Royal Government of Bhutan
Planning Commission
Central Statistical Organization

HOUSEHOLD INCOME AND EXPENDITURE SURVEY 2000

[PHASE 1]

TECHNICAL REPORT

ON
SAMPLING,
DATA COLLECTION,
AND DATA PROCESSING

Thimphu, November 2000

Foreword

Survey implemented by CSO with assistance of ADB (Ta nnnn - xxxx)
Director: K.Tshering, Yeshey Dorji
Staff CSO:
Contribution from consultants:
Mr. Madan Sardana
Mr. V.K. Mahajan
Mr. Elpidio Nogales

ADB EDSD
Abuzar Asra
Olivier Dupriez

Objective of this report is to allow users to analyze data, and critical analysis for improving the data in second phase. Prepared with OD.

TABLE OF CONTENT

Introduction

Objectives

Household Income and Expenditure Survey (HIES) 2000 is the second nationwide survey of households undertaken by Central Statistical Organization (CSO), Planning Commission, Royal Government of Bhutan. The first survey was conducted in 1991-92.

The broad objectives of the survey were:

- To provide useful inputs for the compilation of national accounts of the household sector as well as statistics of distribution of household income and expenditure;
- To provide benchmark information to update weights required in the construction of consumer prices indices (CPI); and
- To provide inputs in the estimation of poverty threshold and its incidences.

Scope and coverage

The survey collected data on household income including, among others, sources of income in cash and in kind and levels of consumption by items of expenditure. In view of the need to update weights used in the current CPI series, detailed items up to 5-digit level were incorporated for collection of consumer expenditure in the survey.

The geographical coverage extended over the entire area of Bhutan excepting a few satellite towns which are neither recognized as urban areas nor or under the administrative control of chupon (block) headman in the rural areas. The population coverage included all households in the country **except** the following:

a. households of expatriates;
b. residents of hotels, boarding and lodging houses, monasteries, school hostels, orphanages, rescue homes, ashram, vagrant houses, and under-trials in jails and indoor patients of hospitals, nursing homes etc.; and
c. Barracks of military and para-military forces including the police.

Sampling

Sampling frame

2.4 The Sample Survey Section (SSS) of CSO maintains a list providing number of households by town and by geog (development block). Maps of towns and geogs are available with the administrative offices of

dzongkhags (district). The maps together with the list of towns and geogs giving number of households provided the frame for designing the HIES 2000.

Determination of sample size

2.5 In determining the overall sample size it is generally advisable to start with the required level of reliability in the estimates expected from the sample and if the field resources and/or budget is a constraint, the precision that can be achieved under the constraint is assessed to decide whether the achievable precision would meet the needs, and, if not whether the budget could be increased is examined.

2.6 HIES covers a large number of data items, some of which like expenditure on durable goods have much more variability than items of frequent occurrence such as expenditure on food. In practice, the sample size required for estimating a few major items with the requisite precision is worked out and the largest of the calculated sizes is taken as the sample for the survey. Since the results of HIES 1991-92 were not available, it was decided to make use of the published results from some of the countries in Asia like India, Bangladesh, Nepal, Thailand, and Philippines.

2.7 A technical study (National Household Survey Capability Programme; Household Income and Expenditure Surveys: A Technical Study, United Nations, 1989) published by the United Nations Statistical Office provides guidance to derive the sample size for HIES. A sample size of 4,000 households for the HIES 2000 was decided on the basis of the technical details and assumptions given in Annex 1. It was also decided to equally distribute this sample between urban and rural areas.

Sample design

2.8 A stratified two-stage sampling design is generally used in the HIES, where stratification is by geographical areas and enumeration areas (EAs) of population census form the primary stage units (PSUs) in the urban areas and villages constitute the PSUs in the rural areas. Households form the second stage units (SSUs) of sampling. The PSUs are generally selected with probability proportion to size (PPS), size being the number of households enumerated in the population census. Equal number of households are selected from within each PSU to make the design self-weighted, which in turn makes the data processing efficient and easy to operate. Since the quality of information relating to number of households with the CSO could not be assessed and comparability of the information over the dzongkhags could not be ensured, it was decided not to use the available information as a measure of size for PPS selection. However, it was decided to use the available information for stratification and also for allocation of sample.

2.9 A stratified multi-stage sampling design was used in HIES 2000. The available information relating to number of households for each town and geog was used to stratify the country into the following four strata:

Stratum 1: Consisting of seven towns each having 850 or more households, viz., Thimphu, Phuentsholing, Gelephu, Punakha, Samdrup Jongkhar, Chhaukha and Wangduephodrang.

Stratum 2: Consisting of remaining 15 towns.

Stratum 3: Consisting of 22 geogs each with at least 750 households.

Stratum 4: Consisting of the remaining 180 geogs.

Selection of Sample in Urban Areas

2.10 The sample of 2,000 hhs for the urban areas was distributed between stratum 1 and stratum 2 in proportion to the number of hhs. This resulted in an allocation of 1,650 hhs to stratum 1 and 350 hhs to stratum 2. All the seven towns in stratum 1 were selected with probability one and the sample of 1,650 hhs was allocated to the seven towns in proportion to the number of hhs. This resulted in an allocation of 800 hhs to Thimphu, 300 to Phuentsholing, 150 to Gelephu, and 100 each to Punakha, Sandrup Jongkhar, Chhukha, and Wangduephodrang.

2.11 The 15 towns in stratum 2 were arranged in descending order of number of hhs and from this list a sample of seven towns was selected as a circular systematic sample. To each of the selected seven towns a sample of 50 hhs was allocated.

2.12 The geographic area in each of the seven towns in stratum 1 and the seven towns selected from stratum 2 was divided into EAs, each of 100-125 hhs and a listing of all the hhs in each EA was undertaken. At the time of listing, for each hh the name of the head of the hh was noted and thereafter it was ascertained whether the head of the hh was an expatriate or not. The EAs enumerated in each town were divided into three socio-economic groups—high, medium, and low, on the basis of known information about the value of real estate and/or the rent of residential accommodation. All the EAs enumerated in each town were arranged in the order starting with "high" followed by "medium" and the "low". In such a list for each town all hhs **excluding those of expatriates** were given a running serial number starting with the first hh in the first EA of the "high" group to the last hh in the last EA of the "low" group. From such an arranged list for each town, the requisite number of sample hhs was selected as a circular systematic sample. Thus the sample design adopted for selection of hhs was uni-stage simple random sample (SRS) in each of the seven towns of stratum 1, and two-stage random sample in case of stratum 2.

Selection of Sample in Rural Areas

2.13 The sample of 2,000 hhs in the rural areas was distributed between stratum 3 and stratum 4 approximately in proportion to the number of hhs, which resulted in an allocation of 880 hhs to stratum 3 and 1,120 hhs to stratum 4. Each of the 22 geogs in stratum 3 was selected with probability one and a sample of 40 hhs was selected from within each geog. For selection of hhs from within each of the 22 geogs in stratum 3, a SRS of four chupons (a group of three to five villages) was selected. All hhs in each selected chupon were listed along with the details relating to name of the head of hh, the size of the hh, prime means of livelihood (PML)—(self-employed in non-agriculture, rural labour, and others), description of activity of any non-agricultural enterprise operated from within the premises of the hh along with a broad industry group code. The list of hhs so prepared for a chupon was rearranged by PML classes and by size of hh within each of the three PML classes. From such an arranged list for each selected chupon in

stratum 3, a circular systematic sample of 10 hhs was selected. Thus a two-stage random sample design was followed for selection of sample hhs in stratum 3.

2.14 The selection of hhs in stratum 4 was done in three stages. In the first stage a sample of 56 geogs was selected from amongst the 180 geogs as a circular systematic sample after arranging the 180 geogs first by dzongkhong and then within each dzongkhong in descending order of the size of the geog in terms of number of hhs. At the second stage, in each selected geog, two chupons were selected as a circular systematic sample after arranging the chupons in descending order of the size of the geog in terms of number of hhs. From within each selected chupon a sample of 10 hhs was selected following the procedure indicated for selection of hhs within each selected chupon in stratum 3.

MAP

Extrapolation

Sampling errors

Data collection

Timing

Questionnaires and manuals

3.1 A questionnaire was used for collection of data for the HIES 1991-92. There are several advantages of using a schedule as an instrument of data collection in situations where trained enumerators are deployed to conduct face-to-face interviews for collection of household based socio-economic data. This allows the enumerator to ask appropriate questions to elicit the requisite information. It has the advantage that having understood the objective of the question, the enumerator can size up the situation and ask questions in a form that is more understandable to the respondent as also more appropriate to the situation. It has also the advantage that the enumerator can adopt a conversional approach for the interview, which is especially

needed in the rural areas of developing countries. It was, therefore, decided to use a schedule as the instrument of data collection for the HIES 2000 in Bhutan. Three sets of schedules, two for listing of hhs (one each for the urban areas and rural areas), and the other for collection of data relating to household income household expenditure, demographic particulars and economic activity of the members of the hh were devised. These schedules were pre-tested in the field, both in urban and rural areas, in January 2000. In the light of the experience of pre-testing , the schedules were finalized.

3.2 Each schedule was divided into a number of homogenous sections, called Blocks, according to the subject/topic and the blocks were divided into sub-blocks, wherever necessary according to nature of the topic covered.

3.3 The listing schedules, 0.1 for the urban areas and 0.2 for the rural areas contained the following three blocks:

Block (0): identification particulars of the EA/ choupan.
Block (1): details of the listing operation.
Block (2): sketch map of the EA/choupan.
Block (3): list of hhs.

Listing Operation

3.4 For listing the hhs all houses in the EA/choupan were surveyed to identify all hhs in that house and at that stage the purpose for which the house was being used, name of the head of the hh, size of hh, whether it was the hh of an expatriate or not were collected, both in the urban areas and rural areas. In addition, for the rural hhs, the prime means of livelihood and whether the hh operated a non-agricultural enterprise at the premises of the hh or without any fixed premises were ascertained. The broad industry group of identified enterprises was also ascertained.

Household Schedule

3.5 In broad terms the household schedule comprised two sections, one dealing with household consumption and the other with household income. As background material for the collection and analysis of household income and

expenditure, the schedule included provision for recording data on size, structure and composition of hh, and activity particulars of members of the hh.

3.6 Household expenditure includes details of common expenditure of the hh for consumption as well as other purposes and personal expenditure of all individual members. Consumption includes not only consumption of items purchased but also consumption out of own production, own business stocks, items received as gifts or in exchange of goods and services, and own housing. The section on hh consumption included a block dealing with sufficiency of food for hh to serve as a rough indicator of poverty.

3.7 Household income includes individual incomes of all hh members as well as the joint and composite income of hh, both in cash and in kind. It also includes incomes from paid employment, entrepreneurial incomes, incomes from property and other sources like current transfers and benefits.

□□ 8□The household schedule comprised the following blocks:

Block (0): identification and operational particulars
Block (1): household characteristics
Block (2): demographic and other particulars of hh members
Block (3): household consumption expenditure
- 3.1 consumption of food, beverages and tobacco
- 3.2 consumption of clothing, bedding and footwear
- 3.3 housing, fuel and light
- 3.4 transport and communication
- 3.5 household operation
- 3.6 education, recreation, entertainment and cultural services
- 3.7 medical care and health services
- 3.8 personal cares and effects
- 3.9 furnishing and equipment
- 3.10 house maintenance and minor repairs
- 3.11 miscellaneous expenses
- 3.12 consumption of selected non-food items from home produced stock
- 3.13 non-consumption expenditure
- 3.14 disbursements other than expenditure
- 3.15 production and consumption from kitchen garden and backyard
- 3.16 sufficiency of food for the household

Block (4): household income
- 4.1 activity particulars of hh members
- 4.2 income from paid employment
- 4.3 checklist for entrepreneurial activities
- 4.4a crop farming and gardening output
- 4.4b crop farming and gardening input

4.4c	computation of net income from crop farming and gardening
4.5a	livestock and poultry farming output
4.5b	livestock and poultry farming input
4.5c	computation of net income from livestock and poultry farming
4.6	computation of net income from other entrepreneurial activities
Block (5):	property and other income
5.1	rental income on real estate
5.2	other incomes received
5.3	other receipts

Reference period

3.10 It is well known that household income and household expenditure are subject to short-term fluctuations. Thus the larger the amount of information gathered in respect of each sample hh, the less would be sampling error. Therefore, from the point of view of accuracy of estimates, the longer the reference period the better it is. On the other hand, larger the reference period, the greater would be the chance of recall lapse. The “end effect” arising from the misplacement of events or transactions would be comparatively higher in case of a short reference period than in case of a long reference period.

3.11 The reference period, therefore, should be short enough but consistent with the requirements of accuracy. There is empirical evidence (**National Household Survey Capability Programme ; Household Income and Expenditure Surveys: A Technical Study, United Nations, 1989**) to show that the shorter the reference period, the higher the estimates of expenditure generated on that

basis, and the larger the reference period, the lower the estimates, especially when the survey data are collected through interviewing.

3.12 Keeping in view the above as also the experience of several countries including India documented in the above mentioned publication of the United Nations, it was decided that for the HIES 2000, a week as also a month be taken as the reference period for recording details relating to consumption of food, drinks, tobacco, and other items of daily requirement. For expenditures on durable items of infrequent purchases, reference period of one year was decided. It was also decided to use a reference period of one month for recording income from paid employment, while for crop farming the reference period for the first round was taken as the last winter crop season. A reference period of one month was taken for recording details relating to livestock and poultry farming. It was also decided to use last one year as the reference period for recording details pertaining to income from other entrepreneurial activities, while property and other incomes were to be recorded with one month as also one year as the reference period.

Methodology, staff, training

5.1 The Central Statistical Organization (CSO) recruited a group of 80 enumerators for conducting the field work of the first round of the survey. Most of the enumerators were students who had written the 12th grade examination in March 2000. CSO provided the services of nine staff members to work as

supervisors on full time basis during the period of the survey. Head Sample Survey Section (SSS), CSO was entrusted with the overall responsibility of organizing and conducting the field work.

5.2 Two comprehensive manuals, one for the enumerators and the other for the supervisors, were prepared. The training manual for the enumerators (TME) included basic concepts and definitions of different items of information, procedure for (i) listing of hhs, (ii) selection of sample hhs, and (iii) recording information in each of the blocks/sub-blocks of schedule 1. The training manual for supervisors (TMS) provided inter-alia detailed instructions for (i) undertaking rationalized supervision of the field work, (iii) ensuring quality of data, and (iii) selection of chupons and sample hhs. A copy of the TME was provided to each enumerator, while each supervisor was given a copy each of TME and TMS. Annexes 5 and 6 contain a copy of TME and TMS respectively**.**

Training of Field and Supervisory Staff

5.3 The nine supervisors designated for the survey underwent training for three days,

29-31 March 2000. The procedure for (i) undertaking demarcation of boundaries of EAs, (ii) preparing sketch maps of EAs, (iii) selection of chupons, (iv) listing of hhs, and (v) selection of hhs, were explained with suitable examples. The approach and procedure to be adopted for (i) enlisting cooperation of the informants, and (ii) for filling the various blocks/sub-blocks of the three schedules 0.1, 0.2, and 1 were also explained to the supervisors with suitable examples. It was decided to use third revised version of International Standard Industrial

Classification (ISIC) and 1988 International Standard Classification of Occupations (ISOC). With suitable examples the supervisors were trained to use ISIC and ISOC for recording codes, and the responsibility of recording the codes in the filled-in schedules was assigned to the supervisors. The enumerators were required to write only the description of the economic activity and type of work done and the supervisors were trained to record the appropriate two-digit codes.

5.4 The enumerators underwent a well planned and organized training program of 11 days form 3 April to 13 April 2000. All the nine supervisors were the basic trainers at the program. Each one of them had been assigned the responsibility of explaining the contents of three/five sub-blocks of the household schedule. Each one of them had prepared transparencies for the presentation. In general, each one of them had done a good job. Head, SSS and the ADB Consultant offered explanations and clarifications to strengthen the presentations of the trainers.

5.5 The first round of classroom lectures was completed on 6 April. Each enumerator was, thereafter, assigned the responsibility of completing the household schedule either for his/her own household or household of anyone else whom they could approach without any problem. Each filled-in schedule was scrutinized by one of the supervisors and necessary clarifications/corrections were given to the concerned enumerator. The after-noon session on 7 April and forenoon session on 8 April were utilized in discussing and resolving the problems noted by each of the supervisors during the course of the scrutiny of the filled-in schedules. Thereafter, the ADB Consultant restated the important

points to be kept in mind while filling different sub-blocks of household expenditure and household income blocks. Each enumerator spent Sunday, 9 April in completing schedule 1 for one more household. The filled-in schedules were scrutinized by the supervisors and problems noted were discussed and resolved during the training program on 10 April.

5.6 Three special documents were prepared by one of the supervisors to help the enumerators in collection of data on some of the items. The first document provided the age conversion from Bhutanese Calendar to English Calendar. The second document was a ready-reckonor for conversion of area under crops and production of crops from local units to standard units of acres and kilograms respectively. The third document provided the translation of names of some important items from English to different dilates of Bhutanese language. Copies of these were provided to each enumerator and each supervisor. These are reproduced in Annex 7.

5.7 The enumerators were taken out for field training in the rural area of Thimphu district on 11 April. They were accompanied by the supervisors, Head SSS and the ADB Consultant. The enumerators were divided into groups of eight/nine for the field training. Each enumerator filled in a household schedule independently. A field demonstration of the procedure of selection of chupon and listing of hhs in selected chupon was organized for the benefit of the supervisors. Each supervisor, thereafter, attempted listing operation in one of the chupons.

5.8 The schedules filled-in by the enumerators during the field training were thoroughly scrutinized by the supervisors in the morning session of 12 April and

deficiencies noted were pointed out to the concerned enumerators. Thereafter, each supervisor presented the deficiencies noted in the schedules filled-in by his group of enumerators. Clarifications and explanations, wherever required, were given by Head, SSS and the ADB Consultant.

Opinion Survey on Payment for Health Services

5.9 At the request of Ministry of Health, CSO agreed to canvass a small questionnaire to seek the opinion of households regarding user payment for health services. It was agreed to canvass a brief questionnaire seeking the above stated opinion from a sub-sample of 400 hhs in the urban areas and a sub-sample of about 50 hhs in some of the geogs known to be using hospital services in urban areas.

5.10 A representative of the Ministry of Health had trained the supervisors for about two hours during the period of training 29-31 March. The enumerators were also trained by the same representative of the Ministry of Health for about two hours on 7 April 2000. Each enumerator had also filled in the questionnaire during the field training on 11 April 2000 from the same hh for which schedule 1 had been filled.

Field Work

5.11The enumerators were divided into nine teams and one supervisor was assigned to each team. To facilitate the work of enumerators as also for effective supervision, each team was provided with a transport. All the teams started the field work on 15 April 2000 and the work was completed by the teams between 4 June and 17 June 2000. On an average an enumerator took about two hours to

collect the prescribed data from a hh.. Each supervisor accompanied each enumerator in his team at least on three occasions for listing of hhs and collection data from selected hhs. Head, SSS and Director, CSO also exercised supervision of the field work.

Publicity and Appeal for Cooperation

5.12 All out efforts were made to seek the cooperation of hhs to provide the data planned to be collected under HIES. As a first step all the Dzongkhag (distict) officers were informed about the schedule of operation of field work under HIES and were in turn requested to issue instructions to all Geog (block) officers to extend full cooperation to the field staff as also advise the heads of selected chupons in the sample to help and assist the field staff in carrying out the field operation. Each enumerator was given a photo identity card and a letter signed by Secretary, Planning Commission introducing him/her and ensuring confidentiality of the information to be provided by the hh. As the second step Director, CSO appealed to people through radio broadcast and press media.

Response rates

5.13 In spite of best efforts made by the enumerators and follow up attempts by the supervisors in most of the cases, there was non-response. As against a planned sample of 4, 000 hhs, the field staff were able to collect data from 3,854

hhs, which works out to a response rate of 96.3 per cent. Stratum 1 (large towns) accounted for about 80 per cent of non-response cases as may be seen from Table 1.

HIES 2000 – Sample size by stratum (table by towncode available at CSO)

Stratum	Sampling design	Sample collected	Rate
1 – Seven towns with more than 850 households (Thimphu, Phuentsholing, Gelephu, Punakha, Samdrup Jongkhar, Chhukha, Wangduephodrang)	1650	1538	93.2%
2 – Remaining 15 towns	350	347	99.1%
3 – Twenty-two geogs with more than 750 households	880	861	97.8%
4 – Remaining 180 geogs	1120	1108	98.9%
Total	**4000**	**3854**	**96.3%**

Reasons for Non-Response

5.14 Failure to establish contact with any adult member in the hh in spite of at least three attempts was the main reason reported by the field staff for non-response and this was so both in the urban and rural areas. There were, of course, some cases of refusal to co-operate, in particular in Thimphu. In most of these cases the concerned supervisor made sincere efforts to convince the head of the hh that data proposed to be collected would not only be of great help to the RGB in devising suitable development programs but also to the industrial units, trading community and the people of Bhutan. The head of the hh was assured that data proposed to be collected will remain confidential and not provided to the Revenue Department or any other organization in the RGB concerned with regulating Acts for industrial or trading activities etc. A few cases of non-response were converted into willing respondents.

5.15 It is proposed to revisit all non-responding hhs during the second round of the survey. Appropriate ways and means would be devised to covert the non-response cases

to responding cases.

Issues

Lack of field scrutiny was a problem in the implementation of the HIES. However, this problem partly arise from:

- The poor design of the questionnaire (in term of layout more than in term of content). A more "interviewer-friendly" questionnaire would have permitted avoiding many mistakes.
- The relatively poor quality of the interviewer and supervisor manuals. Considering the low level of expertise/experience of the data collection staff, training and reference material is of utmost importance and should have been more didactic.

Data processing

Software

Data entry and data editing have been implemented using IMPS (Integrated Microcomputer Processing System) from the US Bureau of Census.

Since some computers used for data entry were operated under DOS environment, version 3.1 of IMPS was used. The following modules of IMPS have been used:

- DATADICT (DATA DICTionary) to design the data dictionary
- CENTRY (CENsus data enTRY) for data entry
- CONCOR (CONsistency and CORrection) for partial editing of the data

! CENVAR

SPSS (version 10) was used for further editing, aggregation, tabulation and analysis.

Staff

Data processing was conducted by the survey and data processing section, with the following staff:

- Yeshey Dordji, head of the section
- Mrs Peden, programmer
- Ms Nima Deki Sherpa, programmer
- Mr. Xxx , assistant

Technical assistance in data processing was also provided by Mr. Olivier Dupriez statistician, ADB (data entry, editing, analysis), Mr. Vijay K. Mahajan, ADB consultant (data editing, tabulation) and Mr. Elpidio Nogales, ADB consultant (data entry).

Data entry

Programs and organization

Data dictionary

Data entry started on July 7, 2000 (using IMPS-CENTRY) and was completed by mid-August. The dataset contains information collected on 3854 households.

The average number of keys per questionnaire has been estimated at 1300 (based on urban questionnaires; expected to be less for rural questionnaires). A security margin of 10% is applied (= 1430 keys per questionnaire). A data entry operator should reach an average of 4000 keys/hour (standard productivity is 4500 keys/hour).

		Planned	Real
1	Number of questionnaires	4000	3854
2	Number of keys per questionnaire (est.)	1430	
3	Total number of keys (= 1 * 2)	5720000	
4	Numbers of keys/hour/operator	4000	
5	Total number of hours (= 3 / 4)	1430	
6	Hours of effective work per day	6.5	
7	Number of days (1 operator) (= 5 / 6)	220	
8	Number of operators	11	
9	**Number of working days (= 7 / 8)**	**20**	

A “processing tracking system” (MS-Excel application) was designed to monitor the data entry and editing activities. Supervisors will be responsible for updating the system on a daily basis. It will automatically generate reports on the progress of the work (% of data entry and editing completed, statistics per operator, daily averages, etc.).

A draft program had been designed to label all variables and values.

Data editing

- It was unfortunately not possible to accommodate all data entry operators in the same room. Operators are working in five different rooms, making it difficult for the two supervisors to ensure a strict control and a permanent assistance to the operators. As a consequence, it has been decided that no editing/corrections will be done by the operators. Only range checks are performed at the time of data entry. Consistency checks and manual corrections were made by the supervisors.
- The current CONCOR application produces lists of errors that can be manually corrected by the supervisors. It doesn’t contain any automatic imputation procedures. Further editing will be required (in particular for income and expenditure variables) after a detailed analytical assessment of the data is done (i.e. at the end of data entry).

After manual editing, the data file will have to be converted into SPSS format. An SPSS conversion program has already been prepared and tested. This program converts the IMPS data file into 41 SPSS data files (one per section of the questionnaire). The program also labels all variables. A draft program has been prepared for labeling values (to be finalized and split in shorter programs).

Further editing will then be done using SPSS (mainly editing of the income and expenditure variables).

SPSS editing: revealed many errors

Issues-discussion

We had to take into consideration the fact that only two staffs were available for supervising inexperienced data entry operators. It would have been very risky if not impossible to implement full consistency checks and corrections at the data entry stage:

- The data entry operators did not have any expertise to decide on what corrections should/could be made.
- Due to the high number of inconsistencies, it would have considerably slowed down the data entry.
- It would have been impossible to globally assess the frequency of inconsistencies and types of corrections made. Corrections could have introduced biases.

I could have designed a “magic” editing program that would have automatically fixed all inconsistencies, but with no guarantee that the resulting data file would correspond to the reality. Therefore, I preferred to minimize automatic imputations (either by cold-deck or hot deck).

Correcting data manually by reference to the filled-in schedules was much safer and had the advantage of providing CSO staff with a very good view and understanding of the data collection problems. There was no “black box” for CSO staff in the editing procedures. I still consider that data editing could have been completed in three weeks after completion of data entry, should we have had a more experienced data processing consultant.

In the next round of the HIES, more checks will be implemented at data entry time. Based on the first round experience and data, it will be possible to define range checks on quantities and values by product, for which no information was available for the first round.

Quantities not processed

No data on prices
Missing cases
Outliers

Aggregation of income and expenditure

Quality control of expenditure data

Reference periods: comparison month versus week*4

Tabulation and analysis

List of documents

1. *Household Income and Expenditure Survey 2000, Operational Manual*
   Central Statistical Organization, Planning Commission, Royal Government of Bhutan
   Thimphu, 2000; prepared by Madan Sardana, ADB Consultant (nn pages)

2. *Household Income and Expenditure Survey 2000, Instruction manual for enumerators*
   Central Statistical Organization, Planning Commission, Royal Government of Bhutan
   Thimphu, 2000 (nn pages)

3. *Household Income and Expenditure Survey 2000, Instruction manual for supervisors*
   Central Statistical Organization, Planning Commission, Royal Government of Bhutan
   Thimphu, 2000 (15 pages)

4. *Household Income and Expenditure Survey 2000, Data Entry Manual*
   Central Statistical Organization, Planning Commission, Royal Government of Bhutan
   Thimphu, 2000; prepared by Elpidio C. Nogales, ADB Consultant (44 pages)

5. *TA 2860 (BHU): Strengthening the Central Statistical Organization, Household Income and Expenditure Survey 2000 - Data Processing – Technical issues and questions*
   Olivier Dupriez, Statistician, ADB – EDSD, 2000 (4 pages)

6. *TA 2860 (BHU): Strengthening the Central Statistical Organization, Household Income and Expenditure Survey 2000 – Technical Note on Data Editing*
   Olivier Dupriez, Statistician, ADB – EDSD, 11 September 2000 (12 pages)

7. *Household Income and Expenditure Survey 2000, Schedule1:Household Schedule, Tabulation Plan*
   Central Statistical Organization, Planning Commission, Royal Government of Bhutan
   Thimphu, 2000; prepared by Madan Sardana, ADB Consultant (6 pages)

8. *Report on Household Income and Expenditure Survey 2000, Preliminary Draft*
   Central Statistical Organization, Planning Commission, Royal Government of Bhutan
   Thimphu, 2000; prepared with Madan Sardana, ADB Consultant (nn pages)

9. *Final Consultancy Report, Bhutan TA*
   Madan Sardana, ADB Consultant, October 2000 (5 pages)

List and description of data files

IMPS Ascii file (not edited); multi-record. See datadict for description

SPSS files:
Contains all data from questionnaire, edited, by section
Also contains calculated variables
HID
Summary file at household level
Summary files for expenditure
Summary file for income

File: b1.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe: all households
Nnn records, nn variables

File: b2.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe: all individuals
Nnn records, nn variables

File: b301.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe: all individuals
Nnn records, nn variables

File: b302.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b303.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b304.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b305.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b306.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b307.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b308.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b309.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b310.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b311.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b312.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b313.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b314.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b315.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b316.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b41.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b42.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b43.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b44a.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b44b.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b44c.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b45a.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b45b.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b45c.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b4601.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b4602.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b4603.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b4604.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b4605.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b4606.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b4607.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b4608.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b4609.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b4610.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b51.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: b52.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients

Universe:
Nnn records, nn variables

File: b53.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: health.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

Constructed files

File: foodexp.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: foodexp2.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: foodexp3.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: foodexp4.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: foodexp5.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: nfoodex.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: nfoodex2.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

File: nfoodex3.sav

Collected variables: all variables from section 1
Computed/constructed variables: HID + weighting coefficients
Universe:
Nnn records, nn variables

List and description of program files

IMPS

Data dictionary
Hiesmain.dd
Hiesmain.lst
Centry
Hiesmain.ap
CONCOR
Hiesedit.cn
Hiesedit.exe
Hiesedit.cob

SPSS – Data editing

Export HIES2000 to SPSS.sps
Converting the dataset from IMPS format (Ascii file) to SPSS format.

Label values HIES2000 – 1 of 5.sps to *Label values HIES2000 – 5 of 5.sps*
Labelling all variables and values
1 of 5: blocks 1 to 3.5
2 of 5: blocks 3.6 to 3.10
3 of 5: blocks 3.11 to 3.16
4 of 5: blocks 4.1 to 4.5c
5 of 5: blocks 4.6 to 5.3

Listing of quantity units.sps and *Quantity units recode.sps*
Standardizing the spelling of each "quantity unit" modality

> The variable "quantity unit" (blocks 3.1 to 3.15 and 4.5.a) was not pre-coded. It had been decided to enter the information exactly as it appears in the questionnaires (as an alphanumeric variable). Due to typo errors and others, the original dataset contained a very long list of modalities (for the same unit, many different spelling were used; for example, Kg was spelled "KG", "Kg", "kg", "kg.", "kilogram", etc). It was necessary to standardize the spelling of each modality. Two SPSS programs have been designed for this: the first one produces a listing of all modalities found in the dataset. Based on the modalities listed by this first program, a second one was designed to automatically recode the modalities when required.

Misc. checks in blocks 1 and 2.sps

> Many consistency checks have already been implemented during data entry (using IMPS-CONCOR program). A SPSS program is used to re-check some key variables and perform a few additional checks.

- No duplicated household identification (in block 3.1): no errors found
- No duplicated individual identification (in block 3.2): no errors found
- Household size declared in block 3.1 is equal to the number of members in block 3.2, for each household: no errors found
- Consistency of information collected on land ownership (in block 3.1): 184 records found where type of land is missing; 108 records found where acres owned is missing (no correction); 18 cases found where “ownership of land’ is “no” but acres owned > 0 (this may be corrected manually or automatically);
- There is one and only one head declared for each household: no errors found
- Age of the head must be >= 16: 3 errors found; must be corrected manually
- If a person is declared as “spouse of the head”, his/her marital status must be “married”: 64 errors found; must be corrected manually
- If the number of spouses found in a household > 0, the head must be declared as “married”: 63 errors found; must be corrected manually
- Consistency between age and marital status: 78 errors found; must be corrected manually
- Consistency between age and relation to the head of household: 195 errors found; must be corrected manually. The relatively high number of errors may be due to the poor legibility of the codes in the questionnaire.

The SPSS program produces listings of errors that can be printed for easy identification of the questionnaires and manual corrections. No errors were found for the key variables.

Remove totals from b301.sps
Remove totals from b302.sps
Remove totals from b303.sps
Remove totals from b304.sps
Remove totals from b305.sps
Remove totals from b306.sps
Remove totals from b307.sps
Remove totals from b308.sps
Remove totals from b309.sps
Remove totals from b310.sps
Remove totals from b311.sps
Remove totals from b313.sps
Remove totals from b314.sps

The consumption blocks (3.1 to 3.14) contain information collected and information computed by the interviewers or supervisors, such as totals of quantities and values by product category (the questionnaires shouldn’t contain such computed data; they are additional sources of errors and unnecessary increase the workload for both data collection and data processing). Further, these computations have not been done systematically. Some questionnaires do not contain any computed total; some contain totals for some sections only. These totals had to be removed from the data files. If not, the quantities and values they represent would be counted twice when data are aggregated. The above mentioned SPSS programs have been produced to identify and automatically remove these totals from blocks 3. 1 to 3.14.

Extreme values in b301 to b315.sps
to identify the records containing such outliers

The dataset contained some extreme and unrealistic values. Most of these cases arise from data entry errors (example: operator entered "10000" instead of "100.00"). The program has been produced to identify the records containing such outliers. This program is relatively basic, although very useful. It may not be considered as the "ultimate" solution to the problem of outliers (further checks will have to be done). It doesn't look for outliers by product. It only identifies extreme values, based on arbitrary "extreme-lines". To set these lines, data files have been sorted (by descending order of quantity/value), and the lines were set after visual observation of the data files. The program may be easily modified if necessary (to use lower/higher lines).

The program produces listings of errors (these listings have been save as an MS-Excel file named *extremes.xls*). It doesn't make any automatic imputation. The corrections have to be made manually, after visual check of the questionnaires. A SPSS program named *Locate extremes in b301 to b315.sps* was produced for easy manual correction. All corrections have been made.

Check duplicated codes for exp.sps

identifying and listing all records containing duplicated item codes.

The survey questionnaire is designed in such a way that the same item codes shouldn't appear twice in the same file (for blocks 3.1 to 3.15). Duplicated codes indicate errors in data entry. The program has been produced for identifying and listing all records containing duplicated item codes. This program doesn't make any automatic correction. All corrections have to be made by visual/manual correction.

Check quantities and values.sps

The questionnaire contains two reference periods for food consumption: last week and last month. In most cases, the information was collected for only one of the two periods, and the interviewers or supervisors had manually computed the equivalent for the other period. When information was collected at the weekly level, the quantity and value were multiplied by 4 to obtain the monthly figures. The exact figures should have been obtained by dividing the numbers by 7 then multiplying the result by 30. When the information was collected for the month, it was divided by 4 to obtain the data for the weekly reference period (instead of being divided by 30 and multiplied by 7). When both reference periods are reported in the data file, it is unfortunately impossible to know whether the household originally reported the information for the week or for the month.

Many records in block 3.1 (and also in blocks 3.2 to 3.14) contain inconsistent information: weekly (monthly) quantity and/or value higher than monthly (annual) quantity and/or value. In some cases, the inconsistencies are due to computation errors made by the interviewers. For example, a consumption of "2 (kg per month)" was divided by 4 to get the information per week, and was reported as "500 (grams per week)" instead of "0.5 kg".

A SPSS program was produced to locate these errors. This program distinguishes different types of errors:

- Errors due to use of two different quantity units (grams/Kg; consistent if a ratio of 1000 is applied to the computed values)
- Records where the information is consistent if a ratio of 10/100/10000 is applied.
- Records where the information is not consistent and cannot be solved by applying a multiplying factor.

Errors of type 1 have been automatically corrected (352 corrections made, out of about 96000 records). Errors of type 2 have been manually corrected (335 records corrected). Errors of type 3 have not been corrected. A high number of errors has been found (about 700 records where quantity last week > quantity last month, and about 450 records where value last week > value last month. This has not been corrected yet. The strategy for correcting these records should be discussed with the survey consultant (Dr. Sardana).

The SPSS program also looks for outliers by product. The program computes the national average and standard deviation per quantity unit for each item/quantity unit. All values lower than the national average – (5xSD) or higher than the national average + (5xSD) are considered as outliers. The program doesn't make any automatic correction. About 370 cases of outliers have been identified before editing of extreme values/consistency between reference periods. Some of these cases will be solved by editing the information on quantity or value. The real number of outliers (based on this identification method) is lower than 370. No correction has been done yet.

A better check will be to compare the price per unit with information on prices collected by the CPI division (see below "Comparison with collected prices"). This may be done only after standardizing the quantity units.

Aggregations of expenditure.sps

A program was produced to aggregate the food consumption per household. **This aggregation program is basic and may be used for editing purposes only**, not for "final aggregations" (see "*4.7 - Aggregations of consumption/expenditure*" below). This program computes the consumption per household per month, by product and category of product. The aggregations are based on the weekly reference period when available; monthly data is used only when information is not available for the week. **This is an arbitrary decision, and there is no reason to consider it as the best solution.**

The program creates 5 working data files (all with data at the household level):

- *Foodexp5.sav* : food expenditure by household and food item (at 5 digit code level)
- *Foodexp4.sav* : food expenditure by household and food item (at 4 digit code level)
- *Foodexp3.sav* : food expenditure by household and food item (at 3 digit code level)
- *Foodexp2.sav* : food expenditure by household and food item (at 2 digit code level)
- *Foodexp.sav* : food expenditure by food item (summary by category of product)

Note: data at the "5 digits level" means data at the "most detailed level available". For some categories of items, it may in fact be the 3 or 4 digits level. This is the reason why foodexp5.sav contains records with 3 or 4 digits codes (it also contains records at the 2 digits level, due to errors in data collection; see below).

For some food items (block 3.1), information had to be collected at the "5 digits" level and others at the "4 digits" level. Unfortunately, the questionnaire also provided space for recording the information at a more aggregated level. As a consequence, data that had to be collected at the 5 digits level is sometimes available only at the 2/3/4 digits level. Also, information that was required at the 4 digits level may be available only at the 2/3 digits level. This clearly appears in file *foodexp5.sav* (after sorting the file by ascending order of variable b301cod5). This remark also applies to other sections of the questionnaire. This can be a problem when aggregating the values and producing tables on detailed structure of consumption. For that reason, it is recommended to "split" these records. For example, if the information on rice consumption for one household was collected at the 3 digits level (one record with code 111) instead of the 4 digits level (1 to 6 records, with codes 1111 to 1115 and 1119), the value and quantity consumed may be divided into 6 records, based on proportions computed at the stratum level. By doing this, **nothing will be changed to the total consumption of the household** and to the consumption by category of product. The only impact of this editing will be to allow the production of "clean" tables on consumption by food item (without "undefined" rows).

This can be done automatically. No manual correction or computation will be required. The SPSS program for doing this has not been produced yet.

Consistency of aggregated exp.sps

Based on the (provisional) aggregated food consumption (file foodexp.sav), one may compute the share of each category of food product in the total food consumption. One may also identify the extreme values of per capita food consumption.

This program is used for:

- **Listing households with very low per capita food expenditure: < 100 Nu. per capita per month (about 15 cases found, including 12 with total food consumption = 0)**
- Listing households with very high per capita food expenditure: > 4000 Nu. per capita per month (few households found)
- Producing a table showing the monthly per capita food expenditure by quintile (min, max, mean and median values)
- Checking the share of each category of food product in the total food consumption. Based on the national mean share and standard deviation, the following rules have been established and used for checking the consistency of these shares (valid range):
  - Cereals: >= 10% and <= 75%
  - Dairy products: <= 35%
  - Eggs: <= 8%
  - Fish: <= 20%
  - Meat: <= 30%
  - Fruits and vegetables: <= 30%
  - Other misc. food: <= 25%
  - Beverages: <= 25%

The SPSS program will produce listings of (potential) errors, identifying the corresponding households. It will not make any automatic correction. Corrections, when required, will have to be made manually after visual check of the questionnaires. **Out of range values are not necessarily errors**. The listings will only identify "suspect" values.

Check rent values.sps

The instruction given to enumerators was to record the amount of actual rent paid if the house is taken out of rent, or the "rental value imputed on the basis of prevailing market rent for a similar house in the locality".

In many cases (about 25% of households), no rent (or rental value) was declared. Imputation will have to be made. Also, the declared rent/rental values will have to be checked. Unfortunately, there is no data on the "prevailing market rent in the locality" and the questionnaire doesn't contain variables describing the house.

The usual procedure to check/impute this information is to run a regression of rental values on housing characteristics and then use the coefficients from such a regression to impute rental values for those who do not rent their dwellings. Since there is no information available on the housing characteristics, a different strategy will have to be implemented. Instructions on how to impute values for missing rent must be provided by the survey consultant (Dr. Sardana).

A program has been written that produces tables showing the (unweighted) percentage and number of households who didn't declare any rent/rental value, and the (unweighted) minimum, maximum, mean and median rent/rental value when declared, by region (towncode).

Aggregations of all expenditure.sps

Aggregations of food expenditure.sps

Aggregations of non-food expenditure.sps

Consistency between ref. Periods b302 to b315.sps

Consistency checks in block 4.1.sps

This program doesn't contain any command for automatic corrections. It will only produce listings of errors for manual correction. The program is divided into 8 parts. Each part must be run separately.

- PART 1: Check that the person has a corresponding record in block 2 (list of records without corresponding record in block 2).
- PART 2: Check that total days (sum of columns 3 to 5)= 365.
- PART 3: Check that col.6 is consistent with info. in cols 3 to 5.
- PART 4A: Check that col.8 to 11 are consistent with col.6. List of persons employed but with no information in cols 8, 10 and 11
- PART 4B: Check that col.8 to 11 are consistent with col.6. List of persons not employed but with information in cols 8, 10 and 11
- PART 5: Compare information with information from block 2. List of persons aged < 10 who are declared employed/available for employment.
- PART 6A: Compare information with information from block 2. If person declared as worker in block 2, col 8, this person must have worked at least one day (col 3 in block 4.1 >= 1).
- PART 6B: Compare information with information from block 2. If person declared as non-worker in block 2, col 8, this person must have worked 0 day (col 3 in block 4.1 = 0).

Consistency checks in block 4.2.sps

This program doesn't make any automatic corrections. The program contains 13 parts that must be executed separately, each one in 3 steps: (1) identification and listing of errors, (2) manual corrections, (3) save/sort file.

- PART 1: Check that the person reported at least one day of employment in col 3, block 4.1
- PART 2: Check that no. of days worked in last month < 31 and <= col.3 block 4.1 (days in last month <= days in last month).
- PART 3: Check that col.10 = col.8 + col.9.
- PART 4: Check that col.16 = col.12 + col.13 + col.14 + col.15.
- PARTS 5 to 13 will list extreme values. For identifying extreme values, some arbitrary "reference values" have been used. You may want to change these reference values if you find them not realistic or if the number of records listed as extreme is too high.
- PART 5: Identification of extreme values for column 8. Reference values: b402c08 expected not to be <500 or >10000
- PART 6: Identification of extreme values for column 9. Reference values: b402c09 expected not to be >2500
- PART 7: Identification of extreme values for column 11. Reference values: b402c11 expected not to be >(b402c10*0.25). In this case, the extreme value has been defined as a percentage of the value found in column 10. I have no idea of what the maximum percentage of deduction at source may be. In the program, I used 25%. You may want to change that (program line 230).

- PART 8: Identification of extreme values for column 12. Reference values: b402c12 expected not to be >500
- PART 9: Identification of extreme values for column 13. Reference values: b402c13 expected not to be >500
- PART 10: Identification of extreme values for column 14. Reference values: b402c14 expected not to be >500
- PART 11: Identification of extreme values for column 15. Reference values: b402c15 expected not to be >500
- PART 12: Identification of extreme values for column 17. Reference values: b402c17 expected not to be >5000
- PART 13: Identification of extreme values for column 18. Reference values: b402c18 expected not to be >5000

Check block 4.3 and related.sps

SPSS – Tabulation and analysis