

## Annex 10: Instructions for data analysis

### 1. Preparing the analysis

- When the data has been entered, make a copy of the file and work only on the copy for the next steps (so you will always have your basic data just in case something gets wrong when you are working with the data).
- Check that the values have been entered correctly (scale, missing values etc.).
- If the study used a smartphone-based data collection system like akvo, the data has to be converted into codes using the search replace function (→ **see Pictures 1 and 2**).
- Categorize open questions without answer categories as well as all the “others” responses in “99” → Add new columns for the new categories in the table.

With open questions or answers in the “others” category, you will get a whole range of written answers. You will have to recode them to be able to analyze them. With this process you examine the responses to a given question and look for ways to categorize them according to their similar meaning.

Example question: B068 “What are the obstacles preventing you from washing your hands with soap?”

1. Try to cluster similar responses.
2. A category name has to be found that reflects the answers of the cluster.
3. A code has to be attributed to each category.
4. Add a new column and name it B068\_kat and enter the relevant new code (**see Picture 3**).

Answer B068	Category	Code
No money for soap	Money	1
Seasonal income		1
No cash income		1
Income too low		1
Local vendors do not have it in the shop	Availability	2
No shop in the village		2
Market is far away		2

- Delete the “-88”, “-99” for all the questions.

## 2. Divide the data sample into "doers" and "non-doers"

To divide the data sample into “doers” and “non-doers” we have to form an indicator based on several questions about the behaviour and define the cut-off point. So if a person reaches a specific score she is classified as a “doer”; below this score she is a non-doer.

The critical questions and the score that define a doer have to be discussed in the group. This discussion should take place while drafting the questionnaire (see Annex 5)

See an example for handwashing → this should be discussed in the group

Example of behaviour questions	Doers	Score Doers	Non-doers	Score Non-doers
Four questions with “Imagine different handwashing situations”	Handwashing with soap is mentioned 4 times	1	Handwashing with soap is mentioned 0-3 times	0
How often did the person wash their hands?	Answer category 5 = always	1	Answer category 1(never), 2 (seldom), 3 (sometimes) and 4 (often)	0
How often did you wash your hands yesterday?	Define a minimum number, e.g. 5 times per day	1	Less than 5	0
Observation: Water and soap available	Yes	1	no	0
<b>Total score</b>		<b>4</b>		<b>0</b>

- Add a column with the following values to all critical questions: “1 = doers” et “0 = non-doers”
- Add a column where you can calculate the total score and decide if the person is a “doer” or a “non-doer”.
- Calculate the percentage of doers and non-doers in the sample.

➔ **See Picture 4**

- Discuss the result:
  - If the percentage of the “doers” in the sample is very low or too high (e.g. below 10 %), discuss if your defined indicator is too ambitious or not ambitious enough. If there really are almost no doers, the question about the intention to perform the behaviour can be added.
- Divide the data sample in “doers” and “non-doers”.

➔ **See Picture 5**

### 3. Analyze the potential intervention for questions measuring the behavioural factors

- For the variable of the data with interval (questions number 1 to 5)
  - Calculate the means for doers and non-doers separately;
  - Calculate the difference in the means for doers and non-doers;
  - Calculate the mean value of the whole sample.
- For yes/no questions:
  - Calculate the means for doers and non-doers separately → the mean value is the percentage of interviewed persons, who said “Yes (1)”;
  - Calculate the difference of the mean between doers and non-doers;
  - Calculate the mean value (percentage of interviewed persons who said “Yes (1)”) for the whole sample.
- For variables with nominal data:
  - Select the most frequent category for doers and non-doers, and calculate the percentage for both doer and non-doer categories;
  - Calculate the differences of percentage between doers and non-doers for both categories;
  - Calculate the percentage for both categories for the whole sample.
  -
- Prepare separate graphs per factor group with the following values:
  - Mean values for doers and non-doers;
  - Difference between the mean values of doers and non-doers;
  - Mean value of the whole sample.

➔ See Picture 6

#### Also:

- Describe the sample using the socioeconomic (revenue, etc.) and socio demographic (age, education, etc.) information and the approximate measures;
- Compare doers and non-doers in relation to this information.

## Picture 1

Excel spreadsheet showing a data table with columns W through AR. A blue callout box points to the 'AE' column header, stating: "Select the colon and the use the fonction find/replace under". A "Find and Replace" dialog box is open, showing "Find what: ou" and "Replace with: 1". The spreadsheet contains various text entries related to water supply and risk assessment.

## Picture 2

Excel spreadsheet showing a data table with columns R through PMH. A red callout box points to the formula bar, stating: "If you have the data for question with multiple answers, you can use the following formula to transfer the answers in the respecting rows". The formula bar shows: `=IF(ISERROR(SEARCH("1";$T11)))0;1`. The spreadsheet contains various numerical and categorical data entries.

## Picture 3

Copy of Annex Example\_data entry sheet\_proofreading\_CORR\_sp - Excel

FILE HOME INSERT PAGE LAYOUT FORMULAS DATA REVIEW VIEW

Q33

Behavior Factor

B022 (pen) What is your husband's main income activity? B022\_kat (pen) What is your husband's main income activity?

B037\_0 In which situations do you wash your hands with water and soap? Open question, check the mentioned answers. Multiple

B037\_1 B037\_2 B037\_3 B037\_4 B037\_5 B037\_6 B037\_99 B037\_Kat\_7

B069 (closed) Do you feel dirty when you don't wash your hands with soap?

B070 (closed) How agreeable is it to wash your hands with soap?

B085 (open) What are the obstacles hindering you from washing your hands with soap?

B085\_kat (open) What are the obstacles hindering you from washing your hands with soap?

B022: Create a column for each answer option. Insert a 1 if the option was mentioned, and a 0 if not.

New column for

B085: Open question answers have to be grouped in new categories and coded. See table

New column for

B037: If answer option 99 "others" was used, the answers have to be categorised.

The new defined categories are entered in an additional column.

B022: Two new categories of income activities have come up for this question in the answer option and

B022	New category B022
Agriculture	1
Cattle raising	2
Business man	3
Artisan	4
employee	5
Fisher men	99
Driver	99

B085	New category B085_kat
No obstacles	1
No moneys	2
No habit	3
Distance to soap shop	4

Picture 4

FILE

HOME

INSERT

PAGE LAYOUT

FORMULAS

DATA

REVIEW

VIEW

N25

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1		Behavior	If the answer is 2 (=handwashing with soap), the person is considered to be a doer and gets the code 1 You can use the following formula: =IF(B4>1;1;0)				If the answer is 5 or more, the person is considered to be a doer and gets the code 1 You can use the following formula: =IF(G4>4;1;0)		Here, you calculate the sum of the 4 sub-indicators. You can use the following formula: =C4+E4+G4+I4		Factors								
2	Question	B030							B031			B061	B069	B070	B071	B072	B073	B074	
	ORD	Imagine you come back from the market in your village, and your neighbour invites you to eat with him. What do you do from the moment you get there until you start to eat? 0= HW not mentioned	Imagine Indicator 1	How often do you wash your hands with soap?	Frequency Indicator 2	How many times did you wash your hands yesterday?	Amount Indicator 3	Observation: Is water and soap available?	Observation Indicator 4	Total Score Indicator Handwashing	Handwashing Doer/Non doer	Imagine if you contract diarrhea. How severe would the impact on your daily life be? (economically, socially)	(closed) Do you feel dirty when you don't wash your hands with soap?	(closed) How agreeable is it to wash your hands with soap?	In your opinion, how many people in your household wash their hands with soap after defecating and before eating?	Do people who are important to you encourage you to always wash your hands with soap after defecating and before eating?	Do you think that you are able to always wash your hand with soap after defecating and before eating?	How important is it to you to always wash your hands with soap and water before handling food?	
3																			
4	1		2	1	5	1	8	1	1	1	4	1	5	5	5	4	4	4	
5	2		2	1	4	0	6	1	1	1	3				3	3	5	3	
6	3		2	1	5	1	5	1	1	1	4	1			5	5	3	4	
7	4		2	1	1	0	5	1	1	1	3	0			3	2	4	2	
8	5		0	0	1	0	2	0	0	0	0	0			2	3	5	3	
9	6		0	0	1	0	1	0	0	0	0	0			4	2	2	4	
10	7		2	1	1	0	6	1	1	1	3	0			3	1	3	2	
11	8		0	0	4	0	3	0	0	0	0	0			2	4	4	3	
12	9		0	0	1	0	4	0	1	1	1	0			5	1	3	2	
13	10		0	0	1	0	3	0	0	0	0	0			4	1	4	4	
14	11		0	0	4	0	0	0	0	0	0	0			5	4	3	2	
15	12		1	0	5	1	7	1	1	1	3	0			4	2	2	3	
16	13		0	0	4	0	2	0	0	0	0	0			3	4	5	4	
17	14		0	0	4	0	6	1	0	0	1	0			4	3	2	3	
18	15		2	1	5	1	5	1	1	1	4	1			4	4	4	4	
19	16		0	0	4	0	1	0	0	0	0	0			4	2	2	2	
20	17		0	0	5	1	3	0	0	0	1	0			5	3	4	3	
21	Average Sample		35.3%		29.4%		47.1%		47.1%		17.6%				4.2	3.2	2.9	3.5	3.1
22			Here, you can calculate the percentage of doers for this specific sub-indicator. You can use the following formula: =AVERAGE(C4:C20)				Check the percentage of doers in your sample. If it is too high or too low, rediscuss your indicator.				Here, you will see the identified doers. Use these scores to separate your sample into doers and non-doers.				Here, you can calculate the mean score of the whole sample for this question. You can use the following formula: =AVERAGE(P4:P20)				
23																			
24																			
25																			
26																			
27																			
28																			
29																			

FILE HOME INSERT PAGE LAYOUT FORMULAS DATA REVIEW VIEW

[illegible]

Picture 6

Mark all the colons and rows you need for your graph and then use the function insert recommended chart

