

GHANA

Strategy Support Program



INTERNATIONAL
FOOD POLICY
RESEARCH
INSTITUTE
Supported by the CGIAR

Short note to users of GAPS 2011/12 minor season data

This Note has been prepared as an output for the Ghana Strategy Support Program and has not been peer reviewed. Any opinions stated herein are those of the author(s) and are not necessarily representative of or endorsed by the International Food Policy Research Institute.

July 2014

CONTENTS

1. Introduction	1
Sampling weights	1
Merging of minor season data files	1
Missing observations	2

TABLES

1—Section specific notes	2
--------------------------	---

I. INTRODUCTION

The data released by the SRID during the pilot of the 2012/13 minor season was analyzed to assess the usefulness of the survey and its relevance to district and national agricultural policy and development planning. The datasets were reviewed and primary analyses were carried out to generate tables noting the limitations on the quality and reliability of the data the pilot survey. Some errors related to data entry lapses were cleaned up through a recall of original completed questionnaires but others remain useful lessons for future surveys. Outstanding issues on the use of the pilot minor season data relate to generalization of data beyond sample characteristics and ability of average data users to combine information from different sections of the questionnaires. Users may call on the survey implementer (SRID) for any further clarification on the use of the data.

Sampling weights

The survey has sampling weights to account for the differential probability of inclusion of districts into the sample for the season. That is the weights vary only at the district level but not at the household or holder levels. This means that there is a limited extent to which results at the household and holder levels can be generalized beyond the sample characteristics because of the unequal probability of inclusion for household and farm holders.

Merging of minor season data files

The other outstanding problem for an effective use of the minor season data is the difficulty for the average user to merge different files from the questionnaires. It is noted that some data files do not have all the unique variables necessary to merge information from different sections of the questionnaires. Some of the sections have all the unique variables but do not have unique values for different sections of the questionnaires. Table 1 shows the specific notes on merging for each data file for the questionnaires. Those with duplicates (significant number of multiple observations for its unique identifiers) are likely to pose problems for users with the interest to analyze information across different sections of the questionnaires and data files.

Table I—Section specific notes

File Name	Unique identifiers	Notes
Health	dist_cod_1; ea_code_1; hh_no; id_013	No duplicates.
Minor shock	dist_code_1; ea_code; hh_no; ho_no; sh_no	No duplicates.
Minor nonjob	dist_code_1; ea_code_1; hh_no; nonag_job_id	No duplicates.
Minor transport	dist_cod; ea_code_1; hh_no	The variable holder was not found in the exported data and thus the three identifiers do not make the file unique. At best household means may be used.
Minor tree	dist_code_1; ea_code; hh_no; hn_017	The variable holder was not found in the exported data and thus the three identifiers do not make the file unique .A few observations have tree codes replaced by NR, meaning not available/no response.
Minorcredit	dist_cod; ea_code_1; hh_no	The variable holder was not found in the exported data and thus the three identifiers do not make the file unique.
Minorcrop	dist_code_1; ea_code_1; hh_no; hn_009; f2b_q25_1	There are a lot of duplicates at the level of the four identifiers. Some observations have missing crop codes.
Minorfeed	dist_code_1; ea_code; hh_no; ho_no; feed_no	No duplicates.
Minorfields	dist_cod_1; ea_code_1; hh_no_1; hn_008; farm_no_1; field_no_1	There are a lot of duplicates at the level of the six identifiers. The variables holder, farm and field start from zero (0) instead of one (1). It is difficult to do any analysis at the field, farm and holder levels because of the duplicates. Just below one third of the fields have measurements and no variables shows why some were not measured.
Minorinput	dist_cod; ea_code; hh_no ho_no; input_no	No duplicates
MinorLivestock	dist_code_1; ea_code; hh_no; hn_016; f2a_q20_1	There are forty (40) duplicates at the level of the identifiers, with surplus of 20 observations.
remaining major	dist_code; ea_code; hh_no; ho_no; f2b_q13_1	There are so many duplicates at the level of the supposed unique variables
MinorMember	dist_code_1; ea_code_1; hh_no_1; id_code	No duplicates

Missing observations

Some questions in the questionnaire do not have data points. Specifically, questions 54 to 71 in Form 2a, do not have data for all respondents. This means that no analysis can be done on aquaculture inputs, practices, outputs and assets using the current data sets. Similarly information on other income generating activities, questions numbered 61 to 67 in Form 2b cannot be analyzed because of missing observations.