

Reading and Understanding DHS Tables

Statistical tables can look intimidating at first glance. This flyer suggests ways to read and understand tables from the 2011-12 Tanzania HIV and Malaria Indicator Survey (THMIS).

Example 1: Knowledge of HIV Prevention Methods A Question Asked of a Subgroup of Survey Respondents

Step 1: Read the title and subtitle. They tell you the topic and the specific population group being described. In this case, the table is about knowledge of HIV prevention methods among women age 15-49 in Tanzania.

Step 2: Scan the column headings—the top horizontal row. They describe how the information is categorized. In this case, each column represents a method of HIV prevention. The last column lists the number of women interviewed.

Step 3: Scan the row headings—the first vertical column. These show the different ways the data are divided into categories based on population characteristics. In this case, the table presents HIV prevention knowledge among women by their age, marital status, urban-rural residence, zone of residence, educational level, and wealth. Most of the tables in DHS reports will be divided into these same categories.

Step 4: Look at the very last row at the bottom of the table. These percentages represent the totals of all women age 15-49 who know each method of HIV prevention. In this case, 69.4% of women age 15-49 know that using condoms prevents HIV infection; 83.8% know that HIV is prevented by limiting sex to one uninfected partner, and 63.2% know both methods of prevention.

Step 5: To find out what percentage of young women (age 15-24) know both HIV prevention methods, draw two imaginary lines, as shown on the table. This shows that 58.7% of young women age 15-24 know that HIV can be prevented by using condoms and by limiting sex to one uninfected partner.

Table 4.2 Knowledge of HIV prevention methods 1

Percentage of women age 15-49 who, in response to prompted questions, say that people can reduce the risk of getting the AIDS virus by using condoms every time they have sexual intercourse, and by having one sex partner who is not infected and has no other partners, by background characteristics, Tanzania 2011-12

Background characteristic	Women			Number of women
	Percentage who say HIV can be prevented by:			
	Using condoms ¹	Limiting sexual intercourse to one uninfected partner ²	Using condoms and limiting sexual intercourse to one uninfected partner ^{1,2}	
Age				
15-24	65.1	79.9	58.7	4,303
15-19	59.4	76.4	52.9	2,414
20-24	72.3	84.5	66.1	1,888
25-29	74.3	86.9	69.1	1,902
30-39	73.2	87.1	67.5	2,932
40-49	68.3	84.4	60.9	1,831
Marital status				
Never married	64.9	81.6	59.3	2,798
Ever had sex	80.8	88.3	75.5	1,303
Never had sex	51.0	75.9	45.3	1,495
Married/living together	69.9	84.5	63.9	6,910
Divorced/separated/widowed	76.4	85.1	68.4	1,258
Residence				
Urban	78.2	92.6	74.7	2,956
Rural	66.1	80.6	59.0	8,011
Zone				
Eastern	81.0	93.7	78.5	1,696
Western	72.3	89.3	67.8	890
Southern	81.6	89.5	76.4	557
Southern Highlands	76.5	85.0	69.5	1,155
Southwest Highlands	55.0	72.5	46.6	1,101
Central	66.1	84.4	62.6	1,100
Northern	55.5	86.0	52.0	1,281
Lake	71.1	76.5	60.5	2,797
Education				
No education	52.4	71.0	44.4	1,955
Primary incomplete	64.0	78.7	55.9	1,380
Primary complete	74.1	86.2	68.1	5,713
Secondary+	76.4	93.3	73.4	1,919
Wealth quintile				
Lowest	58.2	74.6	50.7	1,864
Second	64.9	79.0	56.7	1,974
Middle	68.6	80.9	61.1	1,977
Fourth	72.7	85.5	66.3	2,257
Highest	77.6	93.7	74.9	2,895
Total	69.4	83.8	63.2	10,967

Practice: Use this table to answer the following questions (answers are upside down, below):

- What percentage of married women know that using condoms prevents HIV?
- In which zone are women most likely to know that limiting sex to one uninfected partner reduces HIV risk?
- Compare women with no education to those with secondary or higher levels of education. Who knows more about HIV prevention methods?

(a) 69.9%; (b) Eastern—93.7% (b) More than 90% of women with secondary or higher education know that limiting sex to one uninfected partner reduces the risk of HIV compared to only 71% of women with no education. Knowledge of condoms is also higher among the more educated women.

Example 2: Payment for Sex and Condom Use A Question Asked of a Subgroup of Survey Respondents

Step 1: Read the title and subtitle. In this case, the table is payment for sexual intercourse and condom use among men age 15-49.

Step 2: Scan the column headings—the top horizontal row. In this case, there are two groups being described. The first panel (a) is the percentage of men who paid for sex in the year before the survey among **all** men in the survey. The second panel (b) is about condom use during paid sex and refers only to the men who have paid for sex in the year before the survey.

Step 3: Look at the first panel. The final row indicates that 8.5% of all men paid for sex in the past year. How many men are included in this group? 8,352, or all of the men in the survey.

The second panel (b) asks, among the men who paid for sex in the year before the survey, what percentage reported using a condom the last time they paid for sex? The answer is 52.9%. And how many men are included in this group? Only 708, or 8.5% of 8,352. The second panel is a subgroup of the first.

Step 4: In some zones, there are not many men who report having paid for sex in the year before the survey. When there are fewer than 25 unweighted cases, the figure will be replaced by an asterisk, as in the case of condom use at last sex among men paying for sex in Unguja and Pemba.

When there are between 25 and 49 cases, the figure is placed inside parentheses to warn the reader to use these figures with caution, as in the case of condom use at last sex among men paying for sex in Southern Highlands and Southwest Highlands. Because there are not very men in these zones who reported having paid for sex in the past year it is difficult to get accurate data about them. Use these numbers in parentheses with caution.

Practice: Use this table to answer the following questions (answers are upside down, below):

- Are men in Mainland Tanzania or Zanzibar more likely to report having paid for sex in the past year?
- In what zone are men most likely to report having paid for sex in the past year?
- In urban areas, what percentage of men who paid for sex last year used a condom at last sex?
- What percentage of men in Western Zone used a condom the last time among those who reported paying for sex? Can you use this number with confidence?

Table 6.4 Payment for sexual intercourse and condom use at last paid sexual intercourse 1

Percentage of men age 15-49 who ever paid for sexual intercourse and percentage reporting payment for sexual intercourse in the past 12 months, and among them, the percentage reporting that a condom was used the last time they paid for sexual intercourse, by background characteristics, Tanzania 2011-12

Background characteristic	2 Among all men:		Among men who paid for sex in the past 12 months:	
	Percentage who paid for sexual intercourse in the past 12 months	Number of men	Percentage reporting condom use at last paid sexual intercourse	Number of men
Residence				
Urban	7.0	2,142	66.7	149
Rural	9.0	6,210	49.1	559
Mainland/Zanzibar				
Mainland	a 8.8	8,079	b 52.9	707
Urban	7.2	2,066	66.7	149
Rural	9.3	6,013	49.2	558
Zanzibar	0.3	273	*	1
Unguja	0.3	204	*	1
Pemba	0.5	69	*	0
Zone				
Eastern	5.9	1,363	4	63.2
Western	6.5	736	(60.5)	48
Southern	22.3	371	(57.4)	83
Southern Highlands	9.3	818	(64.4)	76
Southwest Highlands	4.9	851	(44.0)	41
Central	6.6	908	50.4	60
Northern	4.5	855	(57.0)	38
Lake	12.9	2,178	45.6	281
Middle	9.9	1,590	51.6	158
Fourth	8.9	1,749	53.0	156
Highest	5.6	2,123	63.3	119
Total	8.5	8,352	3	52.9
				708

Note: Figures in parentheses are based on 25-49 unweighted cases. An asterisk indicates that a figure is based on fewer than 25 unweighted cases and has been suppressed.
nc = No cases

Note: When parentheses or asterisks are used in a table, the explanation will be noted under the table. If there are no parentheses or asterisks on a table, you can proceed with confidence that enough cases were included in all categories.

For more about weighted and unweighted numbers, see example 4.

a) Mainland—8.8% of men reported having paid for sex compared to <1% in Zanzibar; b) Southern Zone—22.3%; c) 66.7%; d) 60.5%; but the parentheses indicate that this is based on only 25-49 men who reported having paid for sex in the year before the survey, so it should be used with caution.

Example 3: Prevalence of Malaria in Children

Comparing Data and Understanding Patterns

Step 1: Read the title and subtitle. In this case, the table is about prevalence of malaria in children age 6-59 months according to microscopy tests.

Step 2: Scan the column headings- the top horizontal row. In this case there is only one variable, malaria prevalence according to microscopy.

The third column, “number,” indicates how many children were tested for malaria in the survey. In this case, a total of 7,322 children age 6-59 months were tested for malaria.

Step 3: Scan the row headings—the first vertical column. These show the different ways the data are divided into categories based on population characteristics. This table presents malaria among children by age of child, sex of child, residence, region, mother’s education, and wealth quintile of household. The data in these categories will help you understand how malaria prevalence varies throughout the country.

Step 4: In all, 4.1% of children age 6-59 months tested positive for malaria in Tanzania. But this number is a national average; it does not explain how malaria prevalence varies throughout the country and among the population.

Answer the following questions to understand how malaria infection varies throughout the population:

- What are the lowest and the highest malaria rates within the regions? Malaria prevalence ranges from a low of 0.0% in Arusha, Iringa, Mbeya, Singida, Rukwa, and parts of Zanzibar to a high of over 20% in Geita.
- Look for patterns: Does malaria infection increase with age? Mother’s education? Wealth quintile?
- Compare different groups: Are male children more likely to have malaria? What about children living in urban areas?

Step 5: Why is this important? Program managers can use this information to develop effective programmes. For example, certain regions are not very affected by malaria, and therefore targeting a lot of money towards prevention in these regions may not be cost-effective. Similarly, malaria prevalence is highest among children from poor families and among children whose mothers are not well educated. Targeted outreach to these groups will be most effective in preventing malaria.

Table 11.5 Prevalence of malaria in children **1**

Percentage of children age 6-59 months classified by microscopy as having malaria, by background characteristics, Tanzania 2011-12

Background characteristic 3	Malaria prevalence according to microscopy 2	
	Malaria prevalence according to microscopy	Number of children tested
Age (in months)		
6-8	2.1	400
9-11	1.9	400
12-17	2.2	906
18-23	4.0	821
24-35	4.4	1,632
36-47	4.9	1,668
48-59	5.2	1,494
Sex		
Male	4.1	3,694
Female	4.1	3,628
Residence		
Urban	1.0	1,179
Rural	4.7	6,143
Region		
Dodoma	0.5	258
Arusha	0.0	225
Kilimanjaro	0.5	167
Tanga	2.5	405
Morogoro	6.9	242
Pwani	7.4	119
Dar es Salaam	0.3	379
Lindi	4.1	90
Mtwara	2.0	166
Ruvuma	0.8	403
Iringa	0.0	119
Mbeya	0.0	405
Singida	0.0	347
Tabora	3.8	363
Rukwa	0.0	179
Kigoma	9.9	296
Shinyanga	4.4	348
Kagera	5.5	408
Mwanza	5.4	419
Mara	14.4	392
Manyara	0.7	210
Njombe	1.4	167
Katawi	5.7	191
Simiyu	2.0	511
Geita	20.7	316
Kaskazini Unguja	0.0	24
Kusini Unguja	0.6	16
Mjini Magharibi	0.0	97
Kaskazini Pemba	0.0	30
Kusini Pemba	2.1	29
Mother's education²		
No education	6.1	1,647
Primary incomplete	5.2	876
Primary complete	3.3	3,555
Secondary+	1.5	536
Wealth quintile		
Lowest	5.3	1,722
Second	5.5	1,684
Middle	4.3	1,493
Fourth	3.6	1,304
Highest	0.6	1,119
Total	4.1	7,322

Example 4: Understanding Sampling Weights in DHS Tables

A sample is a group of people that have been selected for a survey. In DHS surveys, the sample represents the entire national population. Most countries want to collect data and report information both for the entire country and also for a country's regions or provinces. To estimate geographic differentials for certain demographic indicators data were collected from all of Tanzania's 30 regions.

DHS surveys are designed to provide these national and regional statistics. We want the sample surveyed in each region to resemble the actual population of that region, just as we want the national sample to resemble the actual population of the country. If the regions in a particular country vary in size and especially if some regions have very small populations, then a randomly-drawn sample may not include enough people from each region for analysis.

For example, let's say that you have enough money to interview 10,967 women for a survey that should be representative of both the regions and the entire country (as in the Tanzania table to the right). In Tanzania, the regions are not evenly distributed: some regions are more heavily populated (such as Dar es Salaam), while others have smaller populations (such as Kusini Unguja).

A sampling statistician can determine how many women should be interviewed in each region in order to get reliable statistics. In the case of Tanzania, the **blue column (1)** shows the actual number of women selected and interviewed in each region, ranging from 277 in Dodoma to 629 in Dar es Salaam. With these numbers, there are enough interviews to get reliable results in even sparsely populated regions.

With this distribution of interviews, some regions are overrepresented and some regions are underrepresented. For example, the population of Kusini Unguja in 2010 was less than 1% of the entire Tanzanian population. In contrast, the population of Dar es Salaam was about 10% of the Tanzanian population. But as the blue column shows, the DHS survey has interviewed 289 women in Kusini Unguja, or about 3% of the survey sample. Meanwhile, only 629 women were interviewed in Dar es Salaam, about 6% of the survey sample. This does not accurately represent the population of the country.

Table 3.1 Background characteristics of respondents
Percent distribution of women age 15-49 by selected background characteristics, Tanzania 2011-12

Background characteristic	Women		
	Weighted percent	Weighted number	Unweighted number
Region			
Dodoma	3.8	422	277
Arusha	3.0	331	341
Kilimanjaro	3.5	384	331
Tanga	5.2	566	352
Morogoro	3.6	399	341
Pwani	1.9	213	284
Dar es Salaam	9.9	1,084	629
Lindi	1.7	188	312
Mtwara	3.4	369	301
Ruvuma	6.2	684	364
Iringa	1.8	200	315
Mbeya	6.4	699	378
Singida	3.8	416	386
Tabora	3.9	432	440
Rukwa	1.7	187	342
Kigoma	4.2	458	417
Shinyanga	3.8	415	333
Kagera	4.1	448	340
Mwanza	5.2	570	430
Mara	4.0	433	441
Manyara	2.4	262	385
Njombe	2.5	271	317
Katavi	2.0	214	299
Simiyu	5.7	626	466
Geita	2.8	304	566
Kaskazini Unguja	0.4	42	319
Kusini Unguja	0.2	26	289
Mjini Magharibi	2.1	230	371
Kaskazini Pemba	0.4	47	300
Kusini Pemba	0.4	46	301
Total	100.0	10,967	10,967

In order to get statistics that are representative of the entire country, the distribution of the women in the sample needs to resemble the distribution of the women in the country. Women from a smaller region, like Kusini Unguja, should only contribute a small amount to the national total. Likewise, women from a more populated region, like Dar es Salaam, should contribute more. Therefore, DHS statisticians mathematically adjust or "weight" the number of women from each region so that each region's contribution to the total is proportionate to the actual population of the country. The numbers in the **purple column (2)** represent the "weighted" values. The total sample size of 10,967 women has not changed, but the distribution of the women in the regions has been changed to represent their contribution to the total population size.

How do statisticians weight each category? They recalculate the categories to reflect the real population of the country. If you were to compare the **light red column (3)** to the actual population distribution of Tanzania, you would see that women in each region are contributing to the total sample with the same weight that they contribute to the population of the country. The weighted number in Kusini Unguja has gone down, to represent the small population of that region, while the weighted number in Dar es Salaam has gone up to represent the high population of that region.

With sampling and weighting, it is possible to interview enough women to provide reliable statistics at both the national and provincial level without distorting the overall distribution of the population within the country. In general, only the weighted numbers are shown in each of the DHS tables, so don't be distressed if these numbers seem low: they may actually represent a larger number of women interviewed. And remember, the table will use parentheses and asterisks to warn you if there are too few unweighted cases in any category.