Datasets for ADePT-Food Security Module version 3

by Ana Moltedo and Cristina Álvarez Sánchez

Statistics Division (ESS)

Food and Agriculture Organization of the United Nations

Rome January 2018

This document is an updated version of Chapter 4 within the book *Analyzing Food Security Using Household Survey Data : Streamlined Analysis with ADePT Software* published by the World Bank in 2014. The editors of the book are Ana Moltedo, Nathalie Troubat, Michael Lokshin and Zurab Sajaia.

Table of Contents

List of Tables

Abbreviations and Acronyms

ADePT-FSM	Software Platform for Automated Economic Analysis-ADePT Food Security Module
ASEAN	Association of Southeast Asian Networks
CPI	Consumer Price Index
DHS	Demographic Health Survey
FAO	Food and Agriculture Organization of the United Nations
FCT/FCDB	Food Composition Table/Food Composition Database
FPI	Food Price Index
FS	Food Security
FSM	Food Security Module
HBS	Household Budget Survey
HES	Household Expenditure Survey
HIES	Household Income and Expenditure Survey
ILO	International Labour Organization
INFOODS	International Network of Food Data Systems
LSMS	Living Standards Measurement Study
NCT	Nutrient Conversion Table
HCES	Household Consumption Expenditure Surveys (HBS/HES/HIES/LSMS)
UNICEF	United Nations Children's Fund
USDA	United States Department of Agriculture

Datasets

ADePT-FSM requires four datasets (loaded either in STATA or SPSS format). Three datasets contain data extracted mainly from the original HCES files. They are:

- Dataset 1 (HOUSEHOLD), including household and household head characteristics;
- Dataset 2 (INDIVIDUAL), with household members characteristics;
- Dataset 3 (FOOD), ideally with quantities and monetary values of food items habitually consumed by households.

These three datasets include a household identification code that allows for matching information among them.

The fourth datasets contains data extracted from national/regional food composition tables and food composition databases (FCTs/FCDBs):

- Dataset 4 (COUNTRY_NCT [nutrition conversion table]), with calorie and nutrient values for the food items collected in the survey.

Dataset 1 (HOUSEHOLD)

Dataset 1 has one record for each household and provides information on household characteristics (household size, region and area of residence, total consumption expenditure, income, etc.), household head characteristics (gender, age group, level of education, occupation and economic activity) and price indexes (i.e., he consumer price index [CPI] and the food price index [FPI]).

While household and household head characteristics are extracted from HCES data, the FPI and CPI are provided by national or international organizations, such as the International Labour Organization (ILO). The household and household head characteristics are mainly used to create groupings and produce sub-national estimates. The FPI and CPI are instrumental for deflating the food expenditures and income/expenditure values, respectively, in the presence of one-year surveys.

In addition, the analyst can also define up to five "spare" variables to further disaggregate the food consumption statistics (*hm_var1, hm_var2, ..., hm_var5*). The spare variables can correspond to household/household head characteristics or can be a combination of them.

Table 1 shows the main characteristics of the variables included in dataset 1, the values they can assume, and the associated checks to be performed.

Variable names depicted in the table are not mandatory; however comparison of results intra- and inter- countries is greatly facilitated if a common set of variable names is adopted. Each variable, and each value of categorical variables, has to be described by an appropriate label. Finally, none of the variables are allowed to have missing values.

An important distinction has to be made between *Household member* and *Food partaker*. While only household members share the household income, the food acquired by the household can be distributed to non-household members (such as guests and employees). Therefore, the number of food partakers corresponds to the number of people who actually consumed the food during the reference period.

Example of a reference period for food consumption data for one month:

- A household reported four members
- One member was absent
- One guest and one housekeeper with a child also consumed the food acquired by the household

In this case, the number of partakers for the reference period will be six instead of four: four household members minus the absent member plus the guest, the housekeeper, and the child.

The number of partakers is not always collected in household surveys. However, it is highly recommended to check if this information is available. If so, the variable *Number of partakers* has to be included in dataset 1.

When deriving the statistics related to a variable of analysis such as the location or area of residence or the occupation of the household head, ADePT-FSM excludes all records with missing values in that variable. The consequence of this could be to produce unreliable statistics for that group of analysis. Therefore, it is important to avoid the presence of missing values as much as possible.

Another crucial note regards the variable *Location of the household*. The analyst should always select the geographical domain(s) of which the survey data are representative. For

instance, if the original HCES datasets include both the variable *Province* and the variable *Region*, and the survey was designed to be representative at the province level, then the analyst should select the province.

Finally, only the households that declared food consumption should be in dataset 1; the other ones should be deleted. After the deletion, the household weights should be amended, as follows:

- Sum by enumeration area: (*hh_size* * *hh_wgt*); note that at the national level (*hh_size* * *hh_wgt*) = *population_original* (≈ total country population at the survey year).
- Delete the households that did not declare food from dataset 1.
- Sum by enumeration area: (*hh_size* * *hh_wgt*) = *population_new*.
- Compute: *hh_wgt_adj* = *hh_wgt* * (*population_original* / *population_new*).

Table 1: Dataset 1 (HOUSEHOLD)

Variable name and format	Rationale and values	Remarks and checks
Household number (hh_no) Format: Numeric or String	Identification code of the surveyed household. Sequential numbers or a combination of geographical codes (district, area, village, region, etc). Necessary to link Dataset 1 with Datasets 2 and 3	Each household has to be identified by a unique code. Only the households declaring food consumption should be included in this dataset.
Location of the household (region) Format: Numeric	Identification code of the district, province, or region of residence of the surveyed household. This variable has to *include the labels* corresponding to the geographical groups.	The number of sampled households in each region gives an idea of how reliable the statistics are. The smaller the sample size the higher would be the standard error of an order of 1/sqrt(n). Thus it may be necessary to group some locations into a new one.
Area of residence of the household (urb_rur)	Identification code of the area (urban, rural, semiurban, etc.) of residence of the	The number of sampled households in each area of residence gives an

Format: Numeric	surveyed household. This variable has to *include the labels* corresponding to the areas. Examples: • Code = 1, Label: Urban • Code = 2, Label: Rural	idea of how reliable the statistics are. The smaller the sample size the higher would be the standard error of an order of 1/sqrt(n). Thus it may be necessary to group some areas into a new one (for instance, urban with semiurban).
Household size (hh_size) Format: Numeric	Number of people who usually live together and share the household income.	Excludes • Domestic workers, friends, or relatives who neither live in the house nor share the income • Domestic workers, friends, or relatives who live in the house but don't share the household income
Group household size (hhsizec) Format: Numeric	Identification code of category of household size. This variable has to *include the labels*. Examples: • Code = 1, Label: Less than three • Code = 2, Label: Three or four • Code = 3, Label: Five or six • Code = 4, Label: More than six	The number of sampled households in each category within the group household size gives an idea of how reliable the statistics are. The smaller the sample size the higher would be the standard error of an order of 1/sqrt(n).
Number of food partakers (partakers) Format: Numeric	Average number of people who shared the food during the period of food data collection (reference period)	Partakers are individuals who shared the household food during the reference period. Includes housekeepers, friends, and relatives who may not live in the house but shared the food. Excludes household members who were absent during the food data reference period and therefore did not consume the food.

Household weight (hh_wgt) Format: Decimal	The value of the household weight depends on the sampling frame and is equal to the expansion factor divided by the probability of the household to be sampled. *Household weight* should be *adjusted* for nonresponding households.	The sum of the product *number of household members * household weight* has to be close to the total country population at the year of the survey. Only households declaring food consumption should be included in this dataset. Therefore, after deleting households that did not declare food consumption, household weight should be amended accordingly. Details are provided at the bottom of the table.
Total household consumption expenditure (thh_cexp) Format: Decimal	Sum of household food and nonfood consumption expenditures. Excludes all expenditures not related to household *consumption*, such as investments, life insurance premiums, food for pets or given away, etc.	Monetary values should be expressed in *daily basis.* Each household should have a *positive* value of total consumption expenditure. This value has to be greater than or at least equal to the respective household total food expenditure.
Total household income (thh_inc) Format: Decimal	Sum of the income received by each household member; includes all the possible sources (wages, profit from self-employment, sales of self-produced goods and services, income in kind, transfers, rent received, etc.).[a]	Monetary values should be expressed in *daily basis*. If income data are either not available or not reliable, *total expenditure* can be used as a proxy of income. Total expenditure includes consumption and non-consumption expenditures such as direct taxes, insurance premiums, food given away or animal feed, etc. Each household should have a *positive* value of total income. Also, this value has to be greater than or at least equal to the

		respective household total consumption expenditure.
Month of food data collection (month) Format: Numeric	Identification code of the month during which the food consumption/acquisition data were collected. This variable has to include labels, e.g., values of *1, 2, 3 . . . 12* corresponding to the months January, February, March . . . December.	
Year of the food data collection (year) Format: Numeric	Identification code of the year during which the food consumption/acquisition data were collected. Examples: Values of *1998, 1999, 2000, 2003 . . .* etc.	
Consumer price index (cpi) Format: Decimal	Measures the change in the purchasing power of a currency and the rate of inflation. The consumer price index expresses the current prices of a basket of goods and services in terms of the prices during the same period in a previous year, which shows the effect of inflation on purchasing power. It is one of the best-known lagging indicators. It is used to correct total consumption expenditure and total income for inflation or deflation. All the consumer price indexes should refer to the same base period.[b]	Use the value corresponding to the month and year in which the household food consumption data were collected. If the monetary values are already deflated (or the survey was conducted only over a period of a few months), this variable is not needed.
Food price index (fpi) Format: Decimal	Measures the change in the purchasing power of a currency and the rate of inflation. The food price index expresses the current prices of a food basket in terms of the prices during the same period in a previous year, which shows the effect	Use the value corresponding to the month and year in which the household food consumption data was collected. If the monetary values are already deflated (or the survey was

	of inflation on purchasing power. It is used to correct food monetary values for inflation or deflation. All the food price indexes should refer to the same base period.[c]	conducted only over a period of a few months), this variable is not needed.
Gender of the household head (gender_hh) Format: Numeric	Identification code of the gender of the household head. This variable has to *include labels* corresponding to both sexes. Compulsory value codes: • Code = 1, Label: Male • Code = 2, Label: Female	Missing data on gender are not valid; each household head has to have a value of 1 or 2.
Group household head age (hmagec) Format: Numeric	Identification code of the group to which the household head belongs according to age. This variable has to *include the labels*. Example: • Code = 1, Label: Less than 30 • Code = 2, Label: Between 30 and 44 • Code = 3, Label: Between 45 and 59 • Code = 4, Label: More *than 59*	Records with missing values are deleted by the program, and this may cause unreliable estimates for the variable. The number of sampled households in each category within the group of the household head age gives an idea of how reliable the statistics are. The smaller the sample size the higher would be the standard error of an order of 1/sqrt(n).
Group highest level of education of the household head (hm_edu) Format: Numeric	Identification code of the group to which the household head belongs according to the highest level of education attended by them. This variable has to *include labels*. Examples: • Code = 1, Label: *No education* • Code = 2, Label: Primary school	Records with missing values are deleted by the program and this may cause unreliable estimates for the variable. The number of sampled households in each category within the group highest level of education of the household head gives an idea of how reliable the statistics are. The smaller the

	• Code = 3, Label: Secondary school • Code = 4, Label: *Tertiary education*	sample size the higher would be the standard error of an order of 1/sqrt(n).
Group occupation of the household head (hm_occ) Format: Numeric	Identification code of the group to which the household head belongs according to occupation. It is highly recommended to recode the occupations collected in the survey into major occupation groups defined by the first digit of national/international classifications such as ISCO.[d] This variable has to *include labels*. Examples: • Code = 1, Label: Managers and professionals • Code = 2, Label: Technicians and clerical support • Code = 3, Label: Service and sales workers • Code = 4, Label: Agricultural, forest, fishery workers • Code = 5, Label: Without occupation	Records with missing values are deleted by the program and this may cause unreliable estimates for the variable. The number of sampled households in each category within the group occupation of the household head gives an idea of how reliable the statistics are. The smaller the sample size the higher would be the standard error of an order of 1/sqrt(n).
Group economic activity of the household head (hm_eact) Format: Numeric	Identification code of the group to which the household member belongs according to economic activity. Recode the economic activities collected in the survey into major economic activity groups defined by the first digit of national or international classifications such as ISIC (Rev. 4).[e] This variable has to *include labels*. Examples: • Code = 1, Label: Primary	Records with missing values are deleted by the program and this may cause unreliable estimates for the variable. The number of sampled households in each category of the group economic activity of the household head gives an idea of how reliable the statistics are. The smaller the sample size the higher would be the standard error of an order of 1/sqrt(n).

	(agriculture, fishing, hunting, and mining) • Code = 2, Label: Secondary (manufacturing) • Code = 3, Label: Services • Code = 4, Label: Without an activity	
Additional variables with household/ household head characteristics (hm_var1, …, hm_var5) Format: Numeric	Identification code of the group to which the household or household head belongs according to additional variables. Some examples are: • At household level: Poverty level, Source of drinkable water, Type of access to sanitation, etc. • At household head level: Marital status, Religion, Ethnic group, etc. • At household head and household level: Gender of household head combined with area of residence (urban-rural), etc. These variables should *include the labels*.	Records with missing values are deleted by the program and this may cause unreliable estimates for the variable. The number of sampled households in each category within the population group gives an idea of how reliable the statistics are. The smaller the sample size the higher would be the standard error of an order of 1/sqrt(n).
Note: ISCO - International Standard Classification of Occupations / ISIC - International Standard Industrial Classification of All Economic Activities. a. For detailed information refer to the *Canberra Handbook on Household Income Statistics* (2nd ed., 2011) at http://www.unece.org/index.php?id=28894. b. Sources of data: National or international institutions such as ILO. This information can also be found in FAOSTAT: http://faostat.fao.org/site/683/Default.aspx#ancor. c. Sources of data: National or international institutions such as ILO. This information can also be found in FAOSTAT: http://faostat.fao.org/site/683/Default.aspx#ancor. d. For more information, see http://www.ilo.org/public/english/bureau/stat/isco/index.htm. e. For more information, see http://unstats.un.org/unsd/cr/registry/isic-4.asp.		

Dataset 2 (INDIVIDUAL)

Dataset 2 has one record for each member of the household and provides information on members' characteristics such as sex, age and height. This information is necessary to estimate the dietary energy requirements of the population. Even though some HCES collect data on height, this is usually done only for children under 5 years of age and/or for women in reproductive age. Therefore, the distribution of height across the gender-age groups is usually derived from other sources such as Demographic Health Surveys (DHS), country reference tables or specific publications (e.g., James and Schofield, 1990)[1].

Each variable/variable's value has to be described by an appropriate label explaining its content of information.

Finally, for each household, the number of records in dataset 2 should be equal to the corresponding value of the variable *hh_size* (size of the household) in dataset 1. This means only information about household members is required in this dataset. Therefore, records related to food partakers, such as housekeepers, friends, and relatives who are not household members, should be excluded from dataset 2.

Table 2: Dataset 2 (INDIVIDUAL)

Variable name and format	Rationale and values	Remarks and checks
Household number (hh_no) Format: Numeric or String	Identification code of the household. Sequential numbers or a combination of geographical codes (district, area, village, region, etc.). Necessary for linking dataset 2 with datasets 1 and 3.	Each household has to be identified by a unique code. Only the households declaring food consumption should be included in this dataset.
Relationship between the household member and the head of the household (hm_rel) Format: Numeric	Identification code of the relationship between the household member and the head of the household. This variable has to *include labels.* *Exclude* all individuals who do not share the household income, such as housekeeper, guests, and relatives who do not live in the house or live in the household but do not share the household	*All households* must have a household head. There has to be only *one head* per household.

[1] It is recommended to use the median height of each sex and age group, instead of the mean.

	income. This variable is not needed for the analysis but it is useful to create it for not losing the information.	
Gender (gender) Format: Numeric	Identification code of the gender of the household member. This variable has to **include labels** corresponding to both sexes. **Compulsory value codes are :** • Code 1 = male • Code 2 = female	Missing data on gender are not valid; each household member has to have a value of 1 or 2.
Age (hm_age) Format: Numeric	Values are to be expressed in years. For children less than one year of age, assign the value 0.	Missing data on age are not valid; each household member has to have an age value.
Height (height) Format: Decimal	Values are to be expressed in cm.	Missing data on height are not valid; each household member must have a value greater than 0 in this variable. When height data are not collected in the survey, the median height by age/sex groups obtained from national reference tables, specific publications, or household demographic surveys should be used.

Dataset 3 (FOOD)

Dataset 3 contains information on the household food consumption both in quantity and monetary terms, disaggregated by four main food sources. Each record corresponds to a food item consumed/acquired by the household through a specific source; the dataset may therefore have one or more entries of a given food item per household, depending on the number of sources from which the food item is obtained.[2]

Data in the food dataset should fulfill the following requirements:

- All the food item quantities (including beverages) should be expressed in only one standard unit of measurement to be chosen among kilogram, gram, or pound. For this reason, the analyst has to transform all the food quantities into one unit.
- Food quantities and monetary values should be expressed on a daily basis. It is important to identify the actual reference period for which the households declared food consumption. The recall period is usually clearly stated at the beginning of the food module in the questionnaire, and the enumerators should have had the responsibility to convey the message as clearly as possible. This check is particularly relevant when food data are collected with a diary. In a consumption survey, if a diary is given to the households for a week, households may skip some days. In these cases, the reference period is the actual number of days the diary was filled in. However, in an acquisition survey, the same situation may require a different treatment. If a household is asked to report the food acquired in a week, and the diary is filled in for three days with considerable daily quantities, then it is likely that the food acquired in the three days also covers the four days with missing data. In this case, the most accurate reference period is still seven.
- Food quantities must be related with the variable *Number of food partakers* or *Household size*. Food quantities should be expressed at the household level, not in "per person" amounts. ADePT-FSM automatically calculates the per person values by using the variable *Number of food partakers*, if available; otherwise, it uses the variable *Household size*. Also, food monetary values have to be expressed at the household level so that

[2] For example: a household consumed potatoes; they were partly purchased on the market and partly obtained from own production. This household will have two entries (i.e., two lines) for the food item potatoes: one with quantity and expenditure related to the purchase, and the other with quantity and monetary value related to own production.

ADePT automatically calculates the per person values by using the variable *Household size* (note that the variable *Food partakers* is not taken into consideration when deriving food monetary values at the individual level).

The preparation of the food dataset may require some computational steps to accurately estimate missing quantities of food consumed or monetary values.

Estimate Accurate Quantities of Food Consumption

Since the analysis is focused on the food consumed by the household (HH), the food given away, processed for resale, given to pets/livestock, and wasted has to be excluded. Such detailed data are rarely collected in the HCES, but if they are collected they should be subtracted from the total amount of food acquired. Details are provided in Table 3.

Table 3: Treatment of Food Acquired but Not Consumed by the Household

• Food given away • (e.g., to other households, neighbors)	⇒ Subtract from the household food consumption the food given away
• Food processed for resale (e.g., flour, sugar, eggs used for a cake to be sold)	⇒ Subtract from the household food consumption the food acquired for resale
• Food given to pets or for feeding livestock	⇒ Subtract from the household food consumption the food given to pets or used for feeding livestock
• Food thrown away (e.g., rotted, wasted, etc.)	⇒ Subtract from the household food consumption the food thrown away

It is also important to check if information on the *starting and ending levels of food stock* is available, especially when the survey collects food acquisition data. If data on stocks are collected, they should be used as follows to derive the household food consumption:

HH food consumption = HH food acquired + HH starting food stock – HH ending food stock

Estimate Missing Quantities and Expenditures

For a food item reported by the household, a food quantity with a missing or 0 value is allowed *only* if the food item was consumed away from home (*f_source* = 4). For food expenditure, missing or 0 values are not accepted. Therefore, before loading the data in ADePT, the analyst must estimate the missing/0 values based on median food item unit values. See chapter 2 for a detailed account of such procedures.

Table 4 illustrates the main characteristics of the variables included in dataset 3, the values they can assume, and the associated checks to be performed.

Variable names depicted in the table are not mandatory. Each variable has to be described by an appropriate *label* explaining its content.

Table 4: Dataset 3 (FOOD)

Variable name and format	Rationale and values	Remarks and checks
Household number (hh_no) Format: Numeric or String	Identification code of the household. Sequential numbers or a combination of geographical codes (district, area, village, region, etc.). Necessary to link dataset 3 with datasets 1 and 2.	Each household has to be identified by a unique code. Only the households declaring food consumption should be included in the dataset.
Food item code (item_cod) Format: Numeric	Identification code of the food items listed in the survey. This variable should *include labels*. COICOP[a] or national classification codes can be used.	Include alcoholic beverages and food consumed away from home (canteens, bars, restaurants, etc.). Exclude nonfood items, such as cigars, cigarettes, tobacco, and drugs.
Food item quantity (fd_qty) Format: Decimal	Food quantities should reflect the *food consumption or acquisition* of the household. All food quantities should be expressed on a *daily basis*. All food quantities, *including beverages*, should be expressed in the *same unit of measurement*. The unit of measurement can only be	ADePT estimates the calories and nutrients of missing food quantities *only* for the food consumed away from home. If a household *declared* expenditure for a food item with a food source different from 4 (consumed away), the quantity cannot be missing or 0. *The analyst has to estimate* the missing/0 quantities based on the unit values. The estimation

	grams, kilograms, or pounds. Keep track (by using labels or adding an extra variable) of the unit of measurement used.	has to be carried out *before* loading the dataset in ADePT.
Food item monetary value in local currency (fd_mv) Format: Decimal	Amount paid or estimated for the reported quantity. All food monetary values should be expressed on a *daily basis*.	If a household declared a quantity for a food item, the expenditure/monetary value cannot be missing or 0. *The analyst has to estimate* the missing/0 expenditure/monetary values based on the food item unit values. The estimation has to be carried out *before* loading the dataset in ADePT.
Source of food item (f_source) Format: Numeric	Identification code of the food source. This variable should *include labels*. Compulsory value codes. Examples: • Code = 1, Label: Purchased and consumed at home • Code = 2, Label: Own production • Code = 3, Label: Received in kind • Code = 4, Label: Consumed away from home	ADePT analyzes *four food sources*. If there are *fewer than four food sources, keep this coding structure*. Food sources such as received free or as a gift, from food aid, income in kind, gathering, or fishing should be labeled as *Received in kind* with code 3. No missing values are allowed in this variable.
Note: COICOP = Classification of Individual Consumption According to Purpose. a. For further information see http://unstats.un.org/unsd/cr/registry/regcst.asp?Cl=5.		

Dataset 4 (COUNTRY_NCT)

Dataset 4 contains information on the composition of each food item listed in the survey, in terms of energy and nutrients (nutrient values) per 100 grams *edible*[3] portion. This information is found in national or regional FCTs/FCDBs available either online (e.g., USDA FCT) or in hard copy (e.g., ASEAN FCT). To build dataset 4, the analyst has to match each food item listed in the survey with a food item described in the selected FCT/FCDB.

This section is divided into two parts. The first one describes the variables to be included in the dataset; the second provides detailed guidelines on how to build it.

Variables in Dataset 4

Dataset 4 includes three distinct groups of variables.

- The first group includes calorie and macronutrient values, and it represents the minimum information required to execute the ADePT-Food Security Module.
- The second group includes nutrient values for some vitamins and minerals, necessary to conduct a micronutrient analysis.
- The third group includes nutrient values for essential amino acids, necessary to conduct an analysis of amino acids.

The following tables show the main characteristics of the variables included in dataset 4 and the associated checks to be performed.

Table 5 describes the minimum information required.

Table 6 focuses on the micronutrient analysis.

Finally, Table 7 regards the information needed for the amino acids analysis. Not all the food composition tables have information on amino acids. Information on amino acids can be found in the following sources:

- U.S. Department of Agriculture: http://www.nal.usda.gov/fnic /foodcomp/search/index.html
- FAO website for Amino-Acid Content of Foods and Biological Data on Proteins: http://www.fao.org/docrep/005/AC854T/AC854T00 .HTM

[3] For example, without considering inedible parts such as peels, bones, etc.

- Tanzania Food Composition Table: http://www.fao.org/infoods/infoods/tables-and-databases/africa/en/
- Danish Food Composition Databank (Rev 5.0): http://frida.fooddata.dk

Table 5: Dataset 4 (COUNTRY_NCT): Minimum Information Required

Variable name and format	Rationale and values	Remarks and checks
Food item code (item_cod) Format: Numeric	Identification code of the food item in the survey (e.g., COICOP or national classification codes). This variable has to *include labels* corresponding to the food items.	Includes alcoholic beverages and food consumed away from home (canteens, bars, restaurants, etc.). Excludes cigars, cigarettes, tobacco, and drugs. There has to be one record for each food item collected in the survey. No missing values are allowed for this variable.
Food group (item_grp) Format: Numeric	Identification code of the food group to which the food item belongs. Examples of classifications of food items into food groups could be taken from FBS[a], FAO diet diversity questionnaire[b] or the Minimum Dietary Diversity – Women (MDDW) indicator[c].	No missing values are allowed in this variable. The purpose of this food group is to produce food security statistics derive by food groups.
Food group (diversity_grp) Format: Numeric	Identification code of the 16 food groups defined in the FAO diet diversity questionnaire[b] classification.	The food item that cannot be classified into one of the 16 food groups (e.g. lunch, dinner, meal, hamburger with fries, and pizzas) can have a missing value. The purpose of this food group is to produce a household diet diversity score and food security statistics by terciles of this score.

Refuse factor (refuse) Format: Numeric	Proportion of the nonedible portion of the food item.[d]	The refuse factor has to be expressed in *percentage*: • 0% if the food item is 100% edible (e.g., rice, milk, fillet of fish without spines, meat without bones, and peanuts without shell). • In the case of tea (in leaves) and coffee (in powder) it is suggested to assign 95%. This estimation is based on the assumption of low transference of nutrients to the brewed beverage. • Between 1% and 95% for food items having nonedible portions (e.g., meat with bones, whole fish, peanuts in shell, bananas).
Water (water) Format: Decimal	Grams of water per 100 grams *edible portion* of the food item. Values are compiled from food composition tables.	Missing data are accepted *only* for food items for which it is not possible to define their food composition, such as meals at school or restaurant, lunch, and dinner (food consumed away from home).
Ash (ash) Format: Decimal	Grams of ash per 100 grams *edible portion* of the food item. Values are compiled from food composition tables.	Missing data are accepted *only* for food items for which it is not possible to define their food composition, such as meals at school or restaurant, lunch, and dinner (food consumed away from home).
Protein (fd_pro) Format: Decimal	Grams of protein per 100 grams *edible portion* of the food item. Values are compiled from food composition tables.	Missing data are accepted *only* for food items for which it is not possible to define their food composition, such as meals at school or restaurant, lunch, and dinner (food consumed away from home).
Fats (fd_fat) Format: Decimal	Grams of fats per 100 grams *edible portion* of the food item. Values are compiled from food	Missing data are accepted *only* for food items for which it is not possible to define their food composition, such as meals at

	composition tables.	school or restaurant, lunch, and dinner (food consumed away from home).
Total fibre (fd_fib) Format: Decimal	Grams of total fibre per 100 grams *edible portion* of the food item. Values are compiled from food composition tables.	Missing data are accepted *only* for food items for which it is not possible to define their food composition, such as meals at school or restaurant, lunch, and dinner (food consumed away from home).
Alcohol (fd_alc) Format: Decimal	Grams of alcohol per 100 grams *edible portion* of the food item. Values are compiled from food composition tables.	Missing data are accepted *only* for food items for which it is not possible to define their food composition, such as meals at school or restaurant, lunch, and dinner (food consumed away from home).
Available carbohydrates *(fd_car)* Format: Decimal	Grams of available carbohydrates per 100 grams *edible portion* of the food item. When the food composition table provides information on available carbohydrates, values are compiled from the FCT/FCDB. Otherwise, values are *not compiled* from FCT/FCDB. They are estimated with the formula: *Available carbohydrates =100 – grams of water – grams of ash – grams of protein – grams of fats – grams of alcohol – grams of total fibre.*	Missing data are accepted *only* for food items for which it is not possible to define their food composition, such as meals at school or restaurant, lunch, and dinner (food consumed away from home). *Total* carbohydrates are the sum of *available* carbohydrates and total fibre. Before applying the formula: • Check that none of the nutrient values involved in the formula are missing. • After applying the formula: ○ Ensure that the values of carbohydrates equal to 100 do not come from having missing data on all the nutrient values involved in the formula. Since food items have at least one macronutrient, it is

		impossible to have all missing values. For instance, mineral water has 100 grams of water, and salt has about 99.8 grams of ash. o Check for negative values (only nonnegative values are allowed).
Dietary energy value (fd_kcal) Format: Decimal	Expressed in kilocalories per 100 grams *edible portion* of the food item. Values *are not* compiled from food composition tables. They are calculated using the Atwater system coefficients with the formula: *kilocalories = grams of protein * 4 + grams of fats * 9 + grams of available carbohydrates * 4 + grams of alcohol * 7 + grams of fibre * 2.* If the food item is classified as food consumed away from home and it is not possible to have the nutrient values, the nutrient value of the dietary energy has to be *missing (not 0).*	Missing data are accepted *only* for food items for which it is not possible to define their food composition, such as meals at school or restaurant (food consumed away from home). Nutrient values are available for some food products classified as consumed away from home such as beer, carbonated beverage, roasted maize on the cob, and roasted chicken. Therefore, for these food products consumed away from home, it is possible to obtain the conversion factor for dietary energy using the Atwater system coefficients. Only very few food items have a calorie nutrient value equal 0 (e.g., salt, water, and ice). *To detect errors:* • Check for big differences between the dietary energy values calculated with the formula and those reported in food composition tables (note that there will always be differences between the two variables).

		• Check for big differences in calories among food items belonging to the same food group.
Note: COICOP - Classification of Individual Consumption According to Purpose./ FBS – Food Balance Sheets a. The file FOOD_GROUPS.xls can be downloaded from the FAO webpage of ADePT-FSM: http://www.fao.org/economic/ess/ess-fs/fs-methods/adept-fsn/en/ >> Background information to build the survey specific nutrient conversion table (.zip). b. Refer to the publication Guidelines for measuring household and individual dietary diversity (Kennedy, Ballard and Dop, 2011). c. Refer to the publication Minimum Dietary Diversity for Women: A Guide for Measurement (FAO and FHI360, 2016) d. If no country specific data i available, refer to the file refuse factors.xls on the FAO webpage of ADePT-FSM: http://www.fao.org/economic/ess/ess-fs/fs-methods/adept-fsn/en/ >> Background information to build the survey specific nutrient conversion table (.zip).		

Table 6: Dataset 4 (COUNTRY_NCT): Micronutrient Analysis

Variable name and format	Rationale and values	Remarks and checks
Retinol (retinol) Format: Decimal	Micrograms of preformed retinol (does NOT include the contribution of any pro-vitamin A carotenoids to retinol) per 100 grams *edible portion* of the food item. Values are compiled from food composition tables.	Missing data are accepted *only* for food items for which it is not possible to define their food composition, such as meals at school or restaurant, lunch, and dinner (food consumed away from home).
Beta-carotene equivalents (betacareq) Format: Decimal	Micrograms of beta-carotene equivalents per 100 grams *edible portion* of the food item. If the FCT/FCDB publishes: • Beta-carotene equivalents, values are compiled from the FCT/FCDB;	Missing data are accepted *only* for food items for which it is not possible to define their food composition, such as meals at school or restaurant, lunch, and dinner (food consumed away from home).

	• Beta-carotene, alpha-carotene and beta-cryptoxanthin, values are computed: 1 Beta-carotene equivalent = 1 beta-carotene + 0.5 alpha-carotene + 0.5 beta-cryptoxanthin; • Total vitamin A and retinol, values are computed: 1 Beta-carotene equivalent = total vitamin A – retinol.	
Total vitamin A (vita) Format: Decimal	Micrograms of vitamin A per 100 grams *edible portion* of the food item. The micrograms are expressed in either *retinol activity equivalent (RAE) or retinol equivalent (RE).* The difference between RAE and RE is the formula used to estimate the total amount of vitamin A: *Vitamin A (RAE) = mcg of retinol + (mcg of beta- carotene/12) + (mcg of other carotenoids)/24* *Vitamin A (RE) = mcg of retinol + (mcg of beta-carotene/6) + (mcg of other carotenoids)/12* Values are compiled from food composition tables or estimated from retinol and carotenoids.	If the values are compiled from more than one FCT/FCDB especial attention should be paid to avoid mixing values of vitamin A expressed in different units. In FCTs/FCDBs Vitamin A can be expressed in RE, RAE, or international units (IU). Missing data are accepted *only* for food items for which it is not possible to define their food composition, such as meals at school or restaurant, lunch, and dinner (food consumed away from home).
Vitamin C (vit_c) Format: Decimal	Milligrams of vitamin C per 100 grams *edible portion* of the food item. Values are compiled from food	Missing data are accepted *only* for food items for which it is not possible to define their food composition, such as meals at school or restaurant, lunch, and dinner

	composition tables.	(food consumed away from home).
Vitamin B1 (Thiamine) (vit_b1) Format: Decimal	Milligrams of vitamin B1 per 100 grams *edible portion* of the food item. Values are compiled from food composition tables.	Missing data are accepted *only* for food items for which it is not possible to define their food composition, such as meals at school or restaurant, lunch, and dinner (food consumed away from home).
Vitamin B2 (Riboflavin) (vit_b2) Format: Decimal	Milligrams of vitamin B2 per 100 grams *edible portion* of the food item. Values are compiled from food composition tables.	Missing data are accepted *only* for food items for which it is not possible to define their food composition, such as meals at school or restaurant, lunch, and dinner (food consumed away from home).
Vitamin B6 (vit_b6) Format: Decimal	Milligrams of total vitamin B6 per 100 grams *edible portion* of the food item. Values are compiled from food composition tables.	Missing data are accepted *only* for food items for which it is not possible to define their food composition, such as meals at school or restaurant, lunch, and dinner (food consumed away from home).
Vitamin B12 (Cobalamin) (vit_b12) Format: Decimal	Micrograms of vitamin B12 per 100 grams *edible portion* of the food item. Values are compiled from food composition tables.	Missing data are accepted *only* for food items for which it is not possible to define their food composition, such as meals at school or restaurant, lunch, and dinner (food consumed away from home).
Folate (folate) Format: Decimal	Micrograms of folate per 100 grams *edible portion* of the food item. Values are compiled from food composition tables. The recommended folate intake is given as dietary folate equivalents (DFEs), which accounts for differences in the absorption of naturally occurring food folate and synthetic folic acid obtained from dietary supplements or fortified food.	Missing data are accepted *only* for food items for which it is not possible to define their food composition, such as meals at school or restaurant, lunch, and dinner (food consumed away from home).

	DFEs=mcg folate + 1.7*mcg synthetic folic acid	
Zinc (zinc) Format: Decimal	Milligrams of zinc per 100 grams *edible portion* of the food item. Values are compiled from food composition tables.	Missing data are accepted *only* for food items for which it is not possible to define their food composition, such as meals at school or restaurant, lunch, and dinner (food consumed away from home).
Iron (iron) Format: Decimal	Milligrams of iron per 100 grams *edible portion* of the food item. Values are compiled from food composition tables.	Missing data are accepted *only* for food items for which it is not possible to define their food composition, such as meals at school or restaurant, lunch, and dinner (food consumed away from home).
Iron of animal origin (fe_anim) Format: Decimal	Milligrams of iron from animal origin per 100 grams *edible portion* of the food item. Values of iron are compiled from food composition tables. Then the user classifies the iron as from animal origin if the food item is meat, fish, milk, eggs, or their respective products.	Missing data are accepted *only* for food items for which it is not possible to define their food composition, such as meals at school or restaurant, lunch, and dinner (food consumed away from home).
Iron of non animal origin (fe_nanim) Format: Decimal	Milligrams of iron from non-animal origin per 100 grams *edible portion* of the food item. Values of iron are compiled from food composition tables. Then the analyst classifies the iron as from non-animal origin if the food item is *different* from red or white meat, milk, eggs, or their respective products.	Missing data are accepted *only* for food items for which it is not possible to define their food composition, such as meals at school or restaurant, lunch, and dinner (food consumed away from home).
Haem iron (fe_haem)	Milligrams of iron in its chemical form haem per 100 grams *edible*	Missing data are accepted *only* for food items for which it is not possible to define

Format: Decimal	*portion* of the food item. The percentage of haem iron can be found in the literature.[a] Then the analyst applies the percentage to the total content of iron.	their food composition, such as meals at school or restaurant, lunch, and dinner (food consumed away from home).
Non-haem iron (fe_nhaem) Format: Decimal	Milligrams of iron in its chemical form non-haem per 100 grams *edible portion* of the food item. The percentage of non-haem iron can be found in the literature.[a] Then the analyst applies the percentage to the total content of iron.	Missing data are accepted *only* for food items for which it is not possible to define their food composition, such as meals at school or restaurant, lunch, and dinner (food consumed away from home).
Calcium (calcium) Format: Decimal	Milligrams of calcium per 100 grams *edible portion* of the food item. Values are compiled from food composition tables.	Missing data are accepted *only* for food items for which it is not possible to define their food composition, such as meals at school or restaurant, lunch, and dinner (food consumed away from home).
a. Published data on the percentages of haem iron contained in raw and cooked meat and fish are available in the literature (Carpenter and Clark, 1995, Rangan et al., 1997, Lombardi-Bocciaa et al., 2002, Kongkachuichai et al., 2002, Turhan et al., 2004, Turhan et al., 2006, Cross et al., 2012, Wheal et al., 2016). Moltedo and colleagues (2018) presented the values for each of the cited studies.		

Table 7: Dataset 4 (COUNTRY_NCT): Amino Acids Analysis

Variable name and format	Rationale and values	Remarks and checks
Isoleucine (isoleuc) Format: Decimal	Grams of isoleucine per 100 grams *edible portion* of the food item. Values are compiled from food composition tables.	Missing data are accepted *only* for food items for which it is not possible to define their food composition, such as meals at school or restaurant (food consumed away from home).
Leucine (leucine)	Grams of leucine per 100 grams	Missing data are accepted *only* for food

Format: Decimal	*edible portion* of the food item. Values are compiled from food composition tables.	items for which it is not possible to define their food composition, such as meals at school or restaurant (food consumed away from home).
Lysine (lysine) Format: Decimal	Grams of lysine per 100 grams *edible portion* of the food item. Values are compiled from food composition tables.	Missing data are accepted *only* for food items for which it is not possible to define their food composition, such as meals at school or restaurant (food consumed away from home).
Methionine (methion) Format: Decimal	Grams of methionine per 100 grams *edible portion* of the food item. Values are compiled from food composition tables.	Missing data are accepted *only* for food items for which it is not possible to define their food composition, such as meals at school or restaurant (food consumed away from home).
Phenylalanine (phenyl) Format: Decimal	Grams of phenylalanine per 100 grams *edible portion* of the food item. Values are compiled from food composition tables.	Missing data are accepted *only* for food items for which it is not possible to define their food composition, such as meals at school or restaurant (food consumed away from home).
Threonine (threon) Format: Decimal	Grams of threonine per 100 grams *edible portion* of the food item. Values are compiled from food composition tables.	Missing data are accepted *only* for food items for which it is not possible to define their food composition, such as meals at school or restaurant (food consumed away from home).
Tryptophan (trypto) Format: Decimal	Grams of tryptophan per 100 grams *edible portion* of the food item. Values are compiled from food composition tables.	Missing data are accepted *only* for food items for which it is not possible to define their food composition, such as meals at school or restaurant (food consumed away from home).
Valine (valine) Format: Decimal	Grams of valine per 100 grams *edible portion* of the food item. Values are compiled from food composition tables.	Missing data are accepted *only* for food items for which it is not possible to define their food composition, such as meals at school or restaurant (food consumed away

		from home).
Histidine (histid) Format: Decimal	Grams of histidine per 100 grams *edible portion* of the food item. Values are compiled from food composition tables.	Missing data are accepted *only* for food items for which it is not possible to define their food composition, such as meals at school or restaurant (food consumed away from home).
Cystine (cistyne) Format: Decimal	Grams of cystine per 100 grams *edible portion* of the food item. Values are compiled from food composition tables.	Missing data are accepted *only* for food items for which it is not possible to define their food composition, such as meals at school or restaurant (food consumed away from home).
Tyrosine (tyrosine) Format: Decimal	Grams of tyrosine per 100 grams *edible portion* of the food item. Values are compiled from food composition tables.	Missing data are accepted *only* for food items for which it is not possible to define their food composition, such as meals at school or restaurant (food consumed away from home).

How to Build Dataset 4

Below are some guidelines to build the COUNTRY_NCT input dataset. Steps 7 and 8 can be skipped if micronutrients and amino acids analyses, respectively, are not conducted.

Step 1: Open the template file COUNTRY_NCT_template.xlsx and save it on your computer. This template can be downloaded from the FAO webpage of ADePT-FSM: http://www.fao.org/economic/ess/ess-fs/fs-methods/adept-fsn/en/ >> Background information to build the survey specific nutrient conversion table (.zip).

The template file is composed of different worksheets; one of these is named Archival. Go to Archival and list here all the food items collected in the HCES, inserting their survey code and description in columns A, *Food item code in household survey (item_cod)*, and B, *Food item description in house- hold survey (desc)*. All the food items collected during the survey should be included in the list, including the food items consumed away from home.

Step 2: 2a. Selecting the Food Composition Table Identify the most suitable national or regional food composition table or database (reference file) for matching the food items in the survey with those described in the selected FCT/FCDB. Some criteria that should be taken into consideration in the selection of a FCT/FCDB are the year of publication, the completeness of information (especially for macronutrients), geographic/cultural proximity between the country under study, and those countries/regions for which the food composition table is written.

Some FCTs/FCDBs are available on the web at the following addresses:

- U.S. Department of Agriculture FCT: http://www.nal.usda.gov/fnic/foodcomp/search/index.html
- European FCTs: http://www.eurofir.net/eurofir_knowledge/european_databases
- Latin Foods: http://www.inta.cl/latinfoods/
- INFOODS databases: http://www.fao.org/infoods/infoods/tables-and-databases/en/
- LANGUAL: http://www.langual.org/langual_linkcategory.asp?CategoryID=4&Category=Food+Composition

2b. Food Matching Once the FCT/FCDB is identified, insert its name in column C, *Reference food composition table (FCT).*

After matching a food item listed in the Archival worksheet with a food item in the FCT/FCDB, insert the reference food item's code and description in columns D, *Food code in FCT*, and E, *Food description in FCT*, respectively.[4]

It may happen that a food product listed in the HCES cannot be matched directly with any of the foods in the reference table. Reasons could be: (1) the food item does not exist in the FCT/FCDB or (2) the food product listed in the HCES includes more than one food item of the FCT/FCDB or is broadly described. In the first case, the food matching (step 2b) for that specific item is done using another FCT/FCDB (selected using the criteria mentioned in step 2a) to find out the appropriate food product of reference. In the second case, a weighted average of the nutritional values of all the relevant (i.e., similar, corresponding) food products should be performed. By default, all food items involved have equal weight factor. The best approach for defining the weighting factors is referring to external data collected in national individual dietary

4 For the food matching, consult FAO/INFOODS Guidelines for Food Matching (2012) available at http://www.fao.org/infoods/infoods /standards-guidelines/en/.

surveys (Naska et al, 2007), in case this information is not available the pattern consumption of previous HCES could be used[5]. The use of weights is a better approach than applying equal weight to each food item or considering just one of the listed food items (Naska et al, 2007).

- *Examples*:

a) The food item in the HCES is *broadly described*, for example *rice*. In this case, the color (brown or white) of the rice is not specified. Therefore, a weighted average of the nutritional values of different types of rice is needed. If the food item description in the survey is *rice* and in the list of food items in the survey there is no mention of rice flour, then not only rice grain food items but also rice flour has to be included in the calculation of the average nutritional values.

b) *Different types* of the same food product or different food products are listed together in the HCES as if they were one food item (for example, *white rice, grain or flour, wheat or corn flour*, and *eggplant, cauliflower, broccoli*). If the proportions of consumption are not known, a simple average of the nutritional values is done.

c) *Fresh and dry* food items are listed together (for example, *fresh or powdered milk, whole milk* and *fresh or dried salmon*). If the proportions of consumption are not known, a weighted average of the nutritional values is done assigning a maximum weight factor of 10 percent to the dry product.[6]

For instance, in Table 8 and Table 9, the protein value of the food item collected in the survey is obtained averaging the protein values of similar food items from the FCT/FCDB (the total number of food items from the FCT/FCDB is five). In Table 8, equal weights7 are applied so the weight factor for each food item is 0.2. Table 9 shows an example when the applied weights are different (e.g., they could be obtained from previous analysis of food consumption from household survey data).

Once the matching between the food items in the HCES list and those in the FCT/FCDB is done, insert the food item index matching in column F, *Food Item Index Matching*, of the

[5] For example, from the analysis of previous national consumption surveys in the country, the milk consumption pattern is whole, 90 percent; partially skimmed, 7 percent; skimmed, 3 percent.

[6] The figure 10 percent, though arbitrary, is used to avoid overestimation of nutrient content, as nutrients are more concentrated in dry foods, leading to higher nutrient values per 100 grams edible portion (FAO forthcoming).

[7] A weighted average performed applying equal weight factors is equal to a simple average.

Archival worksheet. The values indicate the type of matching between the food item listed in the survey and the food item selected from the FCT/FCDB. These are the codes:

A = Single, perfect match, no modifications required (apart from edible portion, if indicated)

A2 = Exact match with multiple selections requiring average computation

B = Similar, single match

B2 = Similar match with multiple selections requiring average computation

C = Poor, single match

C2 = Poor match with multiple selections requiring average computation

D = Calories estimated by ADePT using unit calorie cost (applies only to food consumed away from home for which it is not possible to know its composition, such as lunch, dinner or meal, other foods, etc.)

Table 8: Estimating the content of a nutrient applying equal weights

Name of the FCT/FCDB	Item code in the FCT/FCDB	Item description in the FCT/FCDB	Item weight factor	Grams of protein from the FCT/FCDB
USDA	20036	Rice, brown, long-grain, raw	0.2	7.94
USDA	20040	Rice, brown, medium-grain, raw	0.2	7.5
USDA	20444	Rice, white, long-grain, regular, raw, unenriched	0.2	7.13
USDA	20450	Rice, white, medium-grain, raw, unenriched	0.2	6.61
USDA	20052	Rice, white, short-grain, raw	0.2	6.5

Item code in the survey	Item description in the survey	Grams of protein
4002	Rice grain	0.2*7.94+0.2*7.5+0.2*7.13+0.2*6.61+0.2*6.5 = 7.136

Table 9: Estimating the content of a nutrient applying different weights

Name of the FCT/FCDB	Item code in the FCT/FCDB	Item description in the FCT/FCDB	Item weight factor	Grams of protein from the FCT/FCDB
Bolivia FCT	A77	Wheat flour	0.759	8.03
Bolivia FCT	A80	Corn flour	0.241	8.5
Item code in the survey		Item description in the survey	Grams of protein	
4005		Wheat or Corn flour	0.759*8.03+0.241*8.5 = 8.143	

Step 3: In the worksheet Archival, in column G, *Refuse factor (refuse)* insert the food item's refuse factor.[8] In column H, *Item group (item_grp),* insert the food item group to which the food item belongs.

Step 4: In the worksheet Archival, fill all the columns highlighted in gray with the information available in the FCT/FCDB corresponding to each food item, including total carbohydrates for further data-checking purposes. If a nutrient of a food item is missing in the selected FCT/FCDB, look for the respective value in another FCT/FCDB. Insert a comment in the Excel cell of the missing nutrient mentioning the name of the FCT/FCDB from which the value was obtained as well as the food item code and description in the FCT/FCDB.

In the specific case of missing ash content, the value found in another FCT/FCDB has to be adjusted by the total content of solids using the formula:

$$\text{Ash(g)} = \frac{[\text{Ash(g) in other FCT/FCDB} * (100 - \text{Water(g) in the FCT/FCDB})]}{100 - \text{Water(g) in other FCT/FCDB}}$$

[8] In the survey, households report food quantities as purchased/acquired. But many foods have edible and nonedible parts. FCT report nutrients on edible quantities. Therefore, a refuse factor is needed to calculate the edible quantities contained in the quantities reported as purchased/ acquired. Only if we do so, can we apply the nutrients from the households to the food item list.

As for the nutrient values of the food items consumed away from home for which it is not possible to know their composition (meal, lunch, etc.), blank cells are allowed, because their respective nutrient values will be estimated by ADePT-FSM.

The cells of the following columns *should not be filled* in the archival sheet:

- P: *Available carbohydrates by difference (fd_car)*
- R: *Computed calories (kcal) (fd_kcal)*
- U: *Animal iron (milligrams) (fe_anim)*
- V: *Nonanimal iron (milligrams) (fe_nanim)*
- W: *Haem iron (milligrams) (fe_haem)*
- X: *Non-haem iron (milligrams) (fe_haem)*

Once all the required information is inserted in the Archival worksheet, copy it to the Reference worksheet.

Step 5 In the Reference worksheet, compute the grams of *available carbohydrates by difference* in column P, *Available carbohydrates by difference (grams) (fd_car)*, as:

$$\text{fd\_car (column P)} = 100 - \text{Water (column I)} - \text{Ash (column J)} - \text{Protein (column K)}$$

$$- \text{Fat (column L)} - \text{Fibre (column M)} - \text{Alcohol (column N)}$$

Suggested checks:

- The sum of the values in columns M, *fd_fib*, and P, *fd_car*, should be similar to the value in column O, *Carbohydrates including fibre (Total) (grams)*.
- The values in column *fd_car* should be positive or equal to 0. If one value is negative and there was no data entry error in any of the nutrients involved in the computation, assign a value of 0
- Ensure that the values of *fd_car = 100* do not come from having missing data on all the nutrient values involved in the formula. Since food items have at least one macronutrient, it is impossible to have all missing values.[9]

Step 6 In the Reference worksheet, compute the dietary energy value in column R, *Computed calories (kcal) (fd_kcal)*, as:

$$\text{fd\_kcal(column R)} = \text{Protein (column K)} * 4 + \text{Fat (column L)} * 9 + \text{Fiber}$$

[9] For example, mineral water has 100 grams of water; salt has about 99.8 grams of ash, etc.

$$(\text{column M}) * 2 + \text{Alcohol (column N)} * 7 + \text{Available carbohydrates by difference (column P)} * 4$$

Verify that the computed dietary energy values in column R, *Computed calories (kcal) (fd_kcal)*, are similar to those compiled from the FCT/FCDB in column Q, *Calories (kcal)*. There will always be differences between the values of these two columns, but if there are *big* differences, verify that the nutrient values used in the computation are correct.[10] Two of the most common errors are wrong data entry of the food item nutrient content and wrong estimation of available carbohydrates.

Step 7 If the food item in the HCES is of animal origin (as previously defined in this document), in the Reference worksheet, copy the values of column T, *Iron (milligrams) (iron)*, to column U, *Animal iron (milligrams) (fe_anim)*. Similarly, if the food item is not of animal origin, copy the values of column T to column V, *Nonanimal iron (milligrams) (fe_nanim)*.

Step 8 Compile the values of haem and non-haem iron from bibliography and copy them to column W, *Haem iron (milligrams) (fe_haem)* and column X, *Non-haem iron (milligrams) (fe_nhaem)* respectively within the Reference worksheet.

Step 9 When all the above steps were completed copy all the information of the Reference worksheet and paste it to the Upload worksheet. To paste the information, select the function **paste special > values** from the menu.

Only the columns whose variable name is red in the Upload worksheet are needed in dataset 4 and should be uploaded.

An example of a completed COUNTRY_NCT template for a country is available at the FAO webpage of ADePT-FSM: http://www.fao.org/economic/ess/ess-fs/fs-methods/adept-fsn/en/ >> Background information to build the survey specific nutrient conversion table (.zip).

Exogenous parameters to estimate Dietary Energy Requirement

The minimum and average dietary energy requirements (MDER and ADER, respectively) are produced by ADePT-FSM. To estimate the energy requirements, the values for the under-five mortality rate and birthrate are needed. Therefore, ADePT-FSM requires the user

[10] A hypothetical example is that the value of calories for rice white raw published in the FCT is 346 kcal, while the value of calories estimated with the formula is 260 kcal.

to insert these two country-specific parameters. Both parameters are computed at the country level and should refer to the year in which the survey was conducted.

Under-Five Mortality Rate

UNICEF defines the under-five mortality rate as the probability of dying between birth and exactly five years of age expressed per 1,000 live births. Estimates of the under-five mortality rate are available at http://data.unicef.org.

Birthrate

The crude birthrate is the number of births over a given period of time divided by the person–years lived by the population over that period (UN 2011). Estimates of crude birthrate, expressed as the number of births per 1,000 people, are available at http://esa.un.org/unpd/wpp/. In the ADePT module, the value of the birthrate parameter should be expressed *per person*.

References

Carpenter, C. E. & Clark, E. 1995. Evaluation of methods used in meat iron analysis and iron content of raw and cooked meats. Journal of Agricultural and Food Food Chemistry, 43, 1824-1827.

Cross, A. J., Harnly, J. M., Ferrucci, L. M., Risch, A., Mayne, S. T. & Sinha, R. 2012. Developing a heme iron database for meats according to meat type, cooking method and doneness level. Food Nutr Sci, 3, 905-913.

FAO and FHI (Family Health International) 360. 2016. Minimum Dietary Diversity for Women: A Guide for Measurement. Rome: FAO.

FAO, and INFOODS (International Network of Food Data Systems). 2012. Guidelines for Food Matching Version 1.2. Rome: FAO. http://www.fao.org/infoods/infoods/standards-guidelines/en/.

James, W. P. T., and E. C. Schofield. 1990. Human Energy Requirements: A Manual for Planners and Nutritionists. Oxford, UK, Oxford Medical Publications under arrangement with FAO.

Kennedy, G., Ballard, T. & Dop, M. 2011. Guidelines for Measuring Household and Individual Dietary Diversity. Nutrition and Consumer Protection Division, FAO. URL: http://www.fao.org/fileadmin/user_upload/wa_workshop/docs/FAO-guidelines-dietary-diversity2011.pdf

Kongkachuichai, R., Napatthalung, P. & Charoensiri, R. 2002. Heme and Nonheme Iron Content of Animal Products Commonly Consumed in Thailand. Journal of Food Composition and Analysis, 15, 389-398.

Lombardi-Bocciaa, G., Martinez-Dominguez, B. & Aguzzi, A. 2002. Total Heme and Non-heme Iron in Raw and Cooked Meats. Journal of Food Science, 67.

Moltedo A., Álvarez Sánchez C., Troubat N. and Cafiero C. Optimizing the use of ADePT-Food Security Module for Nutrient Analysis. ADePT-FSM Version 3. Statistics Division (ESS). Food and Agriculture Organization of the United Nations. Rome, January 2018.

Naska A., Oikonomou E., Trichopoulou A., Wagnera K. and Gedrich K. 2007. Estimations of daily energy and nutrient availability based on nationally representative household budget

survey data. The Data Food Networking (DAFNE) project. Public Health Nutrition: 10(12), 1422–1429.

Rangan, A., Ho, R., Blight, G. & Binns, C. 1997. Haem iron content of Australian meats and fish. Food Australia, 49, 508-511.

Turhan, S., Ustun, S. N. & Altunkaynak, T. B. 2004. Effect of cooking methods on total and heme iron contents of anchovy (Engraulis encrasicholus). Food Chemistry, 88, 169-172.

Turhan, S., Ustun, S. N. & Bank, I. 2006. Effect of freeze–thaw cycles on total and heme iron contents of bonito (Sarda sarda) and bluefish (Pomatomus saltator) fillets. Food Composition and Analysis, 19, 384-387.

Wheal, M. S., DeCourcy-Ireland, E., Bogard, J. R., Thilsted, S. H. & Stangoulis, J. C. 2016. Measurement of haem and total iron in fish, shrimp and prawn using ICP- MS: Implications for dietary iron intake calculations. Food Chem, 201, 222-9.