Response to ILRI Pre-Interview Task

Sam Larkin

4th September 2019

Abstract

This document briefly describes the procedure followed in the course of responding to the pre-interview task, as well as a graphical summary of the results.

Python 3 was used to process the provided survey data and codebook. The relevant data was extracted from each of the provided csv files and used to determine the (HFIAS) category of each household.

These results were then used to find the prevalence of each category – overall and by region – in order to investigate the relative food (in)security experienced across the four regions covered in the survey data. The results of this comparison can be seen in figure 1 (produced – along with the rest of this report – using LaTeX.

1 Method

The manipulation of the provided survey data and the calculation of the indicator was all carried out using Python 3.

All of the code I used can be found in these two files (here):

- *ilri-pre-interview-task.py*, which calculates the categories and prevalences for the overall survey;
- *ilri-pre-interview-task-v2.py*, which extends the functionality of the original script to include the comparison of the regions.

The procedure followed was composed of the following steps:

1. Import provided csv files;
2. Extract relevant data from the original files;
3. Apply logical conditions described by Coates et. al. to determine the 'categorical HFIAS indicator' for each household;
4. Write the results to my own csv files in 'outputs';
5. Use the results from step 3 to calculate the prevalence of each category (for the overall survey);
6. (script v2 only) Use the results from step 3 and the region data from the survey responses to compare the prevalence of each category within each of the individual regions (for interest – in order to see whether the results were substantially different from region to region);
7. The report and graphical representation of the results were generated using LaTeX.

2 Results

The results produced by the two python scripts can be found here; they have been written to four csv files as follows:

- *"HFIAS_codebook.csv"* is simply a redacted copy of the provided codebook, including only the nine questions relating to the HFIAS indicators;

- *"hh_HFIA_cat.csv"* includes the household number which links the results to the original survey data and the HFIA category for each household (the second script – v2 – adds the region in which each household is located to this file);
- *"HFIA_cat_prevalence.csv"* provides the prevalence (as a percentage) of each category in the overall survey data;
- *"HFIA_cat_prevalence_by_region.csv"* provides the prevalence (as a percentage) of each category within each of the four regions surveyed.

Figure 1 shows the prevalence results for the overall survey data and for the four regions surveyed, such that the proportions for each category can be compared between the regions.

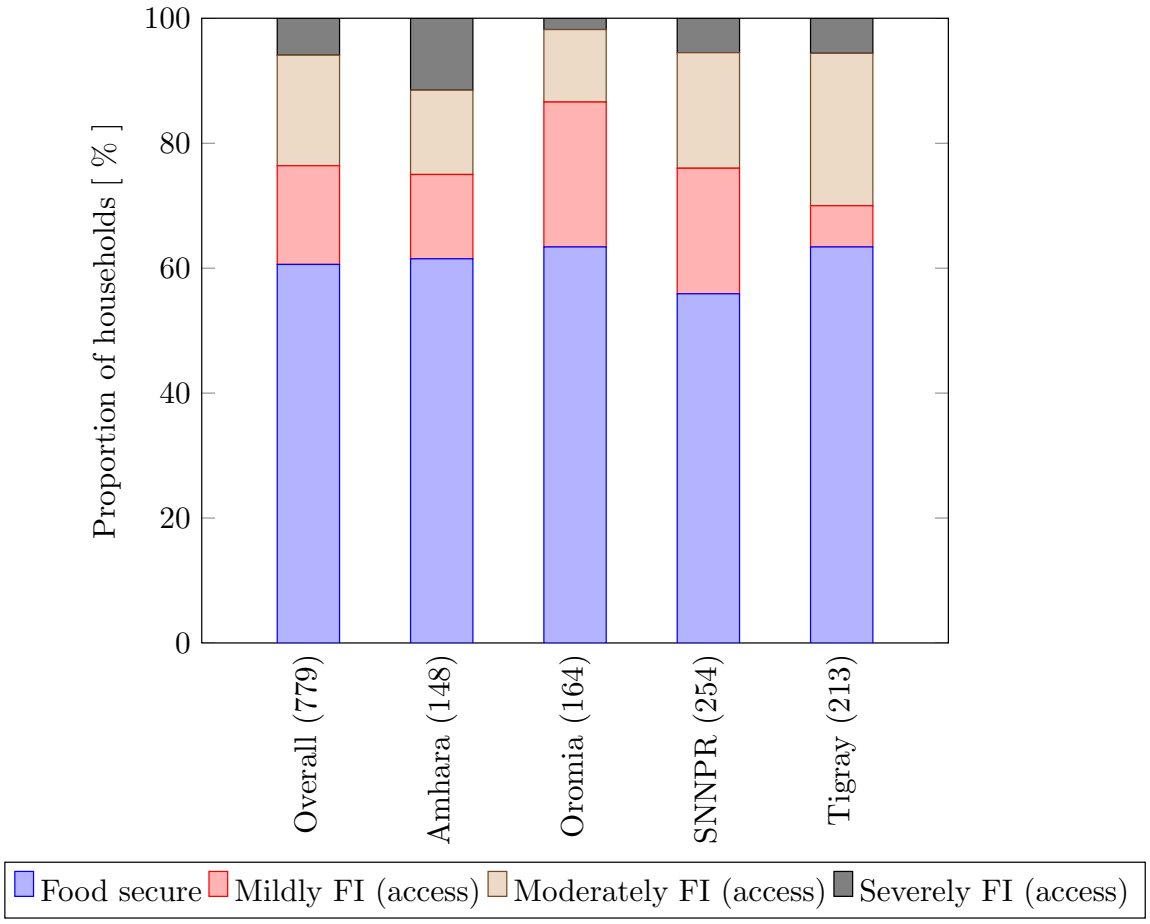

Figure 1: Prevalence of (HFIAS) categories – overall and by region