**September 2022**

**DRYLANDS TRANSFORM - Household survey**

**INSTRUCTIONS FOR DATA CHECKING**

1. **Persons in charge for sections data checking**

|  |  |  |
| --- | --- | --- |
| **SECTION / SUB-SECTION** | **WHO IN CHARGE** | **comment** |
| **SECTION A + SECTION L** - Meta data pre/post-interview | *Cleaning group* |  |
| **SECTION B** - Socio-Demographics | Alice |  |
| **SECTION C** - Livelihoods |  |  |
| 3.01 Land access and ownership | Obj.4 (Göran, Per, James) |  |
| 3.02 Livestock, poultry and other animals | Denis |  |
| 3.03 Farm level questions | Ylva |  |
| **SECTION D** - Migration | Göran |  |
| 4.01 Regular migration with livestock |  |  |
| 4.02 Human migration during the past 6 months |  |  |
| **SECTION E -** Assets / wealth | Alice |  |
| **SECTION F** - Income & Expenditures |  |  |
| 6.01 Expenditures | Alice |  |
| 6.02 Income | Alice |  |
| 6.03 Future perspectives (male) | Gert |  |
| **SECTION G** - Crises (male) |  |  |
| 7.01 Livelihood shocks | Barbara |  |
| 7.02 Conflicts in community and family | Obj.4 |  |
| 7.03 Climate change | Barbara |  |
| **SECTION H -** Support, work and future |  |  |
| 8.01 Social networks and services | Ylva |  |
| 8.02 Woman’s income generating activities | Kristina |  |
| 8.03 Future perspectives (female) | Gert |  |
| **SECTION I** - Crises (female) |  |  |
| 9.01 Livelihood shocks (Loss and diseases) | Barbara |  |
| 9.02 Conflicts within community and family | Obj.4 |  |
| 9.03 Climate change | Barbara |  |
| **SECTION J** - Nutrition | Agneta |  |
| 10.01 Food insecurity A: Household Food Insecurity… |  |  |
| 10.02 Food insecurity B: Months of Adequate… |  |  |
| 10.03 Questions for index child (6-59 months) |  |  |
| 10.04 24h- recall index child + mother/female caretaker |  |  |
| **SECTION K** - Health + Anthropometry |  |  |
| 11.01 Health of household members | Kristina |  |
| 11.02 Child Health + vaccinations (index child) | Kristina |  |
| 11.03 Anthropometry | Agneta |  |

1. **MATERIAL FOR DATA CHECKING**

Generally:

* All files and instructions are available on **Github**
* You must create a Github account.   
  Send your username to Annrose ([arosemwangi@gmail.com](mailto:arosemwangi@gmail.com)), she connects you to the project.
* You need Github only for accessing + downloading the data – no expertise whatsoever reg. Github is needed!
* You will not change any data on Github or elsewhere.

Support files:

* Meta file (word doc) with all questions, variable names and some explanations
* Instructions (word doc), this file

Data files:

* Excel data file “XXX”, one sheet per section. You probably work mostly with this one.
* Excel file “XXX” with summary statistics for continuous variables and frequencies for categorical variables
* SPSS file “XXX”with all data

Reporting file:

* Excel reporting file

1. **INSTRUCTIONS**

**Generally**

* Make sure who in your group is reporting back to us!  
  One team member per section/subsection (see above) is leading but can delegate in the team.   
  If in doubt which data should be corrected: discuss in your group.   
  If necessary, consult with cleaning group.   
  The section leader should finally report to the cleaning group (Annrose et al.)

**Coding explanation**

* *999 missing, 888 n.a.* (not applicable)  
  0=no, 1=yes;  
  See value keys for specific variables/questions (Excel file “xxx”)
* HH code:   
  Kenya 1, Uganda 2;   
  West Pokot 1, Turkana 2; Napak 1, Moroto 2  
  I.e. south/agro-pastoralist = 1; North/pastoralist = 2  
    
  HHs per study site from 001 to …

Format HH code: country\_site\_HH number   
E.g. 1\_1\_001 = first HH in West Pokot, Kenya; 2\_1\_077 = 77th HH in Napak, Uganda

* HH member code:   
  01 male respondent, 02 female respondent, 03 index child, 04-16 for other HH members. 01, 02, 03 not used if the respective person is missing (e.g. no HH member code 03 if there is no index child).

**Checking the data of your section**

* Download the datafiles to your computer, work from there.
* Look closely through the data of your section in Excel.
* Plausibility of values? Too large, too small?   
  Apply filter function in Excel data.  
  Check summary tables by Annrose (Excel file “xxx”)   
  Use SPSS file for doing descriptive statistics.
* Plausibility across variables?   
  E.g. Cattle number missing (empty cell) although HH has cattle? Then should be 999. Cattle number not missing although HH has not cattle? Then should be 888 na.
* Plausibility across sections?  
  E.g. No male respondent present (Section B)   
  –> 6.03 Future perspectives male; 7.02, 7.03 (crisis sub-sections) should be 888 (n.a.)
* Check also comments variable(s) for your section – can help checking other variables  
  (e.g. C\_gen\_comments\_land\_own: “basically communal land”).   
  Then decide whether a variable value has to be changed.
* Check blank cells – Should these be made n.a. (888) (e.g. if whole section does not apply) or missing (999)?

**Reporting in Excel file “XXX”**

* The report file has two sheets per section:   
  One for general issues, one for specific households/observations.  
  Both sheets have headings and examples.
* Be very clear in feedback, do not suggest / ask within this file!   
  Annrose needs to know exactly how to replace a value.
* Report one issue per column.   
  E.g. column “issue 1”: “value 6 occurs but is not possible”. Then in column “correction1”: “replace all values 6 by 5” or “replace all values 6 by missing 888”
* Sheet “HH-specific changes”: Report here those changes that apply only for one specific observation/HH. Name the HH code and necessary changes. One line per issue, you can have several lines for one HH.