```
> setdiff(R03$PID,R01$PID)
character(0)
```

5.4 Sample allocation

Strata of HIES are district (22) and sector (3).

The number of psu selected is as follows;

```
> R23.old<-R23
```

✓ Generated psu identifier including district, sector and psu.

```
> R23["psuid"]<-substr(R23$ID,1,10)
```

✓ Generated data set PSU at psu level consist of DISTRICT, SECTOR and psuid.

```
> PSU<-R23[!duplicated(R23$psuid),c("DISTRICT","SECTOR","psuid")]
> dim(PSU)
[1] 2269    3
> head(PSU)
   DISTRICT SECTOR      psuid
1        11      1 1110007001
11       11      2 1120007010
21       11      1 1110003100
29       11      1 1110010101
37       11      1 1110010102
45       11      1 1110010103
```

✓ Number of psu selected by district and sector;

```
> t<-addmargins(table(PSU$DISTRICT,PSU$SECTOR))
> dist.name<-c("Colombo","Gampaha","Kalutara","Kandy","Matale","Nuwara eliya",
+ "Galle","Matara","Hambantota","Jaffna","Vavuniya","Batticaloa","Ampara",
+ "Trincomalee","Kurunegala","Puttalama","Anuradhapura","Polonnaruwa",
+ "Badulla","Moneragala","Ratnapura","Kegalle")
> t<-addmargins(table(PSU$DISTRICT,PSU$SECTOR))
> rownames(t)<-c(dist.name,"Sri Lanka")
> colnames(t)<-c("Urban","Rural","Estate","Total")
> t
```

	Urban	Rural	Estate	Total
Colombo	166	117	10	293
Gampaha	58	139	3	200
Kalutara	44	104	17	165
Kandy	25	72	22	119
Matale	11	41	10	62
Nuwara eliya	10	29	31	70
Galle	44	104	10	158
Matara	44	83	10	137
Hambantota	21	67	2	90
Jaffna	19	17	0	36
Vavuniya	16	22	0	38
Batticaloa	35	53	0	88
Ampara	32	54	0	86
Trincomalee	17	44	0	61
Kurunegala	13	102	5	120
Puttalama	17	66	4	87
Anuradhapura	15	66	2	83
Polonnaruwa	0	57	2	59
Badulla	12	45	24	81

```
Moneragala      0    52     7    59
Ratnapura      14    59    21    94
Kegalle         7    59    17    83
Sri Lanka     620  1452   197  2269
```

According to the final report of the survey, samples were allocated as in the table 1.1 and 1.2.

Table 1.1 Sample allocation and completion by district

District	Housing units		Households responded
	Selected	Responded	
Sri Lanka	**23,631**	**19,809**	**19,958**
Colombo	3000	2,381	2,404
Gampaha	2030	1,696	1,708
Kalutara	1663	1,355	1,377
Kandy	1214	995	1,010
Matale	630	582	587
Nuwara-eliya	809	604	615
Galle	1622	1,466	1,480
Matara	1368	1,224	1,229
Hambantota	921	801	804
Jaffna	750	329	333
Vavuniya	444	324	328
Batticaloa	930	777	778
Ampara	900	772	772
Trincomalee	640	565	565
Kurunegala	1220	1,061	1,066
Puttalama	867	776	780
Anuradhapura	834	750	750
Polonnaruwa	600	524	525
Badulla	806	728	743
Moneragala	600	492	492
Ratnapura	950	868	871
Kegalle	833	739	741

Table 1.2: Sample allocation for sectors

Sector	Housing units		Households responded
	Selected	Responded	
Sri Lanka	**23,631**	**19,809**	**19,958**
Urban	6,549	5,225	5,273
Rural	15,056	12,862	12,949
Estate	2,026	1,722	1,736

The numbers of sample collected were as follows;

```
# The number of samples by district and survey month
> district.code<-unique(R23$DISTRICT)
> district.code
 [1] 11 12 13 21 22 23 31 32 33 41 43 51 52 53 61 62 71 72 81 82 91 92
> dist.name<-c("Colombo","Gampaha","Kalutara","Kandy","Matale","Nuwara eliya",
```

```
+ "Galle","Matara","Hambantota","Jaffna","Vavuniya","Batticaloa","Ampara",
+ "Trincomalee","Kurunegala","Puttalama","Anuradhapura","Polonnaruwa",
+ "Badulla","Moneragala","Ratnapura","Kegalle")
> length(dist.name)
[1] 22
> t1<-addmargins(table(R23$DISTRICT,R23$MONTH))

> rownames(t1)<-c(dist.name,"Sri Lanka")
> t1[c(23,1:22),]
                 1    2    3    4    5    6    7    8    9   10   11   12   Sum
  Sri Lanka   1643 1700 1821 1520 1718 1599 1716 1484 1573 1736 1705 1743 19958
  Colombo      196  194  166  187  173  165  262  201  200  222  209  229  2404
  Gampaha      126  146  164  113  161  164  171  123  160   93  156  131  1708
  Kalutara      65  147  128  101  111  131  158  147   73  131   96   89  1377
  Kandy         82   74   86   79   72   63  108   87   72  112  111   64  1010
  Matale        37   49   48   43   49   48   40   54   47   45   60   67   587
  Nuwara eliya  56   52   76   46   57   52   52   76   32   10   56   50   615
  Galle        124  118  145   92  129   73  160  138  123  115  145  118  1480
  Matara        92   96  100   88  100   91  123  104  103  120  101  111  1229
  Hambantota    59   59   65   58   60   65   85   58   63   90   62   80   804
  Jaffna        59   43   44   54   45   47    0    0    0    0    0   41   333
  Vavuniya      34   38   36   36   25   28    0    0   19   39   17   56   328
  Batticaloa    95   65   81   66   84   84    0    0   83   80   67   73   778
  Ampara        68   67   65   87   53   94    0    0   93   91   64   90   772
  Trincomalee   57   66   56   50   58   61    0    0   38   59   47   73   565
  Kurunegala    92   91  129   60   92   65   87   98   92   87   87   86  1066
  Puttalama     51   82   92   45  105   44   68   73   42   65   69   44   780
  Anuradhapura  73   56   54   60   55   64   83   53   58   72   53   69   750
  Polonnaruwa   38   47   41   37   42   33   56   37   49   55   41   49   525
  Badulla       45   60   64   49   73   65   75   57   58   71   60   66   743
  Moneragala    44   41   48   43   41   34   31   40   38   48   54   30   492
  Ratnapura     69   73   65   74   61   75   90   75   67   80   69   73   871
  Kegalle       81   36   68   52   72   53   67   63   63   51   81   54   741



# The number of samples by sector and survey month

> t2<-addmargins(table(R23$SECTOR,R23$MONTH))
> rownames(t2)<-c("Urban","Rural","Estate","Sri Lanka")
> t2[c(4,1:3),]
              1    2    3    4    5    6    7    8    9   10   11   12   Sum
  Sri Lanka 1643 1700 1821 1520 1718 1599 1716 1484 1573 1736 1705 1743 19958
  Urban      445  439  453  464  420  429  441  360  433  462  394  533  5273
  Rural     1080 1088 1192  975 1138 1061 1087  996 1036 1130 1125 1041 12949
  Estate     118  173  176   81  160  109  188  128  104  144  186  169  1736
```

Remarks:

According to the delegates from Sri Lanka for the International Workshop,

“HIES is conducted on a week base questionnaire in 52 weeks grouped in 12 consecutive monthly rounds to capture the seasonal variation of household income, consumption and expenditure and living conditions.”

5.5 Population estimates

The estimated number of total households and total household members are as follows;

Un-weighted and weighted number of households and household members

	Un-weighted number	Weighted number
Number of households	19,958	5,079,362
Number of household members	80,872	20,337,761
Household size	4.05	4.00

```
> dim(R01)
[1] 80872    28
> sum(R01$WT)
[1] 20337761
> R23<-hies2009[[23]]
> dim(R23)
[1] 19958    39
> sum(R23$WT)
[1] 5079362
> nrow(R01)/nrow(R23)
[1] 4.052109
> sum(R01$WT)/sum(R23$WT)
[1] 4.003999
```

6. Household Income

6.1 Definition of household income

- **Household income**

According to the final report of the survey, household income is defined as follows;

"The Household Income and Expenditure Survey (HIES) defines the household income as the total income received by all the members of the household, either in cash (monetary income) or in kind (nonmonetary income) from all the sources. The household income sources are investigated and reported under the following 7 main categories in the survey questionnaire.

1. Wages and salaries

2. Agricultural activities (seasonal crops)

3. Agricultural activities (non-seasonal crops)

4. Non agricultural activities

5. Other regular cash receipts such as pensions, dividends, rents, interest amounts received from various types of savings, current remittances and local and foreign transfers

6. Irregular gains or windfall income such as compensations, lottery wins etc. and sales of goods and savings.

7. Income in-kind

Obtaining income information from individuals and households is a difficult task as many people reluctant to disclose many of them and often under report. Therefore to ease the field work, which is the most challenging activity of the survey, and to gather more accurate and reliable data, income information of the household members were collected individually in all the 6 income sections tactically arranged in the HIES questionnaire. The income in kind is mostly the estimated values of the household consumed items such as home grown fruits and vegetables, firewood collected etc. and estimated rental values of owner occupied housing units gathered in the consumption expenditure section of the survey questionnaire. An extra column has been provided at household level in the expenditure section to record estimated values of household consumed goods and services received fully or partially free of charge or purchased on price concessions provided by employers etc. This information of income in-kind along with the monitory income collected in the 6 income sections are aggregated and summarized in order to estimate, average monthly household income (mean income), median income, per capita income, income receivers' income and various other indexes such as, Gini coefficients, shares of income etc. at many different geographic and social domain levels."

And the some results of estimated income in the final report are as follows;

Income

Mean household income per month – Sri Lanka	Rs. 36451
Mean household income per month – Urban sector	Rs. 47783
Mean household income per month – Rural sector	Rs. 35228
Mean household income per month – Estate sector	Rs. 24162
Average per capita income per month – Sri Lanka	Rs. 9104
Average per capita income per month – Urban sector	Rs. 11245
Average per capita income per month – Rural sector	Rs. 8916
Average per capita income per month – Estate sector	Rs. 5782
Income receiver's mean income per month – Sri Lanka	Rs. 20427
Income receiver's mean income per month – Urban sector	Rs. 25419
Income receiver's mean income per month – Rural sector	Rs. 20121
Income receiver's mean income per month – Estate sector	Rs. 11639
Monetary income per month – Sri Lanka	Rs. 31209
Non monetary income per month – Sri Lanka	Rs. 5242

- **Income receivers' income**

Definition of income receiver and income receivers' income;

"In order to obtain the Income receiver's income, the HIES records the household income, received from all the sources, by source and person. The Income receiver's income is the sum of the income values recorded in each income section arranged according to the income source in the survey questionnaire.

If a person is less than 10 years old or his aggregated total monthly income is less than Rs. 200, then he was not defined as an income receiver by the HIES 2009/10 and such income values were added to the income of the heads of the respective households. It is obvious that the household income is so built on the income of the income receivers in the household and thus the total household income of the country is equal to the sum of the income values recorded by the total income receivers at all of the source sections of the survey questionnaire."

Note: It is obvious that non-monetary income is not included in income receiver's income.

The results of income receivers in the final report are as follows;

Table 2.9: Average number of income receivers and household size by sector

Sector/Province	No. of income receivers	Household size
Sri Lanka	1.8	4.0
Urban	1.9	4.3
Rural	1.7	4.0
Estate	2.1	4.2

Table 2.10: Income receivers mean and median monthly income by sector

Sector/Province	Median income (Rs.)	Mean income (Rs.)
Sri Lanka	12500	20427
Urban	15593	25419
Rural	12500	20121
Estate	7644	11639

6.2. Process of estimating household income

The data sets R01, R04, R05, R08, R10, R13, R15, R17 and R19 are used for estimating household income.

The data files to be used for estimating household income;

Type of income		Data file	Reference period	Unit of records
Monetary income	Wage	R08	Wages and allowances for one month, Bonus for 12 month	Individual, Occupation (Pri/Sec)
	Agricultural	R10	12 months,	Individual,

	(Seasonal crops)			Seasonal crop code
	Agricultural (Non-seasonal crops)	R13	One month	Individual, Non-seasonal crop code
	Non-agricultural	R15	One month	Individual, Economic activity code
	Other income	R17	6 items for one month, Remittance for 12 months	Individual
	Windfall	R19	12 months	Individual
Non-monetary income	Income in kind	R04 and R05	One week: CODE<2000 6 months: CODE>3000 & CODE<3300 12 months: CODE>3300 & CODE<3400 | CODE>3500 One month: otherwise	Household, Item code

The process of estimating household income is as follows;

Preparing data files:

1. To filter out records without RESULT=1 (Completed).
   (No such record was found in the given data set.)

2. To filter out records with person number >40. They are not regarded as household members.
   (One record was found in R19.)

3. R04 and R05 have no duplicated item code within the household due to the structure of the questionnaire.
   (No such record was found in the given data set.)

Generating the variable of monthly monetary income

4. To calculate monthly income for each record, that is, for each person and for each activity in R08, R10, R13, R15, R17 and R19. The variable of output includes the estimated total value of the total output of the product sold or to be sold plus consumed or to be consumed.

5. To replace the income in each record with 0 if it is less than 0.

Data file	Variable of monthly income	Definition of monthly monetary income	Remarks
R08	wage.inc	WAGES_SALARIES+ALLOWANCES+ BONUS/12	
R10	crop.inc	(OUTPUT_5_2－HH_CONSUMPTION－ INPUT_5_2)/12	Replace crop.inc=0 if crop.inc<0
R13	livestock.inc	OUTPUT_5_3－INPUT_5_3	Replace livestock.inc=0 if livestock.inc<0
R15	nonagri.inc	OUTPUT_5_4－INPUT_5_4	Replace nonagri.inc=0 if nonagri.inc<0
R17	other.inc	PENSION+DISABILITY_AND_RELIEF+ PROPERTY_RENTS+SAMURDHI+ DIVIDENTS+OTHER+ (ABROAD+LOCAL)/12	
R19	windfall.inc	(LOANS+PAWNING_SELLING+ DEPOSITS_PENSIONS_EPF+ INCOME_WELFARE+SITTU_DEBTS+ COMPENSATION+OTHER_LOTTERY)/12	

Part I: Household income in cash and in kind at household level

Creating data file of household income at household level

6. To aggregate the above monthly income by household, and create data file with household-level records from R08, R10, R13, R15, R17 and R19.

7. To merge the above data files using key of household identifier.

8. To generate the variable of the monthly total monetary household income.

Non-monetary income (income in kind)

Consumption of own production will be captured both at consumption side in section 4 of the

questionnaire and production side in section 5. The former and the latter should be equivalent in the concept, while they are not the same in practice. The next is the treatment for rice and the other products taken by DCS.

For rice (paddy),

to compare the consumption of own production (INKIND_VALUE) of rice in section 4 with the output of paddy consumed by the household (HH_CONSUMPTION) in section 5 at household level, and take the larger one as in-kind income of rice (inkind.income).

9. To aggregate (INKIND_VALUE*(30/7)) of R04 with CODE=<4 by household.
10. To aggregate (HH_CONSUMPTION/12) of R10 with SEAS_CROPS_CODE=1 by household.
11. To take the larger one out of the above two as monthly monetary income of rice (inkind.rice) at household level.

Remarks:

According to the delegates from Sri Lanka for the International Workshop,

"Larger value of freely received rice value and paddy income in kind (household consumption value of the paddy production) is added to the household income as paddy income in kind."

For other products,

to take the consumption of own production (INKIND_VALUE of R04 and NF_INKIND_VALUE of R05) in section 4 as in-kind income.

12. To convert INKIND_VALUE of R04 and NF_INKIND_VALUE of R05 to monthly variable (inkind04 and inkind05).
13. To aggregate the above by household.

Data file	Variable of monthly income	Definition	Remarks
R04	inkind.rice	max (rice.a, rice.b) where, rice.a: Sum of (INKIND_VALUE*(30/7)) of R04 with CODE=<104 within the household rice.b: Sum of (HH_CONSUMPTION/12) of R10 with SEAS_CROPS_CODE=1 within the household	

R04	inkind04	Sum of INKIND_VALUE*(30/7) (CODE>=105) within the household	
R05	inkind05	Sum of the followings within the household NF_INKIND_VALUE/6 if NF_CODE>3000 & NF_CODE<3300, NF_INKIND_VALUE/12 if (NF_CODE>3300 & NF_CODE<3400) or NF_CODE>3500, NF_INKIND_VALUE otherwise	

Generating the variables of non-monetary income and total income at household level

14. To generate the variable of non-monetary income by adding inkind.rice, inkind04 and inkind05.
15. To generate the variable of monthly total income by adding monetary income and non-monetary income.

Part II: Income receivers' income

Data files to be used for estimating income receivers' income

Income source	Data file	Variables to be used	Unit of records
Wage	R08	wage.inc	Individual, Occupation (Pri/Sec)
Agricultural (Seasonal crops)	R10	crop.inc	Individual, Seasonal crop code
Agricultural (Non-seasonal crops)	R13	livestock.inc	Individual, Non-seasonal crop code
Non-agricultural	R15	nonagri.inc	Individual, Economic activity code
Other income	R17	other.inc	Individual
Windfall	R19	windfall.inc	Individual
Individual characteristics	R01	AGE, RELATIONSHIP	Individual

If a person is less than 10 years old or his/her aggregated total monthly income is less than Rs. 200, then he/she was not defined as an income receiver by the HIES 2009/10 and such income values were added to the income of the heads of the respective households.

Note: It is obvious that non-monetary income is not included in income receiver's income.

Generating individual-level income data file

16. To collapse income data files R08, R10, R13 and R15 at individual level
17. To merge all individual-level income data files R08, R10, R13, R15, R17 and R19, as well as individual characteristics file R01 by using PID.
18. To generate the variable of monthly total income.

Transferring the income of persons with age less than 10 years or income less than Rs. 200 to the head of the respective households

19. If a person is less than 10 years old or his/her aggregated total monthly income is less than Rs. 200 and his/her relationship is not a head of household, to add such income values to the income of the heads of the respective households.

Defining the income receivers

20. To select records with age >=10 and total income >=200, or relationship=1. Household heads are qualified as income receivers even if the income of the head is none or less than 200.

6.3 Generating the variable of monthly monetary income

```
> hies2009<-hies2009.old5
> R08<-hies2009[[08]]
> R10<-hies2009[[10]]
> R13<-hies2009[[13]]
> R15<-hies2009[[15]]
> R17<-hies2009[[17]]
> R19<-hies2009[[19]]
> R04<-hies2009[[04]]
> R05<-hies2009[[05]]
> hies2009.i1<-hies2009
```

✓ Confirmed that the following unit variables have no missing value NA.

File	Unit Variable	
R08	PRI_SEC	1 Main occupation 2 Secondary occupation
R10	SEAS_CROPS_CODE	1 Paddy 2 Chillies 3 Onions 4 Vegetables 5 Cereals 6 Yams 7 Tobacco 9 Other
R13	SEASONAL_CROP	1 Tea , Rubber 2 Coconuts 3 Coffee, Pepper Betel etc 4 Banana / Fruits 5 Meat 6 Fish 7 Eggs 8 Milk 9 Other food items 10 Horticulture 19 Other

R15	NON_AGRI	1 Mining & Quarrying. 2 Manufacturing 3 Construction 4 Trade 5 Transport 6 Guest house, restaurants, bars/hotels etc 7 Other services
R17	SERIAL_5_5_1	Person number 1-40
R19	PERSON_5_5_2	Person number 1-40

```
> table(R08$PRI_SEC,useNA="ifany")
    1     2
18159   205
> table(R10$SEAS_CROPS_CODE,useNA="ifany")
   1    2    3    4    5    6    7    9
2782  122   76  695  410   81   23   88
> table(R13$SEASONAL_CROP,useNA="ifany")
   1    2    3    4    5    6    7    8    9   10   19
 989 1355  424  384   68  329  141  216   53   11  224
> table(R15$NON_AGRI,useNA="ifany")
   1    2    3    4    5    6    7
 126  980  152 2282  797  114 1074
> table(R17$SERIAL_5_5_1,useNA="ifany")
   1    2    3    4    5    6    7    8    9   10   11   12   13   16
8408 1344  431  232  180   81   31   12    5    5    3    1    1    1
> table(R19$PERSON_5_5_2,useNA="ifany")
   1    2    3    4    5    6    7    8    9   11
5073 2206  517  228   69   37   11    9    1    1

# Replaced NA of the income variables with 0 in the above data sets
> R08[is.na(R08)]<-0
> R10[is.na(R10)]<-0
> R13[is.na(R13)]<-0
> R15[is.na(R15)]<-0
> R17[is.na(R17)]<-0
```

```
> R19[is.na(R19)]<-0
```

✓ **Generated the monthly income variable for each data set.**

```
== R08 ==
> R08["wage.inc"]<-R08[13]+R08[14]+R08[15]/12
> head(R08[c(17,12:15,19)])
                PID PRI_SEC WAGES_SALARIES ALLOWENCES BONUS  wage.inc
1 11100070010111107       1           7500          0     0  7500.000
2 11100070010211101       1          15000          0     0 15000.000
3 11100070010411101       1           7500       1500  2500  9208.333
4 11100070010511107       1           9000          0     0  9000.000
5 11100070010511108       1          11500          0     0 11500.000
6 11100070010711101       1           6000          0     0  6000.000
```

```
== R10 ==
> R10["crop.inc"]<-(R10[,16]-R10[,17]-R10[,18])/12
> table(R10$crop.inc<0)
FALSE  TRUE
 3262  1015
> R10["crop.inc"]<-ifelse(R10$crop.inc<0,0,R10$crop.inc)
> head(R10[c(20,12,16:18,22)])
                PID SEAS_CROPS_CODE OUTPUT_5_2 HH_CONSUMPTION INPUT_5_2   crop.inc
1 11200070110111103               1      30000          30000     14850    0.00000
2 11200111110311101               4       4000            800       650  212.50000
3 11200111110411102               2       1100            400         0   58.33333
4 11200111150411101               1      25000          25000     16460    0.00000
5 11200070120311101               1      40000          40000     15000    0.00000
6 11200070120511101               1      91000          20000     41000 2500.00000
```

```
== R13 ==
> R13["livestock.inc"]<-R13[,16]-R13[,17]
> table(R13$livestock.inc<0)
FALSE  TRUE
```

```
 4150    44
> R13["livestock.inc"]<-ifelse(R13$livestock.inc<0,0,R13$livestock.inc)
> head(R13[c(19,12,16,17,21)])
                PID SEASONAL_CROP OUTPUT_5_3 INPUT_5_3 livestock.inc
1 11100101010211101             6     385000    260000        125000
2 11200070110111103             2       2200         0          2200
3 11200111100211101             2        650        60           590
4 11200111110311101             3       4000       750          3250
5 11200111110311102             7       1200         0          1200
6 11200111110411103             2      12500      6000          6500

== R15 ==
> R15["nonagri.inc"]<-R15[,13]-R15[,14]
> table(R15$nonagri.inc<0)

FALSE  TRUE
 5463    62
> R15["nonagri.inc"]<-ifelse(R15$nonagri.inc<0,0,R15$nonagri.inc)
> head(R15[c(16,12:14,18)])
                PID NON_AGRI OUTPUT_5_4 INPUT_5_4 nonagri.inc
1 11100070010111104        7       2500         0        2500
2 11100070010211105        7      28000     10000       18000
3 11100070010311101        7      18000         0       18000
4 11100070010611101        5      20000      5000       15000
5 11100070010611102        5      12000      7000        5000
6 11100070010811101        5      25000     11250       13750

== R17 ==
> R17["other.inc"]<-rowSums(R17[,12:17])+(R17[,18]+R17[,19])/12
> head(R17[c(21,12:19,23)])
                PID PENSION DISABILITY_AND_RELIEF PROPERTY_RENTS SAMURDHI DIVIDENDS OTHER
1 11100070010111101       0                  400              0      240         0     0
2 11100070010111102       0                  400              0        0         0     0
3 11100070010511102       0                    0              0      600         0     0
4 11100070010511113   10000                    0              0        0         0     0
5 11100070010611103    7000                    0              0        0         0     0
```

```
6 11200070100111101        0                     0           22000        0          0      0
  ABROAD LOCAL other.inc
1      0     0       640
2      0     0       400
3      0     0       600
4      0     0     10000
5      0     0      7000
6      0 15000     23250

== R19 ==
> R19["windfall.inc"]<-rowSums(R19[,12:18])/12
> head(R19[c(20,12:18,22)])
                PID LOANS PAWNING_SELLING DEPOSITS_PENSIONS_EPF LOTTERY SITTU_DEBTS
1 11200070100222101     0           25000                     0       0           0
2 11200070100511103     0           75000                     0       0           0
3 11200070100611101     0           50000                     0       0           0
4 11200070100711101     0           50000                     0       0           0
5 11100031000211101     0           30000                     0       0           0
6 11100031000211102     0           50000                     0       0           0
  COMPENSATION OTHER_WINDFALL windfall.inc
1            0             0     2083.333
2            0             0     6250.000
3            0             0     4166.667
4            0             0     4166.667
5            0             0     2500.000
6            0             0     4166.667


# Updated hies2009
> hies2009[[8]]<-R08
> hies2009[[10]]<-R10
> hies2009[[13]]<-R13
> hies2009[[15]]<-R15
> hies2009[[17]]<-R17
> hies2009[[19]]<-R19
# Saved hies2009
```

```
> hies2009.i2<-hies2009
```

For reference:

Sum of monthly income in each data set.

File	No. of records	Variable of monthly income	Sum of the variable (un-weighted)	Sum of the variable (weighted)	Weighted average income per household
R08	18,364	wage.inc	250,526,008	62,284,504,589	12,262.3
R10	4,277	crop.inc	13,284,307	3,916,717,127	771.1
R13	4,194	livestock.inc	65,899,155	20,208,439,835	3,978.5
R15	5,525	nonagri.inc	175,260,468	46,762,170,361	9,206.3
R17	10,735	other.inc	87,765,278	21,359,560,030	4,205.2
R19	8,152	windfall.inc	56,655,468	14,142,357,917	2,784.3
		Total	649,390,684	168,673,749,859	33,207.7

Un-weighted sum of the variables

```
> t<-c(t,sum(R10$crop.inc))
> t<-c(t,sum(R13$livestock.inc))
> t<-c(t,sum(R15$nonagri.inc))
> t<-c(t,sum(R17$other.inc))
> t<-c(t,sum(R19$windfall.inc))
> t<-c(t,sum(t))
> names(t)<-c("wage","crop","livestock","nonagri","other","windfall","Total")
> t
     wage      crop livestock   nonagri     other  windfall     Total
250526008  13284307  65899155 175260468  87765278  56655468 649390684
```

Weighted sum of the variables

```
> t<-sum(R08$wage.inc*R08$WT)
> t<-c(t,sum(R10$crop.inc*R10$WT))
> t<-c(t,sum(R13$livestock.inc*R13$WT))
> t<-c(t,sum(R15$nonagri.inc*R15$WT))
> t<-c(t,sum(R17$other.inc*R17$WT))
```

```
> t<-c(t,sum(R19$windfall.inc*R19$WT))
> t<-c(t,sum(t))
> names(t)<-c("wage","crop","livestock","nonagri","other","windfall","Total")
> t
        wage         crop    livestock      nonagri        other     windfall        Total
 62284504589   3916717127  20208439835  46762170361  21359560030  14142357917 168673749859


# Weighted average income per household
> round(t/sum(R23$WT),1)
     wage      crop livestock   nonagri     other  windfall     Total
  12262.3     771.1    3978.5    9206.3    4205.2    2784.3   33207.7


**********************************************************************
# R08: Average monthly income per household from paid employment
> t<-addmargins(tapply(R08$wage.inc*R08$WT,R08$PRI_SEC,sum)/sum(R23$WT))
> names(t)<-c("Main occupation","Secondary occupation","Total")
> round(t[c(3,1,2)],1)
               Total      Main occupation Secondary occupation
             12262.3              12199.4                  62.8

# R10: Average monthly income per household from agricultural activities
> t<-addmargins(tapply(R10$crop.inc*R10$WT,R10$SEAS_CROPS_CODE,sum)/sum(R23$WT))
> names(t)<-c("1 Paddy","2 Chillies","3 Onions","4 Vegetables","5 Cereals",
+ "6 Yams","7 Tobacco","9 Other","Total")
> round(t[c(9,1:8)],1)
       Total      1 Paddy   2 Chillies     3 Onions 4 Vegetables    5 Cereals       6 Yams
       771.1        499.7         27.3         14.5        124.5         63.8         19.0
   7 Tobacco      9 Other
         8.0         14.4


# R13: Average monthly income per household from other agricultural activities
> t<-addmargins(tapply(R13$livestock.inc*R13$WT,R13$SEASONAL_CROP,sum)/sum(R23$WT))
> names(t)<-c("1 Tea, Rubber","2 Coconuts","3 Coffee, Pepper Betel etc",
```

```
+ "4 Banana / Fruits","5 Meat","6 Fish","7 Eggs","8 Milk","9 Other food items",
+ "10 Horticulture","19 Other","Total")
> df<-data.frame(Code=names(t),Average=round(t,1),row.names=NULL)
> df[c(12,1:11),]
                         Code Average
12                      Total  3978.5
1               1 Tea, Rubber  1548.7
2                  2 Coconuts   938.8
3  3 Coffee, Pepper Betel etc   380.5
4           4 Banana / Fruits   365.9
5                       5 Meat   166.0
6                       6 Fish   177.0
7                       7 Eggs    21.2
8                       8 Milk   122.8
9          9 Other food items    53.2
10            10 Horticulture    22.1
11                   19 Other   182.5

# R15: Average monthly income per household from non-agricultural activities
> t<-addmargins(tapply(R15$nonagri.inc*R15$WT,R15$NON_AGRI,sum)/sum(R23$WT))
> names(t)<-c("1 Mining & Quarrying","2 Manufacturing","3 Construction","4 Trade",
+ "5 Transport","6 Guest house, restaurants, bars/hotels etc","7 Other services",
+ "Total")
> df<-data.frame(Code=names(t),Average=round(t,1),row.names=NULL)
> df[c(8,1:7),]
                                      Code Average
8                                    Total  9206.3
1                     1 Mining & Quarrying   140.5
2                          2 Manufacturing  1179.3
3                           3 Construction  2931.7
4                                  4 Trade  2532.9
5                             5 Transport  1318.4
6 6 Guest house, restaurants, bars/hotels etc   171.9
7                           7 Other services   931.7

# R17: Average monthly income per household from other cash receipt
```

```
> r17.inc<-colSums(R17[,12:17]*R17$WT)/sum(R23$WT)
> r17.inc<-c(r17.inc,colSums(R17[,18:19]*R17$WT)/sum(R23$WT)/12)
> r17.inc<-c(r17.inc,sum(r17.inc))
> names(r17.inc)[9]<-"Total"
> round(r17.inc,1)
```

PENSION	DISABILITY_AND_RELIEF	PROPERTY_RENTS	SAMURDHI
1058.3	33.0	501.0	127.4
DIVIDENDS	OTHER	ABROAD	LOCAL
157.5	1103.4	758.3	466.1
Total			
4205.2			

```
# R19: Average monthly income per household by chance or ad hoc gains
> r19.inc<-colSums(R19[,12:18]*R19$WT)/sum(R23$WT)/12
> r19.inc<-c(r19.inc,sum(r19.inc))
> names(r19.inc)[8]<-"Total"
> round(r19.inc,1)
```

LOANS	PAWNING_SELLING	DEPOSITS_PENSIONS_EPF	LOTTERY
1005.4	1311.2	228.5	24.1
SITTU_DEBTS	COMPENSATION	OTHER_WINDFALL	Total
163.4	37.5	14.1	2784.3

6.4 Creating data set of household monetary income at household level

Strategy:

To collapse data files R08, R10, R13, R15, R17 and R19 at household level, and create data frame hhinc with variables of monthly household income by income source.

✓ Generated function ind2hh

```
# Converting individual-level data frame to household-level data frame
# df: individual-level data frame
# variables: to be aggregated by household
# hhid: household identifier
> ind2hh<-function(df,variables,hhid){
+ n<-length(variables)
+ d<-data.frame(ID=unique(hhid))
+ for(j in 1:n){
+ t<-tapply(df[,variables[j]],hhid,sum,na.rm=T)
+ d2<-data.frame(ID=names(t),x=as.vector(t))
+ colnames(d2)[2]<-variables[j]
+ d<-merge(d,d2,by="ID",all=T)
+ }
+ return(d)
+ }

# Example: usage of ind2hh
== Individual-level data set ==
> head(R08[order(R08$ID),c("ID","WAGES_SALARIES","ALLOWENCES","BONUS","wage.inc")])
                 ID WAGES_SALARIES ALLOWENCES BONUS wage.inc
475 111000114803111          25800        4000 10000 30633.33
476 111000114804111          14560        2500 10000 17893.33
477 111000114804111          15000           0     0 15000.00
478 111000114805111          40000        5000 15000 46250.00
479 111000114807111          24560        5000 10000 30393.33
480 111000114808111          16380           0  5000 16796.67
== Collapsed to household-level data set ==
> head(ind2hh(R08,c("WAGES_SALARIES","ALLOWENCES","BONUS","wage.inc"),R08$ID))
               ID WAGES_SALARIES ALLOWENCES BONUS wage.inc
```

```
1 111000114803111           25800        4000 10000 30633.33
2 111000114804111           29560        2500 10000 32893.33
3 111000114805111           40000        5000 15000 46250.00
4 111000114807111           24560        5000 10000 30393.33
5 111000114808111           16380           0  5000 16796.67
6 111000115501111           13000           0     0 13000.00
```

✓ Generated household-level data frame with variables of ID and wage.inc.

```
> hhinc<-R23[c("ID","WT")]
> hhinc<-merge(hhinc,ind2hh(R08,"wage.inc",R08$ID),all=T)
> head(hhinc)
               ID      WT wage.inc
1 111000114801111 30.4971       NA
2 111000114802111 30.4971       NA
3 111000114803111 30.4971 30633.33
4 111000114804111 30.4971 32893.33
5 111000114805111 30.4971 46250.00
6 111000114807111 30.4971 30393.33
> dim(hhinc)
[1] 19958     3
```

✓ Merged hhinc with household-level data frame with variables of ID and crop.inc, and so on.

```
> hhinc<-merge(hhinc,ind2hh(R10,"crop.inc",R10$ID),all=T)
> hhinc<-merge(hhinc,ind2hh(R13,"livestock.inc",R13$ID),all=T)
> hhinc<-merge(hhinc,ind2hh(R15,"nonagri.inc",R15$ID),all=T)
> hhinc<-merge(hhinc,ind2hh(R17,"other.inc",R17$ID),all=T)
> hhinc<-merge(hhinc,ind2hh(R19,"windfall.inc",R19$ID),all=T)
> dim(hhinc)
[1] 19958     8
> colnames(hhinc)<-c("ID","WT","wage","crop","livestock","nonagri","other","windfall")
> hhinc[is.na(hhinc)]<-0
> hhinc["monetary.tt"]<-rowSums(hhinc[,3:8])
> head(hhinc)
               ID      WT     wage crop livestock nonagri other windfall monetary.tt
1 111000114801111 30.4971     0.00    0         0   60000     0        0     60000.00
```

```
2 111000114802111 30.4971     0.00    0       0  80000     0       0  80000.00
3 111000114803111 30.4971 30633.33    0       0      0 32825       0  63458.33
4 111000114804111 30.4971 32893.33    0       0      0     0       0  32893.33
5 111000114805111 30.4971 46250.00    0       0      0  6125       0  52375.00
6 111000114807111 30.4971 30393.33    0       0      0     0       0  30393.33

# Saved hies2009
> hies2009.i3<-hies2009
```

6.5 Non-monetary income (income in-kind)

The variables to be generated;

Data file	Variable of monthly income	Definition	Remarks
R04	inkind.rice	max (rice.a, rice.b) where, rice.a: Sum of (INKIND_VALUE*(30/7)) of R04 with CODE=<104 within the household rice.b: Sum of (HH_CONSUMPTION/12) of R10 with SEAS_CROPS_CODE=1 within the household	
R04	inkind04	INKIND_VALUE*(30/7) (CODE>=105)	
R05	inkind05	NF_INKIND_VALUE/6 if NF_CODE>3000 & NF_CODE<3300, NF_INKIND_VALUE/12 if (NF_CODE>3300 & NF_CODE<3400) or NF_CODE>3500, NF_INKIND_VALUE otherwise	

```
> R04<-hies2009[[4]]
> R05<-hies2009[[5]]
> dim(R04)
[1] 798046     16
> dim(R05)
[1] 568665     16
```

✓ Generated variable rice.a from R04

```
> d<-subset(R04,CODE<=104)
> dim(d)
[1] 25533    16
> table(is.na(d$INKIND_VALUE))
FALSE  TRUE
 2738 22795
> t<-tapply(d$INKIND_VALUE*(30/7),d$ID,sum,na.rm=T)
> rice1<-data.frame(ID=names(t),rice.a=as.vector(t))
> table(rice1$rice.a>0)
FALSE  TRUE
```

```
17095  2375
```

✓ Generated variable rice.b from R10

```
> d2<-subset(R10,SEAS_CROPS_CODE==1)
> dim(d2)
[1] 2782   22
> length(unique(d2$ID))
[1] 2741
> t2<-tapply(d2$HH_CONSUMPTION/12,d2$ID,sum,na.rm=T)
> rice2<-data.frame(ID=names(t2),rice.b=as.vector(t2))
> dim(rice2)
[1] 2741    2
> head(rice2)
               ID    rice.b
1 111000424205111 1666.6667
2 111000526709111  375.0000
3 111000701601111 1416.6667
4 111000701610111  333.3333
5 112000116209111    0.0000
6 112000220507111 2083.3333
```

✓ Generated the variable of inkind.rice after comparing rice.a and rice.b

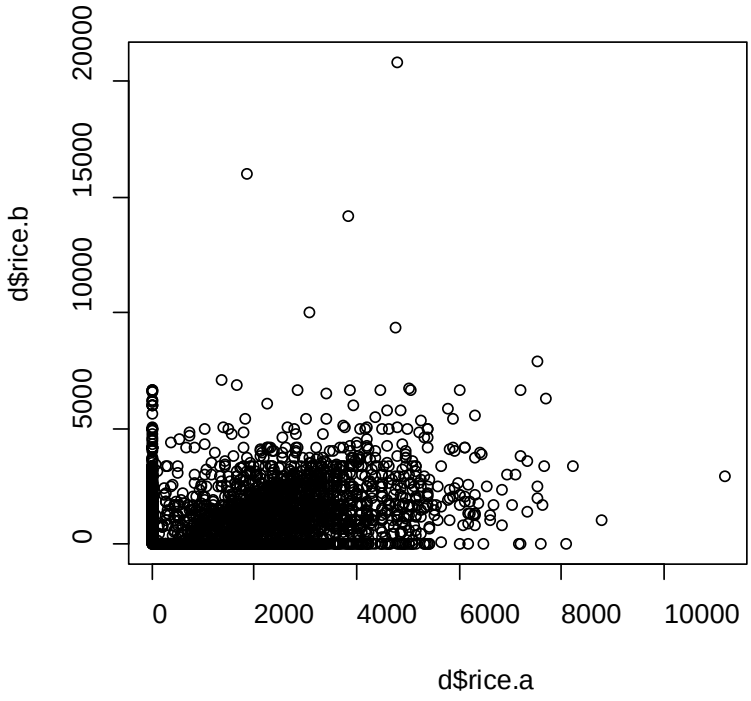
```
> inkind<-R23[c("ID","WT")]
> inkind<-merge(inkind,rice1,all=T)
> inkind<-merge(inkind,rice2,all=T)
> dim(inkind)
[1] 19958     4
> inkind[is.na(inkind)]<-0
> inkind["inkind.rice"]<-pmax(inkind$rice.a,inkind$rice.b)
> head(inkind[inkind$rice.a>0&inkind$rice.b>0,])
                  ID       WT    rice.a    rice.b inkind.rice
330  111000424205111 696.6240 2185.714 1666.6667    2185.714
637  111000701610111 495.3275  900.000  333.3333     900.000
1466 112000220507111 221.5858 2914.286 2083.3333    2914.286
1535 112000321806111 244.8862 3741.429 1050.0000    3741.429
1536 112000321808111 244.8862 2400.000 1260.0000    2400.000
```

```
1607 112000424106111 305.7838 1542.857 4783.3333     4783.333
```

Remarks:

There are 3,277 sample households with rice.a>0 and/or ric.b>0. The value of rice.a is bigger than that of rice.b in 1,989 households.

```
> d<-subset(inkind,rice.a>0|rice.b>0)
> dim(d)
[1] 3277    5
> table(d$rice.a>d$rice.b)
FALSE  TRUE
 1288  1989
> plot(d$rice.a,d$rice.b)
```



✓ Generated the variable inkind04 and inkind05

```
== R04 ==
> d<-subset(R04,CODE>=105)
> dim(d)
[1] 772513     16
> t<-tapply(d$INKIND_VALUE*(30/7),d$ID,sum,na.rm=T)
```

```
> inkind04<-data.frame(ID=names(t),inkind04=as.vector(t))
> dim(inkind04)
[1] 19958     2
> table(inkind04$inkind04>0)
FALSE  TRUE
 5072 14886
```

```
== R05 ==
> d<-subset(R05,!is.na(NF_INKIND_VALUE))
> dim(d)
[1] 43368    16
> d["inkind05"]<-d$NF_INKIND_VALUE
> d["inkind05"]<-ifelse(d$NF_CODE>3000&d$NF_CODE<3300,
+ d$inkind05/6,d$inkind05)
> d["inkind05"]<-ifelse(d$NF_CODE>3300&d$NF_CODE<3400|d$NF_CODE>3500,
+ d$inkind05/12,d$inkind05)
> t<-tapply(d$inkind05,d$ID,sum,na.rm=T)
> inkind05<-data.frame(ID=names(t),inkind05=as.vector(t))
> dim(inkind05)
[1] 19509     2
```

✓ Merged inkind, inkind04 and inkind05.

```
> inkind<-merge(inkind,inkind04,all=T)
> inkind<-merge(inkind,inkind05,all=T)
> dim(inkind)
[1] 19958     7
```

✓ Generated the variable of total non-monetary income

```
> inkind[is.na(inkind)]<-0
> inkind["inkind.tt"]<-rowSums(inkind[,5:7])
> head(inkind[inkind$inkind.rice>0,])
                 ID        WT    rice.a   rice.b inkind.rice    inkind04 inkind05 inkind.tt
37  111000116807111 509.68350  338.5714    0.000    338.5714    42.85714     8000  8381.429
73  111000117405111 342.63580 2721.4286    0.000   2721.4286 26940.00000    51067 80728.429
220 111000310003111  34.93164  428.5714    0.000    428.5714     0.00000     9000  9428.571
330 111000424205111 696.62400 2185.7143 1666.667   2185.7143   171.42857     5000  7357.143
```

```
427 111000526601111 406.64470  257.1429     0.000     257.1429    814.28571       500   1571.429
440 111000526709111 448.65360     0.0000   375.000     375.0000   1714.28571     50000 52089.286
```

✓ Merged hhinc and inkind, and generated the variable of total income.

```
> hhinc.old<-hhinc
> inkind.old<-inkind
> hhinc<-merge(hhinc,inkind[,-2])
> hhinc["ttinc"]<-hhinc$monetary.tt+hhinc$inkind.tt
> head(hhinc[hhinc$inkind.rice>0,])
                 ID        WT  wage     crop livestock nonagri other windfall monetary.tt
37  111000116807111 509.68350 45000    0.000         0        0 12000    0.000        57000
73  111000117405111 342.63580 34500    0.000         0        0     0    0.000        34500
220 111000310003111  34.93164 23000    0.000         0        0     0 3500.000        26500
330 111000424205111 696.62400 30000 2666.667         0        0     0 4583.333        37250
427 111000526601111 406.64470 10000    0.000         0        0     0    0.000        10000
440 111000526709111 448.65360 40000 3375.000    150000        0     0    0.000       193375
       rice.a   rice.b inkind.rice    inkind04 inkind05 inkind.tt     ttinc
37   338.5714    0.000    338.5714    42.85714     8000  8381.429  65381.43
73  2721.4286    0.000   2721.4286 26940.00000    51067 80728.429 115228.43
220  428.5714    0.000    428.5714     0.00000     9000  9428.571  35928.57
330 2185.7143 1666.667   2185.7143   171.42857     5000  7357.143  44607.14
427  257.1429    0.000    257.1429   814.28571      500  1571.429  11571.43
440    0.0000  375.000    375.0000  1714.28571    50000 52089.286 245464.29
```

```
# Average monthly household income per household by income source
> round(colSums(hhinc[,3:16]*hhinc[,"WT"])/sum(hhinc$WT))
       wage        crop   livestock     nonagri       other    windfall monetary.tt
      12262         771        3979        9206        4205        2784       33208
     rice.a      rice.b inkind.rice    inkind04    inkind05   inkind.tt       ttinc
        368         279         477        1147        3955        5579       38786
```

```
# Un-weighted total amount of monthly income
> colSums(hhinc[,3:16])
       wage        crop   livestock     nonagri       other    windfall monetary.tt
```

```
250526008    13284307    65899155   175260468    87765278    56655468   649390684
    rice.a      rice.b inkind.rice     inkind04    inkind05  inkind.tt       ttinc
   5861344     4339848     7536018     22580031    81692825   111808875   761199559
```

```
# Saved hies2009
> hies2009.i4<-hies2009
> hhinc.old2<-hhinc
```

✓ **Summary:**

Average monthly household income per household by income source

	My estimates	Final report of the survey*	
Total income	38,786	36,451	
Monetary income	33,208	30,803	
Wages/Salaries	12,262	12,434	
Agricultural activities	4,750	4,832	
Non-agricultural activities	9,206	6,477	
Other income	4,205	4,252	
Windfall income	2,784	2,808	
Non-monetary income	5,579	5,648	

Note: * The figures of the final report came from Mr. KMR Wickramasinghe, DCS.

For reference:

	Un-weighted estimates	Weighted estimates	
Total number of households	19,958	5,079,362	
Total amount of monthly monetary income	649,390,684	168,673,749,859	
Average monthly monetary income per household	32,538	33,208	

```
> sum(hhinc$monetary.tt)
[1] 649390684
> sum(hhinc$monetary.tt*hhinc$WT)
[1] 168673749859
```

```
> mean(hhinc$monetary.tt)
[1] 32537.86
> sum(hhinc$monetary.tt*hhinc$WT)/sum(hhinc$WT)
[1] 33207.66
```

****
****

Caution: Errors in the distributed data file R15

Through discussion with Mr. KMR Wickramasinghe on the possible causes of large discrepancy of income from non-agricultural activities in the above summary table, the next fact was found.
The following revisions were made during the primary tabulation and the final tabulation in DCS, but it was not reflected on the distributed data file R15 (Section 5.4 Income from non-agricultural activities).

Case (DISTRICT/ PSU/ SAMPLE_N/ NHH/ SERIAL_5_4	Revisions made in DCS
Case 12/38/7/1/2	Output 1,680,000 → 168,000
Case 12/188/1/1/1	Output 3,450,000 → 345,000
Case 61/72/1/1/4	Output 3,500,000 → 350,000
Case 61/93/6/1/1	Output 180,000,000 → 15,000,000 Input 143,680,000 → 12,680,000
Case 91/54/3/1/1	Output 60,000,000 → 5,000,000 Input 50,000,000 → 4,000,000
Case 53/25/1/1/4	The record was deleted because the household size of case 53/25/1/1 is three.

The above revisions were made on the resampled data set. See Chapter 8 Resampling method.

****
****

6.6 Creating individual-level income data file

Strategy for estimating income receivers' income;

1) To collapse income data files R08, R10, R13 and R15 at individual level
2) To merge all individual-level income data files R08, R10, R13, R15, R17 and R19 by using PID.
3) To generate the variable of monthly total income
4) If a person is less than 10 years old or his/her aggregated total monthly income is less than Rs. 200, to add such income values to the income of the heads of the respective households.

Data files to be used for estimating income receivers' income

Income source	Data file	Variables to be used	Unit of records
Wage	R08	wage.inc	Individual, Occupation (Pri/Sec)
Agricultural (Seasonal crops)	R10	crop.inc	Individual, Seasonal crop code
Agricultural (Non-seasonal crops)	R13	livestock.inc	Individual, Non-seasonal crop code
Non-agricultural	R15	nonagri.inc	Individual, Economic activity code
Other income	R17	other.inc	Individual
Windfall	R19	windfall.inc	Individual
Individual characteristics	R01	AGE, RELATIONSHIP	Individual

✓ Defined function col2ind.

```
> col2ind<-function(df,variables,pid){
+ # df: data frame at individual and some unit level
+ # variables: vector of variable names to be applied
+ # indid: individual identifier (grouping factor)
+ # FUN: sum, na.rm=T
+ n<-length(variables)
+ d<-data.frame(PID=unique(pid))
+ for(j in 1:n){
```

```
+ t<-tapply(df[,variables[j]],pid,sum,na.rm=T)
+ d2<-data.frame(PID=names(t),x=as.vector(t))
+ colnames(d2)[2]<-variables[j]
+ d<-merge(d,d2,by="PID",all=T)
+ }
+ return(d)
+ }

# Example: usage of col2ind
> head(R08[R08$PRI_SEC==2,"PID"])
[1] "11200111150911102" "11200070120711101" "11100070130911102" "11300121410611101"
[5] "11100011560811101" "11100011700711101"
== BEFORE ==
> R08[R08$PID=="11200111150911102",
+ c("PID","PRI_SEC","WAGES_SALARIES","ALLOWENCES","BONUS","wage.inc")]
                 PID PRI_SEC WAGES_SALARIES ALLOWENCES BONUS wage.inc
170 11200111150911102       1          20045       4500     0    24545
171 11200111150911102       2          12000          0     0    12000
> df<-col2ind(R08,c("WAGES_SALARIES","ALLOWENCES","BONUS","wage.inc"),R08$PID)
== AFTER ==
> df[df$PID=="11200111150911102",]
                  PID WAGES_SALARIES ALLOWENCES BONUS wage.inc
2245 11200111150911102          32045       4500     0    36545
```

✓ Extracted ID, PID, WT, AGE and RELATIONSHIP from R01, and created data frame pinc.

```
> pinc<-R01[c("ID","PID","WT","AGE","RELATIONSHIP")]
> head(pinc)
              ID               PID       WT AGE RELATIONSHIP
1 111000700101111 11100070010111101 41.42092  77            1
2 111000700101111 11100070010111102 41.42092  52            3
3 111000700101111 11100070010111103 41.42092  47            3
4 111000700101111 11100070010111104 41.42092  41            3
5 111000700101111 11100070010111105 41.42092   6            5
6 111000700101111 11100070010111106 41.42092   2            5
> dim(pinc)
```

```
[1] 80872     5
```

✓ Merged pinc with individual-level R08 with variables of PID and wage.inc.

```
> pinc<-merge(pinc,col2ind(R08,"wage.inc",R08$PID),all.x=T)
> head(pinc[!is.na(pinc$wage.inc),])
                PID              ID      WT AGE RELATIONSHIP wage.inc
13 11100011480311103 111000114803111 30.4971  36            3 30633.33
19 11100011480411101 111000114804111 30.4971  42            1 17893.33
20 11100011480411102 111000114804111 30.4971  18            3 15000.00
24 11100011480511103 111000114805111 30.4971  28            3 46250.00
25 11100011480711101 111000114807111 30.4971  29            1 30393.33
28 11100011480811101 111000114808111 30.4971  29            1 16796.67
> sum(pinc$wage.inc,na.rm=T)
[1] 250526008
```

✓ Merged pinc with individual-level R10 with variables of PID and crop.inc and so on.

```
> pinc<-merge(pinc,col2ind(R10,"crop.inc",R10$PID),all.x=T)
> sum(pinc$crop.inc,na.rm=T)
[1] 17244970

> pinc<-merge(pinc,col2ind(R13,"livestock.inc",R13$PID),all.x=T)
> sum(pinc$livestock.inc,na.rm=T)
[1] 65899155

> pinc<-merge(pinc,col2ind(R15,"nonagri.inc",R15$PID),all.x=T)
> sum(pinc$nonagri.inc,na.rm=T)
[1] 175260468

> pinc<-merge(pinc,R17[c("PID","other.inc")],all.x=T)
> pinc<-merge(pinc,R19[c("PID","windfall.inc")],all.x=T)
> dim(pinc)
[1] 80872    11
> head(pinc)
               PID              ID      WT AGE RELATIONSHIP wage.inc crop.inc
1 11100011480111101 111000114801111 30.4971  44            1       NA       NA
2 11100011480111102 111000114801111 30.4971  41            2       NA       NA
```

```
3 11100011480111103 111000114801111 30.4971  21            3       NA       NA
4 11100011480111104 111000114801111 30.4971  16            3       NA       NA
5 11100011480111105 111000114801111 30.4971  67            4       NA       NA
6 11100011480211101 111000114802111 30.4971  36            1       NA       NA
  livestock.inc nonagri.inc other.inc windfall.inc
1            NA       60000        NA           NA
2            NA          NA        NA           NA
3            NA          NA        NA           NA
4            NA          NA        NA           NA
5            NA          NA        NA           NA
6            NA       80000        NA           NA
```

✓ Replaced NA in income variables with 0.

```
> df<-pinc[,6:11]
> df[is.na(df)]<-0
> pinc<-cbind(pinc[,1:5],df)
> colnames(pinc)[4:11]<-c("age","relation","wage","crop","livestock","nonagri",
+ "other","windfall")
```

✓ Generated the variable ttinc, total monthly income.

```
> pinc["ttinc"]<-rowSums(pinc[,6:11])
> head(pinc)
                PID              ID      WT age relation wage crop livestock
1 11100011480111101 111000114801111 30.4971  44        1    0    0         0
2 11100011480111102 111000114801111 30.4971  41        2    0    0         0
3 11100011480111103 111000114801111 30.4971  21        3    0    0         0
4 11100011480111104 111000114801111 30.4971  16        3    0    0         0
5 11100011480111105 111000114801111 30.4971  67        4    0    0         0
6 11100011480211101 111000114802111 30.4971  36        1    0    0         0
  nonagri other windfall ttinc
1   60000     0        0 60000
2       0     0        0     0
3       0     0        0     0
4       0     0        0     0
5       0     0        0     0
6   80000     0        0 80000
```

Remarks:

Out of 19,958 household heads, the number of those having less than 200 income is 1,832.

```
> addmargins(table(subset(pinc,relation==1)$ttinc<200,useNA="ifany"))
FALSE  TRUE   Sum
18049  1909 19958

# Saved hies2009 and pinc
> hies2009.i5<-hies2009
> pinc.old<-pinc


# For reference:
> colSums(pinc[,6:12],na.rm=T)
     wage      crop livestock   nonagri     other  windfall     ttinc
250526008  13284307  65899155 175260468  87765278  56655468 649390684
```

6.7 Income receivers' income

******* About AGE *******

The variable AGE of R01 has many NA. It is necessary to distinguish persons with age < 10 and persons with age >= 10 in order to define the income receivers.

Firstly, to calculate the age using birth year and birth month if AGE = NA.

Secondly, to estimate age >= 10 unless MARITAL_STATUS =NA.

Lastly, to determine age < 10 considering all persons' information within the household.

```
# Generated the variable age
> d<-R01
> d["smonth"]<-d$MONTH
> d["syear"]<-2009+ifelse(d$MONTH<=6,1,0)
> d["bmonth"]<-d$BIRTH_MONTH
> d["byear"]<-2000+d$BIRTH_YEAR-ifelse(d$BIRTH_YEAR>10,100,0)
> d["m"]<-(d$syear-d$byear)*12+(d$smonth-d$bmonth)
> d["age"]<-ifelse(is.na(d$AGE),floor(d$m/12),d$AGE)
> head(d[is.na(d$AGE),
+ c("PID","AGE","smonth","syear","bmonth","byear","m","age","MARITAL_STATUS")])
                  PID AGE smonth syear bmonth byear  m age MARITAL_STATUS
14  11100070010211107  NA      7  2009     10  2008  9   0             NA
34  11100070010511112  NA      7  2009     12  2008  7   0             NA
138 11100101010411105  NA     10  2009      4  2009  6   0             NA
184 11100101020711104  NA     10  2009     10  2009  0   0             NA
261 11200101040811105  NA     10  2009      6  2009  4   0             NA
267 11200101040911104  NA     10  2009     11  2008 11   0             NA

# 19 NAs found in the variable age
> table(is.na(d$age),useNA="ifany")
FALSE  TRUE
80853    19
> head(d[is.na(d$age),c("PID","AGE","smonth","syear","bmonth","byear","m","age",
+ "MARITAL_STATUS")])
                     PID AGE smonth syear bmonth byear  m age MARITAL_STATUS
15553 12200090430211101  NA      9  2009     NA    NA NA  NA              2
15554 12200090430211102  NA      9  2009     NA    NA NA  NA              2
```

```
15581 12200090430911103  NA      9  2009     NA  2009 NA  NA             NA
16167 12100060580711101  NA      6  2010     NA    NA NA  NA              2
16910 12200110780911103  NA     11  2009     NA    NA NA  NA              2
20181 13200061600711104  NA      6  2010     NA    NA NA  NA              2
```

✓ Defined the variable age10:

0 if age < 10

1 if age >= 10

```
> d["age10"]<-ifelse(d$age>=10|(is.na(d$age)&!is.na(d$MARITAL_STATUS)),1,0)
> table(d$age10,useNA="ifany")
    0     1  <NA>
14181 66687     4
> d[is.na(d$age10),c("PID","AGE","smonth","syear","bmonth","byear","m","age",
+ "MARITAL_STATUS")]
                   PID AGE smonth syear bmonth byear  m age MARITAL_STATUS
15581 12200090430911103  NA      9  2009     NA  2009 NA  NA             NA
31556 23300110400411105  NA     11  2009     NA  2009 NA  NA             NA
36931 31200100370711103  NA     10  2009      8    NA NA  NA             NA
52641 51200110330611106  NA     11  2009     NA    NA NA  NA             NA
```

✓ The top two persons are less than 10 years old, because the birth year is 2009.

The bottom two persons are also considered as age < 10 considering the all household members, as follows;

```
> d[d$ID=="312001003707111",c(27,12,13,19,21,29:35)]
                   PID RELATIONSHIP SEX_LIVING CURR_EDUCATION MARITAL_STATUS
36929 31200100370711101            1          2             7              3
36930 31200100370711102            3          2             7              2
36931 31200100370711103            5          2            NA             NA
      smonth syear bmonth byear   m age age10
36929     10  2009      7  1955 651  54     1
36930     10  2009      7  1981 339  28     1
36931     10  2009      8    NA  NA  NA    NA
> d[d$ID=="512001103306111",c(27,12,13,19,21,29:35)]
                   PID RELATIONSHIP SEX_LIVING CURR_EDUCATION MARITAL_STATUS
52636 51200110330611101            1          1             7              2
52637 51200110330611102            2          2             7              2
```

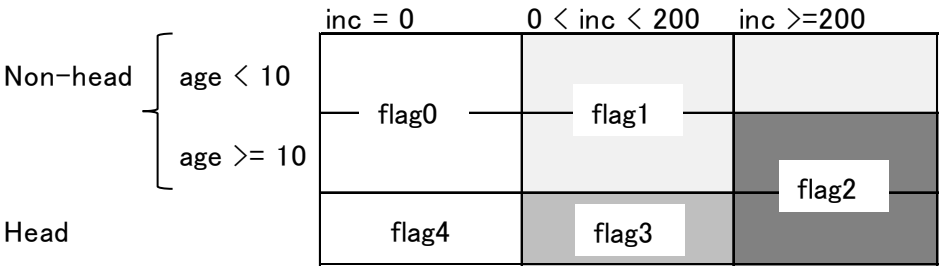
```
52638 51200110330611103             3          1              1              NA
52639 51200110330611104             3          1             NA              NA
52640 51200110330611105             4          2              7               3
52641 51200110330611106             5          2             NA              NA
      smonth syear bmonth byear   m age age10
52636     11  2009      8  1969 483  40     1
52637     11  2009      5  1984 306  25     1
52638     11  2009     10  2005  49   4     0
52639     11  2009      6  2009   5   0     0
52640     11  2009      7  1963 556  46     1
52641     11  2009     NA    NA  NA  NA    NA

# Replaced age10 = NA with age10 = 0.
> d[is.na(d$age10),"age10"]<-0
> table(d$age10,useNA="ifany")
    0     1
14185 66687

# Appended age10 to pinc
> pinc<-merge(pinc,d[c("PID","age10")])

************************************************************
```

Figure: Grouping persons by relationship to household head, age and income



✓ Added flag0=1 to persons in pinc satisfying the next conditions;

1) RELATIONSHIP is not household head.

2) ttinc=0

```
> pinc["flag0"]<-ifelse(pinc$relation!=1&pinc$ttinc==0,1,0)
> addmargins(table(pinc$flag0==1,useNA="ifany"))
FALSE  TRUE   Sum
35655 45217 80872
```

✓ Added flag1=1 to persons in pinc satisfying the next conditions;

1) RELATIONSHIP is not household head.

2) AGE <10 or ttinc <200

3) ttinc>0

```
> pinc["flag1"]<-ifelse(pinc$relation!=1&(pinc$age10==0|pinc$ttinc<200)&
+ pinc$ttinc>0,1,0)
> addmargins(table(pinc$flag1==1,useNA="ifany"))
FALSE  TRUE   Sum
80635   237 80872
> head(pinc[pinc$flag1==1,c("PID","ID","age","relation","ttinc","flag0","flag1")])
                  PID              ID age relation       ttinc flag0 flag1
29   11100011480811102 111000114808111   5        3  8500.00000     0     1
685  11100021950811103 111000219508111   2        3 10500.00000     0     1
941  11100031000511103 111000310005111  28        3   170.00000     0     1
2617 11100070050411104 111000700504111  28        5    83.33333     0     1
3120 11100070260611102 111000702606111  29        2    83.33333     0     1
4905 11100111180711104 111001111807111  81        4   150.00000     0     1
```

✓ Transferred the income of persons with flag=1 to that of the respective household heads.

```
# list of PID with flag=1
> pida<-pinc[pinc$flag1==1,"PID"]
> length(pida)
[1] 237
> head(pida)
[1] "11100011480811102" "11100021950811103" "11100031000511103" "11100070050411104"
[5] "11100070260611102" "11100111180711104"
> table(substr(pida,16,17))
 02  03  04  05  06  07  08
```

```
110  48  35  27  10   5   2

# list of PID of the respective household heads
> pidb<-paste(substr(pida,1,nchar(pida)-2),"01",sep="")
> head(pidb)
[1] "11100011480811101" "11100021950811101" "11100031000511101"
[4] "11100070050411101" "11100070260611101" "11100111180711101"
> pinc["ttinc.before"]<-pinc$ttinc

> pinc.old2<-pinc

# Transfer of income
> for(j in 1:length(pida)){
+ a<-pida[j]
+ b<-pidb[j]
+ pinc[pinc$PID==b,"ttinc"]<-pinc[pinc$PID==b,"ttinc"]+pinc[pinc$PID==a,"ttinc"]
+ pinc[pinc$PID==a,"ttinc"]<-0
+ }

# Example of transfer
> pinc[pinc$ID=="111000114808111",c("PID","age","relation","ttinc","ttinc.before","flag0","flag1")]
                PID age relation    ttinc ttinc.before flag0 flag1
28 11100011480811101  29        1 25296.67     16796.67     0     0
29 11100011480811102   5        3     0.00      8500.00     0     1
> pinc[pinc$ID=="111000310005111",c("PID","age","relation","ttinc","ttinc.before","flag0","flag1")]
                 PID age relation ttinc ttinc.before flag0 flag1
939 11100031000511101  58        1  5170         5000     0     0
940 11100031000511102  37        3     0            0     1     0
941 11100031000511103  28        3     0          170     0     1
942 11100031000511104   9        5     0            0     1     0

> pinc.old3<-pinc
```

✓ Added flag2=1 to persons in pinc satisfying the next conditions;

AGE >=10 and ttinc>=200

```
> pinc["flag2"]<-ifelse(pinc$age10==1&pinc$ttinc>=200,1,0)
> addmargins(table(pinc$flag2,useNA="ifany"))
    0     1   Sum
47354 33518 80872
> head(pinc)
                PID              ID      WT age relation wage crop livestock nonagri other
1 11100011480111101 111000114801111 30.4971  44        1    0    0         0   60000     0
2 11100011480111102 111000114801111 30.4971  41        2    0    0         0       0     0
3 11100011480111103 111000114801111 30.4971  21        3    0    0         0       0     0
4 11100011480111104 111000114801111 30.4971  16        3    0    0         0       0     0
5 11100011480111105 111000114801111 30.4971  67        4    0    0         0       0     0
6 11100011480211101 111000114802111 30.4971  36        1    0    0         0   80000     0
  windfall ttinc age10 flag0 flag1 ttinc.before flag2
1        0 60000     1     0     0        60000     1
2        0     0     1     1     0            0     0
3        0     0     1     1     0            0     0
4        0     0     1     1     0            0     0
5        0     0     1     1     0            0     0
6        0 80000     1     0     0        80000     1
```

✓ Added flag3=1 to household heads with income of 0 < ttinc < 200.

```
> pinc["flag3"]<-ifelse(pinc$relation==1&pinc$ttinc>0&pinc$ttinc<200,1,0)
> table(pinc$flag3,useNA="ifany")
    0     1
80773    99
```

✓ Added flag4=1 to household heads without income.

```
> pinc["flag4"]<-ifelse(pinc$relation==1&pinc$ttinc==0,1,0)
> table(pinc$flag4,useNA="ifany")
    0     1
79071  1801
```

Remarks:

```
Number of household heads without income
> nrow(subset(pinc,relation==1&flag4==1))
[1] 1801
```

✓ Created data set of income receivers consisted of persons with flag2=1 and flag3=1.

```
> receivers<-subset(pinc,flag2==1|flag3==1)
> dim(receivers)
[1] 33617    19

# Estimated number of income receivers
> sum(receivers$WT)
[1] 8469695

# Average number of income receivers per household
> nrow(receivers)/nrow(R23)
[1] 1.684387
> sum(receivers$WT)/sum(R23$WT)
[1] 1.667472

# Average monthly income of income receivers
> sum(receivers$ttinc*receivers$WT)/sum(receivers$WT)
[1] 19914.97

# Total amount of monthly income
> sum(receivers$ttinc)
[1] 649390684
> sum(receivers$ttinc*receivers$WT)
[1] 168673749859
```

✓ **Summary;**

	Un-weighted estimates	Weighted estimates	Final report
Household size	4.05	4.00	4.0
Average number of income receivers per household	1.68	1.67	1.8
Total number of income receivers	33,617	8,469,695	

Income receivers' average monthly income		19,915	20,427
Total amount of monthly income	649,390,684	168,673,749,859	
Number of persons whose income were transferred to the household head	237		
Number of income receivers with ttinc<200	99		

```
> nrow(subset(receivers,relation==1&ttinc<200))
[1] 99
```

Saved hies2009, receivers and pinc

> hies2009.i6<-hies2009

> receivers.old<-receivers

> pinc.old4<-pinc

✓ Sum of ttinc and ttinc.brfore in pinc are the same.

```
> colSums(pinc[,6:19],na.rm=T)
      wage       crop  livestock    nonagri      other   windfall      ttinc
 250526008   13284307   65899155  175260468   87765278   56655468  649390684
     age10      flag0      flag1 ttinc.before      flag2      flag3      flag4
     66687      45217        237  649390684      33518         99       1801
```

✓ Sum of ttinc in receivers is the same as that of pinc..

```
> colSums(receivers[,6:19],na.rm=T)
      wage       crop  livestock    nonagri      other   windfall      ttinc
 250525880   13279031   65893405  175254923   87683520   56634281  649390684
     age10      flag0      flag1 ttinc.before      flag2      flag3      flag4
     33617          0          0  649271039      33518         99          0
```

Remarks:

According to the delegates from Sri Lanka for the International Workshop,

“Household income is calculated at each source at individual level, and minors and minor income receivers are omitted adding their income to the head of the household, and the income receivers income sum up to household income.”

7. Household Expenditure

7.1 Definition of household expenditure

According to the final report of the survey, the definition of household expenditure is as follows;

"Household expenditure data were collected in three major sections of the survey questionnaire.
I. Expenditure on food.
II. Expenditure on non food.
III. Expenditure incurred by boarders and domestic servants.

Under food expenditure, all the food items consumed by the household during the reference period (one week) were collected. For non food expenditure, all non food items and services purchased by the household during the given reference period were collected. Personal expenditure of boarders and domestic servants who live in the household is reported according to the related expenditure group in a separate section of the schedule.

To obtain more accurate date, the first two sections are divided in to 35 sub groups. Of that total 19 sub groups are included under the section of expenditure on food such as cereals , prepared food, vegetables , fish , meat etc. and the rest of 16 sub groups are included under the section of non food expenditure. i.e. housing, fuel and light, health, durable goods etc.

To gather more reliable information on expenditure, food expenditure was collected for 7 consecutive days from each household selected in the sample. But non food expenditure was collected for different reference periods varying from one month to twelve months."

Remarks:

Definition of expenditure on food and non-food

In the survey report, expenditure items are divided into food and non-food as follows;

Food: sum of sub group 01 to 18

Non-food: sum of sub group 19 to 35

Sub group 19 (Liquor, drugs and tobacco) is included in non-food, regardless of the structure of the questionnaire.

The food ratio is also calculated based on the above classification.

The next table of the final report shows the result of average monthly household expenditure;

Table 3.1: Average monthly household expenditure on food and drink and on non-food items by sector, province and district - 2009/10

Sector/ Province/ District	Total expenditure	Expenditure on food & drink	Food ratio	Expenditure on non food items	Non-food ratio
	(Rs)	(Rs)	%	(Rs)	%
Sri Lanka	**31331**	**13267**	**42**	**18064**	**57.7**
Sector					
Urban	44928	16003	36	28925	64.4
Rural	29423	12859	44	16564	56.3
Estate	23988	12201	51	11792	49.2

Remarks:

Reasons why boarders' expenditures are captured in the separate questionnaire

According to the delegates from Sri Lanka for the International Workshop,

"If boarders partake food with the household, he is treated as a household member. As he is not a blood relative or dependent, his outside expenditure is not known to the survey respondent who is a permanent member. Therefore, in order to avert an underestimation of the true household expenditure, he is given a separate section in the HIES questionnaire to report his personal expenditure made in addition to the facilities he enjoys in the household he actually being a member."

Remarks: Process of estimating household expenditure

According to the delegates from Sri Lanka for the International Workshop,

"Household expenditure is calculated for a month at Item level using the following,

Item expenditure per month = Item expenditure x 30 / Reference period in days

And the Item expenditure are summed up to household expenditure."

7.2. Process of estimating household expenditure

Three data sets R04, R05 and R06 are used for estimating household expenditure.

The data files to be used for estimating household expenditure;

Data file	Description	Sub group	Reference period	Remarks
R04	Expenditure on food	01 to 19	7 days	
R05	Expenditure on non food	20 to 29 and 34	1 month	
		30 to 32	6 months	
		33 and 35	12 months	
R06	Expenditure incurred by boarders			4-digit code*
col.3	Food purchased outside	02	7 days	0220
col.4	Fuel	21	1 month	2112
col.5	Clothing	30	6 months	3050
col.6	Non-durable goods	32	1 month	3220
col.7	Household services	29	1 month	2910
col.8	Personal care	22	1 month	2221
col.9	Transport & communication	25	1 month	2510
col.10	Recreation	27	1 month	2720
col.11	Durable goods	33	12 months	3340
col.12	Boarding fees paid	20	1 month	2004
col.13	Sent to family	35	1 month	3510
col.14	Savings provident funds	34	1 month	3411
col.15	Miscellaneous	34	1 month	3412

The process of estimating household expenditure is as follows;

Preparing data files: R04, R05 and R06

1. To filter out records without RESULT=1 (Completed).
   (No such record was found in the given data set.)

2. R04 and R05 have no duplicated item code within the household.
   (No such record was found in the given data set.)

3. Col. 3 to col. 15 of R06 correspond to the 4-digit item codes in the above table.

Note: The 4-digit codes of col. 3 to col.15 are confirmed with Mr. KMR Wickramasinghe, DCS. The summaries of R06 made per month should be added to the household expenditure under the codes given above.

Converting to monthly expenditure

4. To convert VALUE of R04, NF_VALUE of R05 and COL_3 to COL_15 of R06 to monthly expenditure taking into consideration of the reference period.
5. To use the multiplier 30/7 when converting weekly to monthly.
6. To generate the variable of subgroups of expenditure items

Estimating household expenditure by 4-digit item

7. To generate the average monthly household expenditure by item from R04 and R05.
8. To generate the average monthly household expenditure of COL_3 to COL_15 in R06.
9. To add the summaries of R06 to the results of 6 under the codes given in 3.

Creating data on household expenditure at household level

10. To aggregate VALUE of R04 and NF_VALUE of R05 by household and by item subgroup, and create data file with records of household and variables of subgroup.
11. To aggregate COL_3 to COL_15 of R06 by household, and create data file with records of household.
12. To merge the above data files using household identifier.
13. To add COL_3 to COL_15 to the respective subgroups.
14. To generate the variables of total household expenditure, expenditure on food and non-food.

7.3. Converting to monthly expenditure

R04: FOOD_EXP

✓ The variable VALUE and INKIND_VALUE;

VALUE: goods and services consumed in cash and in kind, that is, both bought from the market and homegrown/ freely received.

INKIND_VALUE: a part of VALUE.

Therefore, used only the variable VALUE for estimation of household expenditure.

✓ Generated the variable of monthly expenditure, m.value.

```
> R04["m.value"]<-R04$VALUE*30/7
```

✓ Generated the variable sg of subgroups of 4-digit expenditure item (CODE). Subgroup code is the top-2-digit of 4-digit expenditure item code.

```
> R04["sg"]<-floor(R04$CODE/100)
> head(R04[c(15,11:13,18,19)])
               ID CODE QUANTITY VALUE   m.value sg
1 111000700101111 1105       50     6  25.71429 11
2 111000700101111 1107       20    15  64.28571 11
3 111000700101111 1108       50    18  77.14286 11
4 111000700101111 1109      100    38 162.85714 11
5 111000700101111 1110       50     7  30.00000 11
6 111000700101111 1117        0    20  85.71429 11

# Number of records in R04 by subgroup
> addmargins(table(R04$sg,useNA="ifany"))
    1     2     3     4     5     6     7     8     9    10    11    12
46688 40308 29843 142784 26324  8001 31020 26658  8551 18725 249031  2433
   13    14    15    16    17    18    19   Sum
23599 19087 20581 26171 25178 28152 24912 798046
```

R05: NONFOOD

✓ The variable NF_VALUE and NF_INKIND_VALUE;

NF_VALUE: goods and services consumed in cash and in kind, that is, both bought from the market and freely received/received as a gift.

NF_INKIND_VALUE: a part of NF_VALUE.

Therefore, used only the variable NF_VALUE for estimation of household expenditure.

✓ Generated the variable of sg, subgroup of expenditure item

```
> R05["sg"]<-floor(R05$NF_CODE/100)
> head(R05[c(15,11:13,18)])
               ID NF_CODE NF_QUANTITY NF_VALUE sg
1 111000700101111    2001           0     5000 20
```

```
2 111000700101111     2002            0        25 20
3 111000700101111     2003            0       160 20
4 111000700101111     2101            0       800 21
5 111000700101111     2106         6250       750 21
6 111000700101111     2109            2         7 21

# Number of records in R05 by subgroup
> addmargins(table(R05$sg,useNA="ifany"))
    20     21     22     23     24     25     26     27     28     29     30     31
 32271  76684  85336  17397  37526  22146  22220  12813  36522   3874 115389  22286
    32     33     34     35    Sum
 28868   7783  24913  22637 568665
```

✓ Generated the variable of m.value, monthly expenditure.

```
> R05["m.value"]<-R05$NF_VALUE
> R05["m.value"]<-ifelse(R05$sg>=30&R05$sg<=32,R05$m.value/6,R05$m.value)
> R05["m.value"]<-ifelse(R05$sg==33|R05$sg==35,R05$m.value/12,R05$m.value)
> head(R05[c(15,11:13,18,19)])
                ID NF_CODE NF_QUANTITY NF_VALUE sg m.value
1 111000700101111     2001            0     5000 20    5000
2 111000700101111     2002            0       25 20      25
3 111000700101111     2003            0      160 20     160
4 111000700101111     2101            0      800 21     800
5 111000700101111     2106         6250      750 21     750
6 111000700101111     2109            2        7 21       7
```

R06: BOARDERS

✓ Converted the variables of COL_3, COL_5 and COL_11 of R06 to the monthly expenditure, and generated data frame boarder. Variables of other col. are originally monthly.

```
> boarder<-R06
> boarder[is.na(boarder)]<-0
> boarder$COL_3<-boarder$COL_3*30/7
```

```
> boarder$COL_5<-boarder$COL_5/6
> boarder$COL_11<-boarder$COL_11/12
> head(boarder)
  DISTRICT PSU REC_TYPE SECTOR DSD MONTH SAMPLE_N SERIAL_NO NHH RESULT COL_2    COL_3
1       11 113        7      2   0    11        1         1   1      1     4    0.000
2       11 117        7      2   0    11        1         1   1      1     3 2571.429
3       11 117        7      2   0    11        1         1   1      1     4 2571.429
4       11 143        7      1   0    12        4         1   1      1     4    0.000
5       11 145        7      1   0    12        2         1   1      1     3    0.000
6       11 146        7      1   0    12        9         1   1      1     2    0.000
  COL_4 COL_5 COL_6 COL_7 COL_8 COL_9 COL_10 COL_11 COL_12 COL_13 COL_14 COL_15
1     0     0     0     0     0     0      0      0      0   8000   8000      0
2     0    40     0     0   650   500      0      0   1125      0      0   4500
3     0    50     0     0   650   400      0      0   1125      0      0   4500
4     0     0     0     0     0     0      0      0      0   2500      0      0
5     0     0     0     0     0     0      0      0      0   5000      0      0
6     0     0     0     0     0     0      0      0      0   2000      0      0
               ID              PID        WT
1 112001111301111 11200111130111104 210.42070
2 112001111701111 11200111170111103 269.07250
3 112001111701111 11200111170111104 269.07250
4 111001214304111 11100121430411104 436.39860
5 111001214502111 11100121450211103 192.48930
6 111001214609111 11100121460911102  23.52299
```

7.4. Estimating household expenditure by 4-digit item

- ✓ **Created data frame with variables of 4-digit item code and average monthly expenditure per household.**

- ✓ **Generated monthly household expenditure on food by item from R04**

  (Correspond to Table 5.2 of the final report)

```
> (NHH<-sum(R23$WT))
[1] 5079362
> food.item<-by(R04,R04$CODE,function(df) sum(df$m.value*df$WT)/NHH)
> length(food.item)
```

```
[1] 254
> food.byitem<-data.frame(code=as.integer(names(food.item)),
+ m.value=as.vector(food.item))
> head(food.byitem)
  code     m.value
1  101 975.214735
2  102 478.137585
3  103 839.872898
4  104   5.152743
5  105  44.646871
6  106 189.258824
```

✓ **Generated monthly household expenditure on non-food by item from R5**

(Correspond to Table 5.4 of the final report)

```
> nfood.item<-by(RO5,RO5$NF_CODE,function(df) sum(df$m.value*df$WT)/NHH)
> length(nfood.item)
[1] 183

> nfood.byitem<-data.frame(code=as.integer(names(nfood.item)),
+ m.value=as.vector(nfood.item))
> head(nfood.byitem)
  code      m.value
1 2001 3333.036891
2 2002   13.390113
3 2003   99.752748
4 2101  531.639208
5 2102    9.055037
6 2103   76.389799
```

✓ **Binded two data frames of food.byitem and nfood.byitem together, and generated hhexp.byitem.**

```
> hhexp.byitem<-rbind(food.byitem,nfood.byitem)
> dim(hhexp.byitem)
[1] 437   2
```

Remarks: Out-of-range codes in hhexp.byitem

The below codes in hhexp.byitem are out of range, as compared with the codes on the questionnaire;

Code	monthly value	Number of records	ID
0440	0	1	512001103205111
2112	0.01	1	712000406910111
2506	0.11	2	132001207702111
2510	0	1	612001205701111

✓ Omitted the above codes in hhexp.byitem.

```
> code.err<-c(0440,2112,2506,2510)
> hhexp.byitem.old<-hhexp.byitem
> hhexp.byitem<-hhexp.byitem[!is.element(hhexp.byitem$code,code.err),]
> dim(hhexp.byitem)
[1] 433   2
```

✓ **Generated data frame exp.boarder from boarder**

```
> colnames(boarder)
[1] "DISTRICT"  "PSU"       "REC_TYPE"  "SECTOR"    "DSD"       "MONTH"     "SAMPLE_N"
 [8] "SERIAL_NO" "NHH"       "RESULT"    "COL_2"     "COL_3"     "COL_4"     "COL_5"
[15] "COL_6"     "COL_7"     "COL_8"     "COL_9"     "COL_10"    "COL_11"    "COL_12"
[22] "COL_13"    "COL_14"    "COL_15"    "ID"        "PID"       "WT"
> n.boarder<-rep(0,13)       # number of household-level records with expenditure >0
> av.boarder<-rep(0,13)      # Monthly expenditure per household
> for(j in 1:13){
+ n.boarder[j]<-nrow(subset(hhexp,hhexp[,j+36]>0))
+ av.boarder[j]<-sum(hhexp[,j+36]*hhexp$WT)/NHH
+ }
> n.boarder
 [1] 50  8 80 26  8 81 77 26 17 48 99 50 30
> av.boarder
> av.boarder
 [1] 15.1680377  0.2669037  2.1111812  1.5709600  0.4395956  2.9206370  4.2698707
 [8]  1.9453115  1.1599937  9.9697279 20.4474491  8.9468985  1.4301532
```

```
# Monthly average expenditure per household and the number of household with non-zero expenditure
by item code
> for(j in 1:13){
+ cat(colnames(hhexp)[j+36],":",format(round(av.boarder[j],2),width=8),
+ format(n.boarder[j],width=8),"¥n")
+ }
COL_3 :    15.17       50
COL_4 :     0.27        8
COL_5 :     2.11       80
COL_6 :     1.57       26
COL_7 :     0.44        8
COL_8 :     2.92       81
COL_9 :     4.27       77
COL_10 :    1.95       26
COL_11 :    1.16       17
COL_12 :    9.97       48
COL_13 :   20.45       99
COL_14 :    8.95       50
COL_15 :    1.43       30

# Corresponding 4-digit code, which never overlap the codes of R04 and R05
> code.boarder<-c(0220,2112,3050,3220,2910,2221,2510,2720,3340,2004,3510,3411,3412)

# Generated data frame exp.boarder
> exp.boarder<-data.frame(code=code.boarder,col= colnames(boarder)[12:24],
+ m.value=av.boarder,nhh=n.boarder)
> head(exp.boarder)
  code    col     m.value nhh
1  220 COL_3 15.1680377  50
2 2112 COL_4  0.2669037   8
3 3050 COL_5  2.1111812  80
4 3220 COL_6  1.5709600  26
5 2910 COL_7  0.4395956   8
6 2221 COL_8  2.9206370  81
```

✓ **Binded hhexp.byitem and exp.boarder**

```
> df<-rbind(hhexp.byitem,exp.boarder[,c(1,3)])
> hhexp.byitem<-df[order(df$code),]
> dim(hhexp.byitem)
[1] 446   2

# Average monthly expenditure per household
> sum(hhexp.byitem[,"m.value"])
[1] 31435.24
> sum(hhexp.byitem[hhexp.byitem$code<1900,"m.value"])
[1] 13282.47
> sum(hhexp.byitem[hhexp.byitem$code>=1900,"m.value"])
[1] 18152.77
```

✓ **Made table on average monthly expenditure per household by 4-digit code**

```
> for(j in 1:446){
+ cat(format(hhexp.byitem[j,1],width=4),":",format(round(hhexp.byitem[j,2],2),width=8),"¥n")
+ }
```

Table Average monthly expenditure by 4-digit code

101 :	975.21	302 :	32.44	430 :	21.66
102 :	478.14	303 :	35.95	439 :	16.3
103 :	839.87	304 :	12.73	441 :	63.14
104 :	5.15	305 :	8.97	442 :	41.1
105 :	44.65	306 :	51.97	443 :	16.78
106 :	189.26	309 :	9.56	444 :	12.8
107 :	9.21	401 :	17.84	445 :	5.88
108 :	4.45	402 :	73.89	446 :	5.04
109 :	3.02	403 :	29.25	447 :	3.67
110 :	11.38	404 :	23.43	448 :	5.57
111 :	54.38	405 :	46.48	449 :	4.41
112 :	39.47	406 :	17.74	450 :	4.82
113 :	4.26	407 :	21.34	459 :	29.59
114 :	0.07	408 :	54.24	501 :	42.04
115 :	0.73	409 :	130.86	502 :	18.15
116 :	0.89	410 :	59.97	503 :	173.94
119 :	8.82	411 :	39.78	504 :	9.74
201 :	425.98	412 :	52.22	505 :	25.33
202 :	13.51	413 :	23.44	506 :	3.4
203 :	7.24	414 :	38.06	507 :	2.25
204 :	84.57	415 :	6.94	509 :	2.36
205 :	41.95	416 :	22.68	601 :	357.29
206 :	115.5	417 :	13.06	602 :	116.52
207 :	16.41	418 :	13.86	603 :	16.37
208 :	28.67	419 :	18.75	604 :	11.86
209 :	19.05	420 :	9.53	605 :	9.23
210 :	152.48	421 :	10.01	609 :	5.56
211 :	262.56	422 :	1.57	701 :	115.05
212 :	132.94	423 :	12.87	702 :	14.11
213 :	8.28	424 :	7.59	703 :	11.29
214 :	12.03	425 :	1.85	704 :	74.35
215 :	8.56	426 :	10.51	705 :	66.73
219 :	79.29	427 :	4.95	706 :	128.42
220 :	15.17	428 :	6.28	707 :	46.96
301 :	395.51	429 :	6.47	708 :	106.04

709 :	19.23	1103 :	89.01	1308 :	71.65
710 :	97.38	1104 :	182.28	1309 :	16.82
711 :	7.99	1105 :	71.92	1310 :	24.13
712 :	23.37	1106 :	85.08	1311 :	11.26
713 :	73.92	1107 :	46.24	1312 :	6.6
714 :	111.18	1108 :	62.86	1319 :	11.16
715 :	4.2	1109 :	104.23	1401 :	293.49
716 :	86.71	1110 :	88.44	1402 :	21.92
717 :	22.14	1111 :	40.25	1403 :	3.16
718 :	35.48	1112 :	14.32	1404 :	0.44
719 :	6.53	1113 :	11.2	1409 :	4.5
720 :	12.78	1114 :	17.37	1501 :	452.09
721 :	98.84	1115 :	17.32	1502 :	9.58
801 :	184.62	1116 :	12.37	1503 :	2.92
802 :	27.49	1117 :	24.07	1504 :	0.58
803 :	13.01	1118 :	16.83	1509 :	0.78
804 :	10.38	1119 :	15.67	1601 :	163.15
805 :	3.69	1120 :	60.69	1602 :	13.5
806 :	49.89	1121 :	31.55	1603 :	72.34
807 :	0.78	1122 :	9.63	1604 :	26.28
808 :	0.26	1123 :	1.21	1605 :	32.09
809 :	62.32	1129 :	1.44	1606 :	8.36
810 :	47.77	1201 :	5.21	1607 :	4.62
811 :	3.02	1202 :	10.36	1608 :	11.76
812 :	24.86	1203 :	7.02	1609 :	5.84
813 :	4.48	1204 :	0.91	1610 :	9.1
814 :	0.51	1205 :	0.74	1619 :	20.2
815 :	8.04	1206 :	0.63	1651 :	7.4
816 :	0.13	1209 :	1.4	1652 :	3.54
819 :	51.02	1301 :	20.66	1653 :	6.11
901 :	134.27	1302 :	0.57	1654 :	1.25
909 :	0.18	1303 :	2.14	1659 :	0.34
1001 :	737.32	1304 :	34.7	1701 :	9.06
1002 :	0.99	1305 :	47.03	1702 :	52.1
1101 :	46.41	1306 :	7.12	1703 :	17.08
1102 :	158.96	1307 :	784.61	1704 :	4.49

1705 :	214.06	1920 :	47.88	2303 :	51.75
1706 :	3.27	1921 :	35.48	2304 :	74.24
1707 :	10.06	1922 :	24.52	2305 :	194.43
1708 :	41.75	1923 :	4.77	2306 :	192.26
1709 :	13.32	1924 :	34.68	2307 :	8.43
1710 :	5.56	2001 :	3333.04	2309 :	15.94
1711 :	4.05	2002 :	13.39	2401 :	571.29
1712 :	69.37	2003 :	99.75	2402 :	31.67
1713 :	0.97	2004 :	9.97	2403 :	114.15
1719 :	25.79	2101 :	531.64	2404 :	212.41
1801 :	227.2	2102 :	9.06	2405 :	28.64
1802 :	12.4	2103 :	76.39	2409 :	9.87
1803 :	17.36	2104 :	104.51	2411 :	743.03
1804 :	9.06	2105 :	224.07	2412 :	22.73
1805 :	0.51	2106 :	217.59	2413 :	131.25
1806 :	42.54	2107 :	20.08	2414 :	246.48
1807 :	64.74	2108 :	58.16	2415 :	187.2
1808 :	2.6	2109 :	16.47	2419 :	18.57
1809 :	1.34	2110 :	2.8	2501 :	4.41
1810 :	2.99	2111 :	15.69	2502 :	345.36
1811 :	5.84	2112 :	0.27	2503 :	382.72
1812 :	1.9	2119 :	1.31	2504 :	8.98
1819 :	5.36	2201 :	134.07	2505 :	12.78
1901 :	16.41	2202 :	56.19	2509 :	0.65
1902 :	235.91	2203 :	20.92	2510 :	4.27
1903 :	42.92	2204 :	15.79	2601 :	184.59
1904 :	15.68	2205 :	23.66	2602 :	19.21
1905 :	10.08	2206 :	30.69	2603 :	10.94
1906 :	2.43	2207 :	17.29	2604 :	18.25
1907 :	0.41	2208 :	2.74	2605 :	74.38
1908 :	1.4	2209 :	34.36	2606 :	458.38
1909 :	6.98	2210 :	82.55	2607 :	28.39
1910 :	159.33	2219 :	2.52	2608 :	96.55
1911 :	3.89	2221 :	2.92	2609 :	56.13
1912 :	21.52	2301 :	454.39	2610 :	34.74
1919 :	0.27	2302 :	16.52	2619 :	36.32

2701 :	7.37	3013 :	6.53	3212 :	1.55
2702 :	42.63	3014 :	1.54	3213 :	3.53
2703 :	50.05	3015 :	3.31	3214 :	0.86
2704 :	225.64	3016 :	50.87	3219 :	2.45
2705 :	4.32	3017 :	16.09	3220 :	1.57
2706 :	9.51	3018 :	10.62	3301 :	20.33
2707 :	10.57	3019 :	15.42	3302 :	7.03
2708 :	25.32	3020 :	5.28	3303 :	35.05
2709 :	12.91	3021 :	11.23	3304 :	28.45
2719 :	14.15	3029 :	10.57	3305 :	16.25
2720 :	1.95	3041 :	9.86	3306 :	13.27
2801 :	163.17	3042 :	11.2	3307 :	2.57
2802 :	73.17	3043 :	11.73	3308 :	4.73
2803 :	12.7	3049 :	3.3	3309 :	1.84
2804 :	11.95	3050 :	2.11	3310 :	2.34
2805 :	1.73	3061 :	6.43	3311 :	21.01
2809 :	0.89	3062 :	7.69	3312 :	5.71
2901 :	3.24	3101 :	76.31	3313 :	7.42
2902 :	16.17	3102 :	44.2	3314 :	6.89
2903 :	70.49	3103 :	7.45	3315 :	28.66
2904 :	1.17	3104 :	16.1	3316 :	6.15
2905 :	0.12	3105 :	1.64	3317 :	1.47
2909 :	6.89	3106 :	0.56	3318 :	5.1
2910 :	0.44	3107 :	0.96	3319 :	85.08
3001 :	109.21	3109 :	0.61	3320 :	212.48
3002 :	87.95	3201 :	7.53	3321 :	3.04
3003 :	63.23	3202 :	10.72	3322 :	2.52
3004 :	60.3	3203 :	9.42	3323 :	1.56
3005 :	44.17	3204 :	6.84	3324 :	26.42
3006 :	48.11	3205 :	0.59	3325 :	52.35
3007 :	90.83	3206 :	1.16	3326 :	19.81
3008 :	9.05	3207 :	0.69	3327 :	4.79
3009 :	11.96	3208 :	17.62	3329 :	28.61
3010 :	20.63	3209 :	1.4	3339 :	53.89
3011 :	12.58	3210 :	2.97	3340 :	1.16
3012 :	15.53	3211 :	17.48	3401 :	382.94

3402 : 70.55
3403 : 139.31
3404 : 487.33
3405 : 1607.8
3406 : 66.24
3409 : 238.23
3411 : 8.95
3412 : 1.43
3501 : 400.99
3502 : 158.73
3503 : 14.07
3504 : 87.17
3505 : 711.57
3506 : 253.2
3507 : 24.78
3509 : 90.26
3510 : 20.45

7.5. Household expenditure by subgroup

✓ Generated the variable of subgroup sg in hhexp.byitem

```
> hhexp.byitem["sg"]<-floor(hhexp.byitem$code/100)
```

Monthly expenditure per household on food by subgroup

```
> d<-subset(hhexp.byitem,code<1900)
> t<-addmargins(tapply(d$m.value,d$sg,sum))[c(19,1:18)]
> names(t)<-c("All food items","Cereals","Prepared food","Pulses","Vegitables",
+ "Yams and other","Meat","Fish","Dried fish","Eggs","Coconuts","Condiments",
+ "Other foods","Milk and milk foods","Fats and oils","Sugar, juggery & treacle",
+ "Fruits","Confectionery & short eats","Beverages")
> for(j in 1:19) cat(format(names(t)[j],width=28),format(round(t[j]),width=6),"¥n")
All food items               13282
Cereals                       2669
Prepared food                 1424
Pulses                         547
Vegitables                    1006
Yams and other                 277
Meat                           517
Fish                          1163
Dried fish                     492
Eggs                           134
Coconuts                       738
Condiments                    1209
Other foods                     26
Milk and milk foods           1038
Fats and oils                  324
Sugar, juggery & treacle       466
Fruits                         386
Confectionery & short eats     471
Beverages                      394
```

Monthly expenditure per household on non-food by subgroup

```
> d<-subset(hhexp.byitem,code>=1900)
> t<-addmargins(tapply(d$m.value,d$sg,sum))[c(18,1:17)]
```

```
> names(t)<-c("Total non-food","Liquor,Drugs & Tabacco","Housing",
+ "Fuel & Light","Personal care","Health","Transport","Communication","Education",
+ "Cultural activities & Entertainment","Non-durable household goods",
+ "Household services","Clothing & Textile","Footwear","Durable household goods",
+ "Long durable household goods","Other expenses","Other adhoc expenses")
> for(j in 1:18) cat(format(names(t)[j],width=35),format(round(t[j]),width=6),"¥n")
Total non-food                       18153
Liquor,Drugs & Tabacco                 665
Housing                               3456
Fuel & Light                          1278
Personal care                          424
Health                                1008
Transport                             2317
Communication                          759
Education                             1018
Cultural activities & Entertainment    404
Non-durable household goods            264
Household services                      99
Clothing & Textile                     757
Footwear                               148
Durable household goods                 86
Long durable household goods           706
Other expenses                        3003
Other adhoc expenses                  1761
```

7.6. Creating data set on household expenditure by subgroup at household level

✓ Dropped the next records with out-of-range codes in R04 and R05.

File	Code	monthly value	Number of records	ID
R04	0440	0	1	512001103205111
R05	2112	0.01	1	712000406910111
	2506	0.11	2	132001207702111
	2510	0	1	612001205701111

```
> R04.old<-R04
> R04<-subset(R04,!(ID=="512001103205111"&CODE==440))
> dim(R04)
[1] 798045     19
> R05.old<-R05
> R05<-subset(R05,!(ID=="712000406910111"&NF_CODE==2112))
> R05<-subset(R05,!(ID=="132001207702111"&NF_CODE==2506))
> R05<-subset(R05,!(ID=="612001205701111"&NF_CODE==2510))
> dim(R05)
[1] 568661     19
```

R04

✓ Aggregated the variable of m.value, monthly household expenditure by ID (household) and sg, and created household-level data frame hhexp1 with variables of total monthly expenditure by subgroup.

```
> t<-tapply(R04$m.value,list(R04$ID,R04$sg),sum)
> dim(t)
[1] 19958    19
> t[is.na(t)]<-0
> colnames(t)<-paste("g",1:19,sep="")
> hhexp1<-data.frame(ID=rownames(t),t,row.names=NULL)
> head(hhexp1)
               ID       g1         g2        g3        g4       g5        g6       g7
1 111000114801111 3300.000 1080.0000  385.7143  621.4286 342.8571  942.8571 2850.000
```

```
2 111000114802111 2558.571 2931.4286  685.7143  672.8571 248.5714    0.0000 3000.000
3 111000114803111 4478.571 2121.4286  964.2857 1371.4286 642.8571 1457.1429 2614.286
4 111000114804111 1821.429  698.5714  441.4286  557.1429 231.4286    0.0000 1200.000
5 111000114805111 2708.571  912.8571 2172.8571 1255.7143 300.0000    0.0000 3985.714
6 111000114807111 2091.429  578.5714  120.0000  754.2857 107.1429    0.0000 1200.000
         g8        g9       g10       g11 g12       g13       g14       g15       g16
1  514.2857    0.0000 750.0000 4928.571   0 2657.143 771.4286 557.1429  107.1429
2  424.2857    0.0000 642.8571 4384.286   0 3021.429 621.4286 535.7143  857.1429
3 1234.2857    0.0000 750.0000 2905.714   0 1114.286 642.8571 471.4286  342.8571
4  488.5714  167.1429 535.7143 2442.857   0 1114.286 600.0000 385.7143  278.5714
5  385.7143    0.0000 642.8571 3214.286   0 1692.857 685.7143 385.7143 1842.8571
6  300.0000    0.0000 642.8571 3278.571   0 1328.571 685.7143 394.2857  544.2857
        g17       g18  g19
1    0.0000 900.0000 2700
2 1054.2857 428.5714    0
3 1367.1429 428.5714    0
4  407.1429 257.1429    0
5  300.0000 535.7143    0
6  655.7143 407.1429    0

> hies2009[[04]]<-R04
```

R05

✓ Aggregated the variable of m.value, monthly household expenditure by ID (household) and sg, and created household-level data frame hhexp2 with variables of total monthly expenditure by subgroup.

```
> t<-tapply(R05$m.value,list(R05$ID,R05$sg),sum)
> dim(t)
[1] 19958    16
> t[is.na(t)]<-0
> colnames(t)<-paste("g",20:35,sep="")
> hhexp2<-data.frame(ID=rownames(t),t,row.names=NULL)
> head(hhexp2)
               ID   g20  g21  g22  g23  g24  g25  g26 g27 g28 g29       g30       g31
```

```
1 111000114801111  6558 1387 1540 1200 1000 1000    0 200 510   0 1891.6667 158.33333
2 111000114802111  4620 2220 1065    0 2700 1750 1130   0 480   0 4607.5000 891.66667
3 111000114803111  4238 2275  730  650 1500  950  560   0 480   0  756.6667 233.33333
4 111000114804111  2650  996  640    0  360    0    0   0 310   0  479.1667  63.33333
5 111000114805111 13100 2254 1060    0 1200 1897    0 200 400   0  928.3333 416.66667
6 111000114807111  3650  976  810    0  650  500  260   0 370   0 1595.8333 498.33333
       g32   g33  g34       g35
1 31.66667   0.0    0 333.33333
2 35.00000   0.0    0  83.33333
3 15.00000   0.0  960  41.66667
4 10.83333   0.0  320  41.66667
5 15.00000 812.5 1500 125.00000
6 15.83333   0.0  750  20.83333

> hies2009[[05]]<-R05
```

R06

✓ Collapsed boarder at individual level to household level and created data frame hhexp3.

```
> hhexp3<-data.frame()
> bd.ID<-unique(boarder$ID)
> hhexp3<-bd.ID[order(bd.ID)]
> for(j in 3:15){
+ t<-as.vector(tapply(boarder[,j+9],boarder$ID,sum))
+ hhexp3<-data.frame(hhexp3,t)
+ colnames(hhexp3)[ncol(hhexp3)]<-paste("COL_",j,sep="")
+ }
> colnames(hhexp3)[1]<- "ID"
> dim(hhexp3)
[1] 187  14
> head(hhexp3)
               ID COL_3 COL_4    COL_5 COL_6 COL_7 COL_8 COL_9 COL_10 COL_11
1 111000117002111     0     0 166.6667     0     0     0     0      0      0
2 111000117004111     0     0   0.0000     0     0   300     0      0      0
3 111000117306111     0     0   0.0000   300     0     0     0      0      0
```

```
4 111000117307111      0      0 633.3333      0      0    150      0      0      0
5 111000117405111      0      0    0.0000      0      0      0      0      0      0
6 111000118002111      0      0    0.0000      0      0      0      0      0      0
  COL_12 COL_13 COL_14 COL_15
1      0   5000      0      0
2      0      0   2000      0
3      0   3000      0      0
4      0   3000      0      0
5      0      0      0      0
6      0      0      0      0
```

Merged R04, R05 and R06 at household level

✓ Merged hhexp1, hhexp2 and hhexp3 by using ID

```
> hhexp<-merge(hhexp1,hhexp2)
> dim(hhexp)
[1] 19958    36
> hhexp<-merge(hhexp,hhexp3,all.x=TRUE)
> dim(hhexp)
[1] 19958    49
```

✓ Replaced NA with 0

```
> hhexp[is.na(hhexp)]<-0
> hhexp.old<-hhexp
```

✓ Added COL_3 to COL_15 to the respective subgroups.

```
> hhexp$g2<-hhexp$g2+hhexp$COL_3
> hhexp$g21<-hhexp$g21+hhexp$COL_4
> hhexp$g30<-hhexp$g30+hhexp$COL_5
> hhexp$g32<-hhexp$g32+hhexp$COL_6
> hhexp$g29<-hhexp$g29+hhexp$COL_7
> hhexp$g22<-hhexp$g22+hhexp$COL_8
> hhexp$g25<-hhexp$g25+hhexp$COL_9
> hhexp$g27<-hhexp$g27+hhexp$COL_10
> hhexp$g33<-hhexp$g33+hhexp$COL_11
> hhexp$g20<-hhexp$g20+hhexp$COL_12
```

```
> hhexp$g35<-hhexp$g35+hhexp$COL_13
> hhexp$g34<-hhexp$g34+hhexp$COL_14
> hhexp$g34<-hhexp$g34+hhexp$COL_15
```

✓ Appended weight to hhexp

```
> hhexp<-merge(hhexp,R23[c("ID","WT")])
> colnames(hhexp)
[1] "ID"     "g1"     "g2"     "g3"     "g4"     "g5"     "g6"     "g7"     "g8"
[10] "g9"     "g10"    "g11"    "g12"    "g13"    "g14"    "g15"    "g16"    "g17"
[19] "g18"    "g19"    "g20"    "g21"    "g22"    "g23"    "g24"    "g25"    "g26"
[28] "g27"    "g28"    "g29"    "g30"    "g31"    "g32"    "g33"    "g34"    "g35"
[37] "COL_3"  "COL_4"  "COL_5"  "COL_6"  "COL_7"  "COL_8"  "COL_9"  "COL_10" "COL_11"
[46] "COL_12" "COL_13" "COL_14" "COL_15" "WT"
```

✓ Generated the variable of monthly total expenditure

```
> hhexp["ttexp"]<-rowSums(hhexp[,2:36])
> hhexp["food"]<-rowSums(hhexp[,c(2:19)])
> hhexp["nfood"]<-rowSums(hhexp[,c(20:36)])

# Average monthly household expenditure
> sum(hhexp$ttexp*hhexp$WT)/sum(hhexp$WT)
[1] 31435.24
> sum(hhexp$food*hhexp$WT)/sum(hhexp$WT)
[1] 13282.47
> sum(hhexp$nfood*hhexp$WT)/sum(hhexp$WT)
[1] 18152.77
```

Summary:

	My estimates	Final report of the survey	
Total monthly household expenditure	31,435	31,331	
Food	13,282	13,267	
Non-food	18,153	18,064	

8. Resampling method

Strategies for 80% resampling

1) **To create R25 from hhinc and hhexp as well as R26 from pinc as the next.**
2) **To order hhinc by strata (district and sector), psu and the amount of total household income ttinc, and append the variable sn, the serial number of records to hhinc.**
3) **To select 80% of hhinc, by dropping 20% households using systematic selection with the interval=5 and a random start number, and**
   to create the vector ID.res, a set of ID selected for resampled data set.
4) **For each data set, to select records with ID.res.**
5) **To adjust the weight as the next;**
   Adjusted weight = original weight / 0.8
6) **For each data set, to replace WT with the adjusted weight, and rename R data frame "xxx" as "xxx.80".**
7) **Export R data frames to CSV files with the name "xxx_80.CSV".**

☐ **Data dictionary of R25:**

Summary of household income and expenditure at household level

Number of records: 19,958

Variable	Description	Remarks
ID	Household identifier	
WT	Weight	
ttinc	Monthly total household income	= monetary + inkind
monetary	Monthly total monetary income	= wage + crop + livestock + nonagri + other + widfall
wage	Monthly income from paid employments	
crop	Monthly income from agricultural activities	
livestock	Monthly income from other agricultural activities	
nonagri	Monthly income from non-agricultural activities	
other	Monthly income from other cash receipt	
windfall	Monthly income by chance or ad hoc gains	
inkind	Monthly total non-monetary income	
ttexp	Monthly total household expenditure	= food + nfood
food	Monthly household expenditure on food	= Sum of g1 to g18
nfood	Monthly household expenditure on non-food	= Sum of g19 to g35

g1 to g35	Monthly household expenditure by subgroup	COL_3 to COL_15 are already added to the respective subgroups
COL_3 to COL_15	Monthly household expenditure for boarders	

☐ **Data dictionary of R26:**

Summary of household income at individual level

Number of records: 80,872

Variable	Description	Remarks
ID	Household identifier	
PID	Individual identifier	
WT	Weight	
receiver	Is an income receiver?	1. Yes 0. No
relation	Relationship to the household head	
age10	Is age over 10 years old?	1. age >= 10 0. age < 10
ttinc	Monthly total income after transferred	
transferred	Is income transferred to the head?	1. Yes 0. No
befinc	Monthly total income before transferred	= wage + crop + livestock + nonagri + other + widfall
wage	Monthly income from paid employments	
crop	Monthly income from agricultural activities	
livestock	Monthly income from other agricultural activities	
nonagri	Monthly income from non-agricultural activities	
other	Monthly income from other cash receipt	
windfall	Monthly income by chance or ad hoc gains	

****

== R25 ==

```
> hhinc<-hhinc.old2
> colnames(hhinc)
 [1] "ID"          "WT"          "wage"        "crop"        "livestock"   "nonagri"
```

```
 [7] "other"       "windfall"    "monetary.tt" "rice.a"      "rice.b"      "inkind.rice"
[13] "inkind04"    "inkind05"    "inkind.tt"   "ttinc"
> R25<-hhinc[,c(1,2,16,9,3:8,15)]
> colnames(R25)[4]<-"monetary"
> colnames(R25)[11]<-"inkind"
> head(R25)
               ID      WT    ttinc monetary     wage crop livestock nonagri other windfall
1 111000114801111 30.4971 66850.00 60000.00     0.00    0         0   60000     0        0
2 111000114802111 30.4971 84500.00 80000.00     0.00    0         0   80000     0        0
3 111000114803111 30.4971 67458.33 63458.33 30633.33    0         0       0 32825        0
4 111000114804111 30.4971 35393.33 32893.33 32893.33    0         0       0     0        0
5 111000114805111 30.4971 64875.00 52375.00 46250.00    0         0       0  6125        0
6 111000114807111 30.4971 33893.33 30393.33 30393.33    0         0       0     0        0
  inkind
1   6850
2   4500
3   4000
4   2500
5  12500
6   3500
> colnames(hhexp)
 [1] "ID"     "g1"     "g2"     "g3"     "g4"     "g5"     "g6"     "g7"     "g8"
[10] "g9"     "g10"    "g11"    "g12"    "g13"    "g14"    "g15"    "g16"    "g17"
[19] "g18"    "g19"    "g20"    "g21"    "g22"    "g23"    "g24"    "g25"    "g26"
[28] "g27"    "g28"    "g29"    "g30"    "g31"    "g32"    "g33"    "g34"    "g35"
[37] "COL_3"  "COL_4"  "COL_5"  "COL_6"  "COL_7"  "COL_8"  "COL_9"  "COL_10" "COL_11"
[46] "COL_12" "COL_13" "COL_14" "COL_15" "WT"     "ttexp"  "food"   "nfood"
> R25.old<-R25
> R25<-merge(R25,hhexp[,c(1,51:53,2:49)],by="ID")
> dim(R25)
[1] 19958    62
> colnames(R25)
 [1] "ID"        "WT"        "ttinc"     "monetary"  "wage"      "crop"      "livestock"
 [8] "nonagri"   "other"     "windfall"  "inkind"    "ttexp"     "food"      "nfood"
[15] "g1"        "g2"        "g3"        "g4"        "g5"        "g6"        "g7"
[22] "g8"        "g9"        "g10"       "g11"       "g12"       "g13"       "g14"
```

```
[29] "g15"       "g16"       "g17"       "g18"       "g19"       "g20"       "g21"
[36] "g22"       "g23"       "g24"       "g25"       "g26"       "g27"       "g28"
[43] "g29"       "g30"       "g31"       "g32"       "g33"       "g34"       "g35"
[50] "COL_3"     "COL_4"     "COL_5"     "COL_6"     "COL_7"     "COL_8"     "COL_9"
[57] "COL_10"    "COL_11"    "COL_12"    "COL_13"    "COL_14"    "COL_15"
> R25.old2<-R25


== R26 ==
> pinc<-pinc.old4
> colnames(pinc)
 [1] "PID"          "ID"           "WT"           "age"          "relation"
 [6] "wage"         "crop"         "livestock"    "nonagri"      "other"
[11] "windfall"     "ttinc"        "age10"        "flag0"        "flag1"
[16] "ttinc.before" "flag2"        "flag3"        "flag4"
> R26<-pinc[,c(1:3,5,13,12,16,6:11,15,17,18)]
> colnames(R26)
 [1] "PID"          "ID"           "WT"           "relation"     "age10"
 [6] "ttinc"        "ttinc.before" "wage"         "crop"         "livestock"
[11] "nonagri"      "other"        "windfall"     "flag1"        "flag2"
[16] "flag3"
> colnames(R26)[7]<-"befinc"
> R26["receiver"]<-ifelse(R26$flag2==1|R26$flag3==1,1,0)
> R26["transferred"]<-R26$flag1
> colnames(R26)
 [1] "PID"         "ID"          "WT"          "relation"    "age10"       "ttinc"
 [7] "befinc"      "wage"        "crop"        "livestock"   "nonagri"     "other"
[13] "windfall"    "flag1"       "flag2"       "flag3"       "receiver"    "transferred"
> R26.old<-R26
> R26<-R26[,c(1:3,17,4:6,18,7:13)]
> colnames(R26)
 [1] "PID"         "ID"          "WT"          "receiver"    "relation"    "age10"
 [7] "ttinc"       "transferred" "befinc"      "wage"        "crop"        "livestock"
[13] "nonagri"     "other"       "windfall"
> dim(R26)
[1] 80872    15
```

```
> table(R26$receiver)
    0     1
47255 33617
> table(R26$transferred)
    0     1
80635   237
> R26.old2<-R26

# Appended R25 and R26 to hies2009.
> hies2009<-hies2009.old5
> hies2009<-c(hies2009,list(R25,R26))
> Rnames<-c(Rnames,"R25","R26")
> hies2009.old6<-hies2009
```

✓ Ordered hhinc by strata (province and sector), psu and the amount of household income ttinc.

```
> d<-hhinc
> d["psuid"]<-substr(d$ID,1,10)
> d<-d[order(d$psuid,d$ttinc),c("ID","psuid","ttinc")]
> d["sn"]<-seq(1:nrow(d))
> head(d)
               ID      psuid    ttinc sn
9 111000114810111 1110001148 21700.00  1
7 111000114808111 1110001148 28796.67  2
6 111000114807111 1110001148 33893.33  3
4 111000114804111 1110001148 35393.33  4
5 111000114805111 1110001148 64875.00  5
1 111000114801111 1110001148 66850.00  6
```

✓ Selected 80% of hhinc, by dropping 20% households using systematic selection with the interval=5 and a random start number.

```
> int<-5
> (st<-sample(1:5,1))
[1] 5
> d["flag"]<-ifelse(d$sn%%int==(st-1),0,1)
```

```
> head(d)
               ID      psuid    ttinc sn flag
9 1110000114810111 1110001148 21700.00  1    1
7 1110000114808111 1110001148 28796.67  2    1
6 1110000114807111 1110001148 33893.33  3    1
4 1110000114804111 1110001148 35393.33  4    0
5 1110000114805111 1110001148 64875.00  5    1
1 1110000114801111 1110001148 66850.00  6    1
> table(d$flag)
    0     1
 3991 15967
> nrow(subset(d,flag==1))/nrow(d)
[1] 0.8000301
> ID.res<-subset(d,flag==1)$ID
> length(ID.res)
[1] 15967
```

- ✓ The number of resampled households is 15,967.
- ✓ For each data set, select records with ID.res and confirmed the resampling rate.

```
> hies2009res<-hies2009
> for(j in 1:26){
+ d<-hies2009[[j]]
+ d<-subset(d,is.element(d$ID,ID.res))
+ hies2009res[[j]]<-d
+ }

# The share of resampled households (%)
> for(j in 1:26)
cat(paste(Rnames[j],round(nrow(hies2009res[[j]])/nrow(hies2009[[j]])*100,1)),"¥n")
R01 80
R02 79.9
R03 80
R04 80
R05 80
R06 76.7
```

```
R07 80
R08 79.7
R09 80
R10 80.5
R11 80
R12 80
R13 80.2
R14 80
R15 81.1
R16 80
R17 79.4
R18 80
R19 79.9
R20 80
R21 80
R22 80
R23 80
R24 80
R25 80
R26 80
```

- ✓ Replaced the original weight WT with the adjusted weight wt, and renamed R data frame "xxx" as "xxx.80".

```
> hies2009res.old<-hies2009res
> for(j in 1:26){
+ d<-hies2009res[[j]]
+ d["wt"]<-d$WT/0.8
+ d<-d[,colnames(d)!="WT"]
+ hies2009res[[j]]<-d
+ }
> hies2009res.old2<-hies2009res

> Rnames.res<-paste(Rnames,".80",sep="")
> Rnames.res
 [1] "R01.80" "R02.80" "R03.80" "R04.80" "R05.80" "R06.80" "R07.80" "R08.80" "R09.80"
[10] "R10.80" "R11.80" "R12.80" "R13.80" "R14.80" "R15.80" "R16.80" "R17.80" "R18.80"
```

```
[19] "R19.80" "R20.80" "R21.80" "R22.80" "R23.80" "R24.80" "R25.80" "R26.80"
```

✓ Exported the resampled R data frames to CSV files with the name "xxx_80.CSV".

```
> CSVnames.res<-paste(Rnames,"_80.csv",sep="")
> head(CSVnames.res)
[1] "R01_80.csv" "R02_80.csv" "R03_80.csv" "R04_80.csv" "R05_80.csv" "R06_80.csv"
> for(j in 1:26){
+ cmd<-paste("write.csv(hies2009res[[",j,"]],'",CSVnames.res[j],"',row.names=F)",sep="")
+ eval(parse(text=cmd))
+ }
> list.files()
 [1] "R01_80.csv" "R02_80.csv" "R03_80.csv" "R04_80.csv" "R05_80.csv" "R06_80.csv"
 [7] "R07_80.csv" "R08_80.csv" "R09_80.csv" "R10_80.csv" "R11_80.csv" "R12_80.csv"
[13] "R13_80.csv" "R14_80.csv" "R15_80.csv" "R16_80.csv" "R17_80.csv" "R18_80.csv"
[19] "R19_80.csv" "R20_80.csv" "R21_80.csv" "R22_80.csv" "R23_80.csv" "R24_80.csv"
[25] "R25_80.csv" "R26_80.csv"

# Number of records and variables in resampled data files
> for(j in 1:26){
+ cat(format(Rnames.res[j],width=10),":",format(CSVnames.res[j],width=13),":",
+ format(nrow(hies2009res[[j]]),width=6),":",format(ncol(hies2009res[[j]]),width=3),"¥n")
+ }
R01.80     : R01_80.csv    :  64679 :  28
R02.80     : R02_80.csv    :  16668 :  23
R03.80     : R03_80.csv    :  64674 :  25
R04.80     : R04_80.csv    : 638090 :  16
R05.80     : R05_80.csv    : 454963 :  16
R06.80     : R06_80.csv    :    184 :  27
R07.80     : R07_80.csv    :  15967 :  13
R08.80     : R08_80.csv    :  14635 :  18
R09.80     : R09_80.csv    :  15967 :  13
R10.80     : R10_80.csv    :   3443 :  21
R11.80     : R11_80.csv    :  15967 :  13
R12.80     : R12_80.csv    :  15967 :  13
R13.80     : R13_80.csv    :   3362 :  20
```

```
R14.80     : R14_80.csv     :  15967 :  13
R15.80     : R15_80.csv     :   4479 :  17
R16.80     : R16_80.csv     :  15967 :  13
R17.80     : R17_80.csv     :   8523 :  22
R18.80     : R18_80.csv     :  15967 :  13
R19.80     : R19_80.csv     :   6514 :  21
R20.80     : R20_80.csv     :  15967 :  28
R21.80     : R21_80.csv     :  15967 :  35
R22.80     : R22_80.csv     :  15967 :  47
R23.80     : R23_80.csv     :  15967 :  39
R24.80     : R24_80.csv     :  15967 :  40
R25.80     : R25_80.csv     :  15967 :  62
R26.80     : R26_80.csv     :  64679 :  15
```

****

Revision of resampled micro data regarding errors found in R15

****

The data files of R15.80, R25.80 and R26.80 were revised as follows.

Revision of R15.80, Income from non-agricultural activities

PID	OUTPUT(old)	OUTPUT(rev)	INPUT(old)	INPUT(rev)
12100090380711102	1,680,000	**168,000**	150,000	150,000
12100051880111101	3,450,000	**345,000**	300,000	300,000
61200120720111104	3,500,000	**350,000**	35,000	35,000
61200030930611101	180,000,000	**15,000,000**	143,680,000	**12,680,000**
91300010540311101	60,000,000	**5,000,000**	50,000,000	**4,000,000**

```
> d<-R15.80
> fix(d)
> d[d$PID=="12100090380711102","OUTPUT_5_4"]<-168000
> d[d$PID=="12100051880111101","OUTPUT_5_4"]<-345000
> d[d$PID=="61200120720111104","OUTPUT_5_4"]<-350000
> d[d$PID=="61200030930611101","OUTPUT_5_4"]<-15000000
> d[d$PID=="61200030930611101","INPUT_5_4"]<- 12680000
> d[d$PID=="91300010540311101","OUTPUT_5_4"]<-5000000
> d[d$PID=="91300010540311101","INPUT_5_4"]<- 4000000
```

```
> R15.80.old<-R15.80
> R15.80<-d
```

Revision of R26.80, summary of household income at individual level

PID	nonagri(old)	nonagri(rev)	befinc(old) ttinc(old)	befinc(rev) ttinc(rev)
12100090380711102	1,530,000	**18,000**	1,530,000	**18,000**
12100051880111101	3,150,000	**50,000**	3,166,667	**66,667**
61200120720111104	3,465,000	**315,000**	3,545,863	**395,863**
61200030930611101	36,320,000	**2,320,000**	36,679,000	**2,679,000**
91300010540311101	1e+07	**1,000,000**	10,630,000	**1,630,000**

```
> d<-R26.80
> fix(d)
> d[d$PID=="12100090380711102","nonagri"]<-18000
> d[d$PID=="12100051880111101","nonagri"]<-50000
> d[d$PID=="61200120720111104","nonagri"]<-315000
> d[d$PID=="61200030930611101","nonagri"]<-2320000
> d[d$PID=="91300010540311101","nonagri"]<-1000000
> d[d$PID=="12100090380711102","befinc"]<-18000
> d[d$PID=="12100051880111101","befinc"]<-66667
> d[d$PID=="61200120720111104","befinc"]<-395863
> d[d$PID=="61200030930611101","befinc"]<-2679000
> d[d$PID=="91300010540311101","befinc"]<-1630000
> d[d$PID=="12100090380711102","ttinc"]<-18000
> d[d$PID=="12100051880111101","ttinc"]<-66667
> d[d$PID=="61200120720111104","ttinc"]<-395863
> d[d$PID=="61200030930611101","ttinc"]<-2679000
> d[d$PID=="91300010540311101","ttinc"]<-1630000
> fix(d)
> R26.80.old<-R26.80
> R26.80<-d
```

Revision of R25.80, summary of household income and expenditure at household level

ID	ttinc(old)	ttinc(rev)	monetary(old)	monetary(rev)	nonagri(old)	nonagri(rev)
121000903807111	1,600,418	**88,418**	1,595,418	**83,418**	1,560,000	**48,000**
121000518801111	3,171,667	**71,667**	3,166,667	**66,667**	3,150,000	**50,000**
612001207201111	3,570,198	**420,198**	3,555,863	**405,863**	3,475,000	**325,000**
612000309306111	36,731,136	**2,731,136**	36,707,000	**2,707,000**	36,320,000	**2,320,000**
913000105403111	10,693,837	**1,693,837**	10,680,000	**1,680,000**	1e+07	**1,000,000**

```
> d<-R25.80
> d[d$ID=="121000903807111","nonagri"]<-48000
> d[d$ID=="121000518801111","nonagri"]<-50000
> d[d$ID=="612001207201111","nonagri"]<-325000
> d[d$ID=="612000309306111","nonagri"]<-2320000
> d[d$ID=="913000105403111","nonagri"]<-1000000
> d[d$ID=="121000903807111","monetary"]<-83418
> d[d$ID=="121000518801111","monetary"]<-66667
> d[d$ID=="612001207201111","monetary"]<-405863
> d[d$ID=="612000309306111","monetary"]<-2707000
> d[d$ID=="913000105403111","monetary"]<-1680000
> d[d$ID=="121000903807111","ttinc"]<-88418
> d[d$ID=="121000518801111","ttinc"]<-71667
> d[d$ID=="612001207201111","ttinc"]<-420198
> d[d$ID=="612000309306111","ttinc"]<-2731136
> d[d$ID=="913000105403111","ttinc"]<-1693837
> fix(d)
> R25.80.old<-R25.80
> R25.80<-d
> hies2009res.old<-hies2009res
> hies2009res[[15]]<-R15.80
> hies2009res[[25]]<-R25.80
> hies2009res[[26]]<-R26.80

# Exported CSV files
> write.csv(R15.80,"R15_80.csv",row.names=F)
> write.csv(R25.80,"R25_80.csv",row.names=F)
> write.csv(R26.80,"R26_80.csv",row.names=F)
```

9. R Scripts for some of the Additional Tables of the Final Report

Table A1: Household population by sex and sector

```
> t1<-addmargins(tapply(RO1$WT,list(RO1$SECTOR,RO1$SEX_LIVING),sum))
> colnames(t1)<-c("Male","Female","Total")
> rownames(t1)<-c("Urban","Rural","Estate","Sri Lanka")
> round(t1[c(4,1:3),c(3,1:2)]/1000,1)
            Total   Male  Female
Sri Lanka 20337.8 9645.8 10692.0
Urban      3021.9 1440.1  1581.8
Rural     16289.7 7717.6  8572.1
Estate     1026.1  488.0   538.1
```

Table A1: Household population by sex and province

```
> province<-floor(RO1$DISTRICT/10)
> table(province)
province
    1     2     3     4     5     6     7     8     9
22748  8812 14332  2894  8737  7090  4887  4885  6487
> prov.name<-c("Western","Central","Southern","Northern",
+ "Eastern","North-western","North-central","Uva","Sabaragamuwa")
> t2<-addmargins(tapply(RO1$WT,list(province,RO1$SEX_LIVING),sum))
> colnames(t2)<-c("Male","Female","Total")
> rownames(t2)<-c(prov.name,"Sri Lanka")
> round(t2[c(10,1:9),c(3,1:2)]/1000,1)
                Total   Male  Female
Sri Lanka     20337.8 9645.8 10692.0
Western        5983.2 2823.4  3159.9
Central        2641.7 1222.4  1419.3
Southern       2525.3 1193.5  1331.8
Northern        715.6  338.2   377.4
Eastern        1499.5  729.1   770.4
North-western  2419.8 1160.4  1259.4
```

```
North-central  1237.6  593.5   644.1
Uva            1330.2  630.8   699.4
```

Table A1: Household population by sex and district

```
> dist.name<-c("Colombo","Gampaha","Kalutara","Kandy","Matale","Nuwara eliya",
+ "Galle","Matara","Hambantota","Jaffna","Vavuniya","Batticaloa","Ampara",
+ "Trincomalee","Kurunegala","Puttalama","Anuradhapura","Polonnaruwa",
+ "Badulla","Moneragala","Ratnapura","Kegalle")
> length(dist.name)
[1] 22
> t3<-addmargins(tapply(RO1$WT,list(RO1$DISTRICT,RO1$SEX_LIVING),sum))
> colnames(t3)<-c("Male","Female","Total")
> rownames(t3)<-c(dist.name,"Sri Lanka")
> round(t3[c(23,1:22),c(3,1:2)]/1000,1)
               Total   Male  Female
Sri Lanka    20337.8 9645.8 10692.0
Colombo       2404.5 1167.6  1236.9
Gampaha       2386.0 1108.9  1277.1
Kalutara      1192.8  546.9   645.9
Kandy         1386.9  630.5   756.4
Matale         485.3  229.0   256.3
Nuwara eliya   769.5  362.8   406.7
Galle         1098.8  508.6   590.2
Matara         834.8  387.4   447.4
Hambantota     591.7  297.5   294.2
Jaffna         545.8  256.2   289.6
Vavuniya       169.9   82.0    87.8
Batticaloa     538.3  259.3   279.0
Ampara         618.4  299.8   318.6
Trincomalee    342.8  170.0   172.8
Kurunegala    1585.7  763.1   822.6
Puttalama      834.1  397.3   436.8
Anuradhapura   824.7  399.1   425.6
Polonnaruwa    413.0  194.5   218.5
```

Badulla	853.9	396.9	457.1
Moneragala	476.3	233.9	242.4
Ratnapura	1138.2	552.8	585.4
Kegalle	846.5	401.6	444.9

Table A2: Household population of head of the household by sex and sector

```
> df<-subset(R01,RELATIONSHIP==1)
> dim(df)
[1] 19959    28
> t1<-addmargins(tapply(df$WT,list(df$SECTOR,df$SEX_LIVING),sum))
> colnames(t1)<-c("Male","Female","Total")
> rownames(t1)<-c("Urban","Rural","Estate","Sri Lanka")
> round(t1[c(4,1:3),c(3,1:2)]/1000,1)
           Total   Male Female
Sri Lanka 5079.8 3899.9 1179.9
Urban      711.2  517.6  193.5
Rural     4123.1 3197.8  925.3
Estate     245.5  184.4   61.1
```

Table A3: Percentage distribution of population by selected age groups and sector

Table A34: Percentage distribution of households by principal type of cooking fuel by sector

```
> t1<-tapply(R23$WT,list(R23$SECTOR,R23$COOKING_FUEL),sum)
> t1[is.na(t1)]<-0
> colnames(t1)<-c("Firewood","Gas","Kerosene","Electricity",
+ "Saw dust/Paddy husk","Other")
> rownames(t1)<-c("Urban","Rural","Estate")
> round(t1/1000,1)
       Firewood   Gas Kerosene Electricity Saw dust/Paddy husk Other
Urban     278.8 341.4     81.2         3.8                 1.6   3.9
Rural    3573.5 467.2     44.0         6.0                 4.2  27.1
Estate    238.8   3.7      1.4         0.5                 0.0   0.8
> t2<-cbind(t1[,1:3],Other=rowSums(t1[,4:6]))
> round(t2/1000,1)
       Firewood   Gas Kerosene Other
Urban     278.8 341.4     81.2   9.2
Rural    3573.5 467.2     44.0  37.3
Estate    238.8   3.7      1.4   1.3
> t3<-addmargins(t2)[c(4,1:3),c(5,1:4)]
> round(t3/t3[,1]*100,1)
       Sum Firewood  Gas Kerosene Other
Sum    100      80.6 16.0      2.5   0.9
Urban  100      39.2 48.0     11.4   1.3
Rural  100      86.7 11.3      1.1   0.9
Estate 100      97.3  1.5      0.6   0.5


> round(t1[c(4,1:3),c(7,1:6)]/1000,1)
```

10 Agricultural land

****

Remarks: Unit of area in yard-pound system (UK)

Unit of length and area for measuring land in yard-pound system (UK)

Unit	Definition	Remarks
rod/pole/ perch	1 rod = 5.5 yard = 16.5 feet = 5.0292 m	Typical length of a mediaeval ox-goad
chain	1 chain = 4 rods = 66 feet = 20.1168 m	
furlong	1 furlong = 10 chains = 40 rods = 660 feet = 201.168 m	Distance a team of oxen could plough without resting
mile	1 mile = 8 furlong	
square rod/ square perch/ perch	1 square rod = 30.25 square yards = 25.293 m^2	
rood	1 rood = 1 furlong x 1 rod = 40 square rods = 1,012 m^2	
acre	1 acre = 1 furlong x 1 chain = 4 roods = 160 square rods = 4,840 square yards = 4,046.85 m^2	Amount of land tillable by one man behind one ox in one day
hide/hyde	1 hide = 120 acre = 485.622 m^2	
square mile	1 square mile = 8 furlong x 80 chain = 640 acres = 2.59 km^2	

According to the delegates from Sri Lanka for the International Workshop,

“Metric area is not used for measuring land area in Sri Lanka.

Conversion of it in office is as the land area in perches is given by

Perches + Rood x 40 + Acres x 160 “

****

== R24: Agricultural holdings and livestock==

Out of 19,958 households, 16,524 responded that some household members own agricultural land, but it seems that only few reported the area owned and cultivated.

```
> addmargins(table(R24$IS_AGRILAND_OWNER,useNA="ifany"))
    1     2     3  <NA>   Sum
16524  3432     1     1 19958
```

Number of households reported non-NA for each question of area

```
> for(j in 12:29){
+ n<-sum(!is.na(R24[,j]))
+ cat("Number of non-NAs in",colnames(R24)[j],"= ",n,"¥n")
+ }
Number of non-NAs in PADDY_OWN_ACR =  1515
Number of non-NAs in PADDY_OWN_RT =  1438
Number of non-NAs in PADDY_OWN_PERCH =  208
Number of non-NAs in PADDY_OTHER_ACR =  1105
Number of non-NAs in PADDY_OTHER_RT =  1064
Number of non-NAs in PADDY_OTHER_PERCH =  132
Number of non-NAs in LAND_OWN_ACR =  1110
Number of non-NAs in LAND_OWN_RT =  1100
Number of non-NAs in LAND_OWN_PERCH =  828
Number of non-NAs in LAND_OTHER_ACR =  611
Number of non-NAs in LAND_OTHER_RT =  584
Number of non-NAs in LAND_OTHER_PERCH =  190
Number of non-NAs in HOME_OWN_ACR =  1947
Number of non-NAs in HOME_OWN_RT =  4448
Number of non-NAs in HOME_OWN_PERCH =  9921
Number of non-NAs in HOME_OTHER_ACR =  585
Number of non-NAs in HOME_OTHER_RT =  1220
Number of non-NAs in HOME_OTHER_PERCH =  1003
```

- ✓ Generated variables area1 to area6 in square meters and counted number of households with non-zero area, as follows;

	Owned		Cultivated	
	Variable	No. of non-zero	Variable	No. of non-zero
1. Paddy land	area1	2,548	area2	1,875
2. High land	area3	2,693	area4	1,215
3. Land area with occupied housing units	area5	15,249	area6	2,507

```
> R24.old<-R24
> R24[is.na(R24)]<-0
> variables<-c("area1","area2","area3","area4","area5","area6")
> for(j in 1:6){
+ R24[variables[j]]<-R24[,3*j+9]*4046.85+R24[,3*j+10]*1011.7125+R24[,3*j+11]*25.293
+ }
> nrow(subset(R24,area1>0))
[1] 2548
> nrow(subset(R24,area2>0))
[1] 1875
> nrow(subset(R24,area3>0))
[1] 2693
> nrow(subset(R24,area4>0))
[1] 1215
> nrow(subset(R24,area5>0))
[1] 15249
> nrow(subset(R24,area6>0))
[1] 2507
> nrow(subset(R24,area1+area2+area3+area4+area5+area6>0))
[1] 16521

# Average area per household (m2)
> round(colSums(R24[,41:46]*R24$WT)/sum(R24$WT),1)
 area1  area2  area3  area4  area5  area6
1307.9  699.8  815.6  424.2 1448.4  410.5
```

```
# Average area per household conditional on household with non-zero area
> for(j in 1:6){
+ d<-subset(R24,R24[,variables[j]]>0)
+ cat(variables[j],":")
+ cat(round(sum(d[,j+40]*d$WT)/sum(d$WT),1),"¥n")
+ }
area1 :8155.5
area2 :5836.2
area3 :5246.7
area4 :5687.2
area5 :1841.4
area6 :2723.7
```

✓ Average area per household (m2)

	Owned		Cultivated	
	Average area	Average area conditional on non-zero	Average area	Average area conditional on non-zero
1. Paddy land	1,307.9	8,155.5	699.8	5,836.2
2. High land	815.6	5,246.7	424.2	5,687.2
3. Land area with occupied housing units	1,448.4	1,841.4	410.5	2,723.7

Comparison between R10 and R24

```
== R10: Income from agricultural activities ==
> R10.old<-R10
> R10[is.na(R10)]<-0
```

✓ Generated the variable area in square meter.

```
> R10["area"]<-R10$ACR_5_2*4046.85+R10$RT_5_2*1011.7125+R10$P*25.293
```

✓ Generated R10p as a subset of R10 who have cultivated paddy as employers or own account workers.

```
> R10p<-subset(R10,SEAS_CROPS_CODE==1)
```

```
> dim(R10p)
[1] 2782   22
> table(R10p$area>0)
FALSE  TRUE
   38  2744

# Average cultivated area of paddy per household
> round(sum(R10p$area*R10p$WT)/sum(R24$WT),1)
[1] 1068.9
```

11. Employment

R01 and R08 include information on employment.

== R01 ==

✓ **Number of persons by main activity and employment status**

Main activity	Un-weighted number of persons	Weighted number of persons	Proportion	Proportion of employed
Persons aged 10 and over	66,668	16,763,675	100.0	
1 Employed	27,685	7,052,450	42.1	100.0
1 Government employee	3,410	837,349		11.9
2 Semi government employee	1,190	314,081		4.5
3 Private sector employee	13,435	3,260,912		46.2
4 Employer	540	129,328		1.8
5 Own account worker	7,558	2,042,996		29.0
6 Unpaid family worker	1,545	466,243		6.6
Unknown	7	234		0.0
2 Unemployed	5,155	1,222,140	7.3	
3 Student	12,340	3,091,978	18.4	
4 Household work	14,623	3,662,052	21.8	
5 Unable/ Too old to work	6,429	1,626,131	9.7	
9 Other	409	102,665	0.6	
Unknown	27	6,259	0.0	

```
> d1<-subset(R01,AGE>=10)
> nrow(d1)
[1] 66668
> addmargins(table(d1$MAIN_ACTIVITY,useNA="ifany"))
    1     2     3     4     5     6     9  <NA>   Sum
27685  5155 12340 14623  6429     1   409    26 66668
> t<-tapply(d1$WT,d1$MAIN_ACTIVITY,sum,na.rm=T)
> t<-c(sum(d1$WT),t)
```

```
> round(t)
                1         2         3         4         5         6         9
16763675  7052450  1222140  3091978  3662052  1626131       349    102316
> round(t/t[1]*100,1)
          1     2     3     4     5     6     9
100.0  42.1   7.3  18.4  21.8   9.7   0.0   0.6

> d2<-subset(d1,MAIN_ACTIVITY==1)
> nrow(d2)
[1] 27685
> addmargins(table(d2$EMPLOYMENT_STATUS,useNA="ifany"))
    1     2     3     4     5     6     9  <NA>   Sum
 3410  1190 13435   540  7558  1545     1     6 27685
> t<-tapply(d2$WT,d2$EMPLOYMENT_STATUS,sum,na.rm=T)
> t<-c(sum(d2$WT),t)
> round(t)
              1        2        3        4        5        6        9
7052450  837349  314081 3260912  129328 2042996  466243      234
> round(t/t[1]*100,1)
          1     2     3     4     5     6     9
100.0  11.9   4.5  46.2   1.8  29.0   6.6   0.0
```

✓ For employed, main occupation and main industry are recorded both in 4-digit code.

Variable	Description	Classification	Remarks
MAIN_OCCUPATION	Main occupation	ISCO	See Attachment 6
INDUSTRY	Main industry	ISIC	See Attachment 5

 # Number of employed persons by major occupation

```
> d2["occ"]<-floor(d2$MAIN_OCCUPATION/1000)
> t1<-table(d2$occ,useNA="ifany")
> t1
   0    1    2    3    4    5    6    7    8    9 <NA>
 129 1771 1862 2834 1056 2142 5448 4060 1943 6406   34
> t1<-c(sum(t1),t1[c(2:10,1,11)])
> occ.name<-c("Total","1 Manager, Senior Officials and Legislator",
```

```
+ "2 Professionals","3 Technicians and Associate Professionals",
+ "4 Clerks and Clerical Support Works","5 Services and Sales Workers",
+ "6 Skilled Agricultural, Forestry and Fishery Workers",
+ "7 Craft and Related Trade Workers","8 Plant and machine Operators and Assemblers",
+ "9 Elementary occupations","0 Other","Unknown")
> t2<-tapply(d2$WT,d2$occ,sum,na.rm=T)
> t2<-c(sum(d2$WT),t2[c(2:10,1)],sum(d2$WT)-sum(t2))
> df<-data.frame(Occupation=occ.name,Unweighted=as.vector(t1),
+ Weighted=round(t2),Prop=round(t2/t2[1]*100,1),row.names=NULL)
> df
```

	Occupation	Unweighted	Weighted	Prop
1	Total	27685	7052450	100.0
2	1 Manager, Senior Officials and Legislator	1771	441580	6.3
3	2 Professionals	1862	449305	6.4
4	3 Technicians and Associate Professionals	2834	685194	9.7
5	4 Clerks and Clerical Support Works	1056	276879	3.9
6	5 Services and Sales Workers	2142	518304	7.3
7	6 Skilled Agricultural, Forestry and Fishery Workers	5448	1556740	22.1
8	7 Craft and Related Trade Workers	4060	1067496	15.1
9	8 Plant and machine Operators and Assemblers	1943	516457	7.3
10	9 Elementary occupations	6406	1496287	21.2
11	0 Other	129	37644	0.5
12	Unknown	34	6565	0.1

 # Number of employed persons by major industry

```
> breaks<-c(111,710,1010,3510,3600,4100,4510,4910,5510,5811,6411,6810,6910,
+ 7710,8411,8510,8610,9000,9411,9700,9900,9999)
> d2["ind"]<-cut(d2$INDUSTRY,breaks,right=F,include.lowest=T,
+ labels=c("A","B","C","D","E","F","G","H","I","J","K","L","M","N","O","P",
+ "Q","R","S","T","U"))
> t1<-table(d2$ind,useNA="ifany")
> t1<-c(sum(t1),t1)
> t2<-tapply(d2$WT,d2$ind,sum,na.rm=T)
```

```
> round(t2)
> t2[is.na(t2)]<-0
> t2<-c(sum(d2$WT),t2,sum(d2$WT)-sum(t2))
> ind.name<-c("Total","A Agriculture, Forestry and Fishing","B Mining and Quarrying",
+ "C Manufacturing","D Electricity, gas, steam and air conditioning supply",
+ "E Water Supply and Sewerage",
+ "F Construction","G Wholesale and retail trade",
+ "H Transportation and storage","I Accommodation and food service activities",
+ "J Information and Communication","K Financial and Insurance Activities",
+ "L Real estate activities","M Professional, Scientific and technical activities",
+ "N Administrative and support service activities",
+ "O Public administration and defense",
+ "P Education","Q Human health and social work activities",
+ "R Arts, entertainment and recreation","S Other service activities",
+ "T Activities of households as employers for own use",
+ "U Activities of extraterritorial organizations and bodies","Unknown")
> df<-data.frame(Industry=ind.name,Unweighted=t1,
+ Weighted=round(t2),Prop=round(t2/t2[1]*100,1),row.names=NULL)
> df
```

	Industry	Unweighted	Weighted	Prop
1	Total	27685	7052450	100.0
2	A Agriculture, Forestry and Fishing	8123	2079547	29.5
3	B Mining and Quarrying	0	0	0.0
4	C Manufacturing	4252	1143615	16.2
5	D Electricity, gas, steam and air conditioning supply	21	4779	0.1
6	E Water Supply and Sewerage	420	109720	1.6
7	F Construction	23	7081	0.1
8	G Wholesale and retail trade	1646	440688	6.2
9	H Transportation and storage	3914	921569	13.1
10	I Accommodation and food service activities	537	117856	1.7
11	J Information and Communication	1539	412235	5.8
12	K Financial and Insurance Activities	585	149383	2.1
13	L Real estate activities	0	0	0.0
14	M Professional, Scientific and technical activities	2998	759347	10.8
15	N Administrative and support service activities	1033	246608	3.5
16	O Public administration and defense	0	0	0.0

17	P Education	461	106273	1.5
18	Q Human health and social work activities	0	0	0.0
19	R Arts, entertainment and recreation	466	111508	1.6
20	S Other service activities	1578	422718	6.0
21	T Activities of households as employers for own use	0	0	0.0
22	U Activities of extraterritorial organizations and bodies	46	11314	0.2
23	Unknown	43	8208	0.1

== R08 ==

The questionnaire of R08 is designed to record at most two occupation for paid employees, that is, main occupation and secondary occupation.

If the secondary occupation is recorded, the main occupation should be recorded.

```
# e1: subset of R08 with PRI_SEC=1
> e1<-subset(R08,PRI_SEC==1)
> dim(e1)
[1] 18159    18

# e2: subset of R08 with PRI_SEC=2
> e2<-subset(R08,PRI_SEC==2)
> dim(e2)
[1] 205  18
```

✓ While there are 205 records of PRI_SEC=2, the number of isolated records without the corresponding records of PRI_SEC=1 is 29.

```
> table(is.element(e2$PID,e1$PID))
FALSE  TRUE
   29   176
```

✓ **PRI_SEC=2 of the records with the isolated PID should be read as PRI_SEC=1.**

```
# Isolated PID
```

```
> PID.e<-e2[!is.element(e2$PID,e1$PID),"PID"]
> length(PID.e)
[1] 29
> head(PID.e)
[1] "13200100590411102" "23200070020111101" "33200090260211105"
[4] "33200110380411101" "33200110380611101" "33200110380611102"

> d3<-R08
> d3[is.element(d3$PID,PID.e),"PRI_SEC"]<-1
> e1<-subset(d3,PRI_SEC==1)
> dim(e1)
[1] 18188    18
> e2<-subset(d3,PRI_SEC==2)
> dim(e2)
[1] 176  18
```

✓ Created data frame emp by merging subset of d2 who are employee, e1 and e2 using PID.

```
> emp<-subset(d2,EMPLOYMENT_STATUS<=3)
> emp<-merge(emp,e1[,c(12:15,17)],by="PID",all.x=T)
> colnames(emp)[31:34]<-c("PRI_SEC1","WAGE1","ALLOWANCE1","BONUS1")
> dim(emp)
[1] 18035    34
> emp<-merge(emp,e2[,c(12:15,17)],by="PID",all.x=T)
> dim(emp)
[1] 18035    38
> colnames(emp)[35:38]<-c("PRI_SEC2","WAGE2","ALLOWANCE2","BONUS2")
> colnames(emp)
 [1] "PID"              "DISTRICT"         "PSU"              "REC_TYPE"
 [5] "SECTOR"           "DSD"              "MONTH"            "SAMPLE_N"
 [9] "SERIAL_NO"        "NHH"              "RESULT"           "PERSON_SERIAL_NO"
[13] "RELATIONSHIP"     "SEX_LIVING"       "BIRTH_YEAR"       "BIRTH_MONTH"
[17] "AGE"              "ETHNICITY"        "RELIGION"         "CURR_EDUCATION"
[21] "EDUCATION"        "MARITAL_STATUS"   "MAIN_ACTIVITY"    "MAIN_OCCUPATION"
[25] "INDUSTRY"         "EMPLOYMENT_STATUS" "ID"              "WT"
```

```
[29] "occ"              "ind"              "PRI_SEC1"         "WAGE1"
[33] "ALLOWANCE1"       "BONUS1"           "PRI_SEC2"         "WAGE2"
[37] "ALLOWANCE2"       "BONUS2"
```

```
# Generated monthly wage income from main job
> emp.old<-emp
> emp["WAGE1"]<-ifelse(is.na(emp$WAGE1),0,emp$WAGE1)
> emp["ALLOWANCE1"]<-ifelse(is.na(emp$ALLOWANCE1),0,emp$ALLOWANCE1)
> emp["BONUS1"]<-ifelse(is.na(emp$BONUS1),0,emp$BONUS1)
> emp["wage.inc1"]<-emp$WAGE1+emp$ALLOWANCE1+emp$BONUS1/12
> table(emp$wage.inc1==0)
FALSE  TRUE
17826   209
```

✓ Weighted average monthly wage income of main job by occupation of main job

```
> emp.old2<-emp
> emp<-subset(emp,wage.inc1>0)
> nrow(emp)
[1] 17826
> t<-by(emp,emp$occ,function(df) sum(df$wage.inc1*df$WT)/sum(df$WT))
> t<-as.vector(t)[c(2:10,1)]
> df<-data.frame(Occupation=occ.name[2:11],Wage=round(t))
> df
                                          Occupation  Wage
1            1 Manager, Senior Officials and Legislator 37079
2                                     2 Professionals 23024
3             3 Technicians and Associate Professionals 20994
4                   4 Clerks and Clerical Support Works 18405
5                          5 Services and Sales Workers 14548
6  6 Skilled Agricultural, Forestry and Fishery Workers  7134
7                     7 Craft and Related Trade Workers 11777
8          8 Plant and machine Operators and Assemblers 14298
9                             9 Elementary occupations  9123
10                                          0 Other 25958
```

✓ **Weighted average monthly wage income of main job by industry of main job**

```
> t<-by(emp,emp$ind,function(df) sum(df$wage.inc1*df$WT)/sum(df$WT))
> df<-data.frame(Industry=ind.name[2:22],Wage=round(as.vector(t)))
> df
                                                Industry  Wage
1                          A Agriculture, Forestry and Fishing  7929
2                                      B Mining and Quarrying    NA
3                                           C Manufacturing 12612
4       D Electricity, gas, steam and air conditioning supply 13379
5                                  E Water Supply and Sewerage 13862
6                                           F Construction 21740
7                                G Wholesale and retail trade 13236
8                                H Transportation and storage 13447
9                  I Accommodation and food service activities 12945
10                               J Information and Communication 16092
11                         K Financial and Insurance Activities 26578
12                                    L Real estate activities    NA
13        M Professional, Scientific and technical activities 21575
14            N Administrative and support service activities 19265
15                          O Public administration and defense    NA
16                                          P Education 19819
17                  Q Human health and social work activities    NA
18                       R Arts, entertainment and recreation 12973
19                                  S Other service activities  7258
20        T Activities of households as employers for own use    NA
21 U Activities of extraterritorial organizations and bodies 15401
```

12 Demography

12.1 Ethnicity

Brief outline of ethnic groups in Sri Lanka

Ethnic group	Description
Sinhala	The Sinhalese are an ethnic group native to the island of Sri Lanka. They constitute 75% of the Sri Lanka population. The Sinhalese identity is based on language, historic heritage and religion. The Sinhalese speaks Sinhara and are predominantly Buddhists, although a small percentage of Sinhalese follow branches of Christianity. The Sinhalese are mostly found in North central, Central, South, and West Sri lanka. According to legend Mahavamsa they are the descendants of the exiled Prince Vijaya who arrived from East India to Sri Lanka in 543 BCE.
Sri Lanka Tamil	Sri Lankan Tamils have a very long history in Sri Lanka and have lived on the island since around the 2nd century BCE. Most modern Sri Lankan Tamils claim descent from residents of Jaffna Kingdom, a former kingdom in the north of the island and Vannimai chieftaincies from the east. They constitute a majority in the Northern Province, live in significant numbers in the Eastern Province, and are in the minority throughout the rest of the country. 70% of Sri Lankan Tamils in Sri Lanka live in the Northern and Eastern provinces. Although Sri Lankan Tamils are culturally and linguistically distinct, genetic studies indicate that they are closely related to Sinhalese ethnic group in the island. The Sri Lankan Tamils are mostly Hindus with a significant Christian population. One-third of Sri Lankan Tamils now live outside Sri Lanka. While there was significant migration during the British colonial period, the civil war led to more than 800,000 Tamils leaving Sri Lanka, and many have left the country for destinations such as Canada, India and Europe as refugees.
Indian Tamil	Indian Tamils of Sri Lanka are Tamil people of Indian origin in Sri Lanka. They are also known as Hill Country Tamils, Up-Country Tamils or simply Indian Tamils. They are partly descended from workers sent from South India to Sri Lanka in the 19th and 20th centuries to work in coffee, tea and rubber plantations. Some also migrated on their own as merchants and as other service providers. These Tamil-speakers mostly live in the central highlands. They are instrumental in the plantation sector economy of Sri Lanka. In general socio-economically their standard of living is below that of the national average. In 1948, the United National

	Party government stripped the Indian Tamils of their citizenship. Under the terms of an agreement reached between the Sri Lankan and Indian governments in the 1960s, about forty percent of the Indian Tamils were granted Sri Lankan citizenship, and most of the remainder were repatriated to India. By the 1990s, most Indian Tamils had received Sri Lankan citizenship. Most of Indian Tamils are Hindus with a minority of Christians and Muslims amongst them.
Sri Lanka Moors	Sri Lankan Moors (commonly referred to as Muslims) are the third largest ethnic group in Sri Lanka. They are native speakers of the Tamil language and predominantly followers of Islam. While some sources describe them as a subset of the Tamil people who had adopted Islam as their religion and spoke Tamil as their mother tongue, which they continue to do so, other sources trace their ancestry to Arab traders (Moors) who settled in Sri Lanka sometime between the 8th and 15th centuries. The Arabic language brought by the early merchants is no longer spoken, though many Arabic words and phrases are still commonly used.
Malay	Malays of Sri Lanka originated in Southeast Asia. Their ancestors initially came to the country when both Sri Lanka and Indonesia were colonies of the Dutch, while a second wave (1796–1948) came from the Malay Peninsula, when both Malaya and Sri Lanka were in the British Empire. Most of the early immigrants were soldiers, posted by the Dutch which later continued by the British for colonial administration to Sri Lanka, who decided to settle on the island. Other immigrants were convicts or members of noble houses from Dutch East Indies (present day Indonesia) who were exiled to Sri Lanka and who never left. The main source of a continuing Malay identity is their common Malay language. Many Sri Lankan Malays have been celebrated as courageous soldiers, politicians, sportsmen, lawyers, accountants and doctors.
Burgher	The Burgher people are a Eurasian ethnic group, historically from Sri Lanka, consisting mainly of male-line descendants of European colonists from the 16th to 20th centuries (mostly Portuguese, Dutch, German and British) and local women, with some minorities of Swedish, Norwegian, French and Irish. Today the native language of the Burghers is English. Burghers were defined as those whose father was born in Sri Lanka, with at least one European ancestor on one's direct paternal side, regardless of the ethnic origin of one's mother, or what other ethnic groups may be found on the father's side. Because of this definition, Burghers almost always have European surnames (mostly of Portuguese, Dutch and British origin, but sometimes German, French or

	Russian).

```
> ethnic<-c("Sinhala","Sri Lanka Tamil","Indian Tamil","Sri Lanka Moors",
> "Malay","Burgher","Other")
```

✓ There are 11 NA in ETHNICITY.

```
> table(is.na(RO1$ETHNICITY))
FALSE  TRUE
80861    11
```

 # Number of persons by ethnicity

```
> ethnic<-c("Sinhala","Sri Lanka Tamil","Indian Tamil","Sri Lanka Moors",
+ "Malay","Burgher","Other")
> t<-tapply(RO1$WT,RO1$ETHNICITY,sum,na.rm=T)
> names(t)<-ethnic
> round(addmargins(t))
        Sinhala Sri Lanka Tamil    Indian Tamil Sri Lanka Moors
       15312210         2109879          779465         2070535
          Malay         Burgher           Other             Sum
          25063           17886           22007        20337043
> round(addmargins(t)/sum(t)*100,1)
        Sinhala Sri Lanka Tamil    Indian Tamil Sri Lanka Moors
           75.3            10.4             3.8            10.2
          Malay         Burgher           Other             Sum
            0.1             0.1             0.1           100.0
```

 # Number of persons by ethnicity and sector

```
> m<-matrix(as.vector(by(RO1$WT,list(RO1$ETHNICITY,RO1$SECTOR),sum,na.rm=T)),nrow=7)
> rownames(m)<-ethnic
> colnames(m)<-c("Urban","Rural","Estate")
> m[is.na(m)]<-0
> round(addmargins(m))
                 Urban    Rural  Estate      Sum
Sinhala         1685902 13511797  114511 15312210
Sri Lanka Tamil  705196  1115889  288794  2109879
Indian Tamil      38964   121672  618829   779465
```

```
Sri Lanka Moors  561630  1505006    3900  2070535
Malay             13907    11156       0    25063
Burgher           11963     5923       0    17886
Other              4273    17660      74    22007
Sum             3021834 16289102 1026107 20337043
```

```
> round(prop.table(m,2)*100,1)
               Urban Rural Estate
Sinhala         55.8  82.9   11.2
Sri Lanka Tamil 23.3   6.9   28.1
Indian Tamil     1.3   0.7   60.3
Sri Lanka Moors 18.6   9.2    0.4
Malay            0.5   0.1    0.0
Burgher          0.4   0.0    0.0
Other            0.1   0.1    0.0
```

 # Number of persons by ethnicity and province (unit: thousand persons)

```
> province<-c("Western","Central","Southern","Northern","Eastern","North-western",
+ "North-central","Uva","Sabaragamuwa")
> t<-by(RO1$WT,list(substr(RO1$DISTRICT,1,1),RO1$ETHNICITY),sum,na.rm=T)
> m<-matrix(as.vector(t),nrow=9)
> rownames(m)<-province
> colnames(m)<-ethnic
> m[is.na(m)]<-0
> round(m/1000)
              Sinhala Sri Lanka Tamil Indian Tamil Sri Lanka Moors Malay Burgher Other
Western          5049              365           36              486    19       13    15
Central          1671              174          505              286     1        1     3
Southern         2406               31           15               67     3        0     2
Northern           24              678            0               13     0        0     0
Eastern           359              567            3              566     0        4     0
North-western    1960               77            2              380     0        0     0
North-central    1117                4            0              116     0        0     0
Uva              1058               84          132               55     0        0     1
Sabaragamuwa     1667              129           87              101     1        0     0
```

 # Composition of ethnic groups by province (unit: %)

```
> round(prop.table(m,1)*100,1)
```

	Sinhala	Sri Lanka Tamil	Indian Tamil	Sri Lanka Moors	Malay	Burgher	Other
Western	84.4	6.1	0.6	8.1	0.3	0.2	0.3
Central	63.3	6.6	**19.1**	10.8	0.0	0.0	0.1
Southern	95.3	1.2	0.6	2.7	0.1	0.0	0.1
Northern	3.4	**94.8**	0.0	1.8	0.0	0.0	0.0
Eastern	23.9	**37.8**	0.2	**37.8**	0.0	0.3	0.0
North-western	81.0	3.2	0.1	15.7	0.0	0.0	0.0
North-central	90.3	0.3	0.0	9.4	0.0	0.0	0.0
Uva	79.5	6.3	9.9	4.1	0.0	0.0	0.1
Sabaragamuwa	84.0	6.5	4.4	5.1	0.1	0.0	0.0

 # Regional distribution by ethnic groups (unit: %)

```
> round(prop.table(m,2)*100,1)
```

	Sinhala	Sri Lanka Tamil	Indian Tamil	Sri Lanka Moors	Malay	Burgher	Other
Western	33.0	17.3	4.6	**23.5**	77.2	70.3	70.4
Central	10.9	8.2	**64.8**	13.8	4.2	6.7	14.2
Southern	15.7	1.5	1.9	3.3	12.2	0.0	9.3
Northern	0.2	**32.2**	0.0	0.6	0.0	0.0	0.0
Eastern	2.3	**26.9**	0.4	**27.4**	0.0	22.2	0.0
North-western	12.8	3.6	0.3	18.4	0.9	0.8	1.7
North-central	7.3	0.2	0.0	5.6	1.2	0.0	0.0
Uva	6.9	4.0	16.9	2.7	0.0	0.0	4.1
Sabaragamuwa	10.9	6.1	11.1	4.9	4.3	0.0	0.3

Summary:

- Shinhalese are the majority ethnic group of Sri Lanka. They occupy 75.3%, followed by 10.4% of Sri Lanka Tamil and 10.2% of Sri Lanka Moors.
- The ethnic group shows a distinct geographical distribution. Sri Lanka Tamil is dominant in Northern province and major in Eastern. Sri Lanka Moors live in Eastern and Western. Many Indian Tamil live in Estate sector of Central and Uva provinces.

12.2 Religion

✓ Religion is closely related to ethnicity.

```
> table(is.na(RO1$RELIGION))
FALSE  TRUE
80863     9
> table(RO1$RELIGION)
    1     2     3     4     9
50341 13146  9946  7424     6

> religion<-c("Buddhism","Hindu","Islam","Catholic","Other")
> t<-tapply(RO1$WT,RO1$RELIGION,sum,na.rm=T)
> names(t)<-religion
> round(addmargins(t))
Buddhism    Hindu    Islam Catholic    Other      Sum
14359512  2456608  2098303  1422117      539 20337079

> round(addmargins(t)/sum(t)*100,1)
Buddhism    Hindu    Islam Catholic    Other      Sum
    70.6     12.1     10.3      7.0      0.0    100.0

# Number of persons by religion and sector
> m<-matrix(as.vector(by(RO1$WT,list(RO1$RELIGION,RO1$SECTOR),sum,na.rm=T)),nrow=5)
> rownames(m)<-religion
> colnames(m)<-c("Urban","Rural","Estate")
> m[is.na(m)]<-0
> round(addmargins(m))
           Urban    Rural  Estate      Sum
Buddhism 1391905 12851017  116590 14359512
Hindu     535896  1093447  827266  2456608
Islam     574923  1514761    8619  2098303
Catholic  519125   829420   73572  1422117
Other         16      464      60      539
Sum      3021865 16289107 1026107 20337079
```

```
> round(prop.table(m,2)*100,1)
               Urban Rural Estate
         Urban Rural Estate
Buddhism  46.1  78.9   11.4
Hindu     17.7   6.7   80.6
Islam     19.0   9.3    0.8
Catholic  17.2   5.1    7.2
Other      0.0   0.0    0.0

# Number of persons by ethnicity and religion
> m<-matrix(as.vector(by(RO1$WT,list(RO1$ETHNICITY,RO1$RELIGION),sum,na.rm=T)),nrow=7)
> colnames(m)<-religion
> rownames(m)<-ethnic
> m[is.na(m)]<-0
> round(addmargins(m))
                Buddhism   Hindu   Islam Catholic Other      Sum
Sinhala         14323068   13139    7767   968071   165 15312210
Sri Lanka Tamil    25667 1746464   19258   318430    60  2109879
Indian Tamil        5917  685965    5192    82390     0   779465
Sri Lanka Moors     2479    4157 2041476    22408    16  2070535
Malay                753     587   21750     1674   299    25063
Burgher              291       0       0    17594     0    17886
Other               1336    6295    2830    11545     0    22007
Sum             14359512 2456608 2098273  1422112   539 20337043

> round(addmargins(m,2)/rowSums(m)*100,1)
                Buddhism Hindu Islam Catholic Other Sum
Sinhala             93.5   0.1   0.1      6.3   0.0 100
Sri Lanka Tamil      1.2  82.8   0.9     15.1   0.0 100
Indian Tamil         0.8  88.0   0.7     10.6   0.0 100
Sri Lanka Moors      0.1   0.2  98.6      1.1   0.0 100
Malay                3.0   2.3  86.8      6.7   1.2 100
Burgher              1.6   0.0   0.0     98.4   0.0 100
Other                6.1  28.6  12.9     52.5   0.0 100
```

12.3 Marriage

☐ **Polygamy**

Islamic law allows up to four wives, and Sri Lanka has about 10% of Islamic population.
In HIES, only Muslim mails were allowed to have more than one wife. There are one such case in the data set of HIES 2009.

Remarks:
According to the delegates from Sri Lanka for the International Workshop,
“Only Muslim ethnicities can enjoy multiple wives legally in Sri Lanka, and even then it is captured if they live in the same HIES household in the HIES.”

```
# Number of spouses within the household
> nsp<-tapply(R01$RELATIONSHIP==2,R01$ID,sum,na.rm=T)
> table(nsp)
nsp
    0     1     2
 5736 14220     2
> names(nsp)[nsp==2]
[1] "131000413703111" "532000201104111"
> R01[R01$ID=="131000413703111",]
      DISTRICT PSU REC_TYPE SECTOR DSD MONTH SAMPLE_N SERIAL_NO NHH RESULT
19272       13 137        1      1   0     4        3         1   1      1
19273       13 137        1      1   0     4        3         1   1      1
19274       13 137        1      1   0     4        3         1   1      1
19275       13 137        1      1   0     4        3         1   1      1
19276       13 137        1      1   0     4        3         1   1      1
19277       13 137        1      1   0     4        3         1   1      1
19278       13 137        1      1   0     4        3         1   1      1
19279       13 137        1      1   0     4        3         1   1      1
      PERSON_SERIAL_NO RELATIONSHIP SEX_LIVING BIRTH_YEAR BIRTH_MONTH AGE
19272                1            1          1         45           3  65
19273                2            2          2         47           2  63
19274                3            2          2         47          12  62
19275                4            3          1         78          12  31
```

```
19276               5            5          2          81          11  28
19277               6            5          2           6           3   4
19278               7            5          1           9           2   1
19279               8            5          1           9           2   1
      ETHNICITY RELIGION CURR_EDUCATION EDUCATION MARITAL_STATUS
19272         4        3             7         6              2
19273         4        3             7        19              2
19274         4        3             7        19              2
19275         4        3             7         7              2
19276         4        3             7         3              2
19277         4        3             1        NA             NA
19278         4        3            NA        NA             NA
19279         4        3            NA        NA             NA
      MAIN_ACTIVITY MAIN_OCCUPATION INDUSTRY EMPLOYMENT_STATUS
19272             1            3424     5211                5
19273             4              NA       NA               NA
19274             4              NA       NA               NA
19275             1            7411      601                3
19276             4              NA       NA               NA
19277            NA              NA       NA               NA
19278            NA              NA       NA               NA
19279            NA              NA       NA               NA
                  ID              PID       WT
19272 131000413703111 13100041370311101 28.45289
19273 131000413703111 13100041370311102 28.45289
19274 131000413703111 13100041370311103 28.45289
19275 131000413703111 13100041370311104 28.45289
19276 131000413703111 13100041370311105 28.45289
19277 131000413703111 13100041370311106 28.45289
19278 131000413703111 13100041370311107 28.45289
19279 131000413703111 13100041370311108 28.45289
> R01[R01$ID=="532000201104111",]
      DISTRICT PSU REC_TYPE SECTOR DSD MONTH SAMPLE_N SERIAL_NO NHH RESULT
58319       53  11        1      2   0     2        4         1   1      1
58320       53  11        1      2   0     2        4         1   1      1
58321       53  11        1      2   0     2        4         1   1      1
```

```
      PERSON_SERIAL_NO RELATIONSHIP SEX_LIVING BIRTH_YEAR BIRTH_MONTH AGE
58319                1            1          1         76           3  33
58320                2            2          2         81           3  28
58321                3            2          2          5          11   4
      ETHNICITY RELIGION CURR_EDUCATION EDUCATION MARITAL_STATUS
58319         4        3              7        12              2
58320         4        3              7         8              2
58321         4        3              1        NA             NA
      MAIN_ACTIVITY MAIN_OCCUPATION INDUSTRY EMPLOYMENT_STATUS
58319             1            9162      161                 3
58320             4              NA       NA                NA
58321            NA              NA       NA                NA
                   ID               PID       WT
58319 532000201104111 53200020110411101 272.5964
58320 532000201104111 53200020110411102 272.5964
58321 532000201104111 53200020110411103 272.5964
```

There are two cases of multiple wives within the household.
The first case might be possible, because the religion of the household is Islam.
However, the second case is coding error. RELATIONSHIP=2 of PID=" 53200020110411103" should be read as 3.

☐ **Ethnicities and religions of married couples**

Most of married couples of heads and spouses are within the same ethnic groups.

Remarks:
According to the delegates from Sri Lanka for the International Workshop,
"Mixed marriages are rare and rarely found in the HIES kind of sample."

```
> table(R01$RELATIONSHIP)
> d1<-subset(R01,RELATIONSHIP==1)[c("ID","ETHNICITY","RELIGION","SEX_LIVING")]
> d2<-subset(R01,RELATIONSHIP==2)[c("ID","ETHNICITY","RELIGION","SEX_LIVING")]
> d<-merge(d1,d2,by="ID")
> t<-table(d$ETHNICITY.x,d$ETHNICITY.y)
> rownames(t)<-ethnic
```

```
> colnames(t)<-ethnic
> t
 Head             Spouse
                   Sinhala Sri Lanka Tamil Indian Tamil Sri Lanka Moors Malay Burgher Other
  Sinhala            10000              25            13              12     1       6    10
  Sri Lanka Tamil       49            1926             2               7     0       0     0
  Indian Tamil           8               7           585               3     0       0     0
  Sri Lanka Moors       20               3             0            1478     4       0     0
  Malay                  5               0             1               6    32       0     0
  Burgher                4               0             0               0     0       9     0
  Other                  2               0             0               0     0       1     4

> t<-table(d$RELIGION.x,d$RELIGION.y)
> rownames(t)<-religion
> colnames(t)<-religion
> t
Head        Spouse
            Buddhism Hindu Islam Catholic Other
  Buddhism      9069    20     9      112     1
  Hindu           23  2115     5       28     0
  Islam            8     1  1534        4     0
  Catholic        87    13     3     1189     0
  Other            0     0     0        1     1
```

Summary

- HIES allows to have more than one wife. However, such cases are rare in the data set.
- Most of marriages are conducted within the same ethnic group.

12.4 Family type

Family type based on the number of persons by generation

Data file for determining the family type (1)

```
# Maximum household size is 17.
> hhsize<-tapply(R01$PID,R01$ID,length)
> max(hhsize)
[1] 17
```

✓ Designed household-level data file consist of number of persons by relationship as well as data of relationship, marital status, sex and age of each person.

Layout of record

ID	Number of persons by relationship										Relationship of each person					Marital status of each person					Sex of each person					Age of each person				
	T	1	2	3	4	5	6	7	x	9	1	2	3	…	17	1	2	3	…	17	1	2	3	…	17	1	2	3	…	17

```
> df<-R01[,c(26:28,11:13,16,21)]
> colnames(df)[4:8]<-c("no","relation","sex","age","marital")
> head(df)
                ID               PID       WT no relation sex age marital
1 111000700101111 11100070010111101 41.42092  1        1   2  77       3
2 111000700101111 11100070010111102 41.42092  2        3   1  52       2
3 111000700101111 11100070010111103 41.42092  3        3   1  47       1
4 111000700101111 11100070010111104 41.42092  4        3   2  41       5
5 111000700101111 11100070010111105 41.42092  5        5   1   6      NA
6 111000700101111 11100070010111106 41.42092  6        5   2   2      NA
```

✓ Created data frame ft: family type

```
> ft<-data.frame(ID=names(hhsize),hhsz=hhsize,row.names=NULL)
```

```
# Number of persons by relationship
> for(j in c(1:7,9)){
+ d<-subset(df,relation==j)
+ nj<-tapply(d$PID,d$ID,length)
+ dj<-data.frame(ID=names(nj),nj)
+ ft<-merge(ft,dj,by.x="ID",by.y="ID",all.x=T)
+ colnames(ft)[ncol(ft)]<-paste("n",j,sep="")
+ }
> dim(ft)
[1] 19958    10
> head(ft)
               ID hhsz n1 n2 n3 n4 n5 n6 n7 n9
1 111000114801111    5  1  1  2  1
2 111000114802111    5  1  1  3 NA NA NA NA NA
3 111000114803111    8  1  1  1 NA  5 NA NA NA
4 111000114804111    3  1 NA  1 NA  1 NA NA NA
5 111000114805111    3  1  1  1 NA NA NA NA NA
6 111000114807111    3  1  1  1 NA NA NA NA NA
> ft.old<-ft


# Relationship, marital status, sex and age by person number
> for(j in 1:17){
+ d<-subset(df,no==j)[,c(1,5,8,6,7)]
+ colnames(d)<-c("ID",paste(c("r","m","s","a"),j,sep=""))
+ ft<-merge(ft,d,by.x="ID",by.y="ID",all.x=T)
+ }
> dim(ft)
[1] 19958    78
> ft.old2<-ft
> colnames(ft)
[1] "ID"   "hhsz" "n1"   "n2"   "n3"   "n4"   "n5"   "n6"   "n7"   "n9"   "r1"   "m1"
[13] "s1"   "a1"   "r2"   "m2"   "s2"   "a2"   "r3"   "m3"   "s3"   "a3"   "r4"   "m4"
[25] "s4"   "a4"   "r5"   "m5"   "s5"   "a5"   "r6"   "m6"   "s6"   "a6"   "r7"   "m7"
[37] "s7"   "a7"   "r8"   "m8"   "s8"   "a8"   "r9"   "m9"   "s9"   "a9"   "r10"  "m10"
[49] "s10"  "a10"  "r11"  "m11"  "s11"  "a11"  "r12"  "m12"  "s12"  "a12"  "r13"  "m13"
```

```
[61] "s13"  "a13"  "r14"  "m14"  "s14"  "a14"  "r15"  "m15"  "s15"  "a15"  "r16"  "m16"
[73] "s16"  "a16"  "r17"  "m17"  "s17"  "a17"


# Ordered the variables as designed
> ft<-ft[,c(1:10,seq(11,75,by=4),seq(12,76,by=4),seq(13,77,by=4),seq(14,78,by=4))]
> colnames(ft)
[1] "ID"   "hhsz" "n1"   "n2"   "n3"   "n4"   "n5"   "n6"   "n7"   "n9"   "r1"   "r2"
[13] "r3"   "r4"   "r5"   "r6"   "r7"   "r8"   "r9"   "r10"  "r11"  "r12"  "r13"  "r14"
[25] "r15"  "r16"  "r17"  "m1"   "m2"   "m3"   "m4"   "m5"   "m6"   "m7"   "m8"   "m9"
[37] "m10"  "m11"  "m12"  "m13"  "m14"  "m15"  "m16"  "m17"  "s1"   "s2"   "s3"   "s4"
[49] "s5"   "s6"   "s7"   "s8"   "s9"   "s10"  "s11"  "s12"  "s13"  "s14"  "s15"  "s16"
[61] "s17"  "a1"   "a2"   "a3"   "a4"   "a5"   "a6"   "a7"   "a8"   "a9"   "a10"  "a11"
[73] "a12"  "a13"  "a14"  "a15"  "a16"  "a17"

> ft.old3<-ft
> ft[is.na(ft)]<-0
> head(ft)
              ID hhsz n1 n2 n3 n4 n5 n6 n7 n9 r1 r2 r3 r4 r5 r6 r7 r8 r9 r10 r11 r12 r13 r14
1 111000114801111    5  1  1  2  1  0  0  0  0  1  2  3  3  4  0  0  0  0   0   0   0   0   0
2 111000114802111    5  1  1  3  0  0  0  0  0  1  2  3  3  3  0  0  0  0   0   0   0   0   0
3 111000114803111    8  1  1  1  0  5  0  0  0  1  2  3  5  5  5  5  5  0   0   0   0   0   0
4 111000114804111    3  1  0  1  0  1  0  0  0  1  3  5  0  0  0  0  0  0   0   0   0   0   0
5 111000114805111    3  1  1  1  0  0  0  0  0  1  2  3  0  0  0  0  0  0   0   0   0   0   0
6 111000114807111    3  1  1  1  0  0  0  0  0  1  2  3  0  0  0  0  0  0   0   0   0   0   0
  r15 r16 r17 m1 m2 m3 m4 m5 m6 m7 m8 m9 m10 m11 m12 m13 m14 m15 m16 m17 s1 s2 s3 s4 s5 s6 s7 s8
1   0   0   0  2  2  1  1  3  0  0  0  0   0   0   0   0   0   0   0   0  1  2  2  2  1  0  0  0
2   0   0   0  2  2  1  1  0  0  0  0  0   0   0   0   0   0   0   0   0  1  2  1  2  1  0  0  0
3   0   0   0  2  2  2  2  0  0  0  0  0   0   0   0   0   0   0   0   0  1  2  1  2  2  1  1  2
4   0   0   0  3  2  2  0  0  0  0  0  0   0   0   0   0   0   0   0   0  2  1  2  0  0  0  0  0
5   0   0   0  2  2  2  0  0  0  0  0  0   0   0   0   0   0   0   0   0  1  2  1  0  0  0  0  0
6   0   0   0  2  2  0  0  0  0  0  0  0   0   0   0   0   0   0   0   0  1  2  1  0  0  0  0  0
  s9 s10 s11 s12 s13 s14 s15 s16 s17 a1 a2 a3 a4 a5 a6 a7 a8 a9 a10 a11 a12 a13 a14 a15 a16 a17
1  0   0   0   0   0   0   0   0   0 44 41 21 16 67  0  0  0  0   0   0   0   0   0   0   0   0
2  0   0   0   0   0   0   0   0   0 36 31 15 12  7  0  0  0  0   0   0   0   0   0   0   0   0
3  0   0   0   0   0   0   0   0   0 65 63 36 36  9  6  6  6  0   0   0   0   0   0   0   0   0
```

```
4  0   0   0   0   0   0   0   0   0 42 18 17  0  0  0  0  0  0   0   0   0   0   0   0   0   0
5  0   0   0   0   0   0   0   0   0 63 62 28  0  0  0  0  0  0   0   0   0   0   0   0   0   0
6  0   0   0   0   0   0   0   0   0 29 25  4  0  0  0  0  0  0   0   0   0   0   0   0   0   0

> ft.old4<-ft

# Confirmed that all household heads have person number=1.
> table(ft$r1)
    1
19958

# Number of sample households consisted of a head' s couple only
> nrow(subset(ft,hhsz==2&n2==1))
[1] 1386
> 1386/19958
[1] 0.06944584

# Number of sample household consisted of a head' s couple and their children
> nrow(subset(ft,n2==1&n3>=1&hhsz==(n1+n2+n3)))
[1] 8663
> 8663/19958
[1] 0.4340615

# Number of sample household consisted of a father and his children
> nrow(subset(ft,s1==1&n2==0&n3>=1&hhsz==(n1+n2+n3)))
[1] 380

# Number of sample household consisted of a mother and her children
> nrow(subset(ft,s1==2&n2==0&n3>=1&hhsz==(n1+n2+n3)))
[1] 1720
```

✓ Summary:

The share of nuclear family household in the sample household is 61%.

✓ Generated variables g1 to g17 for household members. The values of g1 to g17 are as follows;

Value	Description	Definition
1	Grandparents' generation	relation=5 & age.diff>=35
2	Parents' generation	relation=4
3	Head's generation	relation=1 or 2, relation=5 & abs(age.diff)<=15
4	Children's generation	relation=3, relation=5 & -14=<age.diff<=-34
5	Grandchildren's generation	relation=5 & age.diff<=-35
6	Other relative	relation=5 but not classified in the above
7	Other	relation=6, 7 or 9

where,

age.diff = age – head's age

```
> for(j in 1:17){
+ gen<-rep(0,nrow(ft))
+ relation<-ft[,j+10]
+ age.diff<-ft[,j+61]-ft[,62]
+ gen<-ifelse(relation==1|relation==2,3,gen)
+ gen<-ifelse(relation==3,4,gen)
+ gen<-ifelse(relation==4,2,gen)
+ gen<-ifelse(relation==5,6,gen)
+ gen<-ifelse(relation==5&age.diff>=35,1,gen)
+ gen<-ifelse(relation==5&abs(age.diff)<=15,3,gen)
+ gen<-ifelse(relation==5&age.diff<=-14&age.diff>=-34,4,gen)
+ gen<-ifelse(relation==5&age.diff<=-35,5,gen)
+ gen<-ifelse(relation>=6,7,gen)
+ ft["gen"]<-gen
+ colnames(ft)[ncol(ft)]<-paste("g",j,sep="")
+ }
> head(ft[,c(1:18,62:69,79:86)])
              ID hhsz n1 n2 n3 n4 n5 n6 n7 n9 r1 r2 r3 r4 r5 r6 r7 r8 a1 a2 a3 a4 a5
1 111000114801111    5  1  1  2  1  0  0  0  0  1  2  3  3  4  0  0  0 44 41 21 16 67
2 111000114802111    5  1  1  3  0  0  0  0  0  1  2  3  3  3  0  0  0 36 31 15 12  7
3 111000114803111    8  1  1  1  0  5  0  0  0  1  2  3  5  5  5  5  5 65 63 36 36  9
4 111000114804111    3  1  0  1  0  1  0  0  0  1  3  5  0  0  0  0  0 42 18 17  0  0
5 111000114805111    3  1  1  1  0  0  0  0  0  1  2  3  0  0  0  0  0 63 62 28  0  0
```

```
6 111000114807111    3  1  1  1  0  0  0  0  0  1  2  3  0  0  0  0  0 29 25  4  0  0
  a6 a7 a8 g1 g2 g3 g4 g5 g6 g7 g8
1  0  0  0  3  3  4  4  2  0  0  0
2  0  0  0  3  3  4  4  4  0  0  0
3  6  6  6  3  3  4  4  5  5  5  5
4  0  0  0  3  4  4  0  0  0  0  0
5  0  0  0  3  3  4  0  0  0  0  0
6  0  0  0  3  3  4  0  0  0  0  0

> ft.old5<-ft
```

Grouping sample households based on the number of persons by generation

```
# Defined variable pi, the number of persons in generation i
> for(j in 1:7){
+ ft["pi"]<-rowSums(ft[,79:95]==j)
+ colnames(ft)[ncol(ft)]<-paste("p",j,sep="")
+ }
> head(ft[,c(1,79:87,96:102)])
               ID g1 g2 g3 g4 g5 g6 g7 g8 g9 p1 p2 p3 p4 p5
1 111000114801111  3  3  4  4  2  0  0  0  0  0  1  2  2  0
2 111000114802111  3  3  4  4  4  0  0  0  0  0  0  2  3  0
3 111000114803111  3  3  4  4  5  5  5  5  0  0  0  2  2  4
4 111000114804111  3  4  4  0  0  0  0  0  0  0  0  1  2  0
5 111000114805111  3  3  4  0  0  0  0  0  0  0  0  2  1  0
6 111000114807111  3  3  4  0  0  0  0  0  0  0  0  2  1  0

# Number of persons by generation
> colSums(ft[,96:102])
   p1    p2    p3    p4    p5    p6    p7
  169  1555 35606 35505  7052   628   357

> ft.old6<-ft
```

Table Type of households by number of household members' generations

Number of generation	Household consisted of the below generation	Code	Number of households	Percentage
	All households		19,958	100.0
1	Total		2,593	12.9
	Single	11	902	4.5
	Head and spouse only	12	1,386	6.9
	Head, spouse and others	13	61	0.3
	Other	14	244	1.2
2	Total		12,472	62.5
	Head – Children	21	12,271	61.5
	Parents - Head	22	201	1.0
3	Total		4,671	23.3
	Head – Children - Grandchildren	31	3,181	15.9
	Head – Grandchildren	32	400	2.0
	Parents – Head - Children	33	1,067	5.3
	Grandparents – Parents - Head	34	5	0.0
	Grandparents - Head	35	18	0.1
4	Total		218	1.1
	Parents – Head – Children - Grandchildren	41	83	0.4
	Parents – Head – Grandchildren	42	3	0.0
	Grandparents – Parents – Head – Children	43	15	0.1
	Grandparents – Head - Children	44	117	0.6
5	Grandparents – (Parents) – Head – (Children)- Grandchildren	50	4	0.0

✓ Summary:

The household consisted of two generations occupies 63%, followed by 23% of three-generation household and 13% of one-generation household.

```
# Grouped households using p1 to p5, regardless of the value of p6 and p7.
> h<-rep(0,nrow(ft))
> h<-ifelse(ft$p1==0&ft$p2==0&ft$p3>0&ft$p4==0&ft$p5==0,14,h)
```

```
> h<-ifelse(ft$hhsz==1,11,h)
> h<-ifelse(ft$hhsz==2&ft$n2==1,12,h)
> h<-ifelse(ft$p1==0&ft$p2==0&ft$p3>=3&ft$p4==0&ft$p5==0&ft$n2==1,13,h)
> h<-ifelse(ft$p1==0&ft$p2==0&ft$p3>0&ft$p4>0&ft$p5==0,21,h)
> h<-ifelse(ft$p1==0&ft$p2>0&ft$p3>0&ft$p4==0&ft$p5==0,22,h)
> h<-ifelse(ft$p1==0&ft$p2==0&ft$p3>0&ft$p4>0&ft$p5>0,31,h)
> h<-ifelse(ft$p1==0&ft$p2==0&ft$p3>0&ft$p4==0&ft$p5>0,32,h)
> h<-ifelse(ft$p1==0&ft$p2>0&ft$p3>0&ft$p4>0&ft$p5==0,33,h)
> h<-ifelse(ft$p1>0&ft$p2>0&ft$p3>0&ft$p4==0&ft$p5==0,34,h)
> h<-ifelse(ft$p1>0&ft$p2==0&ft$p3>0&ft$p4==0&ft$p5==0,35,h)
> h<-ifelse(ft$p1==0&ft$p2>0&ft$p3>0&ft$p4>0&ft$p5>0,41,h)
> h<-ifelse(ft$p1==0&ft$p2>0&ft$p3>0&ft$p4==0&ft$p5>0,42,h)
> h<-ifelse(ft$p1>0&ft$p2>0&ft$p3>0&ft$p4>0&ft$p5==0,43,h)
> h<-ifelse(ft$p1>0&ft$p2==0&ft$p3>0&ft$p4>0&ft$p5==0,44,h)
> h<-ifelse(ft$p1>0&ft$p3>0&ft$p5>0,50,h)
> t<-addmargins(table(h))
> t
h
   11    12    13    14    21    22    31    32    33    34    35    41    42    43
  902  1386    61   244 12271   201  3181   400  1067     5    18    83     3    15
   44    50   Sum
  117     4 19958
> round(t/t[length(t)]*100,1)
h
   11    12    13    14    21    22    31    32    33    34    35    41    42    43
  4.5   6.9   0.3   1.2  61.5   1.0  15.9   2.0   5.3   0.0   0.1   0.4   0.0   0.1
   44    50   Sum
  0.6   0.0 100.0


# Family type by sector
> df<-merge(R23[,c("ID","DISTRICT","SECTOR","WT")],ft[,c("ID","hhsz","type")])
> dim(df)
[1] 19958     6
> head(df)
              ID DISTRICT SECTOR      WT hhsz type
```

```
1 111000114801111        11       1 30.4971     5    33
2 111000114802111        11       1 30.4971     5    21
3 111000114803111        11       1 30.4971     8    31
4 111000114804111        11       1 30.4971     3    21
5 111000114805111        11       1 30.4971     3    21
6 111000114807111        11       1 30.4971     3    21
> m<-matrix(as.vector(tapply(df$WT,list(df$SECTOR,df$type),sum,na.rm=T)),
+ nrow=3)
> rownames(m)<-c("Urban","Rural","Estate")
> colnames(m)<-c(11,12,13,14,21,22,31,32,33,34,35,41,42,43,44,50)
```

 # Weighted number of households by family type and sector

```
> round(m)
           11      12    13     14       21     22      31     32      33  34    35     41  42
Urban   26628   49928 1636 13869  432982   7427 118456 13692   37747 517 1014   2105  NA
Rural  189488 292136 8374 39872 2584633 40457 622561 81081 217751 534 2592 16166 695
Estate  10539   19187  614  2197  132267   3264  50545  6763   16066 274  102   2012  NA
         43     44  50
Urban   456   4711  NA
Rural  2594 23180 543
Estate  798    794 119
> m[is.na(m)]<-0
> round(prop.table(m,1)*100,1)
        11  12  13  14   21  22   31  32  33  34  35  41 42  43  44 50
Urban  3.7 7.0 0.2 2.0 60.9 1.0 16.7 1.9 5.3 0.1 0.1 0.3  0 0.1 0.7  0
Rural  4.6 7.1 0.2 1.0 62.7 1.0 15.1 2.0 5.3 0.0 0.1 0.4  0 0.1 0.6  0
Estate 4.3 7.8 0.2 0.9 53.9 1.3 20.6 2.8 6.5 0.1 0.0 0.8  0 0.3 0.3  0
```

```
# Family type by ethnicity of household head
> df<-merge(df,subset(RO1,RELATIONSHIP==1)[,c("ID","ETHNICITY","RELIGION")])
> dim(df)
[1] 19958     8
> head(df)
              ID DISTRICT SECTOR      WT hhsz type ETHNICITY RELIGION
1 111000114801111       11      1 30.4971    5   33         1        4
```

```
2 111000114802111        11       1 30.4971     5    21          1         1
3 111000114803111        11       1 30.4971     8    31          1         4
4 111000114804111        11       1 30.4971     3    21          1         1
5 111000114805111        11       1 30.4971     3    21          1         4
6 111000114807111        11       1 30.4971     3    21          1         1
> m<-matrix(as.vector(tapply(df$WT,list(df$ETHNICITY,df$type),sum,na.rm=T)),
+ nrow=7)
> colnames(m)<-c(11,12,13,14,21,22,31,32,33,34,35,41,42,43,44,50)
> rownames(m)<-c("Sinhala","Sri Lanka Tamil","Indian Tamil","Sri Lanka Moors",
+ "Malay","Burgher","Other")
```

 # Weighted number of households by family type and ethnicity

```
> round(m)
                    11     12   13    14      21    22     31    32     33  34   35
Sinhala         186422 291899 8436 46250 2423201 38790 620660 80082 214380 631 3195
Sri Lanka Tamil  20690  31691 1311  6180  320502  3976  64706 10323  24098 421   79
Indian Tamil      6059  14306  319  1342   97528  3382  38935  4442  13396 274  161
Sri Lanka Moors  12171  21442  557  2127  298060  4645  65176  6629  19162  NA  272
Malay              516   1868   NA    39    4817    39    636    60     52  NA   NA
Burgher            798     45   NA    NA    1906   317    668    NA    476  NA   NA
Other               NA     NA   NA    NA    3861    NA    779    NA     NA  NA   NA
                   41  42   43    44  50
Sinhala         14969 695 1793 19664 543
Sri Lanka Tamil  1727  NA  979  6195  NA
Indian Tamil     2508  NA  730  1567 119
Sri Lanka Moors  1067  NA  347  1187  NA
Malay              13  NA   NA    NA  NA
Burgher            NA  NA   NA    73  NA
Other              NA  NA   NA    NA  NA
> m[is.na(m)]<-0
> round(prop.table(m,1)*100,1)
                   11   12  13  14   21  22   31  32   33  34  35  41 42  43  44  50
Sinhala           4.7  7.4 0.2 1.2 61.3 1.0 15.7 2.0  5.4 0.0 0.1 0.4  0 0.0 0.5 0.0
Sri Lanka Tamil   4.2  6.4 0.3 1.3 65.0 0.8 13.1 2.1  4.9 0.1 0.0 0.4  0 0.2 1.3 0.0
Indian Tamil      3.3  7.7 0.2 0.7 52.7 1.8 21.0 2.4  7.2 0.1 0.1 1.4  0 0.4 0.8 0.1
Sri Lanka Moors   2.8  5.0 0.1 0.5 68.9 1.1 15.1 1.5  4.4 0.0 0.1 0.2  0 0.1 0.3 0.0
```

```
Malay            6.4 23.2 0.0 0.5 59.9 0.5  7.9 0.7  0.6 0.0 0.0 0.2  0 0.0 0.0 0.0
Burgher         18.6  1.1 0.0 0.0 44.5 7.4 15.6 0.0 11.1 0.0 0.0 0.0  0 0.0 1.7 0.0
Other            0.0  0.0 0.0 0.0 83.2 0.0 16.8 0.0  0.0 0.0 0.0 0.0  0 0.0 0.0 0.0
```

```
# Average household size by sector and by ethnicity of household head
> t<-as.vector(by(df,df$SECTOR,function(df) sum(df$hhsz*df$WT)/sum(df$WT)))
> names(t)<-c("Urban","Rural","Estate")
> round(t,2)
 Urban  Rural Estate
  4.25   3.95   4.18

> t<-as.vector(by(df,df$ETHNICITY,function(df) sum(df$hhsz*df$WT)/sum(df$WT)))
> names(t)<- c("Sinhala","Sri Lanka Tamil","Indian Tamil","Sri Lanka Moors",
+ "Malay","Burgher","Other")
> round(t,2)
       Sinhala Sri Lanka Tamil    Indian Tamil Sri Lanka Moors           Malay
          3.87            4.29            4.26            4.78            3.90
       Burgher           Other
          3.74            4.05
```

Remarks:

According to the delegates from Sri Lanka for the International Workshop,

“Sinhala HIES households are mostly nuclear families and others are not.”

- ✓ Summary:

  Average household size of Sinhala is smaller than that of Sri Lanka Tamil and Sri Lanka Moors. Consequently, average household size is smaller in rural than in urban.

Married couple in parents' generation

The objective is to identify the married couple of a father and a mother within the household.

✓ Generated the next variables in data frame ft;
npm: number of male persons with relation=4 (parents)
npf: number of female persons with relation=4
npmm: number of male persons with relation=4 and marital=2
npmf: number of female persons with relation=4 and marital=2
npc: number of married couples of parents

```
> ft["npm"]<-0
> ft["npf"]<-0
> ft["npmm"]<-0
> ft["npmf"]<-0
> ft["npc"]<-0
> for(j in 1:17){
+ ft["npm"]<-ifelse(ft[,j+10]==4&ft[,j+44]==1,ft$npm+1,ft$npm)
+ ft["npf"]<-ifelse(ft[,j+10]==4&ft[,j+44]==2,ft$npf+1,ft$npf)
+ ft["npmm"]<-ifelse(ft[,j+10]==4&ft[,j+27]==2&ft[,j+44]==1,ft$npmm+1,ft$npmm)
+ ft["npmf"]<-ifelse(ft[,j+10]==4&ft[,j+27]==2&ft[,j+44]==2,ft$npmf+1,ft$npmf)
+ }
```

✓ The number of fathers is 385 and the number of mothers is 1170. Of which, the number of married fathers is 246 and the number of married mothers is 318.
There is no household with two fathers or two mothers.

```
> table(ft$npm)
    0     1
19573   385
> table(ft$npf)
    0     1
18788  1170
> table(ft$npmm)
    0     1
19712   246
```

```
> table(ft$npmf)
    0     1
19640   318
```

✓ If there are a married father and a married mother in the household, they are estimated as a married couple.

Number of parents couple in the household

```
> ft["npc"]<-ifelse(ft$npmm==1&ft$npmf==1,1,0)
> table(ft$npc)
    0     1
19777   181
```

Summary: Parents

Sex	Number of parents	Of which, married	Number of married couple within the household
Total	1,555	564	181
Male	385	246	
Female	1,170	318	

Married couple in children' generation

The objective is to identify the married couple of children within the household.
The relationship code of "spouses of children" is assumed to be 5 "other relative" in this analysis.

✓ Generated the next variables in data frame ft;
ncm: number of male persons with relation=3
ncf: number of female persons with relation=3
ncmm: number of male persons with relation=3 and marital=2
ncmf: number of female persons with relation=3 and marital=2
ncum: number of male persons with relation=3 and marital<=1
ncuf: number of female persons with relation=3 and marital<=1
n45mm: number of male person with generation=4 and marital=2 and relation=5

n45mf: number of female person with generation=4 and marital=2 and relation=5

ncc: number of married couples of children

```
> ft["ncm"]<-0
> ft["ncf"]<-0
> ft["ncmm"]<-0
> ft["ncmf"]<-0
> ft["ncum"]<-0
> ft["ncuf"]<-0
> ft["n45mm"]<-0
> ft["n45mf"]<-0
> ft["ncc"]<-0
> for(j in 1:17){
+ ft["ncm"]<-ifelse(ft[,j+10]==3&ft[,j+44]==1,ft$ncm+1,ft$ncm)
+ ft["ncf"]<-ifelse(ft[,j+10]==3&ft[,j+44]==2,ft$ncf+1,ft$ncf)
+ ft["ncmm"]<-ifelse(ft[,j+10]==3&ft[,j+27]==2&ft[,j+44]==1,ft$ncmm+1,ft$ncmm)
+ ft["ncmf"]<-ifelse(ft[,j+10]==3&ft[,j+27]==2&ft[,j+44]==2,ft$ncmf+1,ft$ncmf)
+ ft["ncum"]<-ifelse(ft[,j+10]==3&ft[,j+27]<=1&ft[,j+44]==1,ft$ncum+1,ft$ncum)
+ ft["ncuf"]<-ifelse(ft[,j+10]==3&ft[,j+27]<=1&ft[,j+44]==2,ft$ncuf+1,ft$ncuf)
+ ft["n45mm"]<-ifelse(ft[,j+10]==5&ft[,j+27]==2&ft[,j+44]==1&ft[,j+78]==4,ft$n45mm+1,ft$n45mm)
+ ft["n45mf"]<-ifelse(ft[,j+10]==5&ft[,j+27]==2&ft[,j+44]==2&ft[,j+78]==4,ft$n45mf+1,ft$n45mf)
+ }
> sum(ft$ncm)
[1] 17005
> sum(ft$ncf)
[1] 15924
> sum(ft$ncum)
[1] 14994
> sum(ft$ncmm)
[1] 1925
> sum(ft$ncuf)
[1] 13876
> sum(ft$ncmf)
[1] 1746
> sum(ft$n45mm)
[1] 899
```

```
> sum(ft$n45mf)
[1] 996
```

✓ Defined function child.couple.

```
# Arguments: ncmm, ncmf, n45mm, n45mf as same as the above
# Return value: number of married couples of children within the household
> child.couple<-function(ncmm,ncmf,n45mm,n45mf){
+ n<-0
+ while(ncmm>0&n45mf>0){
+ n<-n+1
+ ncmm<-ncmm-1
+ n45mf<-n45mf-1
+ }
+ while(ncmf>0&n45mm>0){
+ n<-n+1
+ ncmf<-ncmf-1
+ n45mm<-n45mm-1
+ }
+ return(n)
+ }
```

✓ Applied function child.couple for all households.

```
> for(j in 1:nrow(ft)){
+ ft[j,"ncc"]<-child.couple(ft[j,"ncmm"],ft[j,"ncmf"],ft[j,"n45mm"],ft[j,"n45mf"])
+ }

# Number of married couples of children
> sum(ft$ncc)
[1] 1566

# Frequency of number of children' s married couples within the household
> table(ft$ncc)
    0     1     2     3
18466  1421    68     3

> ft.old7<-ft
```

Summary: Children

Sex	Number of children	Of which, unmarried	Of which, married	generation=4 and marital=2 and relation=5	Number of married couple within the household
Total					1,566
Male	17,005	14,994	1,925	899	
Female	15,924	13,876	1,746	996	

Characteristics of household heads

✓ Sex and marital status of household heads

```
> d<-subset(R01,RELATIONSHIP==1)
> t<-table(d$SEX_LIVING,d$MARITAL_STATUS)
> rownames(t)<-c("Male","Female")
> colnames(t)<-c("Never","Married","Widowed","Divorced","Separated")
> t
        Never Married Widowed Divorced Separated
  Male    249   14093     648       27       140
  Female  196    1560    2737       48       260
> round(prop.table(t,1)*100,1)
        Never Married Widowed Divorced Separated
  Male    1.6    93.0     4.3      0.2       0.9
  Female  4.1    32.5    57.0      1.0       5.4
```

✓ Age and marital status of female household heads

```
> d2<-subset(d,SEX_LIVING==2)
> age10<-c(0,20,30,40,50,60,70,99)
> d2["age.gr"]<-cut(d2$AGE,breaks=age10,right=F,include.lowest=T,
+ labels=c("0-19","20-29","30-39","40-49","50-59","60-69","70-99"))
> t2<-table(d2$age.gr,d2$MARITAL_STATUS)
> colnames(t2)<- c("Never","Married","Widowed","Divorced","Separated")
> t2
```

```
       Never Married Widowed Divorced Separated
0-19       9       5       0        0         1
20-29     15     199       7        3         7
30-39     11     524      84       13        47
40-49     28     481     348       12        81
50-59     54     241     712       14        79
60-69     58      84     800        5        37
70-99     21      25     785        1         8
> round(prop.table(t2,1)*100,1)
       Never Married Widowed Divorced Separated
0-19    60.0    33.3     0.0      0.0       6.7
20-29    6.5    86.1     3.0      1.3       3.0
30-39    1.6    77.2    12.4      1.9       6.9
40-49    2.9    50.6    36.6      1.3       8.5
50-59    4.9    21.9    64.7      1.3       7.2
60-69    5.9     8.5    81.3      0.5       3.8
70-99    2.5     3.0    93.5      0.1       1.0
```

Household consisted of a mother and her unmarried children only

```
> d<-subset(ft,s1==2&hhsz>=2&hhsz==(1+ncum+ncuf))
> nrow(d)
[1] 1577
```

Household consisted of a father and his unmarried children only

```
> d<-subset(ft,s1==1&hhsz>=2&hhsz==(1+ncum+ncuf))
> nrow(d)
[1] 350
```

13. References

[1] “Household Income and Expenditure Survey – 2009/10 Final Report”, DCS, Sri Lanka, August 2011

[2] “Household Income and Expenditure Survey – 2009/10 Preliminary Report”, DCS, Sri Lanka, January 2011

[3] Ms. Dilhanie Deepawansa, DCS, “Determinants of the Poverty in Western Province in Sri Lanka”, Research - based Training Program, AY 2010, UNSIAP

203

(For office use)

Month		Sector	District		DS Division	
PSU Number						
SSU Number & Household Number						

CONFIDENTIAL
The information collected in this survey will be strictly confidential according to the survey ordinance & individual level information will not be divulged to any person or agency.

HOUSEHOLD INCOME AND EXPENDITURE SURVEY - 2009/10

-

NATIONAL HOUSEHOLD SAMPLE SURVEY PROGRAMME

SURVEY SCHEDULE

Department of Census & Statistics
Ministry of Finance and Planning
Sri Lanka

IDENTIFICATION INFORMATIONS

1

1. Address (Location) :-.....

.....

2. Province :- .....

3. District :- .....

4. DS Division :- .....

5. Name of MC / UC.....

(If urban sector only)

Ward No :- .....

6. GN Division : Number :- ..... Name :- .....

7. Name of Village :- .....

(If rural sector only)

8. Name of Estate :- .....

(If estate sector only)

9. C.B.No. :- .....

*** * Result Code**

Completed	1
Deferred	2
Not competent respondant at home	3
Refused	4
Household is temporarily closed	5
Household is demolished / Vacant	6
Other (Specify)	9

10

Survey Month		Sector		District	DS Division	

11

PSU Number			SSU Number		Household Number

12 Number of Households in this unit ☐

13 Result ☐ *

14 Name of the Head of the Household :- .....

.....

15 Interviewer's

Name :- .....

16 Supervising Officer's

Name :- .....

Signature :- ..... Date :- .....

*** For Office Use Only**

2

Codes for Section 1

Col. 3 - Relationship to head of the household

Head of the household.....1
Wife/ Husband 2
Son/ Daughter 3
Parents 4
Other Relative 5
Domestic Servants 6
Boarder 7
Other 9

Col. 4 - Sex

Male 1
Female 2

Col. 7 - Ethnicity

Sinhala.....1
Sri LankaTamil.....2
Indian Tamil.....3
Sri Lanka Moors 4
Malay 5
Burgher 6
Other 9

Col. 8 - Religion

Buddhist 1
Hindu 2
Islam 3
Roman Catholic/
Other Christian 4
Other 9

Col. 9 -Attendance at School or Other Educational Institution

Pre school 1
School 2
University 3
Other educational institution 4
Vocational/ Technical institution 5
Pending results G.C.E. (O.L/A.L) 6
Does not attend 7

Col. 10 -Level of Education

Studing in Grade 1..... 00
Passed Grade 1.....01
Passed Grade 2.....02
Passed Grade 3.....03
Passed Grade 4.....04
Passed Grade 5.....05
Passed Grade 6.....06
Passed Grade 7.....07
Passed Grade 8.....08
Passed Grade 9.....09
Passed Grade 10.....10
Passed G.C.E. (O/L) or
equivalent 11
Passed Grade 12.....12
Passed G.C.E.(A/L) .or
equivalent 13
Passed GAQ/GSQ 14
Passed Degree 15
Passed Post Graduate Degree / Diploma....16
Special Education unit 17
No Schooling 19

Col. 11 - Marital Status

Never Married 1
Married 2
Widowed 3
Divorced 4
Separated 5

Col. 12 -Current Activity

Employed.....1
Unemployed.....2
Student 3
Household work.....4
Unable / Too old to work.....5
Other 9

Col. 15 - Employment status

Government employee.....1
Semi government employee 2
Private sector employee 3
Employer 4
Own account worker 5
Unpaid family worker.....6

Section 1 - Demographic Characteristics

3

Serial Number	Name of all persons who usually live in this household	Relationship to head of the household	Sex	Date of Birth		Age as at last birthday (years)	Ethnicity	Religion	Attendance at school or other Educational institution (3 years and over)	Level of Education (5 Years and over)	For Persons 10 years and over				
				Year	Month						Marital Status	Current activity	For employed persons only (Column 12 code = 1)		
													* Main Occupation	Main Industry **	Employment status
1	2	3	4	5		6	7	8	9	10	11	12	13	14	15
01															
02															
03															
04															
05															
06															
07															
08															
09															
10															
11															
12															
13															
14															
15															

	Name the persons who are member of this household and usually live elsewhere in the country or abroad		Usual residence In the country - 1 Abroad - 2
41			
42			
43			
44			

Serial Number	* Main Occupation (Describe)	** Main Industry (Describe)

4

Codes for Section 2

For Column 4 and 5

Grade 1..... 01
Grade 2.....02
Grade 3.....03
Grade 4.....04
Grade5.....05
Grade 6.....06
Grade 7.....07
Grade 8.....08
Grade 9.....09
Grade 10.....10
Grade 11.....11
Grade 12.....12
Grade 13.....13
Special Education Unit....14
Not relevant 19

Column 7

Walk 1
Bicycle 2
Motor bicycle / Three-wheeler/Car 3
School hiring Van/ Bus 4
By Bus 5
By Train 6
Other (Specify) 9

Column 9

School too far away.....1
Financial problems 2
Had to help house keeping activities /
Family business 3
Disability / Illness 4
Civil disturbance 5
Not willing to attend / Poor academic progress 6
Incompletion of 5 years at the begining of the school year.....7
Other (specify) 9

Column 10

Further schooling not available or too far away.....1
Finalcial Problems.....2
Had to help house keeping activities / Family business.....3
Disability / Illness.....4
Civil disturbance 5
Not willing to attend / Poor academic progress.....6
Pending results (G.C.E. (O/L) / G.C.E. (A/L)) 7
Completed G.C.E. (A/L) / Grade 13 8
Other (Specify) 9

Section 2 . School Education (for persons aged 5-19 years)

5

Name of persons (age 5-19 years) who usually live in this household (as given in col.2 of sec.1)	Serial number as given in column 1 of section 1		School Education * (code)	For Persons = code 1 in col 3					persons = code 2 in col 3	For persons = code 3 in col 3	
				Grade (current year)	Grade (Previous year)	Distance to school from house (km)	Mode of travel to school (main code)	Time taken to school (minutes)	Why did you never attend school (code)	Why did you leave school (code)	Which year did you leave school
1	2		3	4	5	6	7	8	9	10	11
1											
2											
3											
4											
5											
6											
7											
8											
9											
10											
11											
12											

* Currently attending school..1.....→..... . go to col. 4 Never attended school ....2.....→.....go to col. 9 Attended school in the past... 3.....→.....go to col. 10

6

Codes for Section 3

Column 4

Treatment for illness.....1

Treatment for injury.....2

Medical checkup/ Consultation 3

Immunization 4

Treatment for infectious diseases (Injections ect).....5

Other (specify) 9

Column 6

Treatment for illness 1

Treatment for injury 2

Operation / Surgey.....3

Child delivery.....4

Treatment for infectious diseases.....5

An accident.....6

Other (specify) 9

Column 8

Heart Conditions /Diseases 01

Blood pressure 02

Diabetics 03

Asthma 04

Epilepsy 05

Cancer 06

Stomach diseases / Gastritis.....07

Diseases related to Eyes 08

Arthritis 09

Mental retardation.....10

Haemorrhoids.....11

Catarrh.....12

Severe headache.....13

Naturally Disabled 14

Disabled by an accident 15

Other 19

Section 3 -Health

7

Name of all persons who usually live in this household (as given in col. 2 of sec.1)	Serial number as given in column 1 of section 1	During the last month did you visit any of the Government / Private hospital orMedical / Health Center to obtain out patient health Yes - 1 No - 2 → go to col.5	Resons for visit (write the relevant code)	During the last 12 month have you stayed at a Government / Private hospital as an in patient Yes - 1 No - 2 → go to col.7	Resons for Stay (write the relevant code)	Do you suffer from Chronic illness/ Disability Yes - 1 No - 2 → go to sec.4	What chronic illness / Disability do you suffer from (write the relevant code)	Have your job effected /caused for your illness / Disability Yes - 1 No - 2	How long have you been suffering from your illness/ disability Year / Month	Did you have to stop doing your usual activities because of this illness / disability Yes - 1 No - 2 → go to sec.4	If so how many days did you stop your usual activities during last month
1	2	3	4	5	6	7	8	9	10	11	12
1											
2											
3											
4											
5											
6											
7											
8											
9											
10											
11											
12											
13											
14											
15											
16											
17											
18											

Section 4 -Expenditure

8

4.1 Weekly Consumptions on Food & Drink

Item	Code	unit	Consumed		Home grown/ Freely received (Rs.) *
			Qty.	Value (Rs.)	
1	2	3	4	5	6
1. Cereals 01					
1. Rice - (Kekulu)	0101	grams			
2. Rice - (Samba)	0102	grams			
3. Rice - (Nadu)	0103	grams			
4. Rice - (Other)	0104	grams			
5. Rice flour	0105	grams			
6. Wheat flour	0106	grams			
7. Kurakkan flour	0107	grams			
8. Ulundu flour	0108	grams			
9. Maize	0109	grams			
10. Samaposha	0110	grams			
11. Noodles	0111	grams			
12. Papadam	0112	grams			
13. Infant Cerial food/ Malt etc.	0113	grams			
14. Barley	0114	grams			
15. Sago	0115	grams			
16. Corn flakes	0116	grams			
17. Other (specify)	0119	grams			
2. Prepared foods 02					
1. Bread (Normal)	0201	grams			
2. Bread (Special)	0202				
3. Rospaan	0203	number			
4. Buns / Spanchi	0204	number			
5. Hoppers	0205	number			
6. String hoppers	0206	number			
7. Pittu	0207	number			

Items	Code	Unit	Consumed		Home grown/ Freely received (Rs.) *
			Qty.	Value (Rs.)	
1	2	3	4	5	6
8. Rotti	0208	number			
9. Thosa /Itly	0209	number			
10. Rice (meat & vegetables)	0210	number			
11. Rice (fish/ egg & vegetables)	0211	number			
12. Rice (vegetables only)	0212	number			
13. Curry (meat)	0213				
14. Curry (fish)	0214				
15. Curry (vegetables)	0215				
16. Other	0219				
3. Pulses 03					
1. Dhal	0301	grams			
2. Green gram	0302	grams			
3. Gram	0303	grams			
4. Cowpea	0304	grams			
5. Soya	0305	grams			
6. Soyameat	0306	grams			
7. Other	0309	grams			
4. Vegetables 04					
1. Ash plantain	0401	grams			
2. Brinjal	0402	grams			
3. Bandakka	0403	grams			
4. Bitter gourd	0404	grams			
5. Long beans	0405	grams			
6. Snake gourd	0406	grams			
7. Ridge gourd	0407	grams			
8. Sweet pumpkin	0408	grams			

**** The estimated value cf items which are consumed from home grown / freely received should be included in col.6***

Item	Code	Unit	Consumed		Home grown/ Freely received * (Rs.)
			Qty.	Value (Rs.)	
1	2	3	4	5	6
9. Beans	0409	grams			
10. Carrot	0410	grams			
11. Beetroot	0411	grams			
12. Cabbage	0412	grams			
13. Tomatoes	0413	grams			
14. Leeks	0414	grams			
15. Knolkhol	0415	grams			
16. Capsicum	0416	grams			
17. Dambala	0417	grams			
18. Raddish	0418	grams			
19. Drumstick	0419	grams			
20. Cucumber	0420	grams			
21. Kekiri	0421	grams			
22. Ash pumpkin	0422	grams			
23. Elabatu	0423	grams			
24. Kohila yams	0424	grams			
25. Lotus stemo	0425	grams			
26. Plantain flower	0426	grams			
27. Ambarella	0427	grams			
28. Cadjunuts (raw)	0428	grams			
29. Mushrooms	0429	grams			
30. Jack immatured(Polos)	0430				
31. Other	0439				
Leafy vegetables					
1. Mukunuwenna	0441	bundles			
2. Gotukola	0442	bundles			
3. Kankun	0443	bundles			

Item	Code	Unit	Consumed		Homegrown/ Freely received * (Rs.)
			Qty.	Value (Rs.)	
1	2	3	4	5	6
4. Katurumurunga	0444	bundles			
5. Nivithi	0445	grams			
6. Thampala	0446	bundles			
7. Sarana	0447	bundles			
8. Kohila leaves	0448	bundles			
9. Onion leaves	0449	grams			
10. Cabbage leaves	0450	grams			
11. Other leaves	0459				
5. Yams and Other 05					
1. Jak / Jak seeds	0501				
2. Bread fruit	0502	number			
3. Potatoes	0503	grams			
4. Sweet potatos	0504	grams			
5. Mannioc	0505	grams			
6. Kiriala	0506	grams			
7. Innala	0507	grams			
8. Other yams	0509	grams			
6. Meat 06					
1. Chicken	0601	grams			
2. Beef	0602	grams			
3. Mutton	0603	grams			
4. Pork	0604	grams			
5. Sausages/ Meat balls	0605	grams			
6. Other	0609	grams			

* ***The estimated Value of items which are consumerd from home grown / freely received should be included in col.6***

10

Item	Code	Unit	Consumed		Home grown / Freely received (Rs.) *
			Qty.	Value (Rs.)	
1	2	3	4	5	6
7. Fish 07					
1. Balaya	0701	grams			
2. Seer	0702	grams			
3. Mora	0703	grams			
4. Paraw	0704	grams			
5. Talapath	0705	grams			
6. Kelawalla	0706	grams			
7. Mullet	0707	grams			
8. Other large fish	0708	grams			
9. Sprats	0709	grams			
10. Hurulla	0710	grams			
11. Karalla / Katuwalla	0711	grams			
12. Kumbalawa / Angila	0712	grams			
13. Salaya / Sudaya	0713	grams			
14. Other small fish	0714	grams			
15. Lula	0715	grams			
16. Teppili/Telapiya/Korali	0716	grams			
17. Other fresh water fish	0717	grams			
18. Prawns	0718	grams			
19. Crabs	0719	grams			
20. Cuttle fish	0720	grams			
21. Canned fish (salmon)	0720	grams			
8. Dried fish 08					
1. Sprats	0801	grams			
2. Keeramin	0802	grams			
3. Salaya	0803	grams			
4. Hurulla	0804	grams			
5. Seer	0805	grams			
6. Katta	0806	grams			
7. Koduwa	0807	grams			
8. Anjila	0808	grams			
9. Balaya	0809	grams			
10. Shark / Keelan	0810	grams			
11. Paraw	0811	grams			
12. Anguluwa	0812	grams			
13. Prawns	0813	grams			
14. Cuttle fish	0814	grams			
15. Fresh water fish	0815	grams			
16. Jadi	0816	grams			
17. Other	0819	grams			
9. Eggs 09					
1. Hen	0901	number			
2. Others	0909	number			
10. Coconuts 10					
1. Nuts	1001	number			
2 . Coconut milk powder	1002	grams			
11. Condiments 11					
1. Dried chilllies	1101	grams			
2. Chilly powder	1102	grams			
3. Red onions	1103	grams			
4. Bombey onions	1104	grams			
5. Garlic	1105	grams			
6. Maldive fish	1106	grams			
7. Pepper	1107	grams			

* ***The estimated Value of items which are consumed from home grown / freely received should be included in col. 6***

11

Item	Code	Unit	Consumed		Home grown / Freely received * (Rs.)
			Qty.	Value (Rs.)	
1	2	3	4	5	6
8. Turmeric /Turmeric powder	1108	grams			
9. Curry power (sarakku)	1109	grams			
10. Green chillies	1110	grams			
11. Limes	1111	number			
12. Cumin seeds	1112				
13. Fennel seeds	1113				
14. Coriander	1114				
15. Mathe seeds	1115				
16. Mustard	1116				
17. Goraka	1117				
18. Tamarind	1118				
19. Cinnamon	1119				
20. Salt	1120				
21. Rampe/Karapincha	1121				
22. Ginger	1122				
23. Vinegar	1123				
24. Other	1129				
12 .Other foods 12					
1. Sauce	1201				
2. Marmite/Vegemite	1202				
3. Soup cubes	1203				
4. Lime pickle	1204				
5. Chutney	1205				
6. Canned fruits	1206				
7. Others	1209				

Item	Code	Units	Consumed		Home grown / Freely received * (Rs.)
			Qty.	Value (Rs.)	
1	2	3	4	5	6
13 .Milk and Milk foods 13	1 Bottle =750 ml				
1. Cow milk	1301	ml			
2. Goat milk	1302	ml			
3. Sterilized milk	1303	ml			
4. Curd	1304				
5. Yoghurt / Moru	1305				
6. Condensed milk	1306	ml			
7. Milk powder	1307	ml			
8. Infant milk powder	1308	ml			
9. Butter	1309	ml			
10. Margarine	1310	ml			
11. Cheese	1311	ml			
12. Milk packets (liquid)	1312	ml			
13. Other	1319				
14 . Fats and oils 14	1 Bottle = 750 ml				
1. Coconut oil	1401	ml			
2. Vegetable oil	1402	ml			
3. Gingerley oil	1403	ml			
4. Ghee	1404	ml			
5. Other	1409	ml			
15. Sugar, Juggery & Treacle 15					
1. Sugar	1501	grams			
2. Juggery	1502	grams			
3. Treacle	1503	ml			
4. Honey	1504				
5. Other	1509				

* ***The estimated value of items which are consumed from home grown / freely received should be included in col.6***

12

Item	Code	Unit	Consumed		Home grown /Freely re-ceived *
			Qty.	Value (Rs.)	(Rs.)
1	2	3	4	5	6
16. Fruits 16					
1. Banana	1601	number			
2. Pineapple	1602	number			
3. Papaw	1603	number			
4. Mangoes	1604	number			
5. Apple	1605	number			
6. Avacado	1606	number			
7. Wood apple	1607	number			
8. Orange	1608	number			
9. King coconut / Kurumba	1609	number			
10. Grapes	1610	number			
11. Other	1619				
Dried fruits					
1. Dates	1651	grams			
2. Cajunuts	1652	grams			
3. Ground nuts	1653	grams			
4. Plums	1654	grams			
5. Other	1659	grams			
17. Confectionery & short eats 17					
1. Jam	1701	grams			
2. Ice -Cream	1702				
3. Chocolates	1703	grams			
4. Toffees	1704	number			
5. Biscuits	1705	grams			
6. Jelly packets	1706				
7. Snacks	1707	grams			
8. Cake	1708	grams			
9. Kewum/Kokis etc.	1709	number			
10. Musket/Kaludodal etc.	1710	grams			
11. Thalabola/Rulan etc.	1711	number			
12. Cutlets/Patis/Wade Pestry etc.	1712	number			
13. Palmyrah Products	1713	number			
14. Other	1719				
18. Beverages (non alcoholic) 18					
1. Tea dust/leaves	1801	grams			
2. Coffee powder/seeds	1802	grams			
3. Soft drinks	1803				
4. Fruit drink/cordial	1804				
5. Saruwath	1805	number			
6. Milk tea	1806	number			
7. Plain tea	1807	number			
8. Milk coffee	1808	number			
9. Coffee	1809	number			
10. Ice packets	1810	number			
11. Gruel	1811	number			
12. Botteled water	1812	number			
13. Other	1819				

* ***The estimated value of items which are consumed from home grown / freely received should be included in Col. 6***

13

Item	Code	Unit	Consumed		Home grown /Freely received * (Rs.)
			Qty.	Value (Rs.)	
1	2	3	4	5	6
19.Liquor,Drugs and Tobacco 19					
1. Toddy	1901				
2. Arrack	1902				
3. Kassippu	1903				
4. Beer/Stout	1904				
5. Whisky/Brandy	1905				
6. Gin	1906				
7. Wine	1907				
8. Other liquor	1908				
9. Drugs (Ganja,Heroin)	1909				
10. Cigarettes	1910	number			
11. Cigars	1911	number			
12. Beedi	1912	number			
13. Pipe tobacco	1919				
14. Betel leaves	1920	number			
15. Arecanuts	1921	number			
16. Tobacco chewing	1922				
17. Chunam	1923				
18. Bulathvita/ Saravita	1924	number			

**** The estimated value of items which are consumed from home grown / freely received should be included in col. 6***

14

4. 2 Household expenditure on Housing, Fuel & Light, Non-durable goods, Services & Consumer durable for main household.

Item	Code	Unit	Main household Qty.	Main household Value (Rs.)	Freely received / Received as a gift(Rs.)
1	2	3	4	5	6
20 1. Housing (Average per month)					
Rent (Estimated rent of owner occupied house should be included in col. 5 and col.6	2001				
Taxes	2002				
Water bills	2003				
21 2.1 Fuel & Light (Average per month)					
Electricity	2101				
Solar power	2102				
Kerosene oil	2103	ml			
Fire wood (Purchased)	2104	kg			
Fire wood (free)	2105				
L.P.Gas	2106	grams			
2.2 Fuel & Light (Last month)					
Normal blubs	2107	number			
Energy saving blubs	2108	number			
Matches	2109	number			
Candles	2110	number			
Batteries (used for lighting purposes)	2111	number			
Others	2119				

Item	Code	Unit	Main household Qty.	Main household Value (Rs.)	Freely received/ Received as a gift(Rs.)
1	2	3	4	5	6
22 3. Personal care expense (Last month)					
Toilet Soap	2201				
Tooth paste	2202				
Tooth brush	2203				
Cosmetics powder	2204				
Perfumes	2205				
Face Cream / Lotions	2206				
Hair oil, Cream and hair dye	2207				
Lipstick / Cutex	2208				
Shaving equipments	2209				
Hair cut/ Hair dressing & Shaving charges	2210				
Others	2219				
23 4 . Health Expenses (Last month)					
Fees to Private medical practitioners (Included cost of medicine)	2301				
Ayurvedic consultation fees (Included cost of medicine)	2302				
Consultation fees to specialist	2303				
Payments to medical laboratories for test analysis. etc.	2304				
Payment to private hospitals and Nursing homes	2305				
Purchased of medical and pharmaceutical products	2306				
Spectacles	2307				
Others	2309				

Main household - include all membes who usually live in this household and exclude servants and boarders
The estimated value of items which are received freely / fringe benefit / relief payment/ donation / as a gift should be included in col. 6

Item	Code	unit	Main house hold		Freely received/ Received as a gift
			Qty.	Value (Rs.)	
1	2	3	4	5	6
5. Transport (Last month) Transport fees 24					
Train /Bus (Other than schooling)	2401				
Van (Other than schooling)	2402				
Taxi car/Threewheeler (Other than schooling)	2403				
Transport charges for schoolig / Preschooling	2404				
Ships & air lines	2405				
Other	2409				
Maintanence of private vehicals					
Petrol & other fuel	2411				
Oil	2412				
Tyre,tube & spare parts	2413				
Repair charges of spare parts accessaries	2414				
Licence & Insurance (Average per month)	2415				
Other	2419				
6. Communication (Last month) 25					
Postal & Telegraph Charges	2501				
Telephone Charges (Domestic) (Average per month)	2502				
Telephone charges (Moble)	2503				
Telephone Charges (Taken from outside)	2504				
E-Mail / Internet fees	2505				
Other	2509				

Item	Code	unit	Main house hold		Freely received/ Received as a gift
			Qty.	value (Rs.)	
1	2	3	4	5	6
7. Education (Last month) 26					
Exercise books/ Stationeries	2601				
News papers / Magazines(related to edu.)	2602				
School text books	2603				
School facility fees (Government)	2604				
School fees (private)	2605				
Tution fees	2606				
Boarding fees	2607				
Fees for higher education (External degree / Post graduate degree, Diploma etc.)	2608				
Fees for professional & technical education	2609				
Fees for pre school	2610				
Other	2619				
8. Recreation, entertainment & Cultural activities (Last month) 27					
Cinamas/ Drama / Video films	2701				
Books / News papers / Magazines	2702				
Lotteries / Bettings	2703				
Excursions / Pilgrimages	2704				
Arts / Music / Dancing	2705				
Sports	2706				
Pets / Aquarium	2707				
Toys	2708				
Maintanence of radio recivers, television & Video deck	2709				
Other	2719				

Main household - include all membes who usually live in this household & exclude servants & boarders

The estimated value of items which are received freely / fringe benefit / relief payment/ donation / as a gift should be included in col. 6

16

Item	Code	unit	Main household		Freely received/ Received as a gift
			Qty.	Value (Rs.)	
1	2	3	4	5	6
9. Non-durable household goods (Last month) 28					
Washing Soap	2801				
Washing powder	2802				
Mosquito Coils	2803				
Detergents	2804				
Insecticiders (for home use)	2805				
Other	2809				
10. Household Services (Last month) 29					
Laundry Charges	2901				
Grinding charges	2902				
Wages to Servants / Chauffers	2903				
Charges for Day Care Centers	2904				
Charges for Elders home etc	2905				
Payments for other household sevices	2909				
11. Clothing & Textiles (Last 6 month) 30					
Trousers	3001	number			
Shirts	3002	number			
Sarees	3003	number			
Tea Shirts	3004	number			
Sarongs / Vetties	3005	number			
Frocks	3006	number			

Itme	Code	Unit	Main household		Freely received/ Received as a gift
			Qty.	Value (Rs.)	
1	2	3	4	5	6
Skirts / Blouses	3007	number			
Banians	3008	number			
Men's Under wears	3009	number			
Women's Under wears	3010	number			
Housecoat / Kimonas / Night dresses	3011	number			
Brassieres	3012	number			
Socks / Stockings	3013	number			
Ties	3014	number			
Handkerchiefs	3015	number			
Children's dresses	3016	number			
School uniforms	3017	number			
Towels	3018	number			
Bed sheets	3019	number			
Pillow cases	3020	number			
Mosquito nets	3021	number			
other	3029	number			
Materials purchased in meters (Last 6 month)					
For Clothing	3041	cm			
For School Uniforms	3042	cm			
For Curtains	3043	cm			
Other	3049	cm			
Tailoring Chrges (Last 6 month)					
For clothing	3061				
For school uniform	3062				

Main household - include all membes who usually live in this household & exclude servants boarders.
The estimated value of items which are received freely / fringe benefit / relief payment/ donation / as a gift should be included in col. 6

Item	Code	Unit	Main household		Freely received/ Received as a gift
			Qty.	Value (Rs)	
1	2	3	4	5	6
12.Footwear & Other personal effect 31 (Last 6 months)					
Shoes	3101	number			
Sandals / Slippers	3102	number			
Umbrellas	3103	number			
Hand bags/ Travelling bags	3104	number			
Hats / Helmets	3105	number			
Rain coats	3106	number			
Repair charges for shoes, bags, umbrellas etc.	3107				
Others	3109	number			
13. Durable household goods 32 (Last 6 months)					
Glass ware	3201	number			
Plasticware	3202	number			
Aluminiumware	3203	number			
Clayware / pots	3204	number			
Cane ware	3205	number			
Cutlery	3206	number			
Other	3207	number			

Item	Code	Unit	Main household		Freely received/ Received as a gif
			Qty.	Value (Rs)	
1	2	3	4	5	6
Brooms / Brushes / Ekels	3208	number			
Mats/ Pillows	3209	number			
Rugs / Carpets	3210	number			
Mettress	3211	number			
Torches	3212	number			
Thermos flasks	3213	number			
Twines / Ropes	3214	number			
Others	3219	number			
14.Durablehouseholdgoods 33 (Last 12 month)					
Chairs (Plastic / Wood)	3301	number			
Tables	3302	number			
Settees	3303	number			
Almyrahs	3304	number			
Beds	3305	number			
Other furniture	3306	number			
Clocks / Wrist watches	3307	number			
Cookers (Kerosene / gas / Electric)	3308	number			
Mixingmachines,Grinders, Beaters	3309	number			
Irons	3310	number			
Refrigerators	3311	number			
Fans	3312	number			
Sewing machines	3313	number			

Main household - include all membes who usually live in this household & exclude servants & boarders.
The estimated value of items which are received freely / fringe benefit / relief payment/ donation / as a gift should be included in col. 6

Item	Code	Unit	Main household		Freely recived/ Received as a gift
			Qty.	Value (Rs.)	
1	2	3	4	5	6
Radio, Record players & Tape recorders	3314	number			
Television / Video decks	3315	number			
Washing machines	3316	number			
Musical Instruments	3317	number			
Bicycles	3318	number			
Motor cycles/ Scooters	3319	number			
Motor cars / Vans	3320	number			
Prams / Gocarts / Tricycles	3321	number			
Garderning tools & equipments	3322	number			
Cameras & Projectors	3323	number			
Computers	3324	number			
Jewelleries	3325	number			
Telephones / Cellular phones	3326	number			
Cookers , Toasters, Hot plates, Heaters etc.	3327	number			
Other	3329	number			
Durable goods purchased in instalments	3339				

Item	Code	Unit	Main household		Freely recived/ Received as a gift
			Qty.	Value (Rs.)	
1	2	3	4	5	6
34 15. Other expenses (Last month)					
Provident fund / W & O.P. fund	3401				
Contribution to trade unions, Thriff Societies etc..	3402				
Insurance /Agrahara	3403				
Other savings (Including seettu)	3404				
Payment for debits	3405				
Income tax	3406				
Other	3409				
35 16. Other adhoc (rarely) expenses (Last 12 month)					
Expenditure on weddings/ funerals for family members	3501				
Social activities/ Ceremonies	3502				
Litigation	3503				
Gift, Donation, Similar transfers	3504				
Maintenance / Repairing (Houses)	3505				
Purchased properties Houses	3506				
Loans given	3507				
Other	3509				

Main household - include all membes who usually live in this household & exclude servants & boarders
The estimated value of items which are received freely / fringe benefits / relief payment / donation / as a gift should be included in col. 6

4.3 Expenditure on Housing, Fuel & Light, Non-durable goods, Services & Consumer durables for boarders & domestic servants.

Are there Boarders & / or Domestic Servants in this house hold Yes [1] No [2] → Go to section 5

Name of boarders & domestic servants who has spent money on any of the following columns 3 -15	Serial number as given in Column 1 of section 1	Food purchased outside the household .(Last week)	Fuel & Light (Last month)	Clothing & Textiles (Last 6 Months)	Non -durable household goods (Last month)	Household services. (Last month)	Personal care & Health expenses (Last month)	Transport & Communication (Last month)	Recreation, Entertainments Education & Cultural activities etc.	Durable household goods (Last 12 months)	Boarding fees paid (Last month)	Amount sent to family/ Parents (Last month)	Amount transfered as savings provident funds, W & O.P., taxes,donations etc. (Last month)	Miscellaneous (Last month)
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1														
2														
3														
4														
5														
6														
7														

If any boarder or domestic servant given in column 1 of section 1, not spent any of the above Please not down here.

20

Section 5 - Income

5.1 Income from paid employments / During last weeks / Last calendar month

Did any of the household members (usually live) engage as an employee during last 4 weeks / last calendar month

Yes [1] No [2] → go to section 5.2

List all household members (usually live in this household) including boarders & domestic servants who were paid employees during last four weeks / last calendar month	Serial number as given in Column 1 of section 1		Principal / Secondary occupation Main 1 Secondary 2		Last month Wages / Salaries (Rs.)	Last month Tips, Commissions Overtime pay etc. (Rs.)	Last 12 months Bonus, Arrears Payment (Rs.)
1	2		3		4	5	6
1			1				
			2				
2			1				
			2				
3			1				
			2				
4			1				
			2				
5			1				
			2				
6			1				
			2				
7			1				
			2				
8			1				
			2				

5.2 Income from agricultural activities - (Paddy, Other seasonal crops)

Did any of the household members cultivate paddy, other seasonal crops as an employer or own account worker for sale and / or household consumption during last cultivation year. **

Yes [1] No [2] → 5 go to section 5.3

List all household members (usually live) including boarders & domestic servants who have cultivated paddy & / or other seasonal crops as employers or own account workers.	Serial number as given in column 1 of section 1	Paddy / Other seasonal crops (Describe)	Code +	Cultivated area			** Lastcultivationyear		
							Value of output (Rs.)	Value of amount Consumed by this household (Rs.)	* Cost (Rs.)
1	2	3	4	5	6	7	8	9	10
1.									
2.									
3.									
4.									
5.									
6.									
7.									
8.									
9.									
10.									

\+ Code for column 3

Paddy.....1
Chillies. 2
Onions 3
Vegetables.....4
Cereals.....5
Yams..... .....'6
Tobacco.....7
Other.... 9

** Last Cultivation year-

Last two cultivation seasons

(yala / Maha or Maha / yala)

*

Hired Labour			
Seeds			
Fertilizer			
Chemicals			
Transport / Tractor			
Other			
Total			

22

5.3 Income from other agricultural activities.

Did any of the household members engage in other agricultural activities /Livestock as an employer or own account worker for sale during last 12 months

Yes [1] No [2] → go to section 5.4

List all household members (usually live) including boarders & domestic servants who have engaged any other agricultural activites / Livestock as employers or own account workers	Serial number as given in column 1 of section 1	Agricultural product (Describe)	+ Code	Lastmonth/Averagepermonth				
				Cultivated area			Value of Out put (Rs.)	* Cost (Rs.)
1	2	3	4	5	6	7	8	9
1								
2								
3								
4								
5								
6								
7								
8								
9								
10'								

*

Hired labour		
Seeds		
Fertilizer		
Chemicals		
Transport/ Tractor charges		
Other		
Total		

\+ Codes for column

Tea , Rubber 01
Coconuts 02
Coffee, Pepper Betel etc.....03
Banana / Fruits 04
Meat 05
Fish 06
Eggs 07
Milk.....08
Other food items.....09
Horticulture.....10
Other.....19

Not necessary for cultivated area under the codes from 05-10 in column 4.

5.4 Income From Non- agricultural activities

Did any of the household members engage in any non- agricultural activities as an employer or own account worker for sale during last calendar month.

Yes [1] No [2] → go to section 5.5.1

List all household members (usually live) including boarders & domestic servants who have engaged Non- agricultural activities as employers or own account workers.	Serial number as given in coloumn 1 of section 1	Economic activity (Describe)	+ Code	Lastcalendarmonth	
				Value of output (Rs.)	* Cost (Rs.)
1	2	3	4	5	6
1.					
2					
3					
4					
5					
6					
7					
8					
9					
10					

\+ Codes for coloumn 3

Mining & Quarrying. 1
Manufacturing.... 2
Construction. 3
Trade.....4
Transport. 5
Guest house, restaurants, bars , hotels etc.....6
Other services.. 7

*

Hired labour				
Inputs				
Transport				
Other				
Total				

5.5.1. Income from Other cash receipt during last calendar month / last 12 month

yes [1] No [2] → go to 5.5.2

List all household members (usually live) including boarders & servants who have received any other cash receipt.	Income form last calendar month								Last 12 month	
	Serial number as given in column 1 of section 1		Pension Payment (Rs.)	Disability/ relief payments (Rs.)	Rents from properties/ boarding fees etc. (Rs.)	Samurdhi (Rs.)	Dividents / Interests (Rs.)	Other (Rs.)	Current remittance & transfers	
									outside the country (Rs.)	within the Country (Rs.)
1	2		3	4	5	6	7	8	9	10
1										
2										
3										
4										
5										
6										
7										

5.5.2 Income by Chance or adhoc gains during last 12 months

Yes [1] No [2] → go to section 6

List all household members (usually live) excluding boarders & servants who have received income by chance or rarely.	Serial number as given in colomn 1of section 1	Loans taken from banks/ money lenders etc. (Rs.)	Sales of assets (Land , house, jewellary) (Rs.)	Withdrawal from savings & bank deposits. (Rs.)	Income receives from births, deaths, marriages / welfare society, etc (Rs.)	Seettu / Repayments of loans given (Rs.)	Compersation health / insurance etc (Rs.)	Other lottery & other adhoc gains (Rs.)
1	2	3	4	5	6	7	8	9
1								
2								
3								
4								
5								
6								
7								

Section 6A - Inventry of durable goods

Items	Code	Have - 1 Have not - 2
Household Equipments		
Radio / Cassette player	5001	
Television	5002	
V.C.D. / D.V.D.	5003	
Sewing machines	5004	
Washing machines	5005	
Refrigerators	5006	
Cookers (Gas, Kerosene, Electric)	5007	
Electric fans	5008	
Telephone (Domestic)	5009	
Telephone (mobile)	5010	
Personal Computers	5011	
For Transport purpose		
Bicycles	5012	
Motor cycles / Scooters	5013	
Three wheelers	5014	
Motor cars / Vans	5015	
Bus / Lorry	5016	
Agricultural Equipments		
Tractors (Two wheel)	5017	
Tractors (Four wheel)	5018	
Sprayers	5019	
Threshers	5020	
Water pumps (Only for agricultural purpose)	5021	
Fishing Equipments		
Boats	5022	
Fishing nets	5023	

26

Section 6 B . Debts (Excluding boarders & Servants)

	Debited to	Have any of the household members debited to the followings Yes -1 go to col. 3 No -2	Amount to pay (Rs.)
	1	**2**	**3**
1	Banks (Government / Private)		
2	Finance Companies / Leasing Companies		
3	Own Place of work (Departments ,Boards , Private Companies etc.)		
4	Money Lenders		
5	Retail outlets		
6	Sales of assets (Land , houses, jewelleries etc.)		
7	Durable goods purchased in instalments		
9	Other (Specify)		

Section 7 - Access to facilities

Section 7.1

Serial No.	Place of Facilits	Distance from your house to this closest facility (km)	Time taken from your house to this closest facility (Minutes)
01	Bus Halt (nearest)		
02	Pre school / Montessori		
03	Primary School		
04	Maha Vidyalaya/ M.M.V. / Jathika Pasala		
05	Base Hospital / Teaching Hospital /District Hospital		
06	Maternity Home		
07	Government Dispensary		
08	Private Dispensary		
09	Clinic (Maternity / Infancy)		
10	M.O.H. office		
11	M.C. / U.C. / P. S.		
12	Divisional Secretariat office		
13	G.N.Office		
14	Post office / Sub post office		
15	Bank (Govt. / Private)		
16	Agrarian Service Center		

Section -7.2

	Yes	No
1. Do you have electricity supply (main line) nearby your area.	1	2
2. Do you have telephone facilities in your area.	1	2
3. Do you have pipe borne line (main line) nearby your area	1	2

Section 8 - Housing Information

1 . Type of Structure

Single house	1
Flat	2
Attached house /Annex	3
Line room / Row house	4
Slum / Shanty	5
Other (specify)	9

2. Number of bed rooms

3. Total floor area (sq. feet)

less than 100	1
100 - less than 250	2
250 - less than 500	3
500 - less than 750	4
750 or more than 750	5

4. Principal materials of Construction

wall

Brick	1
Cabook	2
Cement block	3
Pressed soil blocks	4
Mud	5
Plank / Metal sheet	6
Cadjan / Palmyrah	7
Other (specify)	9

(B) Floor

Cement	1
Terrazo / Tile	2
Mud	3
Other (specify)	9

(C) Roof

Tile	1
Asbestos	2
Concrete	3
Metal sheet	4
Cadjan / Palmyrah / Straw	5
Other (specify)	9

5. Tenure

Constructed / Purchased by an occupant	1
Inherited	2
Freely received / Received as a gift	3
Compensated	4
Rent free (Employer / Other)	5
Relief Payment (Employer / Other)	6
Rent / Lease	7
Encroached	8
Other (specify)	9

6. A Main source of drinking water

Protected well within premises	1
Protected well outside premises	2
Unprotected well	3
Tube well	4
Tap within unit / Premises (main line)	5
Tap outside premises (main line)	6
Stream water collected & distributed by pipe lines	7
River / Tank / Streams	8
Other (specify)	9

B. Distance to take source of drinking water (Meters)

Within premises	1	
Outside premises	2	

C. Was there enough water to drink, bath & wash during last year

		Yes	No
1.	Drink	1	2
2.	Bath /wash	1	2

7. A. Availability of toilet

Exclusive for the household	1
Sharing with another household	2
Public convenience	3
None	4 → go to quection 8

B. Type of Toilet

Water seal	1
Pour flush	2
Pit	3
Other (specify)	9

8. Disposal of Garbage

Collected by garbage truck	1
Buried / Burned	2
Process for fertilizer	3
Dumped within premises	4
Dumped / Throw away out side premises	5
Other (specify)	9

9. Principal Type of Lighting

Kerosene	1
Electricity	2
Solar energy	3
Generator / Battery	4
Other (specify)	9

10. Principal Type of cooking fuel

Fire wood	1
Gas	2
Kerosene	3
Electricity	4
Saw dust / Paddy husk	5
Other (specify)	9

11. Did any of the household member collect fire wood during last month

Yes [1]

No [2] → go to Q.12

(B) Distance & place of collecting firewood

		(Meters)
Own land	1	
Forest	2	
Other land.....	3	

12. (A) Was this housing unit effected by any natural disaster during last year

Yes [1]

No [2] - go to section 9

(B) Nature of disaster

	Yes	No
1. Flood	1	2
2. Drought	1	2
3. Earthslip.....	1	2
4. Civil disturbance.....	1	2
5. Attacked by wild animals.....	1	2
6. Other (specify)	1	2

Section 9 - Agricultural Holdings & Livestock

1. Does any member / s of your household own any agricultural land / lands

Yes [1] No [2] → go to Q 3

2. Land area

	Owned			Cultivated. (Rent, lease, joined own, ande etc.)		
	1			**2**		
	A.	R.	P.	A.	R.	P.
1. Paddy land						
2. High land						
3. Land area with occupied housing units						

3. Livestock (owned)

Category	Yes 1 No 2	* Number of livestock (code)
	1	**2**
1.Cattle / Buffaloes		
2. Goats / Sheep		
3. Swine (pigs)		
4. Poultry		
5. Other		

*

(1) 5 or less than 5

(2) 6 to 10

(3) 11 to 50

(4) More than 50

Attachment 2

Sri Lanka HIES 2009/10 Data Dictionary

NOTE:

1. For data set from R02 to R24, the top 9 variables, that is, DISTRICT, SECTOR, DSD, MONTH, PSU, SAMPLE_N, SERIAL_NO, NHH and RESULT are skipped in the following sheets, because they are the same as those in R01.

2. The next variables are added to data files.

Variable	Description
ID	Household identifier
PID	Individual identifier, if the data set includes person number
wt	Adjusted weight for the resampled micro data set

3. The next item labels are revised.

File	Item label (old)	Item label (revised)
R10	Roots	Roods
R13	Acrs	Acres
	Roots	Roods

4. R25 and R26 are compiled by the author, as written in Chapter 6, 7 and 8. It might include errors. Please use it at your own risk.

5. The sequence of the following sheets is by the type value, but not by the file name.

6. As for crop code in R13 (Section 5.3), it is not clear why “tea” and “rubber” are allocated in the same code 1, regardless of totally different crops.

Level Label / Record Label	Level Name / Record Name	Type Value	Req	Max	Rec Len
Sri Lanka HIES2009 Questionnaire	HIES2009_QUEST				
Section - 1 - Demography	SEC_1_DEMOGRAPHIC	01	Yes	40	43
Section - 2 - School Education	SEC_2_SCHOOL_EDUCATION	02	No	40	36
Section - 3 - Health	SEC_3_HEALTH	03	No	40	34
Section - 4.1 - Food expenditure	SEC_4_1_FOOD_EXP	04	No	240	41
Section - 4.2 - Non food	SEC_4_2_NONFOOD	05	No	175	41
Section - 4.3 - Is Boaders Servent	SEC_4_3_IS_BOADERS	06	No	1	18
Section - 4.3 - Boaders Servent	SEC_4_3_BOARDERS	07	No	40	89
Section - 5.1 - Is employment Income	SEC_5_1_IS_EMP_INCOME	08	No	1	18
Section - 5.1 - Employment Income	SEC_5_1_EMP_INCOME	09	No	40	40
Section - 5.2 - Is aggricultural Income	SEC_5_2_IS_AGRI_INCOME	10	No	1	18
Section - 5.2 - Aggricultural Income	SEC_5_2_AGRI_INCOME	11	No	40	53
Section - 5.3 - Is other aggricultural income	SEC_5_3_IS_OTHER_AGRI_INCOME	12	No	1	18
Section - 5.3 - Other aggricultural income	SEC_5_3_OTHER_AGRI_INCOME	13	No	40	45
Section - 5.4 - Is non_agri income	SEC_5_4_IS_NON_AGRI_INCOME	14	No	1	18
Section - 5.4 - Non agri income	SEC_5_4_NON_AGRI_INCOME	15	No	40	38
Section - 5.5.1 - Is other income	SEC_5_5_1_IS_OTHER_INCOME	16	No	1	18
Section - 5.5.1 - Other income	SEC_5_5_1_OTHER_INCOME	17	No	40	72
Section - 5.5.2 - Is windfall income	SEC_5_5_2_IS_WINDFALL_INCOME	18	No	1	18
Section - 5.5.2 - Windfall income	SEC_5_5_2_WINDFALL_INCOME	19	No	40	75
Section - 6A - Durable goods	SEC_6A_DURABLE_GOODS	20	No	1	40
Section - 6B - Debtness	SEC_6_B_DEBTNESS	21	No	1	81
Section - 7 - Basic facilities	SEC_7_BASIC_FACILITIES	22	No	1	100
Section - 8 - Housing	SEC_8_HOUSING	23	No	1	50
Section - 9 - Land and animal	SEC_9_LAND_ANIMAL	24	No	1	63

R01 Record: Section - 1 - Demography

Item Label		Item Name	Start	Len	Data Type	Item Type	Occ	Dec	Char	Zero Fill
(record type)			1	2	AN	I	1	0	No	No
District	(id)	DISTRICT	3	2	N	I	1	0	No	Yes
11:13 Western										
21:23 Central										
31:33 Southern										
41:45 Northern										
51:53 Eastern										
61:62 North-western										
71:72 North-central										
81:82 Uva										
91:92 Sabaragamuwa										
(Blank)Blank										
Sector	(id)	SECTOR	5	1	N	I	1	0	No	Yes
1:3										
(Blank)										
DS division	(id)	DSD	6	2	N	I	1	0	No	Yes
3:99										
(Blank)										
Month	(id)	MONTH	8	2	N	I	1	0	No	Yes
1:12										
(Blank)										
PSU Number	(id)	PSU	10	3	N	I	1	0	No	Yes
1:300										
(Blank)										
Sample No.	(id)	SAMPLE_N	13	2	N	I	1	0	No	Yes
1:10										
(Blank)										
Household Serial No.	(id)	SERIAL_NO	15	1	N	I	1	0	No	Yes
1:9										
(Blank)										
Number of households	(id)	NHH	16	1	N	I	1	0	No	Yes
1:9										
(Blank)										

Result Code (id)	RESULT	17	1	N	I	1	0	No	Yes
1 Completed									
2 Deferred									
3 Not competent respondant at home									
4 Refused									
5 Household is temporarily closed									
6 Household is demolished / Vacant									
7 Rejected									
9 Other (Specify)									
(Blank)									
1Person	PERSON_SERIAL_NO	18	2	N	I	1	0	No	Yes
1:45									
(Blank)									
3Relationship	RELATIONSHIP	20	1	N	I	1	0	No	Yes
1 Head									
2 Spouse									
3 Son/Daughter									
4 Parents									
5 Other Ralative									
6 Domestic Servent									
7 Boarders									
9 Other									
(Blank)									
4Sex	SEX_LIVING	21	1	N	I	1	0	No	Yes
1 Male/Local									
2 Female/Abroad									
(Blank)									
5Birth Year	BIRTH_YEAR	22	2	N	I	1	0	No	Yes
0:99									
(Blank)									
5Birth Month	BIRTH_MONTH	24	2	N	I	1	0	No	Yes
1:12									
(Blank)									
6Age	AGE	26	2	N	I	1	0	No	Yes
0:99									
(Blank)									

7Race	ETHNICITY	28	1	N	I	1	0	No	Yes
1 Sinhala 2 Sri Lanka Tamil 3 Indian Tamil 4 Sri Lanka Moors 5 Malay 6 Burger 9 Othe r (Blank)									
8Reliqion	RELIGION	29	1	N	I	1	0	No	Yes
1 Buddhist 2 Hindu 3 Islam 4 Roman Ctholic/Christian 9 Other (Blank)									
9Current Education	CURR_EDUCATION	30	1	N	I	1	0	No	Yes
1 Pre school 2 School 3 University 4 Other educational institution 5 Vocational/ Technical institution 6 Pending results G.C.E. (O.L/A.L) 7 Does not attend (Blank)									
10Education	EDUCATION	31	2	N	I	1	0	No	Yes
0 Studingin Grade 1 1 Passed Grade 1 2 Passeed Grade2 3 Passed Grade 3 4 Passed Grade 4 5 Passed Grade 5 6 Passed Grade 6 7 Passed Grade 7 8 Passed Grade 8 9 Passed Grade 9 10 Passed Grade 10 11 Passe G.C .E (O/L) or Equivalent 12 Passed Grade 12 13 Passed G.C.E.(A/L) or Equivalent 14 Passed GAQ/GSQ 15 Passed Degree									

16 Passed Post Graduate Dgree/Diploma
17 Special Education
19 No
Schooling
(Blank)

11 Marital — MARITAL_STATUS — 33 1 N I 1 0 No Yes
1 Never Married
2 Married
3 Widowed
4 Divorced
5 Separated
(Blank)

12 Usual Activity — MAIN_ACTIVITY — 34 1 N I 1 0 No Yes
1 Employed
2 Unemployed
3 Student
4 Household work
5 Unable / Too old to work
6 Othe
r
(Blank)

13 Main Occupation — MAIN_OCCUPATION — 35 4 N I 1 0 No Yes (Blank)
(na)

14 Industry — INDUSTRY — 39 4 N I 1 0 No Yes (Blank)
(na)

15 Employment Status — EMPLOYMENT_STATUS — 43 1 N I 1 0 No Yes
1 Government employee
2 Semi government employee
3 Private sector employee
4 Employer
5 Own account worker
6 Unpaid family worker
(Blank)

R02 Record: Section - 2 - School Education

Item Label	Item Name	Start	Len	Data Type	Item Type	Occ	Dec	Char	Zero Fill
2 Sr. No	R2_PERSON_SERIAL Yes 1:40	18	2	N	I	1	0	No	
(Blank)									
3 School Education 1 Currently attending school 2 Never attend school 3 Attended school in the past (Blank)	R2_SCHOOL_EDUCATION	20	1	N	I	1	0	No	Yes
4CurrentGrade 1 Grade 1 2 Grade 2 3 Grade 3 4 Grade 4 5 Grade 5 6 Grade 6 7 Grade 7 8 Grade 8 9 Grade 9 10 Grade 10 11 Grade 11 12 Grade 12 13 Grade 13 19 Not relevant (Blank)	GRADE_THIS_YEAR	21	2	N	I	1	0	No	Yes
5GradePreviouse	GRADE_LAST_YEAR	23	2	N	I	1	0	No	Yes
1 Grade 1 2 Grade 2 3 Grade 3 4 Grade 4 5 Grade 5 6 Grade 6 7 Grade 7 8 Grade 8 9 Grade 9 10 Grade 10 11 Grade 11 12 Grade 12 13 Grade 13 19 Not relevant (Blank)									
6 DistanceToSchool	DISTANCE	25	2	N	I	1	0	No	

Yes 0:99	DistanceToSchool								
(Blank)									
7Transport	TRANSPORT_MEDIUM	27	1	N	I	1	0	No	Yes
1 Walk									
2 Bicycle									
3 Motor bicycle / Three-wheeler/Car									
4 School hiring Van/ Bus									
5 By Bus									
6 By Train									
9 Other									
(Specify) (Blank)									
8TimeTo School	TIME_TO_SCHOOL	28	3	N	I	1	0	No	Yes
0:120 TimeTo									
School (Blank)									
9WhyNeverSchooling	NOSCHOOLING_REASON	31	1	N	I	1	0	No	Yes
1 School too far away									
2 Financial problems									
3 Had to help house keeping activities /Family Business									
4 Disability / Illness									
5 Civil disturbance									
6 Not willing to attend / Poor academic progress									
7 Incompletion of									
9 Other (specify)									
(Blank)									
10 ReasonsNotGoingSchool	REASON_NOT_GOING	32	1	N	I	1	0	No	Yes
1 Further schooling not available or too far away									
2 Finalcial Problems									
3 Had to help house keeping activities / Family business									
4 Disability / Illness									
5 Civil disturbance									
6 Not willing to attend / Poor academic progress									
7 Pending results (G.C.E. (O/L) / G.C.E. (A/L)									
8 Completed A/L									
9 Other									
11WhenStopSchool	WHEN_STOP_SCHOOLING	33	4	N	I	1	0	No	Yes
0:9999 WhenStopSchool									
(Blank)									

R03 Record: Section - 3 - Health

Item Label	Item Name	Start	Len	Data Type	Item Type	Occ	Dec	Dec Char	Zero Fill
2 Person Sr. 1:40 (Blank)	R3_PERSON_SERIAL	18	2	N	I	1	0	No	Yes
3 Did attend hospital 1:2 Did attend hospital (Blank)	DID_ATTEND_HOSPITAL	20	1	N	I	1	0	No	Yes
4 Reason for attending 1 Treatment for illness 2 Treatment for injury 3 Medical checkup/ Consultation 4 Immunization 5 Treatment for infectious diseases(Injection etc) 9 Other (specify) (Blank)	REASON_HOSPITAL	21	1	N	I	1	0	No	Yes
5 Did stay in hospital 1:2 (Blank)	IS_STAY_HOSPITAL	22	1	N	I	1	0	No	Yes
6 Reason for hospitalise 1 Treatment for illness 2 Treatment for injury 3 Operation / Surgey 4 Child delivery 5 Treatment for infectious diseases 6 An accident 9 Other (specify) (Blank)	REASON_STAY	23	1	N	I	1	0	No	Yes
7 Is ill or dissable 1:2 (Blank)	IS_ILL_DISABLE	24	1	N	I	1	0	No	Yes
8 Illnee or dissability 1 Heart Conditions /Diseases 2 Blood pressure 3 Diabetics 4 Asthma 5 Epilepsy 6 Cancer 7 Stomach diseases / Gastritis 8 Diseases related to Eyes 9 Arthritis 10 Mental retardation 11 Haemorrhoids 12 Catarrh	WHAT_ILL_DISABLE	25	2	N	I	1	0	No	Yes

13 Severe headache
14 Naturally Disabled
15 Disabled by an accident
19 Other
(Blank)

9 Is employmet caused	IS_EMPL_REASON	27	1	N	I	1	0	No Yes

1:2
(Blank)

10 Duration of illness disa	DURATION_YEARS	28	2	N	I	1	0	No Yes

0:99 Duration of illness disa
(Blank)

10 Months duration	DURATION_MONTHS	30	2	N	I	1	0	No Yes

0:12 Months duration
(Blank)

11 Did absent for normal	IS_ABSENT_ACT	32	1	N	I	1	0	No Yes

1:2
(Blank)

12 Duration absent days	DAYS_ABSENT	33	2	N	I	1	0	No Yes

0:99 Duration absent days
(Blank)

R04 Record: Section - 4.1 - Food expenditure

Item Label	Item Name	Start	Len	Data Type	Item Type	Occ	Dec	Dec Char	Zero Fill
2. Code Yes 101:1924	CODE	18	4	N	I	1	0	No	
4. Quantity 0:999999 (Blank)	QUANTITY	22	6	N	I	1	0	No	Yes
5. ValueVALUE 0:9999999 (Blank)	28	7	N	I	1	0	No	Yes	
6. Inkind value 0:9999999 (Blank)	INKIND_VALUE	35	7	N	I	1	0	No	Yes

R05 Record: Section - 4.2 - Non food

Item Label	Item Name	Start	Len	Data Type	Item Type	Occ	Dec	Char	Zero Fill
2. Code Yes 2001:3509 (Blank)	NF_CODE	18	4	N	I	1	0	No	
4. Quantity 0:999999 (Blank)	NF_QUANTITY	22	6	N	I	1	0	No	Yes
5. ValueNF_VALUE 0:9999999 (Blank)	28	7	N	I	1	0	No	Yes	
6. Inkind value 0:9999999 (Blank)	NF_INKIND_VALUE	35	7	N	I	1	0	No	Yes

R07 Record: Section - 4.3 - Is Boaders Servent

Item Label	Item Name	Start	Len	Data Type	Item Type	Occ	Dec	Dec Char	Zero Fill
Is Boarders and Domestic Servents	IS_BOARDERS_SERVENTS	18	1	N	I	1	0	No	Yes
1:2									
(Blank)									

R06 Record: Section - 4.3 - Boaders Servent

Item Label	Item Name	Start	Len	Data Type	Item Type	Occ	Dec	Dec Char	Zero Fill
2 Person 1:40 SerNo (Blank)	COL_2	18	2	N	I	1	0	No	Yes
3 Food 2:70000 (Blank)	COL_3	20	5	N	I	1	0	No	Yes
4 Fuel and lite 2:5000 (Blank)	COL_4	25	5	N	I	1	0	No	Yes
5 Clothing 10:7000 (Blank)	COL_5	30	5	N	I	1	0	No	Yes
6 Non durable goods 10:5000 (Blank)	COL_6	35	5	N	I	1	0	No	Yes
7 HH services 10:5000 (Blank)	COL_7	40	5	N	I	1	0	No	Yes
8 Personal effects 10:5000 (Blank)	COL_8	45	5	N	I	1	0	No	Yes
9 Transport 10:5000 (Blank)	COL_9	50	6	N	I	1	0	No	Yes
10 Ent., Edu and cult 10:3000 (Blank)	COL_10	56	5	N	I	1	0	No	Yes
11 Durable goods 10:5000 (Blank)	COL_11	61	6	N	I	1	0	No	Yes
12 Bording fees 10:10000 (Blank)	COL_12	67	5	N	I	1	0	No	Yes
13 Send to family 100:10000 (Blank)	COL_13	72	6	N	I	1	0	No	Yes
14 Savins etc 50:20000 (Blank)	COL_14	78	6	N	I	1	0	No	Yes

15 Other COL_15 84 6 N I 1 0 No Yes
10:6000
(Blank)

R09 Record: Section - 5.1 - Is employment Income

Item Label	Item Name	Start	Len	Data Type	Item Type	Occ	Dec	Dec Char	Zero Fill
Is employment income 1:2 (Blank)	IS_EMPLOYMENT_INCOME	18	1	N	I	1	0	No	Yes

R08 Record: Section - 5.1 - Employment Income

Item Label	Item Name	Start	Len	Data Type	Item Type	Occ	Dec	Dec Char	Zero Fill
2Scr.No	SERIAL_NO_SEC_1	18	2	N	I	1	0	No	Yes
	1:40								
	(Blank)								
3Principal/Second	PRI_SEC	20	1	N	I	1	0	No	Yes
1:2									
	(Blank)								
4 WagesSalary	WAGES_SALARIES	21	6	N	I	1	0	No	Yes
50:300000									
	(Blank)								
5 Allowences	ALLOWENCES	27	7	N	I	1	0	No	Yes
100:100000									
	(Blank)								
6 Bonus	BONUS	34	7	N	I	1	0	No	Yes
10:250000									
	(Blank)								

R11 Record: Section - 5.2 - Is aggricultural Income

Item Label	Item Name	Start	Len	Data Type	Item Type	Occ	Dec	Dec Char	Zero Fill
Is agricultural income	IS AGRICULTURAL INCOME	18	1	N	I	1	0	No	Yes
1:2									
(Blank)									

R10 Record: Section - 5.2 - Aggricultural Income

Item Label	Item Name	Start	Len	Data Type	Item Type	Occ	Dec	Dec Char	Zero Fill
2Person 1:40 (Blank)	SER_NO_SEC_5_2	18	2	N	I	1	0	No	Yes
4Code 1 Paddy 2 Chillies 3 Onions 4 Vegetables 5 Cereals 6 Yams 7 Tobacco 9 Other (Blank)	SEAS_CROPS_CODE	20	1	N	I	1	0	No	Yes
5Acres 0:25 (Blank)	ACR_5_2	21	3	N	I	1	0	No	Yes
6Roods 0:3 (Blank)	RT_5_2	24	1	N	I	1	0	No	Yes
7Perchs 0:39 (Blank)	P	25	2	N	I	1	0	No	Yes
8Output value 50:1000000 (Blank)	OUTPUT_5_2	27	9	N	I	1	0	No	Yes
9Household consumption 10:50000 (Blank)	HH_CONSUMPTION	36	9	N	I	1	0	No	Yes
10Input value 10:1000000 (Blank)	INPUT_5_2	45	9	N	I	1	0	No	Yes

R12 Record: Section - 5.3 - Is other aggricultural income

Item Label	Item Name	Start	Len	Data Type	Item Type	Occ	Dec	Dec Char	Zero Fill
Is other agrri income	IS OTHER AGRRI INCOME	18	1	N	I	1	0	No	Yes
1:2									
(Blank)									

R13 Record: Section - 5.3 - Other aggricultural income

Item Label	Item Name	Start	Len	Data Type	Item Type	Occ	Dec	Dec Char	Zero Fill
2Person 1:40 (Blank)	SER_NO_SEC_5_3	18	2	N	I	1	0	No	Yes
4Other crop code 1 Tea , Rubber 2 Coconuts 3 Coffee, Pepper Betel etc 4 Banana / Fruits 5 Meat 6 Fish 7 Eggs 8 Milk 9 Other food items 10 Horticulture 19 Other	SEASONAL_CROP	20	2	N	I	1	0	No	Yes
5Acres 0:25 (Blank)	ACRES_5_3	22	3	N	I	1	0	No	Yes
6Roods 0:3 (Blank)	ROOTS_5_3	25	1	N	I	1	0	No	Yes
7Perchs 0:39 (Blank)	PERCHS_5_3	26	2	N	I	1	0	No	Yes
8Output value 50:1000000 (Blank)	OUTPUT_5_3	28	9	N	I	1	0	No	Yes
9Input value 50:1000000 (Blank)	INPUT_5_3	37	9	N	I	1	0	No	Yes

R14 Record: Section - 5.4 - Is non_agri income

Item Label	Item Name	Start	Len	Data Type	Item Type	Occ	Dec	Dec Char	Zero Fill
Is non agri income	IS_NON_AGRI_INCOME	18	1	N	I	1	0	No	Yes
1:2									
(Blank)									

R15 Record: Section - 5.4 - Non agri income

Item Label	Item Name	Start	Len	Data Type	Item Type	Occ	Dec	Dec Char	Zero Fill
2 Person 1:40 (Blank)	SERIAL_5_4	18	2	N	I	1	0	No	Yes
4 Non agri code 1 Mining & Quarrying. 2 Manufacturing 3 Construction 4 Trade 5 Transport 6 Guest house, restaurants, bars/hotels etc 7 Other services (Blank)	NON AGRI	20	1	N	I	1	0	No	Yes
5 Output value 50:500000 (Blank)	OUTPUT_5_4	21	9	N	I	1	0	No	Yes
6 Input value 50:500000 (Blank)	INPUT_5_4	30	9	N	I	1	0	No	Yes

R16 Record: Section - 5.5.1 - Is other income

Item Label	Item Name	Start	Len	Data Type	Item Type	Occ	Dec	Dec Char	Zero Fill
Is Other Income	IS_OTHER_INCOME	18	1	N	I	1	0	No	Yes
1:2									
(Blank)									

R17 Record: Section - 5.5.1 - Other income

Item Label	Item Name	Start	Len	Data Type	Item Type	Occ	Dec	Dec Char	Zero Fill
2Person 1:40 (Blank)	SERIAL_5_5_1	18	2	N	I	1	0	No	Yes
3Pention 100:60000 (Blank)	PENSION	20	6	N	I	1	0	No	Yes
4Disability 50:7000 (Blank)	DISABILITY_AND_RELIEF	26	5	N	I	1	0	No	Yes
5Rent 50:250000 (Blank)	PROPERTY_RENTS	31	7	N	I	1	0	No	Yes
6Samurdhi 50:3500 (Blank)	SAMURDHI	38	5	N	I	1	0	No	Yes
7Profits 10:50000 (Blank)	DIVIDENDS	43	7	N	I	1	0	No	Yes
8Other 10:250000 (Blank)	OTHER	50	7	N	I	1	0	No	Yes
9Foreign 100:700000 (Blank)	ABROAD	57	8	N	I	1	0	No	Yes
10Local 50:600000 (Blank)	LOCAL	65	8	N	I	1	0	No	Yes

R18 Record: Section - 5.5.2 - Is windfall income

Item Label	Item Name	Start	Len	Data Type	Item Type	Occ	Dec	Dec Char	Zero Fill
Is windfall income 1:2 (Blank)	IS_WINDFALL_INCOME	18	1	N	I	1	0	No	Yes

R19 Record: Section - 5.5.2 - Windfall income

Item Label	Item Name	Start	Len	Data Type	Item Type	Occ	Dec	Dec Char	Zero Fill
2Person 1:40 (Blank)	PERSON_5_5_2	18	2	N	I	1	0	No	Yes
3Loans WD 100:1000000 (Blank)	LOANS	20	8	N	I	1	0	No	Yes
4Selling pawning 500:500000 (Blank)	PAWNING_SELLING	28	8	N	I	1	0	No	Yes
5Deposits 100:600000 (Blank)	DEPOSITS_PENSIONS_EPF	36	8	N	I	1	0	No	Yes
6Income from Welfare 10:1000000 (Blank)	INCOME_WELFARE	44	8	N	I	1	0	No	Yes
7Rcvd loans seettu 100:100000 (Blank)	SITTU_DEBTS	52	8	N	I	1	0	No	Yes
8Compenstions 100:500000 (Blank)	COMPENSATION	60	8	N	I	1	0	No	Yes
9Other Lottery_games 100:500000 (Blank)	OTHER_LOTTERY	68	8	N	I	1	0	No	Yes

R21 Record: Section - 6A - Durable goods

Item Label	Item Name	Start	Len	Data Type	Item Type	Occ	Dec	Dec Char	Zero Fill
Radio	RADIO	18	1	N	I	1	0	No	Yes
1:2 Radio / Cassette player									
(Blank)									
TV	TV	19	1	N	I	1	0	No	Yes
1:2 Television									
(Blank)									
VCD	VCD	20	1	N	I	1	0	No	Yes
1:2 V.C.D. / D.V.D.									
(Blank)									
Sewing Mechine	SEWING_MECHINE	21	1	N	I	1	0	No	Yes
1:2 Sewing machines									
(Blank)									
Washing mechine	WASHING_MECHINE	22	1	N	I	1	0	No	Yes
1:2 Washing machines									
(Blank)									
Fridge	FRIDGE	23	1	N	I	1	0	No	Yes
1:2 Refrigerators									
(Blank)									
Cookers	COOKERT	24	1	N	I	1	0	No	Yes
1:2 Cookers (Gas, Kerosene, Electric)									
(Blank)									
Electric Fans	ELECTRIC_FANS	25	1	N	I	1	0	No	Yes
1:2 Electric fans									
(Blank)									
Telephone	TELEPHONE	26	1	N	I	1	0	No	Yes
1:2 Telephone (Domestic)									
(Blank)									
Telephone Mobile	TELEPHONE_MOBILE	27	1	N	I	1	0	No	Yes
1:2 Telephone (mobile)									
(Blank)									
Computers	COMPUTERS	28	1	N	I	1	0	No	Yes
1:2 Personal Computers									
(Blank)									
Bicycle	BICYCLE	29	1	N	I	1	0	No	Yes
1:2 Bicycles									
(Blank)									
Motor Bicycle	MOTOR_BICYCLE	30	1	N	I	1	0	No	Yes
1:2 Motor cycles / Scooters									
(Blank)									
Three wheeler	THREE_WHEELER	31	1	N	I	1	0	No	Yes

1:2 Three wheelers								
(Blank)								
Motor car van	MOTOR_CAR_VAN	32	1	N	I	1	0	No	Yes
1:2 Motor cars / Vans								
(Blank)								
Bus lorry	BUS_LORRY	33	1	N	I	1	0	No	Yes
1:2 Bus / Lorry								
(Blank)								
Tractor 2 wheel	TRACTOR_2_WHEEL	34	1	N	I	1	0	No	Yes
1:2 Tractors (Two wheel)								
(Blank)								
Tractor 4 wheel	TRACTOR_4_WHEEL	35	1	N	I	1	0	No	Yes
1:2 Tractors (Four wheel)								
(Blank)								
Pesticider	PESTICIDER	36	1	N	I	1	0	No	Yes
1:2 Sprayers								
(Blank)								
Paddy blower	PADDY_BLOWER	37	1	N	I	1	0	No	Yes
1:2 Threshers								
(Blank)								
Water pumps	WATER_PUMPS	38	1	N	I	1	0	No	Yes
1:2 Water pumps								
(Blank)								
Boats	BOATS	39	1	N	I	1	0	No	Yes
1:2 Boats								
(Blank)								
Fishing nets	FISHING_NETS	40	1	N	I	1	0	No	Yes
1:2 Fishing nets								
(Blank)								

R20 Record: Section - 6B - Debtness

Item Label	Item Name	Start	Len	Data Type	Item Type	Occ	Dec	Dec Char	Zero Fill
Banks	BANKS	18	1	N	I	1	0	No	Yes
1:2 Banks									
(Blank)									
Bank amount	BANK_AMOUNT	19	7	N	I	1	0	No	Yes
100:1000000 Bank Amount									
(Blank)									
Finance	FINANCE	26	1	N	I	1	0	No	Yes
1:2 Finance									
(Blank)									
Finance amount	FINANCE_AMOUNT	27	7	N	I	1	0	No	Yes
100:1000000 Finance Amount									
(Blank)									
Employer	EMPLOYER	34	1	N	I	1	0	No	Yes
1:2 Own Place of work									
(Blank)									
Employer amount	EMPLOYER_AMOUNT	35	7	N	I	1	0	No	Yes
100:300000 Employer amount									
(Blank)									
Lender	LENDER	42	1	N	I	1	0	No	Yes
1:2 Money Lenders									
(Blank)									
Lender amount	LENDER_AMOUNT	43	7	N	I	1	0	No	Yes
100:100000 Money Lenders Amount									
(Blank)									
Retail shops	RETAIL_SHOPS	50	1	N	I	1	0	No	Yes
1:2 Retail outlets									
(Blank)									
Retail shop amount	RETAIL_SHOP_AMOUNT	51	7	N	I	1	0	No	Yes
100:3000 Retail outlets Amount									
(Blank)									

Pawning	PAWNING	58	1	N	I	1	0	No	Yes
1:2 Pawning									
(Blank)									
Pawning amount	PAWNING_AMOUNT	59	7	N	I	1	0	No	Yes
100:100000 Pawning Amount									
(Blank)									
Instalment goods	INSTALMENT_GOODS	66	1	N	I	1	0	No	Yes
1:2 Durable goods purchased in instalments									
(Blank)									
Instalement amount	INSTALEMENT_AMOUNT	67	7	N	I	1	0	No	Yes
100:100000 Instalement amount									
(Blank)									
Other	OTHER_DEBTS	74	1	N	I	1	0	No	Yes
1:2 Other									
(Blank)									
other amount	OTHER_AMOUNT	75	7	N	I	1	0	No	Yes
100:100000 Other mount									
(Blank)									

R22 Record: Section - 7 - Basic facilities

Item Label	Item Name	Start	Len	Data Type	Item Type	Occ	Dec	Dec Char	Zero Fill
Bus halt 0:99 Bus Halt (Blank)	BUS_HALT	18	2	N	I	1	0	No	Yes
Bus halt time 1:120 BHTime (Blank)	BUS_HALT_TIME	20	3	N	I	1	0	No	Yes
Pre school 0:99 Pre School (Blank)	PRE SCHOOL	23	2	N	I	1	0	No	Yes
Pre school time 1:120 Pre SchTime (Blank)	PRE_SCHOOL_TIME	25	3	N	I	1	0	No	Yes
Primery school 0:99 Primary School (Blank)	PRIMERY_SCHOOL	28	2	N	I	1	0	No	Yes
Primery school time 1:120 Primary SchoolTime (Blank)	PRIMERY_SCHOOL_TIME	30	3	N	I	1	0	No	Yes
Secondery school 0:99 Secondery School (Blank)	SECONDERY_SCHOOL	33	2	N	I	1	0	No	Yes
Sec school time 1:120 SeconderySchoolTime (Blank)	SEC SCHOOL TIME	35	3	N	I	1	0	No	Yes
Hospital 0:99 Base Hospital (Blank)	HOSPITAL	38	2	N	I	1	0	No	Yes
Hospital time 1:120 Base Hospital Time (Blank)	HOSPITAL_TIME	40	3	N	I	1	0	No	Yes

Matrenity home	MATRENITY_HOME	43	2	N	I	1	0	No	Yes
0:99 Maternity Home									
(Blank)									
Maternity home time	MATERNITY_HOME_TIME	45	3	N	I	1	0	No	Yes
1:120 Maternity HomeTime									
(Blank)									
Gov dispensary	GOV_DISPENSARZ	48	2	N	I	1	0	No	Yes
0:99 Government Dispensary									
(Blank)									
Gov dispensary time	GOV_DISPENSARY_TIME	50	3	N	I	1	0	No	Yes
1:120 Government DispensaryTime									
(Blank)									
Private dispensary	PRIVATE DISPENSARY	53	2	N	I	1	0	No	Yes
0:99 Private Dispensary									
(Blank)									
Private dispensary time	PRIVATE_DISPENSARY_TIME	55	3	N	I	1	0	No	Yes
1:120 Private DispensaryTime									
(Blank)									
Maternity clinic	MATERNITY_CLINIC	58	2	N	I	1	0	No	Yes
0:99 MatenityClinic									
(Blank)									
Maternity clinic time	MATERNITY_CLINIC_TIME	60	3	N	I	1	0	No	Yes
1:120 MaternityclinicTime									
(Blank)									
DMO	DMO	63	2	N	I	1	0	No	Yes
0:99 DMO									
(Blank)									
DMO time	DMO_TIME	65	3	N	I	1	0	No	Yes
1:120 DMOTime									
(Blank)									
MCUCPC	MCUCPC	68	2	N	I	1	0	No	Yes
0:99 MCUCPC									
(Blank)									
MCUCPC time	MCUCPC_TIME	70	3	N	I	1	0	No	Yes
1:120 MCUCPCTime									
(Blank)									
DS office	DS_OFFICE	73	2	N	I	1	0	No	Yes
0:99 DS Office									
(Blank)									
DS office time	DS_OFFICE_TIME	75	3	N	I	1	0	No	Yes
1:120 DSOfficeTime									
(Blank)									

GN office	GN_OFFICE	78	2	N	I	1	0	No	Yes
0:99 GNOffice									
(Blank)									
GN office time	GN_OFFICE_TIME	80	3	N	I	1	0	No	Yes
1:120 GNOfficeTime									
(Blank)									
Post office	POST OFFICE	83	2	N	I	1	0	No	Yes
0:99 Post office									
(Blank)									
Post office time	POST_OFFICE_TIME	85	3	N	I	1	0	No	Yes
1:120 PostOfficeTime									
(Blank)									
Bank	BANK	88	2	N	I	1	0	No	Yes
0:99 Bank									
(Blank)									
Bank time	BANK_TIME	90	3	N	I	1	0	No	Yes
1:120 BankTime									
(Blank)									
Agri office	AGRI_OFFICE	93	2	N	I	1	0	No	Yes
0:99 AgriOffice									
(Blank)									
Agri office time	AGRI_OFFICE_TIME	95	3	N	I	1	0	No	Yes
1:120 AgriOfficeTime									
(Blank)									
Is power lines near	IS_POWER_LINES_NEAR	98	1	N	I	1	0	No	Yes
1:2									
(Blank)									
Is tel lines near	IS_TEL_LINES_NEAR	99	1	N	I	1	0	No	Yes
1:2									
(Blank)									
Is water service near	IS WATER SERVICE NEAR	100	1	N	I	1	0	No	Yes
1:2									
(Blank)									

R23 Record: Section - 8 - Housing

Item Label	Item Name	Start	Len	Data Type	Item Type	Occ	Dec	Dec Char	Zero Fill
1 Structure	STRUCTURE	18	1	N	I	1	0	No	Yes
1 Single house									
2 Flat									
3 Attached house /Annex									
4 Line room / Row house									
5 Slum / Shanty									
9 Other (specify)									
(Blank)									
2 Bed rooms	BED_ROOMS	19	1	N	I	1	0	No	Yes
0:9 Bedroom									
(Blank)									
3 Area	AREA	20	1	N	I	1	0	No	Yes
1 less than 100									
2 100 - less than 250									
3 250 - less than 500									
4 500 - less than 750									
5 750 or more than 750									
(Blank)									
4 Walls	WALLS	21	1	N	I	1	0	No	Yes
1 Brick									
2 Cabook									
3 Cement block									
4 Pressed soil blocks									
5 Mud									
6 Plank / Metal sheet									
7 Cadjan / Palmyrah									
9 Other (specify)									
(Blank)									

4 Floor	FLOOR	22	1	N	I	1	0	No	Yes
1 Cement 2 Terrazo / Tile 3 Mud 9 Other (specify) (Blank)									
4 Roof	ROOF	23	1	N	I	1	0	No	Yes
1 Tile 2 Asbestos 3 Concrete 4 Metal sheet 5 Cadjan / Palmyrah / Straw 9 Other (specify) (Blank)									
5 Tenure	TENURE	24	1	N	I	1	0	No	Yes
1 Constructed / Purchased by an occupant 2 Inherited 3 Freely received / Received as a gift 4 Compensated 5 Rent free 6 Relie payment 7 Rent / Lease 8 Encroached 9 Other (specify) (Blank)									
6a Drinking water	DRINKING_WATER	25	1	N	I	1	0	No	Yes
1 Protected well within premises 2 Protected well outside premises 3 Unprotected well 4 Tube well 5 Tap within unit / Premises (main line) 6 Tap outside premises (main line) 7 Stream water collected & distributed by pipe lines 8 River / Tank / Streams 9 Other (specify) (Blank)									
6b1 Own Other water	OWN_WATER	26	1	N	I	1	0	No	Yes
0:2 OwnOtherWater (Blank)									
6b2 Water distance	WATER_DISTANCE	27	4	N	I	1	0	No	Yes
0:9999 WaterDistance (Blank)									

6c1 Water sufficency	WATER_SUFFICENCY	31	1	N	I	1	0	No	Yes
1:2 Drink									
(Blank)									
6c2 Other water sufficency	OTHER_WATER_SUFFICENCY	32	1	N	I	1	0	No	Yes
1:2 Bath/Wash									
(Blank)									
7a Availability of Tioilet	AVAILABILITY_TIOILET	33	1	N	I	1	0	No	Yes
1 Exclusive for the household									
2 Sharing with another household									
3 Public Convenience									
4 None									
7b Toilet type	TOILET_TYPE	34	1	N	I	1	0	No	Yes
1 Water seal									
2 Pour flush									
3 Pit									
4 Other (specify)									
(Blank)									
8 Garbage dumping	GARBAGE_DUMPING	35	1	N	I	1	0	No	Yes
1 Collected by garbage truck									
2 Buried / Burned									
3 Process for fertilizer									
4 Dumped within premises									
5 Dumped / Throw away out side premises.									
9 Other (specify)									
(Blank)									
9 Lite source	LITE_SOURCE	36	1	N	I	1	0	No	Yes
1 Kerosene									
2 Electricity									
3 Solar energy									
4 Generator / Battery									
9 Other (specify)									
(Blank)									
10 Cooking fuel	COOKING_FUEL	37	1	N	I	1	0	No	Yes
1 Fire wood									
2 Gas									
3 Kerosene									
4 Electricity									
5 Saw dust / Paddy husk									
9 Other (specify)									
(Blank)									

11a Is collect firewood 1:2 (Blank)	IS_COLLECT_FIREWOOD	38	1	N	I	1	0	No	Yes
11b Fire wood ownForestOther 0:3 Fire wood ownForestOther (Blank)	FIRE_WOOD_OWN	39	1	N	I	1	0	No	Yes
11b FWD distance 0:9999 FWDDistance (Blank)	OTHER_DISTANCE	40	4	N	I	1	0	No	Yes
12a Natural calamity 1:2 Natural calamity (Blank)	NATURAL_CALAMITY	44	1	N	I	1	0	No	Yes
12b1 Flooding 1:2 Floodind (Blank)	FLOODING	45	1	N	I	1	0	No	Yes
12b2 Drought 1:2 Drought (Blank)	DROUGHT	46	1	N	I	1	0	No	Yes
12b3 Land slides 1:2 Land slides (Blank)	LAND_SLIDES	47	1	N	I	1	0	No	Yes
12b4 Civil unrest 1:2 Civil unrest (Blank)	CIVIL_UNREST	48	1	N	I	1	0	No	Yes
12b5 Wild animal 1:2 Wild animal (Blank)	WILD_ANIMAL	49	1	N	I	1	0	No	Yes
12b6 Other calamity 1:2 Other calamity (Blank)	OTHER_CALAMITY	50	1	N	I	1	0	No	Yes

R24 Record: Section - 9 - Land and animal

Item Label	Item Name	Start	Len	Data Type	Item Type	Occ	Dec	Dec Char	Zero Fill
Is agriland owner 1:2 (Blank)	IS_AGRILAND_OWNER	18	1	N	I	1	0	No	Yes
Paddy own acr 0:25 Paddy Own Acr (Blank)	PADDY_OWN_ACR	19	3	N	I	1	0	No	Yes
Paddy own rt 0:3 Paddy own rt (Blank)	PADDY_OWN_RT	22	1	N	I	1	0	No	Yes
Paddy own perch 0:39 Paddy own perch (Blank)	PADDY_OWN_PERCH	23	2	N	I	1	0	No	Yes
Paddy other acr 0:25 Paddy other acr (Blank)	PADDY_OTHER_ACR	25	3	N	I	1	0	No	Yes
Paddy other rt 0:3 PaddyOtherrt (Blank)	PADDY OTHER RT	28	1	N	I	1	0	No	Yes
Paddy other perch 0:39 Paddy other perch (Blank)	PADDY_OTHER_PERCH	29	2	N	I	1	0	No	Yes
Land own acr 0:25 Land own acr (Blank)	LAND_OWN_ACR	31	3	N	I	1	0	No	Yes
Land own rt 0:3 LandOwnrt (Blank)	LAND OWN RT	34	1	N	I	1	0	No	Yes
Land own perch 0:39 Land own perch (Blank)	LAND_OWN_PERCH	35	2	N	I	1	0	No	Yes

Land other acr	LAND_OTHER_ACR	37	3	N	I	1	0	No	Yes	
0:25 Land other acr										
(Blank)										
Land other rt	LAND_OTHER_RT	40	1	N	I	1	0	No	Yes	
0:3 LandOtherrt										
(Blank)										
Land other perch	LAND_OTHER_PERCH	41	2	N	I	1	0	No	Yes	
0:39 Land other perch										
(Blank)										
Home own acr	HOME_OWN_ACR	43	3	N	I	1	0	No	Yes	
0:25 HomeOwnAcr										
(Blank)										
Home own rt	HOME_OWN_RT	46	1	N	I	1	0	No	Yes	
0:3 Home own rt										
(Blank)										
Home own perch	HOME_OWN_PERCH	47	2	N	I	1	0	No	Yes	
0:39 Home own perch										
(Blank)										
Home other acr	HOME_OTHER_ACR	49	3	N	I	1	0	No	Yes	
0:25 Home other acr										
(Blank)										
Home other rt	HOME_OTHER_RT	52	1	N	I	1	0	No	Yes	0:3
Home other rt										
(Blank)										
Home other perch	HOME_OTHER_PERCH	53	2	N	I	1	0	No	Yes	
0:39 Home other perch										
(Blank)										
Cows buffalows	COWS_BUFFALOWS	55	1	N	I	1	0	No	Yes	
1:2 Cows buffalows										
(Blank)										
Cows count	COWS_COUNT	56	1	N	I	1	0	No	Yes	
1:4 Cows count										
(Blank)										
Goats sheeps	GOATS_SHEEPS	57	1	N	I	1	0	No	Yes	
1:2 Goats sheeps										
(Blank)										
Goat count	GOAT_COUNT	58	1	N	I	1	0	No	Yes	
1:4 Goat count										
(Blank)										
Pigs	PIGS	59	1	N	I	1	0	No	Yes	
1:2 Pigs										
(Blank)										

Pigs count	PIGS_COUNT	60	1	N	I	1	0	No	Yes
1:4 Pigs count									
(Blank)									
Chickens	CHICKENS	61	1	N	I	1	0	No	Yes
1:2 Chickens									
(Blank)									
Chicken count	CHICKEN_COUNT	62	1	N	I	1	0	No	Yes
1:4 Chicken count									
(Blank)									
Other	OTHER_ANIMALS	63	1	N	I	1	0	No	Yes
1:2 OtherAnimal									
(Blank)									

☐ **Data dictionary of R25:**

Summary of household income and expenditure at household level

Number of records: 19,958

Variable	Description	Remarks
ID	Household identifier	
wt	Weight	
ttinc	Monthly total household income	= monetary + inkind
monetary	Monthly total monetary income	= wage + crop + livestock + nonagri + other + widfall
wage	Monthly income from paid employments	
crop	Monthly income from agricultural activities	
livestock	Monthly income from other agricultural activities	
nonagri	Monthly income from non-agricultural activities	
other	Monthly income from other cash receipt	
windfall	Monthly income by chance or ad hoc gains	
inkind	Monthly total non-monetary income	
ttexp	Monthly total household expenditure	= food + nfood
food	Monthly household expenditure on food	= Sum of g1 to g18
nfood	Monthly household expenditure on non-food	= Sum of g19 to g35
g1 to g35	Monthly household expenditure by subgroup	COL_3 to COL_15 are already added to the respective subgroups
COL_3 to COL_15	Monthly household expenditure for boarders	

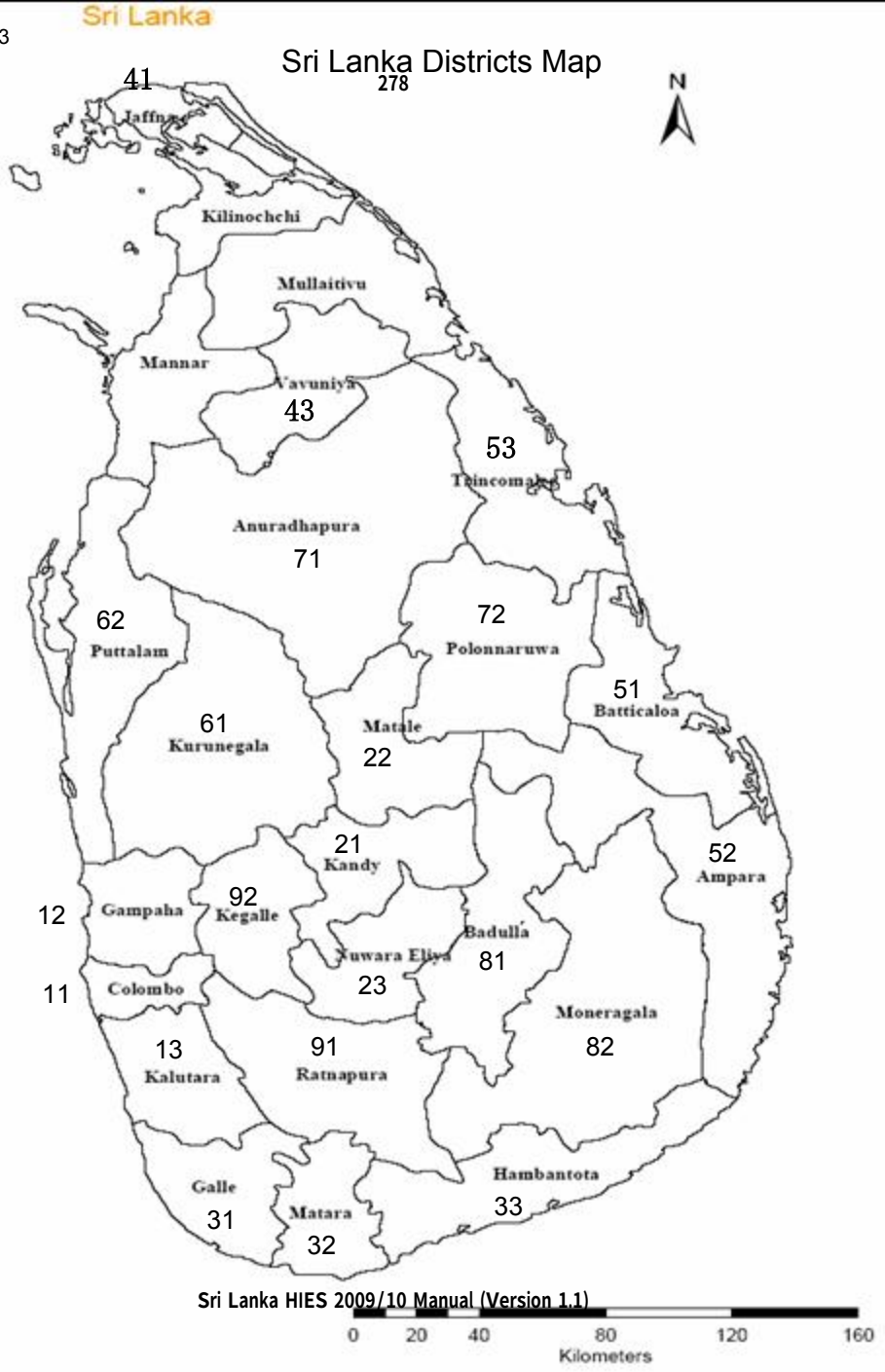
☐ **Data dictionary of R26:**

Summary of household income at individual level

Number of records: 80,872

Variable	Description	Remarks
ID	Household identifier	
PID	Individual identifier	
wt	Weight	
receiver	Is an income receiver?	1. Yes 0. No
relation	Relationship to the household head	
age10	Is age over 10 years old?	1. age >= 10 0. age < 10
ttinc	Monthly total income after transferred	
transferred	Is income transferred to the head?	1. Yes 0. No
befinc	Monthly total income before transferred	
wage	Monthly income from paid employments	
crop	Monthly income from agricultural activities	
livestock	Monthly income from other agricultural activities	
nonagri	Monthly income from non-agricultural activities	
other	Monthly income from other cash receipt	
windfall	Monthly income by chance or ad hoc gains	

Attachment 3



Attachment 4

Administrative structure, and urban, rural and estate

This attachment came from the research paper on “Determinants of the Poverty in Western Province in Sri Lanka”, Research - based Training Program, AY 2010, UNSIAP by Ms. Dilhanie Deepawansa, Department of Census and Statistics, Sri Lanka.

****

3. Data and Methodology

3.1 Introduction

Poverty is not a separate phenomenon that can be studied without considering socio economic factors such as education, health, available of drinking water, electricity etc. The Household Income and Expenditure Survey conducted by Department of Census and Statistics collect data on different dimensions. This study is based on the results of household Income and expenditure survey 2006/07 data.

Figure 3.1 provides a schematic presentation of the administrative structure of Sri Lanka.

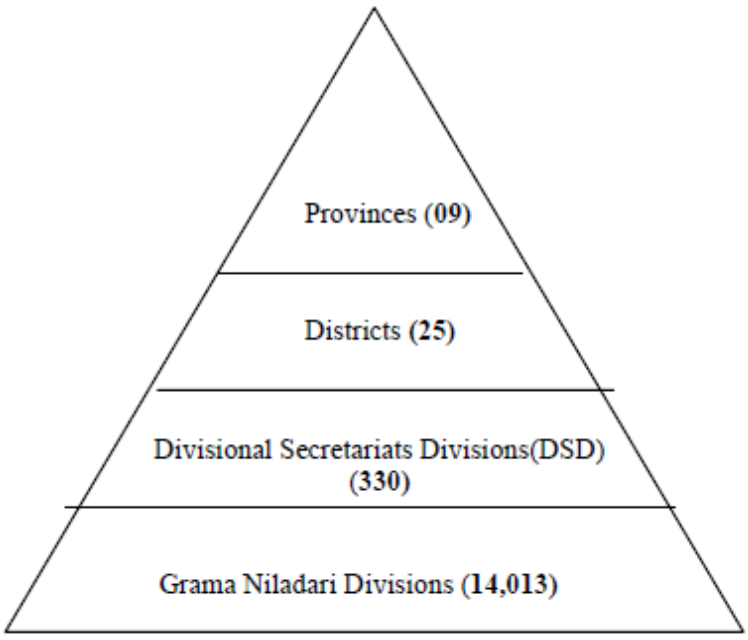


Figure 3.1: Administrative Structure of Sri Lanka

Three kinds of local government authorities function at the lowest level of the administrative structure. These are Municipal council, urban councils and Divisional councils (Pradeshiya sabha). The constituents of the urban and rural sectors are determined by the kind of local governance the lowest-level administrative area units are subject to. The areas governed by Municipal councils are called urban sector and the areas governed by Divisional councils are called rural sector. In Sri Lanka there is a special sector called Estate sector which is demarcated by the plantation areas in the country and they are governed by Divisional Councils. The definition for three sectors is as follows.

Urban Sector : Area governed by either Municipal Council or Urban Council is considered as Urban Sector.
Estate Sector : Plantation areas, which are more than 20 acres of extent and having not less than 10 residential labourers, are considered as estate sector.
Rural Sector : Residential areas, which do not belong to urban sector or estate sector, are considered as rural sector.

The population representation in these three sectors are as follows;

Table 3.1: Distribution of population by sector in Sri Lanka-2006/07

Sector	Population (%)
Urban	14.9
Rural	79.7
Estate	5.4
Total	100.0

****

Annex 03

Estate sector

In addition to the Urban and Rural sectors, Sri Lankan economy has a special sub-economic sector, called Estate sector. . This sector is somewhat unique in its characteristics in all forms, ranging from composition of household units to organization of political establishments. It is also found unique when compared with other general economics around the world. However, the community settlement and its economic behavior and characteristics can be compared with other overseas Indian migrant communities in Fiji, Mauritius, Malaysia, Trinidad and Guyana, as Sri Lankan Estate

Sector community have adapted to the “alien environment and conditions provided within the plantation set up”. Many researchers have found that the estate sector communities of Sri Lanka share many features and similarities with such overseas communities.

British colonization introduced radical economic changes in Sri Lanka which caused the transition from an agriculture-based subsistence economy to a plantation-based market economy. The origin of Estate sector can be traced specially to setting up of tea plantation on the island by the British colonists. In 1867, a Scotsman named James Taylor cleared 19 acres of forest and planted the first seeds. By the mid 1870s, the estate had grown to nearly 100 acres and several other estates were found to have come up.

With the expansion of tea plantation from 1890s, labourers became somewhat scarce, demanding high wages and not to be depended on. To solve the problem of labour scarcity, the British planters decided to supply the estates with immigrant labours from South India. The bulk of the regular labour force was drawn from the Tamil Districts of southern India, mainly Madurai, Tinnelvelly and Tanjore. This process was assisted by the proximity and condition of poverty prevailing in these districts of Southern India. Poverty, various forms of agrestic servitude which had engendered an agricultural labouring class, low wages, and near famine conditions combined to induce segments of southern Indian population to accept employment on the plantation in Sri Lanka. As a result, immigration to Sri Lanka increased more than emigration towards a more permanent and semipermanent

settlement of Tamil labours from South India. More needs to be known, however,

regarding the dynamic of recruitment and the social composition of the Indian Tamil labour force in Sri Lanka. The majority of who probably carried low caste status in their homeland. Unlike village life in South India, caste identity does not emerge as the most dominants or imperative status among the Estate Tamils, but certain aspects of caste are still important for their social organization.

Even today in plantations (Estate sector) not only in tea plantation but also in rubber and coconut, Tamil workers are employed in large numbers.

Geographically and socially isolated, the plantations led the Estate Tamils to emerge as a relatively “closed” community limited to the territorial boundaries of the plantations. Work, housing, and condition of life, including recreation, take place within the same physical and organizational territory of the estate. As a result, the Tamil estate workers’ view of outside world is often limited to the boundaries of estate where they spent their almost or entire life.

Typically a plantation provides rent-free housing, free medicines from its surgeries (dispensary), a maternity ward and midwife(on large estate) estate school, nursery, day care centers for kids, temples, churches, transports etc., for the workers of the estates and their families. Most of the services and needs related to the production and social life are provided within the estate.

In Tea estate, tea pickers are always females and male workers are engaged in manual work and

most of the factory work such as processing tea.

All successive governments elected since independence introduced various economic and social developmental programs to address the issues persisting in the Estate sector. The estate workers did not enjoy the benefits or have the same rights as other citizens in the country. For example, they did not have the right to vote until they are properly registered.

There was also a pressure from various influential social groups to expel them to India as they view the Estate workers as a burden that was imposed on Sri Lankan economy by the British for their economic gains. As a result, hundreds of migrant workers were repatriated while the same number of workers were offered Sri Lankan citizenship under Srima – Shastri fact signed between Sri Lanka and India (1960-65)

The Estate Tamils franchise and citizenship was extended in 1986, which resulted in their participation in political representation that influenced the successive governments to introduce various developmental programs for the Estate Sector. This has had rather positive outcome in their social, educational and political spheres.

The present government has pledged in its election manifesto to provide more radical developmental programs to the Estate workers. Among them, the land ownership, proper housing facilities, promising educational, health, employment and skill development and upward social mobility can be highlighted as remarkable steps overcome miseries experienced by the Estate Workers in general.

****

Attachemnt 5

Classification of industry (4-digit)

111 Growing of cereals (except rice), leguminous crops and oil seeds
112 Growing of rice
113 Growing of vegetables and melons, roots and tubers
114 Growing of sugar cane
115 Growing of tobacco
116 Growing of fibre crops
119 Growing of other non-perennial crops
121 Growing of grapes
122 Growing of tropical and subtropical fruits
123 Growing of citrus fruits
124 Growing of pome fruits and stone fruits
125 Growing of other tree and bush fruits and nuts
126 Growing of oleaginous fruits
127 Growing of beverage crops
128 Growing of spices, aromatic, drug and pharmaceutical crops
129 Growing of other perennial crops
130 Plant propagation
141 Raising of cattle and buffaloes
142 Raising of horses and other equines
143 Raising of camels and camelids
144 Raising of sheep and goats
145 Raising of swine/pigs
146 Raising of poultry
149 Raising of other animals
150 Mixed farming (both crop and livestock)
161 Support activities for crop production
162 Support activities for animal production
163 Post-harvest crop activities (Preparation for primary market)
164 Seed processing for propagation
170 Hunting. Trapping and related service activites
210 Silviculture and other forestry activities
220 Logging
230 Gathering of non-wood forest products
240 support services to forestry
311 Marine Fishing
312 Freshwater fishing
321 Marine aquaculture
322 Freshwater aquaculture
710 Mining of iron ores
729 Mining of other non ferrous metal ores
810 Quarrying of stone, sand and clay
891 Mining of chemical and fertilizer minerals
892 Extraction of peat
893 Extraction of salt
899 Other mining and quarrying n.e.c.
910 Support activities for petroleum and natural gas extraction (Provided on Contract Basis)
990 Support activities for other mining and quarrying
1010 Processing and preserving of meat
1020 Processing and preserving of fish, crustaceans and molluscs
1030 Processing and preserving of fruit and vegetable
1040 Manufacture of vegetable and animal oils and fats
1050 Manufacture of dairy products
1061 Manufacture of grain mill products
1062 Manufacture of starches and starch products
1071 Manufacture of bakery products
1072 Manufacture of sugar
1073 Manufacture of cocoa, chocolate and sugar confectionery
1074 Manufacture of macaroni, noodles, couscous and similar farinaceous products
1075 Manufacture of prepared meals and dishes
1079 Manufacture of other food products n.e.c.
1080 Manufacture of prepared animal feeds
1101 Distilling, rectifying and blending of spirits
1102 Manufacture of wines
1103 Manufcture of malt liquors and malt
1104 Manufacture of soft drinks; production of mineral waters and other bottled waters
1200 Manufacture of tobacco products
1311 Preparation and spinning of textile fibres
1312 Weaving of textiles
1313 Finishing of textile
1391 Manufacture of knitted and crochted fabrics
1392 Manufacture of madeup textile articles, except apparel
1393 Manufacture of carpets and rugs
1394 Manufacture of twine, cordage, rope and cables of coir fibres
1399 Manufacture of other textile n.e.c.
1410 Manufacture of wearing apparel except fur apparel
1420 Manufacture of articles of fur
1430 Manufacture of knitted and crocheted apparel
1511 Tanning and dressing of leather, dressing and dyeing of leather
1512 Manufacture of luggage, handbags and the like saddlery and harness
1520 Manufacture of footwear
1610 Saw-milling and planning of wood
1621 Manufacture of veneer sheets and wood based panels

1622 Manufacture of Builder's carpentry and joinery
1623 Manufacture of wooden containers
1629 Manufacture of other products of wood, manufacture of articles of cork straw and plaiting materials
1701 Manufacture of pulp, paper and paper board
1702 Manufacture of corrugated paper and paper board and of containers of paper and paper board
1709 Manufacture of other articles of paper and paper board
1811 Printing
1812 Service activities related to printing
1820 Reproduction of recorded media
1910 Manufacture of coke oven products
1920 Manufacture of refined petroleum products
2011 Manufacture of basic chemicals
2012 Manufacture of fertilizers and nitrogen compounds
2013 Manufacture of plastics and synthetic rubber in primary forms
2021 Manufacture of pesticides and other agro chemical products
2022 Manufacture of paints, varnishes and similar coatings printing ink and mastics
2023 Manufacture of soap and detergents, cleaning and polishing preparations, perfumes and toilet preparations
2029 Manufacture of other chemical products n.e.c.
2030 Manufacture of man-made fibres
2100 Manufacture of pharmaceuticals, medicinal, chemical and botanical products
2211 Manufacture of rubber tyres and tubes, retreading and rebuilding of rubber tyres
2219 Manufacture of other rubber products
2220 Manufacture of plastic products
2310 Manufacture of glass and glass products
2391 Manufacture of refractory products
2392 Manufacture of clay building materials
2393 Manufacture of other porcelain and ceramic products
2394 Manufacture of cement, lime and plaster
2395 Manufacture of articles of concrete, cement and plaster
2396 Cutting, shaping and finishing of stones
2399 Manufacture of non-metallic mineral products n.e.c.
2410 Manufacture of basic iron and steel
2420 Manufacture of basic precious and other non-ferrous metals
2431 Casting of iron and steel
2432 Casting of non-ferrous metals
2511 Manufacture of structural metal products
2512 Manufacture of tanks, reservoirs and containers of metal
2513 Manufacture of steam generators, except central heading hot water boilers
2520 Manufacture of weapons and ammunition
2591 Forging, pressing, stamping and roll-forming of metal, powder metallurgy
2592 Treatment and coating of metal, machining
2593 Manufacture of cutlery, hand tools and general hardware
2599 Manufacture of other fabricated metal products n.e.c.
2610 Manufacture of electronic components & boards
2620 Manufacture of computers and peripheral equipment
2630 Manufacture of communication equipment
2640 Manufacture of consumer electronics
2651 Manufacture of measuring, testing, navigating and control equipments
2652 Manufacture of watch and clocks
2660 Manufacture of irradiation,electro medical and electrotherapeutic equipment
2670 Manufacture of optical instruments and photographic equipment
2680 Manufacture of magnetic and optical media
2710 Manufacture of electric motors, generators, transformers and electricity distribution & control apparatus
2720 Manufacture of batteries and accumulators
2731 Manufacture of fibre optic cables
2732 Manufacture of other electronic and electric wire and cables
2733 Manufacture of wiring devices
2740 Manufacture of electric lighting equipments
2750 Manufacture of domestic appliance
2790 Manufacture of other electrical equipment
2811 Manufacture of engine and turbines, except aircraft, vehicle and cycle engine
2812 Manufacture of fluid power equipments
2813 Manufacture of pumps, compressors,taps and valves
2814 Manufacture of bearings, gears, gearing and driving elements
2815 Manufacture of ovens, farnaces& furnace burners
2816 Manufacture of lifting and handling equipments
2817 Manufacture of office machinery and equipment (except computers and peripheral equipment)
2818 Manufacture of power-driven hand tools
2819 Manufacture of other general- purpose machinery
2821 Manufacture of agricultural and forestry machinery
2822 Manufacture of metal-forming machinery and machine tools
2823 Manufacture of machinery for metallurgy

2824 Manufacture of machinery for mining,quarrying and construction
2825 Manufacture of machinery for food, beverages and tobacco processing
2826 Manufacture of machinery for textile, apparel and leather production
2829 Manufacture of other special-purpose machinery
2910 Manufacture of motor vehicles
2920 Manufacture of bodies for motor vehicles; Manufacture of trailers & semi-trailers
2930 Manufacture of parts and accessories for motor vehicles
3011 Building of ships and floating structures
3012 Building of pleasure & sporting boats
3020 Manufacture of railway locomotive & rolling stock
3030 Manufacture of air & spacecraft & related machinery
3040 Manufacture of military fighting vehicles
3091 Manufacture of motorcycles
3092 Manufacture of bicycles & invalid carriages
3099 Manufacture of other transport equipment n.e.c.
3100 Manufacture of furniture
3211 Manufacture of jewellery & related articles
3212 Manufacture of imitation jewellery and related articles
3220 Manufacture of musical instruments
3230 Manufacture of sports goods
3240 Manufacture of games & toys
3250 Manufacture of medical & dental instruments & supplies
3290 Other manufacturing n.e.c.
3311 Repair of fabricated metal products
3312 Repair of machinery
3313 Repair of electronic & optical equipment
3314 Repair of electrical equipment
3315 Repair of transport equipment ; except motor vehicles
3319 Repair of other equipment
3320 Installation of industrial machinery & equipment
3510 Electric power generation, transmission and distribution
3520 Manufacture of gas, distribution of gaseous fuels through mains
3530 Steam and air conditioning supply
3600 Water collection,treatment and supply
3700 Sewerage
3811 Collection of non-hazardous waste
3812 Collection of hazardous waste
3821 Treatment and disposal of non-hazardous waste
3822 Treatment and disposal of hazardous waste
3830 Materials recovery
3900 Remediation activities and other waste management services
4100 Construction of Buildings
4210 Construction of Roads and Railways
4220 Construction of Utility Projects
4290 Construction of other Civil Engineering Projects
4311 Demolition
4312 Site Preparation
4321 Electrical Installation
4322 Plumbing, Heat and Air-conditioning Installation
4329 Other Construction installation
4330 Building Completion and Finishing
4390 Other Specialized Construction Activities
4510 Sale of motor vehicles
4520 Maintenance and repair of motor Vehicles
4530 Wholesale and retail sale of motor vehicle part and accessories
4540 Sale, maintenance and repair of motor cycles, and related parts and accessories
4610 Wholesale on a fee or contract basis
4620 Wholesale of agricultural raw materials and live animals
4630 Wholesale of food, beverages and tobacco
4641 Wholesale of textile clothing and Foot wear
4649 Wholesale of other household goods
4651 Wholesale of computers, computers peripheral equipment and software
4652 Wholesale of electronic and telecommunication equipment and parts
4653 Wholesale of agricultural machinery, equipment and supplies
4659 Wholesale of other machinery and equipment
4661 Wholesale of solid, liquid and gaseous fuels and related products
4662 Wholesale of metals and metal ores
4663 Wholesale of construction materials, hardware, plumbing and heating equipment and supplies
4669 wholesale of waste and scrap and other products n.e.c
4690 Non - specialized wholesale trade
4711 Retail sale in non-specialized stores with food, beverages or tobacco predominating
4719 Other retail sale in non-specialized stores
4721 Retail sale of food in specialized stores
4722 Retail sale of beverages in specialized stores
4723 Retail sale of tobacco products in specialized stores
4730 Retail sale of automotive fuel in specialized stores
4741 Retail sale of computers, peripheral units, software and telecommunication equipment in specialized stores
4742 Retail sale of audio and video equipment in specialized stores

4751 Retail sale of textile in specialized stores
4752 Retail sale of hardware, paints and glass in specialized stores
4753 Retail sale of carpets, rugs, wall and floor coverings in specialized stores
4759 Retail sale of electrical household appliances, furniture, lighting equipment and other household articles in specialized stores
4761 Retail sale of books, newspapers and stationary in specialized stores
4762 Retail sale of music and video recording in specialized stores
4763 Retail sale of sporting equipment in specialized stores
4764 Retail sale of games and toys in specialized Stores
4771 Retail sale of clothing, footwear and leather articles in specialized stores
4772 Retail sale of pharmaceutical and medical goods, cosmetic and toilet articles in specialized stores
4773 Other retail sale of new goods in specialized stores
4774 Retail sale of second-hand goods
4781 Retail sale via stalls and markets of food, beverages and tobacco products
4782 Retail sale via stalls and markets of textiles, clothing and footwear.
4789 Retail sale via stalls and markets of other goods
4791 Retail sale via mail order house or via internet
4799 Other retail sale not in stores, stalls or markets.
4911 passenger transport by inter-urban railways
4912 Freight rail transport
4921 Urban and suburban passenger land transport
4922 Other passenger land transport
4923 Freight transport by road
5011 Sea and costal passenger water transport
5012 Sea and coastal freight water transport
5021 Inland passenger water transport
5022 Inland freight water transport
5110 Passenger air transport
5120 Freight air transport
5210 Warehousing and storage
5221 Service activities incidental to land transportation
5222 Service activities incidental to water transportation
5223 Service activities incidental to air transportation
5224 Cargo handling
5229 Other transportation support activities
5310 Postal activities
5320 Courier activities
5510 Short term accommodation activities
5520 Camping grounds, recreational vehicle, parks and trailer parks
5590 Other accommodation
5610 Restaurants and mobile food service activities
5621 Event catering
5629 Other food service activities
5630 Beverage Serving activities
5811 Book publishing
5812 Publishing of Directories and mailing lists
5813 Publishing of Newspapers Journals and Periodicals
5819 Other publishing Activities
5820 Software publishing
5911 Motion picture, video and television programmed production activities
5912 Motion picture, video and television programmed post-production activities
5913 Motion picture, video and television programmed distribution activities
5914 Motion picture projection activities
5920 Sound recording and music publishing activities
6010 Radio broadcasting activities
6020 Television programming and broadcasting activities
6110 Wired telecommunications activities
6120 Wireless tele communications activities
6130 Satellite telecommunications activities
6190 Other telecommunications activities
6201 Computer programming activities
6202 Computer consultancy and computer facilities management activities
6209 Other information technology and computer service activities
6311 Data processing, hosting and related activities
6312 Web portals
6391 News agency activities
6399 Other information service activities n.e.c
6411 Central Banking
6419 Other monetary Intermediation (with or without current account)
6420 Activities of holding companies
6430 Trusts, funds and similar financial entities
6491 Financial leasing
6492 Other Credit Granting
6499 Other financial service activities, except insurance and pension funding activities n.e.c
6511 Life Insurance
6512 Non-life insurance
6520 Reinsurance
6530 Pension funding
6611 Administration of finanical markets
6612 Security and commodity contracts brokerage (security dealing on behalf of others)

6619 Other activities auxiliary to financial service activities
6621 Risk and damage evaluation
6622 Acitivities of insurance agents and brokers
6629 Other activities auxiliary for insurance and pension funding
6630 Fund management activities
6810 Real estate activities with own or leased property
6820 Real estate activities on a fee or contract basis
6910 Legal activites
6920 Accounting, book keeping and auditing activities, tax consultancy
7010 Activities of head offices
7020 Management consultancy activities
7110 Architectural and engineering activites and related technical consultancy
7120 Technical testing and analysis
7210 Research and experimental development on natural science and engineering
7220 Research and experimental development on social science and humanities
7310 Advertising
7320 Market research and public opinion polling
7410 Specialized design activities
7420 Photographic activities
7490 Other professional,scientific and technical activities n.e.c
7500 Veterinary activities
7710 Renting and leasing of motor vechicles
7721 Renting and leasing of recreational and sports goods
7722 Renting of video tapes and disks
7729 Renting and leasing of other personal and household goods
7730 Renting and leasing of other machinery, equipment and tangible goods
7740 Leasing of intellectual property and similar products, except coypyrighted works
7810 Activities of employment placement agencies
7820 Temporary employment agency activities
7830 Other human resources provision
7911 Travel agency activities
7912 Tour operator activities
7990 Other reservation service and related activities
8010 Private security activities
8020 Security systems service activities
8030 Investigation activities
8110 Combined facilities support activities
8121 General cleaning of building
8129 Other building and industrial cleaning activities
8130 Landscape care and maintenance service activities
8211 Combined office administrative service activities
8219 Photocopying, doucment preparation and other specialized office support activities
8220 Activities of call centres
8230 Organization of conventions and trade shows
8291 Activities of collection agencies and credit bureaus
8292 Packaging activities
8299 Other business support service activities n.e.c
8411 General public administration activities
8412 Regulation of the activities of providing health care, education, cultural services and other social services, excluding social security
8413 Regulation of and contribution to more efficient operation of businesses
8421 Foreign affairs
8422 Defence activities
8423 Public order and safety activities
8430 Compulsory social security activities
8510 Pre-primary and primary education
8521 General secondary education
8522 Technical and vocational secondary education
8530 Higher education
8541 Sports and recreation education
8542 Cultural education
8549 Other education n.e.c.
8550 Educational support activities
8610 Hospital activities
8620 Medical and dental practice activities
8690 Other human health activities
8710 Residential nursing care facilities
8720 Residential care activities for mental retardation, mental health and substance abuse
8730 Residential care activities for the elderly and disabled
8790 Other residential care activities
8810 Social work activities without accommodation for the elderly and disabled
8890 Other social work activities without accommodation
9000 Creative, arts and entertainment activities
9101 Library and archives activities
9102 Museums activities and operation historical sites and buildings
9103 Botanical and zoological gardens and natural reserves activities
9200 Gambling and betting activities
9311 Operation of sports facilities
9312 Activities of sports clubs
9319 Other sports activities
9321 Activities of amusement parks and theme parks
9329 Other amusement and recreation activities

9411 Activities of business, employers and professional membership organizations
9412 Other activities of business and employers membership organizations
9420 Activities of trade unions
9491 Activities of religious organizations
9492 Activities of political organizations
9499 Activities of Membership Organizations
9511 Repair of computers and communication equipment
9512 Repair of communication equipment
9521 Repair of personal and household goods
9522 Repair of household appliences and garden equipment
9523 Repair of footwear and leather goods
9524 Repair of furniture and home furnishings
9529 Repair of other personal and household goods
9601 Washing and (dry-) cleaning of textile and fur products
9602 Hair dressing and other beauty treatment
9603 Funeral and related activities
9609 Other personal service activities n.e.c.
9700 Activities of households as employers of domestic Personnel
9810 Undifferentiated goods producing activities of private households for own use
9820 Undifferentiated service-producing activities of private households for own use
9900 Activities of extraterritorial Organizations and bodies

Major industry code

Code	4-digit	Description
A	111-322	Agriculture, Forestry and Fishing
B	710-990	Mining and Quarrying
C	1010-3320	Manufacturing
D	3510-3530	Electricity, gas, steam and air conditioning supply
E	3600-3900	Water Supply; sewerage, water management and remediation activities
F	4100-4390	Construction
G	4510-4799	Wholesale and retail trade; repair of motor vehicles and motor cycles
H	4910-5320	Transportation and storage
I	5510-5630	Accommodation and food service activities
J	5811-6399	Information and Communication
K	6411-6630	Financial and Insurance Activities
L	6810-6820	Real estate activities
M	6910-7500	Professional, Scientific and technical activities
N	7710-8299	Administrative and support service activities
O	8411-8430	Public administration and defense; compulsory social security
P	8510-8550	Education
Q	8610-8890	Human health and social work activities
R	9000-9329	Arts, entertainment and recreation
S	9411-9609	Other service activities
T	9700-9820	Activities of households as employers, undifferentiated goods and services producing activities of households for own use
U	9900	Activities of extraterritorial organizations and bodies

Attachment 6

Classification of occupation (4-digit)

110 Commissioned Armed Forces Officers
120 Non-commissioned Armed Forces Officers
130 Armed Forces Occupations - Other Ranks
210 Workers Reporting Occupations Unclassifiable
220 Workers Reporting Occupations Unidentifiable or Inadequately Described
1111 Legislators
1112 Assistant Registrar of Companies
1113 Traditional Chiefs and Heads of Village
1114 Senior Officials of Special-interest Organizations
1120 Managing Directors and Chief Executives
1211 Cooperate Managers
1212 Finance Managers
1213 Human Resource Managers
1214 Policy and Planning Managers
1219 Business Services and Administration Managers NEC
1221 Sales and Marketing Managers
1222 Advertising and Public Relations Managers
1223 Research and Development Managers
1311 Agricultural and Forestry Production Managers
1312 Aquaculture and Fisheries Production Managers
1321 Manufacturing Managers
1322 Mining Managers
1323 Construction Managers
1324 Supply, Distribution and Related Managers
1330 Information and Communications Technology Service Managers
1341 Child Care Services Managers
1342 Health Services Managers
1343 Aged Care Services Managers
1344 Social Welfare Managers
1345 Education Managers
1346 Financial and Insurance Services Branch Managers
1349 Other Professional Services Managers
1411 Hotel Managers
1412 Restaurant Managers
1420 Shop Managers
1431 Sports, Recreation and Cultural Centre Managers
1439 Services Managers NEC
2111 Physicists and Astronomers
2112 Meteorologists
2113 Chemists
2114 Geologists and Geophysicists
2120 Mathematicians, Actuaries and Statisticians
2131 Biologists, Botanists, Zoologists and Related Professionals
2132 Farming, Forestry and Fisheries Advisors
2133 Environmental Protection Professionals
2141 Industrial and Production Engineers
2142 Civil Engineers
2143 Environmental Engineers
2144 Engineer, Industrial Machinery and Tools
2145 Chemical Engineers
2146 Mining Engineers, Metallurgists and Related Professionals
2149 Engineering Professionals NEC
2151 Electrical Engineers
2152 Electronic Engineers
2153 Telecommunication Engineers
2161 Building Architects
2162 Landscape Architects
2163 Product and Garment Designers
2164 Town and Traffic Planners
2165 Cartographers and Surveyors
2166 Graphic and Multimedia Designers
2211 Generalist Medical Practitioners
2212 Specialist Medical Practitioners
2221 Nursing Professionals
2222 Midwifery Professionals
2230 Traditional and Complementary Medicine Professionals
2240 Paramedical Practitioners
2250 Veterinarians
2261 Dentists
2262 Pharmacists
2263 Environmental and Occupational Health and Hygeine Professionals
2264 Physiotherapists
2265 Dieticians and Nutritionists
2266 Audiologists and Speech Therapists
2267 Optometrists and Ophthalmic Opticians
2269 Health Professionals NEC
2310 University and Higher Education Teachers
2320 Vocational Education Teachers
2330 Secondary Education Teachers
2341 Primary School Teachers
2342 Early Childhood Educators
2351 Education Methods Specialists
2352 Special Needs Teachers
2353 Other Language Teachers
2354 Other Music Teachers
2355 Other Arts Teachers
2356 Information Technology Trainers

2359 Teaching Professionals NEC
2411 Accountants
2412 Financial and Investment Advisors
2413 Financial Analysts
2414 Assessors
2421 Management and Organization Analysts
2422 Policy Administration Professionals
2423 Personnel and Careers Professionals
2424 Training and Staff Development Professionals
2431 Advertising and Marketing Professionals
2432 Public Relations Professionals
2433 Technical and Medical Sales Professionals (Excluding Information Technology)
2434 Information and Communications Technology Sales Professionals
2511 Systems Analysts
2512 Software Developers
2513 Web and Multimedia Developers
2514 Applications Programmers
2519 Software and Applications Developers and Analysts NEC
2521 Database Designers and Administrators
2522 System Administrators
2523 Computer Network Professionals
2529 Database and Network Professionals NEC
2611 Lawyers
2612 Judges
2619 Legal Professionals NEC
2621 Archivists and Curators
2622 Librarians and Related Information Professionals
2631 Economists
2632 Sociologist, Anthropologist and Related Professionals
2633 Philosophers, Historians and Political Scientists
2634 Psychologists
2635 Social Work and Counselling Professionals
2636 Religious Professionals
2641 Authors and Related Writers
2642 Journalists
2643 Translators, Interpreters and Other Linguists
2651 Visual Artists
2652 Musicians, Singers and Composers
2653 Dancers and Choreographers
2654 Film, Stage and Related Directors and Producers
2655 Actors
2656 Announcers on Radio, Television and Other Media
2659 Creative and Performing Artists NEC
3111 Chemical and Physical Science Technicians
3112 Civil Engineering Technicians
3113 Electrical Engineering Technicians
3114 Electronics Engineering Technician
3115 Mechanical Engineering Technicians
3116 Chemical Engineering Technician (General)
3117 Mining and Metallurgical Technicians
3118 Draughtspersons
3119 Physical and Engineering Science Technicians NEC
3121 Mining Supervisors
3122 Manufacturing Supervisors
3123 Construction Supervisors
3131 Power Production Plant Operators
3132 Incinerator and Water Treatment Plant Operators
3133 Chemical Processing Plant Controllers
3134 Petroleum and Natural Gas Refining Plant Operators
3135 Metal Production Process Controllers
3139 Process Control Technicians NEC
3141 Life Science Technicians (Excluding Medical)
3142 Agricultural Technicians
3143 Forestry Technicians
3151 Ships' Engineers
3152 Ships' Deck Officers and Pilots
3153 Pilots and Related Associated Professionals
3154 Air Traffic Controllers
3155 Air Traffic Safety Electronic Technicians
3211 Medical Imaging and Therapeutic Equipment Technicians
3212 Medical and Pathology Laboratory Technicians
3213 Pharmaceutical Technicians and Assistants
3214 Medical and Dental Prosthetic Technicians
3221 Nursing Associate Professionals
3222 Midwifery Associate Professionals
3230 Traditional and Complementary Medicine Associate Professionals
3240 Veterinary Technicians and Assistants
3251 Dental Assistants and Therapists
3252 Medical Records and Health Information Technicians
3253 Community Health Workers
3254 Dispensing Opticians
3255 Physiotherapy Technicians and Assistants
3256 Medical Assistants
3257 Environmental and Occupational Health Inspectors and Associates
3258 Ambulance Workers
3259 Health Associate Professionals NEC
3311 Securities and Finance Dealers and Brokers
3312 Credit and Loans Officers
3313 Accounting Associate Professionals
3314 Statistical, Mathematical and Actuarial Associate Professionals
3315 Valuers and Loss Assessors

3321 Insurance Representatives
3322 Commercial and Sales Representatives
3323 Buyers
3324 Trade Brokers
3331 Clearing and Forwarding Agents
3332 Conference and Event Planners
3333 Employment Agents and Contractors
3334 Real Estate Agents and Property Managers
3339 Business Services Agents NEC
3341 Office Supervisors
3342 Legal Secretaries
3343 Administrative and Executive Secretaries
3344 Medical Secretaries
3349 Administrative and Specialized Secretaries NEC
3351 Cargo Officer
3352 Government Tax and Excise Officials
3353 Government Social Benefits Officials
3354 Government Licensing Officers
3359 Customs, Tax and Related Government Associate Professionals NEC
3360 Other Government Associate Professionals
3411 Police Inspectors and Detectives
3412 Legal and Related Associate Professionals
3421 Social Work Associate Professionals
3422 Religious Associate Professionals
3431 Athletes and Sports Players
3432 Sports Coaches, Instructors and Officials
3433 Fitness and Recreation Instructors and Program Leaders
3441 Photographers
3442 Interior Designers and Decorators
3443 Gallery, Museum and Library Technicians
3449 Other Artistic and Cultural Associate Professionals
3511 Information and Communication Technology Operations Technicians
3512 Information and Communications Technology User Support Technicians
3513 Computer Network and System Technicians
3514 Web Technicians
3521 Broadcasting and Audio-visual Technicians
3522 Telecommunication Engineering Technicians
4110 General Office Clerks
4120 Secretaries (General)
4131 Typists and Word Processing Operators
4132 Data Entry Clerks
4211 Bank Tellers and Related Clerks
4212 Bookmakers, Croupiers and Related Gaming Workers
4213 Pawnbrokers and Money-lenders
4214 Debt Collectors and Related Workers
4221 Travel Consultants and Clerks
4222 Contact Centre Information Clerks
4223 Telephone Switchboard Operators
4224 Hotel Receptionists
4225 Enquiry Clerks
4226 Receptionists (General)
4227 Survey and Market Research Interviewers
4229 Client Information Workers NEC
4311 Accounting and Book Keeping Clerks
4312 Statistical, Finance and Insurance Clerks
4313 Payroll Clerks
4321 Stock Clerks
4322 Production Clerks
4323 Transport Clerks
4411 Library Clerks
4412 Mail Carriers and Sorting Clerks
4413 Coding, Proof-reading and Related Clerks
4414 Scribes and Related Workers
4415 Filing and Copying Clerks
4416 Personnel Clerks
4419 Clerical Services Related Workers NEC
5111 Flight Attendants and Travel Stewards
5112 Railway Sleeping Car Attendant
5113 Travel Guides
5121 Chefs
5122 Cooks
5131 Waiters
5132 Bartenders
5141 Hairdressers
5142 Beauticians and Related Workers
5151 Cleaning and Housekeeping Supervisors in Offices, Hotels and Other Establishments
5152 Domestic House Keepers
5153 Building Caretakers
5161 Astrologers, Fortune-tellers and Related Workers
5162 Companions and Valets
5163 Undertakers and Embalmers
5164 Pet Groomers and Animal Care Workers
5165 Driving Instructors
5169 Personal Service Workers NEC
5211 Stall and Market Salespersons
5212 Street Food Salespersons
5221 Shop Keepers
5222 Shop Supervisors
5223 Shop Sales Assistants
5230 Cashiers and Ticket Clerks
5241 Fashion and Other Models
5242 Sales Demonstrators
5243 Door to Door Salespersons

5244 Contact Centre Salespersons
5245 Service Station Attendants
5246 Food Service Counter Attendants
5249 Sales Workers NEC
5311 Child Care Workers
5312 Teachers' Aides
5321 Health Care Assistants
5322 Home-based Personal Care Workers
5329 Personal Care Workers in Health Services NEC
5411 Fire Fighters
5412 Police Officers
5413 Prison Guards
5414 Security Guards
5415 Home Guards
5419 Protective Service Workers NEC
6111 Field Crops and Vegetable Growers (Excluding Cereal Crops, Estate Crops and Minor Export Crops) - Market -oriented
6112 Cereal Crop Cultivators (Excluding Paddy) (Market -oriented)
6113 Paddy Cultivators (Market -oriented)
6114 Tea Cultivators (Market -oriented)
6115 Rubber Cultivators (Market -oriented)
6116 Coconut Cultivators (Market -oriented)
6117 Minor Export Crops Cultivators (Market -oriented)
6118 Gardeners, Horticultural and Nursery Growers (Market -oriented)
6119 Mixed Crops Growers (Market -oriented)
6121 Livestock and Dairy Producers (Market -oriented)
6122 Poultry Producers (Market -oriented)
6123 Apiarists and Seri Culturists (Market -oriented)
6129 Animal Producers NEC (Market -oriented)
6130 Mixed Crops and Animal Producers (Market -oriented)
6210 Forestry and Related Workers (Market -oriented)
6221 Aquaculture Workers (Market -oriented)
6222 Ornamental Fish Workers (Market -oriented)
6223 Inland Water Fishery Workers (Market -oriented)
6224 Coastal Water Fishery Workers (Market -oriented)
6225 Deep-sea Fishery Workers (Market -oriented)
6226 Hunters and Trappers (Market -oriented)
6311 Agricultural Farm Supervisors
6312 Forestry Supervisors
6320 Fishery Supervisors
6330 Livestock and Dairy Farm Supervisors
6411 Subsistence Crop Farmers
6412 Subsistence Crop Farmers NEC
6420 Subsistence Livestock Farmers
6430 Subsistence Mixed Crop and Livestock Farmers
6440 Subsistence Fishers, Hunters, Trappers and Gatherers
7111 House Builders
7112 Brick Layers and Related Workers
7113 Stone Masons, Stone Cutters, Stone Splitters and Stone Carvers
7114 Concrete Placers, Concrete Finishers and Related Workers
7115 Carpenters and Joiners
7116 Masons (General)
7119 Building Frame and Related Trade Workers NEC
7121 Roofers
7122 Floor Layers and Tile Setters
7123 Plasterers
7124 Insulation Workers
7125 Glaziers
7126 Plumbers and Pipe Fitters
7127 Air Conditioning and Refrigeration Mechanics
7131 Painters and Related Workers
7132 Spray Painters and Varnishes
7133 Building Structure Cleaners
7211 Metal Molders and Core makers
7212 Welders and Flame Cutters
7213 Sheet-metal Workers
7214 Structural-metal Preparers and Erectors
7215 Riggers and Cable Splicers
7221 Blacksmiths, Hammersmiths and Forging Press Workers
7222 Toolmakers and Related Workers
7223 Metal Working Machine Tool Setters and Operators
7224 Metal Polishers, Wheel Grinders and Tool Sharpeners
7231 Motor Vehicle Mechanics and Repairers
7232 Aircraft Engine Mechanics and Repairers
7233 Agricultural and Industrial Machinery Mechanics and Repairers
7234 Bicycle and Related Repairers
7310 Precision-instrument Makers & Repairers
7311 Musical Instrument Makers and Tuners
7312 Jewellery and Precious-metal Workers
7313 Potters and Related Workers
7314 Glass Makers, Cutters, Grinders and Finishers
7315 Sign Writers, Decorative Painters, Engravers and Etchers
7316 Handicraft Workers in Wood, Cane and Related Materials
7317 Handicraft Workers in Textile, Leather and Related Materials
7318 Coir Industry Workers
7319 Other Handicraft Workers
7321 Pre-Press Technicians
7322 Printers

7323 Print Finishing and Binding Workers
7411 Building and Related Electricians
7412 Electrical Mechanics and Fitters
7413 Electrical Instruments Repairer
7421 Electronics Mechanics and Servicers
7422 Information and Communications Technology Servicers and Installers
7511 Butchers, Fishmongers and Related Food Preparers
7512 Bakers, Pastry-cooks and Confectionery Makers
7513 Dairy-products Makers
7514 Fruit, Vegetable and Related Preservers
7515 Food and Beverage Tasters and Graders
7516 Tobacco Preparers and Tobacco Products Makers
7517 Other Products Makers
7521 Wood Treaters
7522 Cabinet-makers and Related Workers
7523 Woodworking-machine Tool Setters and Operators
7531 Tailors, Dressmakers, Furriers and Hatters
7532 Textile, Leather and Related Pattern-Makers and Cutters
7533 Sewers, Embroiders and Related Workers
7534 Upholsterers and Related Workers
7535 Pelt Dressers, Tanners and Fell mongers
7536 Shoemakers and Related Workers
7541 Underwater Divers
7542 Shot firers and Blasters
7543 Product Graders and Testers (Except Foods and Beverages)
7544 Fumigators and Other Pest and Weed Controllers
7549 Craft and Related Workers NEC
8111 Mining and Quarries
8112 Mineral Ore and Stone Processing Plant Operators
8113 Well Drillers and Borers and Related Workers
8114 Cement, Stone and Other Mineral Products Machine Operators
8121 Metal Processing Plant Operators
8122 Metal Finishing, Plating and Coating Machine Operators
8131 Chemical Products Plant and Machine Operators
8132 Photographic Products Machine Operators
8141 Rubber Products Machine Operators
8142 Plastic Products Machine Operators
8143 Paper Products Machine Operators
8151 Fibre Preparing, Spinning and Winding Machine Operators
8152 Weaving and Knitting Machine Operators
8153 Sewing Machine Operators
8154 Bleaching, Dying and Fabric Cleaning Machine Operators
8155 Fur and Leather Preparing Machine Operators
8156 Shoemakers and Related Machine Operators
8157 Laundry Machine Operators
8159 Textile, Fur and Leather Products Machine Operators NEC
8160 Food and Related Products Machine Operators
8171 Pulp and Paper Making Plant Operators
8172 Wood Processing Plant Operators
8181 Glass and Ceramics Plant Operators
8182 Steam Engine and Boiler Operators
8183 Packing, Bottling and Labelling Machine Operators
8189 Stationary Plant and Machine Operators NEC
8211 Mechanical Machinery Assemblers
8212 Electrical Equipment Assemblers
8213 Electronic Equipment Assemblers
8214 Metal, Rubber and Plastic Products Assemblers
8215 Wood and Related Products Assemblers
8216 Paperboard, Textile and Related Products Assemblers
8219 Assemblers NEC
8311 Locomotive Engine Drivers
8312 Railway Brakers, Signallers, Shunters and Switch Operators
8321 Motor Cycle Drivers
8322 Car, Taxi and Van Drivers
8323 Three Wheeler Drivers
8331 Bus Drivers
8332 Heavy Truck and Lorry Drivers
8341 Mobile Farm and Forestry Plant Operators
8342 Earth-moving and Related Plant Operators
8343 Crane, Hoist and Related Plant Operators
8344 Lifting Truck Operators
8350 Ships' Deck, Crews and Related Workers
9111 Domestic Cleaners and Helpers
9112 Cleaners and Helpers in Offices, Hotels and Other Establishments
9121 Hand Launderers and Pressers
9122 Vehicle Cleaners
9123 Window Cleaners
9129 Other Cleaning Workers
9211 Crop Farm Labourers
9212 Tea Pluckers
9213 Rubber Tappers
9214 Cinnamon Tappers
9215 Livestock Farm Labourers
9216 Mixed Crops and Livestock Farm Labourers
9217 Garden and Horticultural Labourers
9218 Forestry Labourers
9219 Fishery and Aquaculture Labourers
9311 Mining and Quarrying Labourers

9312 Civil Engineering Labourers
9313 Building Construction Labourers
9321 Hand Packers
9322 Manufacturing Laborers NEC
9331 Hand and Pedal Vehicle Drivers
9332 Drivers of Animal-drawn Vehicles and Machinery
9333 Freight Handlers
9334 Shelf Fillers
9335 Transport Labourers
9411 Fast Food Preparers
9412 Kitchen Helpers
9510 Street and Related Service Workers
9520 Street Vendors
9611 Garbage Collectors
9612 Refuse Sorters
9613 Sweepers and Related Labourers
9621 Messengers, Package Deliverers and Luggage Porters
9622 Odd Job Persons
9623 Meter Readers and Vending-machine Collectors
9624 Water and Firewood Collectors
9629 Elementary Service Workers NEC

Major occupation code

Code	4-digit	Description
1	1111-1439	Manager, Senior Officials and Legislator
2	2111-2659	Professionals
3	3111-3522	Technicians and Associate Professionals
4	4110-4419	Clerks and Clerical Support Works
5	5111-5419	Services and Sales Workers
6	6111-6440	Skilled Agricultural, Forestry and Fishery Workers
7	7111-7549	Craft and Related Trade Workers
8	8111-8350	Plant and machine Operators and Assemblers
9	9111-9629	Elementary occupations
0	0110-0220	Other