MACRO-MODULE 5

GENDER STATISTICS AND WATER FROM DATA TO INFORMATION

OUTLINE

Module 5.1: Data recording

- Data entry process

- Build a dataset

- Create a recording code

- Transform perceptive data into quantitative data
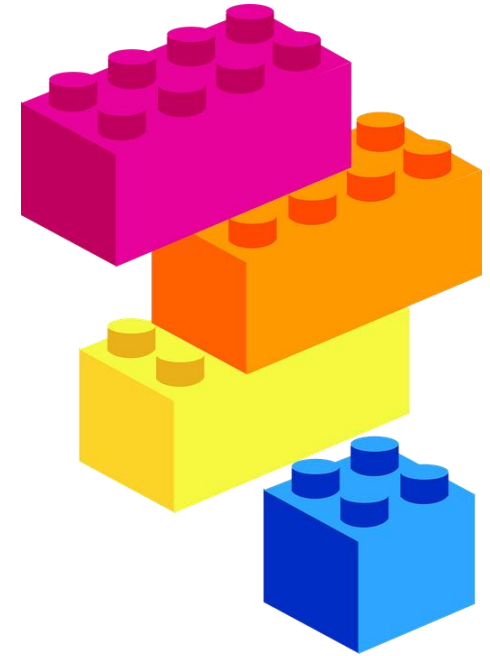
Module 5.2: Data analysis

- Information we can obtain from our database

- Basic analysis of sex-disaggregated data and graphic representation

Module 5.1: Data recording
Introduction

Scope:

- Data harmonization

- Data comparison among different regions, context or survey
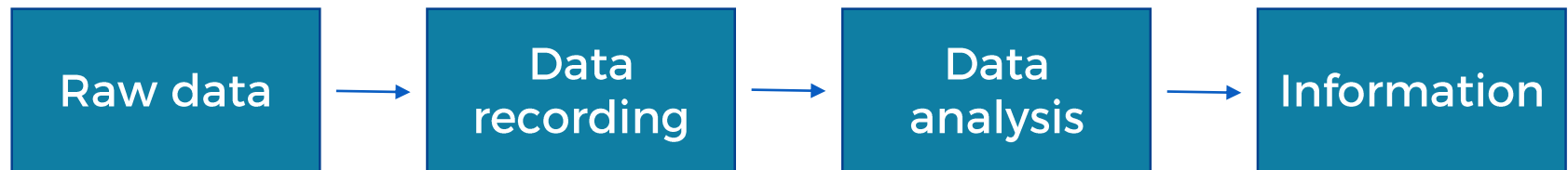
Process:

Standardization of the recording process

Building database

Moving from raw data to elaborated information

Raw data	→	Data recording	→	Data analysis	→	Information

code the answers of the survey   ->   insert them into a matrix

Data must be recorded in a systematic way.

Good data analysis depends from good data recording

Module 5.1: Data recording

Data entry process

Recording the data means preparing the data for successive elaboration and production of information.

The data entry involves the conversion of the questionnaires filled up during the survey into a **'format'**.

What does 'formatting' the questionnaire mean?

-Use the same measure unit for the same data (liters, buckets, etc)

-Use always the same code that has been established for a certain datum.

(i.e. if the answer of the question: 'Which is the purification water system?' has been codified as "F/T/C/B/N", therefore only the letters F, T, C, B or N will be used for all the recorded data as input value.

The dataset is made with the collected data

Module 5.1: Data recording
Build a database

Usually, the database is graphically represented by a matrix

The rows contain the recorded observations (interviewees' answers), and the columns the questions of the questionnaire: the cells contain the answers of the respondents

Identification	Sex (M/F)	Age	Education level (1-2-3)	Activity	Marital status (S-M-W-D)
1	M	35	2	Farmer	M
2	F	28	2	Farmer	M
3	M	10	1	Student	S
4	……..				

The database should:
Maintain a structured order
follow the structure of the questionnaire
maintain, as far as possible, the sequence of the questions

Module 5.1: Data recording
Build a database

Gender–sensitive database

GENERAL INFORMATION			HH INFORMATION							WATER															
0	0	3	4	5	6	7	8	9	10	11	11A	11B	11C	11D	11E	12A	12B	12C	12D	12E	13A	13B	14A	14E
ID	B/NB	Location	HH siz	bers	pers	resp	he res	leve	dren	sour	rce t	e co	er q	r qu	s in t	the sour	ce prot	nditi	er q	r qu	H by	HH b	water	to tr
1	B	Shalla, Citica Kuntufa, Citica	6	2		F	35	2		N						4	U	0	0	0	40	10	N	
2	B	Shalla, Citica Kuntufa, Kala	4	0	0	F	50	1		N						5	U	0	0	0	60	40	N	
3	B	Shalla, Citica Kuntufa, Hadapo	5	2	0	M	23	10		N						4	U	0	0	0	60	40	N	
4	B	Shalla, Citica Kuntufa	9	5	0	M	45	2	4	N						5	U	0	0	0	40	20	N	
5	B	Shalla, Citica Kuntufa, Citicha	17	4	0	M	38	2	1	N						4	U	0	0	0	60	40	N	
6	B	Shalla, Citica Kuntufa, Hadapo	11	1	0	M	35	5	3	N						8	U	0	0	0	80	40	N	
7	B	Shalla, Citica Kuntufa, Kala	3	0	0	F	50	0		N						8	U	0	0	0	80	40	N	
8	B	Shalla, Citica Kuntufa, Kala	12	2	0	M	38	5	5	N						4	U	0	0	0	80	40	N	
9	B	Shalla, Citica Kuntufa, Citicha	7	2	0	M	30	2	1	N						4	U	0	0	0	40	20	N	
10	NB	Shalla, Citica Kuntufa, Citicha	12	0	0	M	45	0	8	N						7	U	0	0	0	40	20	N	
11	B	Shalla, Citica Kuntufa, Citicha	6	3	0	M	26	0	0	N						6	U	0	0	0	60	20	N	
12	B	Shalla, Citica Kuntufa, Kalla	6	2	0	M	28	5	3	N						4	U	0	0	0	60	20	N	
13	B	Shalla, sadacha Kamale	12	5	0	M	36		2	N						0,3	P	1	1	0	65	25	Y	disti
14	B	Shalla, sadacha Kamale	6	2	0	M	27	6	0	N						4	P	1	0	0	100	50	Y	disti
15	B	Shalla, sadacha Kamale	14	2	0	M	42	4	1	N						3	P	1	0	0	70	30	Y	disti
16	B	Shalla, sadacha Kamale	16	3	0	F	40	3	7	N						3	P	1	0	0	100	50	Y	disti
17	B	Shalla, sadacha Kamale	12	4	0	M	46	0	1	N						1	P	1	0	0	150	50	Y	disti
18	NB	Shalla, sadacha Kamale	7	2	0	M	21	0	1	N						2	P	1	0	0	75	16	Y	disti
19	B	Shalla, sadacha Kamale	11	2	0	M	36	0		N						3	P	1	0	0	60	20	Y	disti
20	B	Shalla, sadacha Kamale	12	3	0	M	29	1	2	N						0,15	P	1	1	1	100	50	Y	disti
21	B	Shalla, sadacha Kamale	9	3	0	M	35	0	4	N						4	P	1	1	0	75	50	Y	disti
22	B	Shalla, sadacha Kamale	4	2	0	F	25	0	0	N						0,4	P	0	0	0	50	20	Y	disti
23	NB	Shalla, sadacha Kamale	16	4	0	M	40	0	6	N						0,2	P	0	0	0	90	30	Y	disti

Module 5.1: Data recording

Create a recording code

To simplify and accelerate the data-entry process, it is useful to code the answers of the respondents.

In case of multi-choice questions, we identify each answer with a number.
Ex. In this case the number is related to an 'intensity' figure (where '0' means insufficient and the highest number means excellent).

Which is your perception regarding water quantity?	
Answer	Code
☐ Insufficient	→ 0
☒ Sufficient	→ 1
☐ Good	→ 2
☐ Excellent	→ 3

	questions	
Id	Which is your perception regarding water quality?	Which is your perception regarding water quantity?
1	2	1
2	3	
3	2	

Module 5.1: Data recording
Transform perceptive data to quantitative data

In order to translate perceptive data into numeric ones, we need "to scale" the perception provided in the answers, using a coding graduation system. This process starts from the choice of the methodology.

If we use multi-choice questions, we select and address the possible answers **before** the submission of the questionnaire.

1	2	3	4	5
Strongly negative	Negative	Neutral	Positive	Strongly positive
🙁	😕	😐	🙂	😊

Module 5.1: Data recording
Transform perceptive data to quantitative data

Tips

- The data entry step requires time, precision and attention to avoid errors and complication during the analysis;

- In case of no answer questions (the so-called missing data), do not leave the cell empty, but insert a code for "question without answer" or write "na", this expedient will help you during the data analysis;

- Use meaningful field names (directly linked to the question);

- Organize the scheme matrix and standardized it as much as possible

Module 5.2: Data analysis

Information distilled from our database

- Data analysis consists in the use of the recorded data to give responses to our questions

- The data are recorded in databases, which are simply "containers" of numbers and letters

- A database has no meaning without a data elaboration that transforms data into useful information.

Module 5.2: Data analysis
Information distilled from our database

The graphs provide an immediate 'picture' and understanding of the data

Most common and useful graphs:
- bar charts
- histograms
- tacked columns

Less useful are pie charts that must be carefully used

Module 5.2: Data analysis

Information distilled from our database

The database is our baseline to apply statistic tools to distill information through gender analysis

The starting point is usually to group the raw data into categories: for example, if you are interested in differences by sex, the first thing to do is to group your data per sex.
Gender statistics requireat least two disaggregation types.

	A	B	C	D	E	F	G	H	I
1		1	2	3	4	5	6	7	8
2	ID	sex	age	educational level	recidence	marital sectors	children	state no	training on gender issues
28	25	f	2	2					n
29	26	f	2	2	p	s	y	1	y
30	27	m	2	3	p	s	y	1	y
31	28	f	4	0	p	s	y	2	n
32	29	m	4	1	p	m	y	8	n
33	30	f	3	3	t		y	3	n
34	31	f	2	3	t	s	n		n
35	32	m	2	2	p	s	y	1	n
36	33	f	3	1	p	s	y	5	y
37	34	m	3	1	p	s	y	3	n
38	35	f	2	3	p	s	y	2	n
39	36	f	3	0	p	m	y	6	n
40	37	f	2	2	p	s	y	1	n
41	38	f	3	2	t	s	y	6	n
42	39	f	3	0	p	s	y	2	n
43	40	m	3	1	p	n	y	7	n

DATA | female data | male data | DATA_ELABORATION_FEMALE | DAT

Module 5.2: Data analysis
Information distilled from our database

CLASS ACTIVITY

Analyze the different disaggregation categories to perform a correct gender analysis concerning the question: 'What is the perception of the transboundary water commission staff regarding women's involvement/contribution in decision making process?

Module 5.2: Data analysis
Information distilled from our database

CLASS ACTIVITY
possible answer

The subject of the question is the transboundary water commission staff. We want to analyze the perception of the subject regarding a specific aspect: women's involvement/contribution in decision making process. The first step is to separately analyze the females' and males' answers. Another key point could be the job category of the respondent. For this reason a further disaggregation parameter is the job category.

| ID	Sex (M/F)	Age	Job category (E-executive; C-consultant, T-technician; Administrative)	Perception of women's involvement (from 1 to 5)
1	M	45	T	4
2	F	23	A	2
3	F	56	E	2
4	M	34	C	3
5	M	57	E	5
6	M	49	C	4
7	F	28	A	1
8	M	26	T	3

Module 5.2: Data analysis

Basic analysis of sex-disaggregated data

A basic descriptive analysis of sex-disaggregated data involves the calculation of simple measures, as **percentages, rates, ratios and averages.**

These measures are useful to calculate the **gender indicators**, that show how women perform in comparison to men in the different studied areas.

- Relative frequency

- Percentage

- Absolute frequency

- Ratio

- Central tendency

- Average

- Median
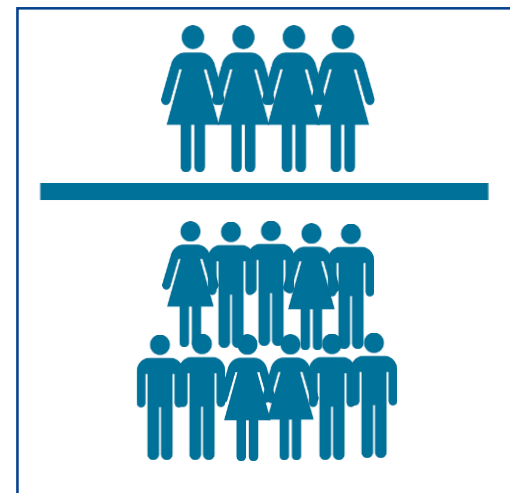
Relative frequency

Module 5.2: Data analysis

Basic analysis of sex-disaggregated data

Suppose to analyze the question: 'Who is responsible for water collection and/or distribution in the household?"

If we want to summarize this information stressing the differences between women and men, it is necessary to separately analyze the two sets of answers. The formatted results could be presented in a table or graphically with a histogram.
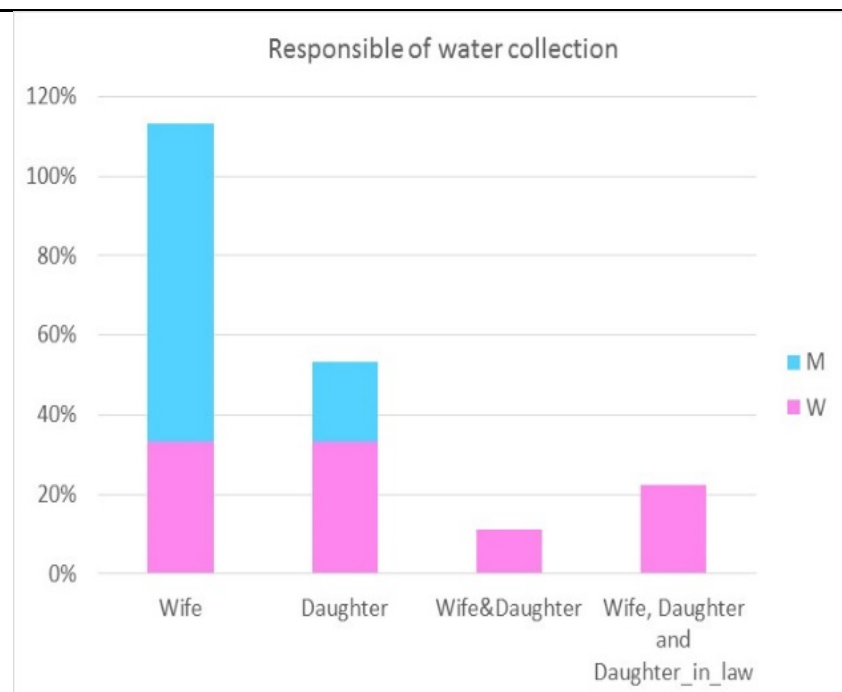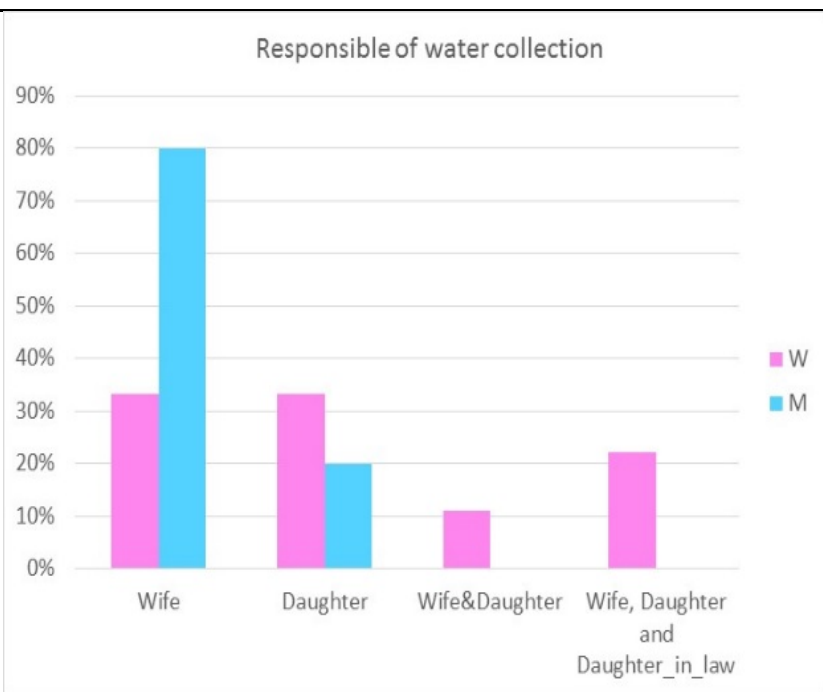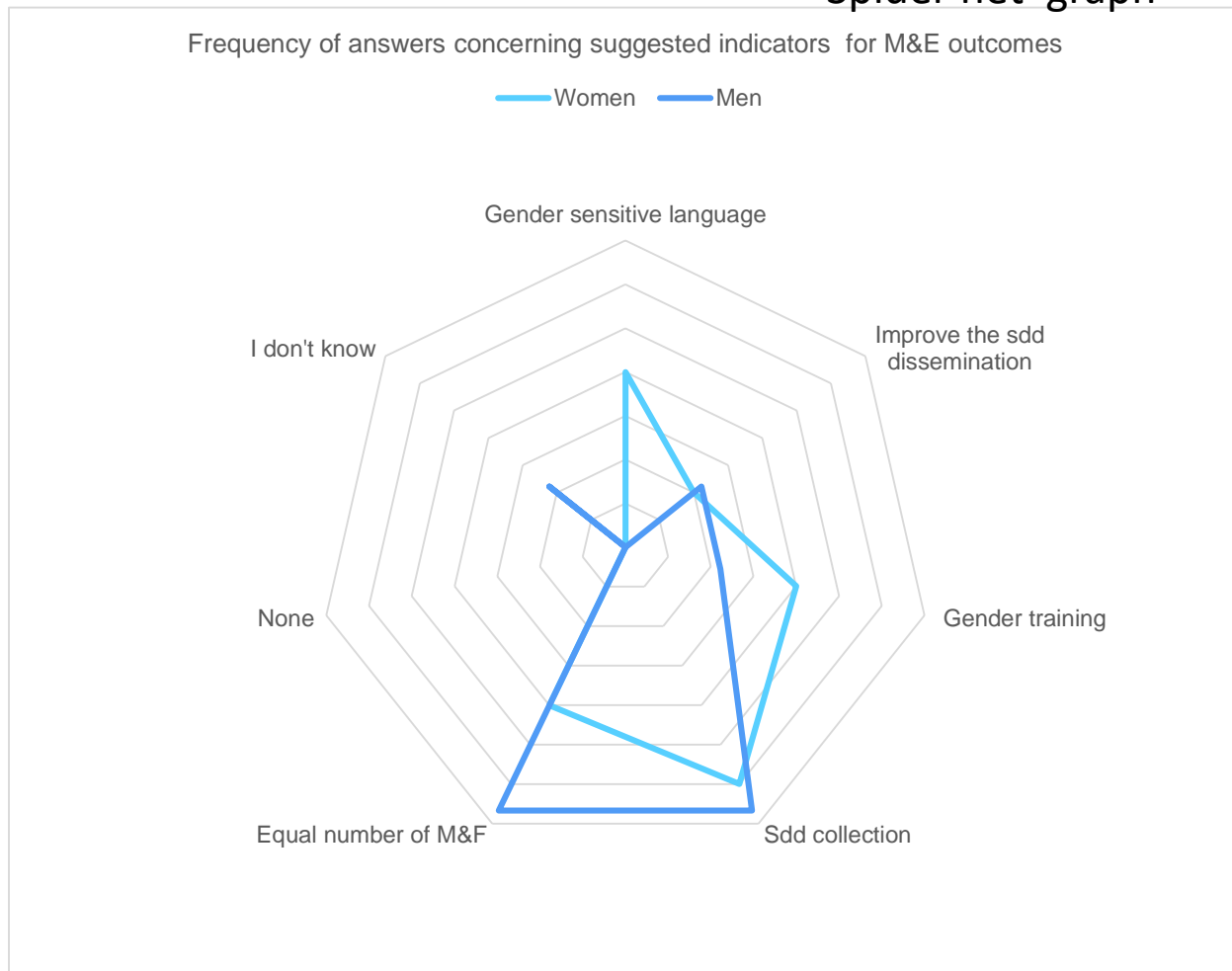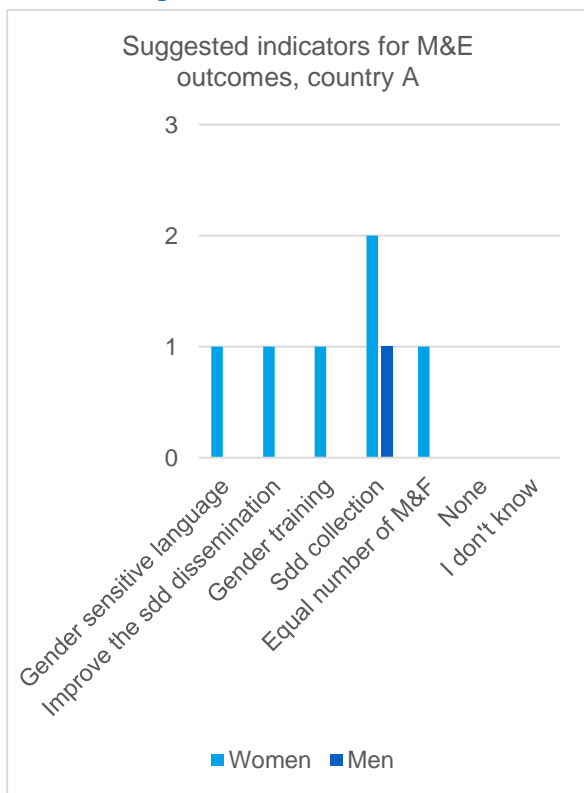
Module 5.2: Data analysis

Basic analysis of sex-disaggregated data

Spider net graph

Data representation: Graphs bar chart

Suggested indicators for M&E outcomes, country A

Frequency of answers concerning suggested indicators for M&E outcomes

Module 5.2: Data analysis
Basic analysis of sex-disaggregated data

TIP:
do not use
too many
graphics!
….and careful
to use the pie
chart….

Children responsible of the water collection
in the households

Son
2%

Daughter
98%
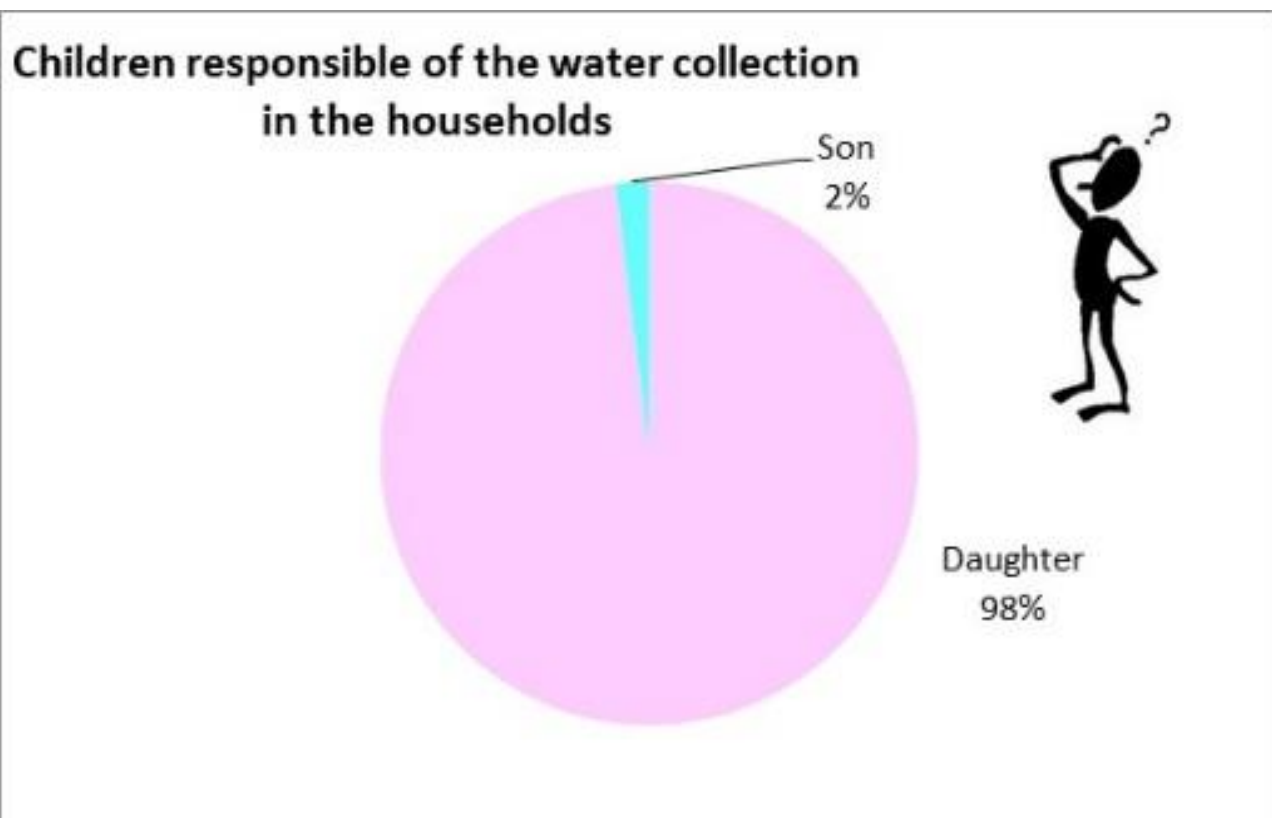
MACRO-MODULE 5

GENDER STATISTICS AND WATER FROM DATA TO INFORMATION

THANK YOU