Guiding Document to Navigate Endline Data

The endline data (IAPP_endline_data_for_publicUse.dta), hereafter called Data, is a cleaned and de-identified version of endline survey data that was accumulated as part of Integrated Agricultural Productivity Project (IAPP) in Bangladesh in 2015. The Data is ready made for public use.

There are 1,920 surveyed households across two districts in north and south of Bangladesh. Data has been de-identified for privacy of survey subjects. Unique identifier is household ID number (hhid). Seasonal focus on the survey is Boro, Aus, and Aman seasons for crop and fisheries modules. There are nine modules covered in the survey (please refer to survey form), i.e., crops; agricultural inputs; livestock and poultry; fisheries; farmer groups; assets, income, expenditures and consumption; finances; negative events; and food security. Variables in the Data are ordered based on the ordering of survey – crop variables appear first and so on.

Each of the nine modules are explained below in terms of their variable naming structure. The general format of variable naming is as follow:

Prefix_variablename_subloop(s)_outerloop

Prefix refers to survey module section

Variablename is the core name of variable

Subloop refers to inner loops, for instance, crops loops and season loops

Outerloop is the most outer loop encapsulating the subloop(s)

Module Crop [prefix - c1]:

The variables in crop module generally follow the structure: *prefix_variablename_loop1_loop2_loop3*. Loop 3 being the most outer loop is a *plot* loop, numbered 1 to 10. Loop 2 is the second most outer loop and it is the *season* loop, numbered 1 to 3. Loop 1 is the third most outer loop and it is *crop* loop, numbered 1 to 6. When seed varieties of crops are dealt with, another loop, which becomes the fourth most outer loop, also exists. The varieties of seeds are numbered 1 to 3.

An example of a variable from crop module is: *c1_var_slct_wht_vr1_c4_s3_p10*. C1 is the prefix for module; *var_slct_wht* is the variablename; *vr1* is the seed variety 1, *c4* is crop number 4, *s3* is season 3, and *p10* is plot number 10.

Module Crop Inputs [prefix – c2]:

The variables in crop input module have the structure: *prefix_variablename_loop1_loop2*. Loop 2 is the crop input loop. There are a maximum of up to 12 inputs that some farmers have used, so the loop runs for 12 rounds. Loop 1 is the season loop. There are three loops for seasons.

Some of the core variables in crop module had been reshaped from a 'wide' format to a 'long' which means that a single variable contains all information instead of being spread over 12 variables. For instance, the amount of Urea used in season 1 was previously scattered over 12 variables, since some farmers would mention Urea usage as their 5th or 9th input. Now, you can have all such information in one variable, i.e., c2_amt_n_s1_Urea.

List of reshaped *variablename*: amt_n, amt_u, inp_free_n, inp_free_u, any_free, inp_spend, prop_pad, prop_whe, prop_mas, prop_mun, prop_ses, prop_oth, prop_sum

Module Livestock and Poultry [prefix – d1]:

The general modified variable structure in this module appears as: *prefix_animalname_variablename* (some variables in the beginning of this module don't have a prefix). The animalname portion has been added to the variable name structure during a reshaping step to have the variables separated by each animal type. For instance, *d1_goat_a_q2* variable only tells you about number of goats owned by a family. Previously this information was spread over multiple variables.

Module Fisheries [prefix – d2]:

The main prefix of this module is d2, however the variables will also start with a prefix `pd'. There are two main loops for the variables in this module, i.e., pd#_f#, where *pd* is the pond loop, and *f* is the fish type loop. Number of loops for ponds and fish types are 3 and 6, respectively. Pd# and f# loops will show up either at the very end of a variable, or right after the prefix.

Module Farmers Group [prefix – e]:

The main prefix of this module is letter *e*. There are no loops in this module.

Module Assets, Income, Consumption [prefix – asset, income, consume]:

This module is simplified in terms of its prefixes as every set of variables is prefixed with *asset, income*, or *consume*.

Module Finances [prefix – k]:

There are no loops in this module, and the prefix is letter *k*.

Module Negative Events [prefix – drought, flood, cold, heat, hail, price, health]:

Variables in this module are prefixed with the above prefixes. There are no loops.

Module Food Security [prefix – didConsume, noFood, hunger]:

Variables in this module are prefixed with the above prefixes. There are no loops.