



Food and Agriculture
Organization of the
United Nations



How to develop Socio-Economic surveys for baseline assessments

Questionnaire & Implementation

Action Against Desertification programme

Workshop on socio-economic baseline assessment and M&E planning
Ouagadougou, Burkina Faso, 21st – 25th March 2016



INTRODUCTION

- Collecting baseline data is the crucial step to develop a prospective evaluation of the project
 - Baseline data are collected prior to the implementation of the project activities
 - Baseline data can be compared to the evaluation of the project outcomes during the project
- The most common and effective way to collect socio-economic baseline data is household survey
 - Identifying the relevant information to be collected – suited for calculating the desired indicators – and drafting the questionnaire
 - Hiring enumerators to administer the household survey and implement the interviews
 - Filling in a data matrix that, after check and clean, can be elaborated with multivariate and inferential statistics



OUTLINE OF THE PRESENTATION

- Planning the socio-economic baseline assessment
- Sampling strategies
- Preparing the questionnaire
- Implementing the household survey
- Data elaboration

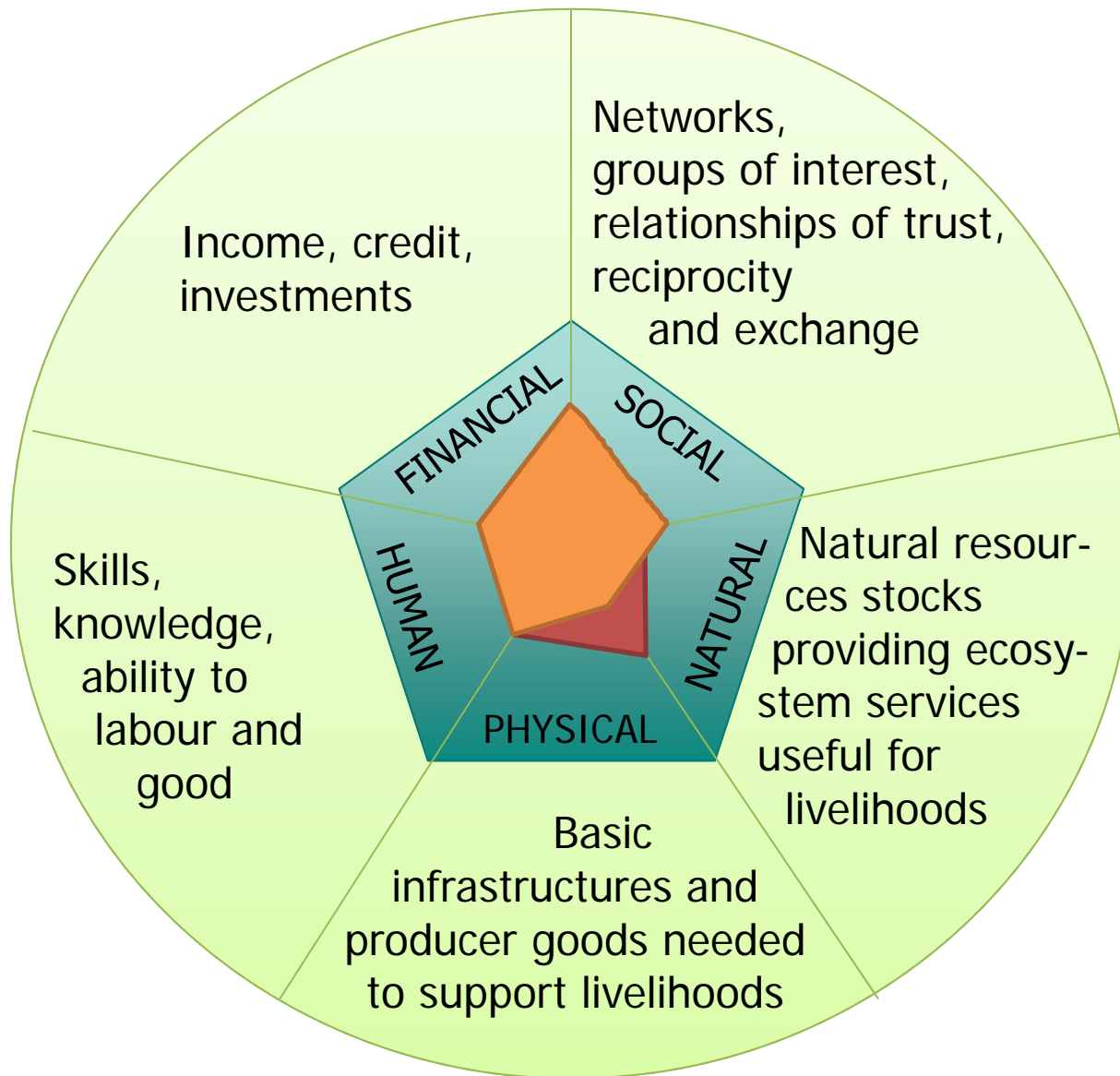


PREPARING THE QUESTIONNAIRE [1]

- The questionnaire is the tool we use to collect data on the socio-economic situation of the households prior to the project activities (with the objective to draw the baseline)
- A consolidated framework to follow to organize the questions is the **sustainable livelihood framework (SLF)**
 - An approach seeking to gain an accurate understanding of people strengths and how they convert these into livelihood outcomes
 - SLF is designed along the asset pentagon (human capital, social capital, natural capital, physical capital, financial capital)
 - The centre of the pentagon indicates zero access to the assets
 - The combination of the different levels of the assets define the inter-relationship between the assets and can be represented with pentagons of different shapes



PREPARING THE QUESTIONNAIRE [2]



PREPARING THE QUESTIONNAIRE [3]

- Following the SLF, the questionnaire can be divided in **five sections**, one for each capital
- Within each section, all questions must have a purpose and collect useful information, in order to limit the length of the questionnaire
- In household surveys all the questions are usually **closed-ended**, with the following possible answers
 - Yes/no (dummy variables)
 - Numbers (quantitative variables)
 - List of options to be ticked (qualitative variables)



PREPARING THE QUESTIONNAIRE [4]

Number of household members by gender	Male <input type="text"/>	Female <input type="text"/>	Total <input type="text"/>
---------------------------------------	---------------------------	-----------------------------	----------------------------

Household members are involved in the Community Forest Committee	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> N/A
--	---

Marital status of household head	<input type="checkbox"/> Never married <input type="checkbox"/> Married (monogamous) <input type="checkbox"/> Married (polygamous) <input type="checkbox"/> Divorced/separated <input type="checkbox"/> Widowed
----------------------------------	---

- When appropriate, questions may have an additional option
 - N/R: question not relevant for the situation of the household
 - N/A: the respondent does not know the answer



PREPARING THE QUESTIONNAIRE [5]

- An important advantage of closed-ended questions is that answers can easily be coded to better compile the data matrix
 - For dummy variables, YES usually is coded with 1 and NO is coded with 0
 - Numbers do not need to be coded, the figure is reported in the matrix as such
 - Options can be coded with subsequent numbers

Marital status of household head	<input type="checkbox"/> Never married	1
	<input type="checkbox"/> Married (monogamous)	2
	<input type="checkbox"/> Married (polygamous)	3
	<input type="checkbox"/> Divorced/separated	4
	<input type="checkbox"/> Widowed	5

- N/A and N/R answers can be coded with pre-determined and easily recognizable figures like 999, 888 etc.
- Coding should always be consistent all along the questionnaire



PREPARING THE QUESTIONNAIRE: : THE GAMBIA CASE

➤ The questionnaire is adapted to the local context

Q16	Coping strategy to deal with food and/or cash shortages		Yes	No
		<input type="checkbox"/> Sale of livestock	2	1
		<input type="checkbox"/> Fuelwood collection and sale	2	1
		<input type="checkbox"/> Barter trade	2	1
		<input type="checkbox"/> Remittances	2	1
		<input type="checkbox"/> Reduced meals	2	1
		<input type="checkbox"/> Loans	2	1
		<input type="checkbox"/> Petty business	2	1
		<input type="checkbox"/> Gardening	2	1
		<input type="checkbox"/> Food harvesting from forests for household-consumption and/or sale	2	1
		<input type="checkbox"/> Borrowing food from family or friends	2	1
		<input type="checkbox"/> Harvesting and selling products from forests and/or grassland (incense, etc.)	2	1
		<input type="checkbox"/> Fishing	2	1
		<input type="checkbox"/> Other:		

➤ and to specific project objectives, and is gender-sensitive

Q56	Household women are actively involved in forest management and protection		2
		<input type="checkbox"/> Yes	1
		<input type="checkbox"/> No	99
		<input type="checkbox"/> N/A	



IMPLEMENTING THE HOUSEHOLD SURVEY [1]

- Once a draft of the questionnaire has been constructed, the questionnaire should be **tested** to ensure that local people fully understand each question
 - Administer the questionnaire to a small group of respondents (not really a sample)
 - Check that all questions are clear and answers consistent
 - Let the interviewers practice with the administration of the questionnaire
 - If needed, revise the questionnaire after the test
- Logistics should be arranged (travels and transportation for the enumerators) in order to make the administration of the questionnaire efficient



IMPLEMENTING THE HOUSEHOLD SURVEY [2]

- In the administration of the questionnaire, the **role of the interviewers** is crucial
 - They deliver the survey to the local population, thus they are the first responsible of the goodness of the data collected
 - They should be able to communicate with the respondents in the local language
 - They should be acquainted with the tools used in the study and the structure of the study
- In order to reduce the possible influence on the respondents' answers, comments and judgments by the interviewer should always be avoided
- The questions should always be administered following the same order and formulation as in the final version of the questionnaire



IMPLEMENTING THE HOUSEHOLD SURVEY [3]

- The interviewers should closely **follow the sampling scheme** defined for the survey and interview the right households (failing to do so compromise the validity of the results)
- The quality of the work done by the interviewers should be actively monitored by a **supervisor**
- The national consultant in charge of the socio-economic baseline assessment should be responsible for the work done by the interviewers



DATA ELABORATION [1]

- After administering the questionnaire, the interviewers should compile the results in the data matrix, following the coding scheme adopted
- In the data matrix, the data collected in one household surveyed are compiled in the same row; for each question, the code is put in a column
 - The final matrix will have N rows and K columns
 - N is the number of the household surveyed
 - K is the number of questions administered
- It is recommended that the matrix is filled after each day of the survey, so that any problem in the questionnaires can be readily recognized and fixed



DATA ELABORATION [2]

- Once the data matrix is constructed, it needs **cleaning**
- Cleaning seeks to identify mistakes that could have been done in compiling the data
- Cleaning should be done for each column, by browsing the data in search of possible mistaken codes
- Once the matrix has been cleaned, **descriptive analysis** can be performed
 - Tables for dummy and qualitative variables, reporting the figures of each option
 - Charts (pie charts or histograms) for dummy and qualitative variables, to show the quota of answers for each option
 - Basic descriptive statistics for quantitative variables (average value, median, standard deviation)



DATA ELABORATION [3]

- **Multivariate and inferential analyses** of the data are then performed, with the aim of
 - Making inference to generalize the data retrieved at the sample level to the whole population
 - Highlighting possible inter-correlations among the variables
 - Identifying causal connections among the variables, i.e. drivers of specific phenomena
- Main statistical tools to develop such analyses are
 - Estimation of the population features from the sample data
 - Correlation
 - Regression analysis
 - Principal component analysis, correspondence analysis and other multivariate techniques to reduce a large body of data into a few meaningful summary measures



DATA ELABORATION: : THE GAMBIA CASE

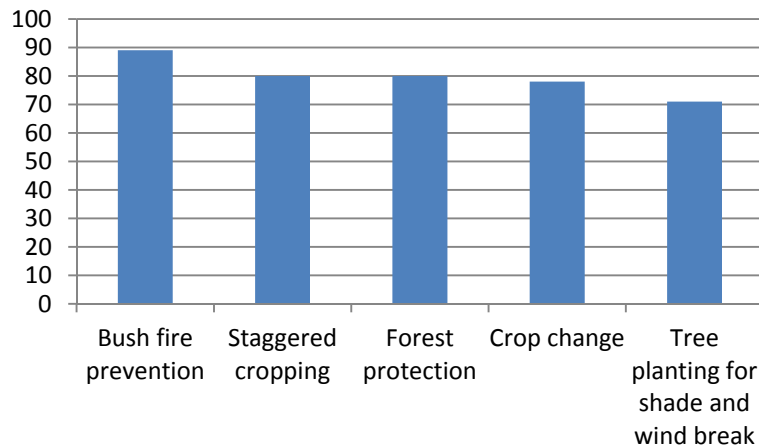
- Data from the matrix are elaborated as percentages and organized in tables (STATA)

- and used for producing charts

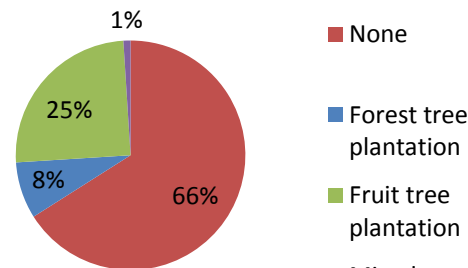
Percentage distribution of agriculture land accessible to HHs for farming by type and project cluster

Cluster	Home garden		Upland		Low land	
	Yes	No	Yes	No	Yes	No
Cluster 1	82	18	99	1	95	5
Cluster 2	81	19	95	5	67	33
Cluster 3	34	66	96	4	11	89
Cluster 4	64	36	97	3	73	27
Cluster 5	58	42	92	8	78	22
Overall	65	35	96	4	68	32

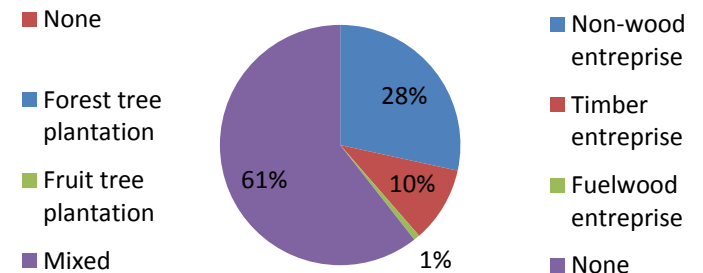
CC adaptation strategies (%HH)
Human capital



Households own private plantations (%HH)
Natural capital



Households engaged in small forest-based enterprise (%HH)
Financial capital





Food and Agriculture
Organization of the
United Nations



Thanks for your attention !

