



Reading and Understanding DHS Tables

Statistical tables can look intimidating at first glance. These examples, from the Zambia DHS 2007, highlight ways to read and understand tables.

Example 1: Knowledge of HIV Prevention Methods: A question asked of all 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 women age 15-49 in Zambia. This represents the entire female survey population in most DHS surveys.

Step 2: Scan the column headings (the top horizontal row). They describe how the information is categorized. In this case, each column represents one aspect of knowledge of prevention the women report to have. Note that the last column lists the (weighted) number of women in each group. For more on weighting, see example 4.

Step 3: Scan the row headings (the first vertical column). These show the different categories the data are divided into based on the population characteristics. In this case, the table presents knowledge of HIV prevention methods by age, marital status, urban/rural residence, province of residence, educational level, and wealth. Most of the tables in DHS reports will be divided up in the 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, 73.2% of women age 15-49 know that HIV can be prevented by using condoms.

Step 5: To find out what percentage of women in Lusaka know that HIV can be prevented by using condoms and limiting sex to one uninfected partner, draw 2 imaginary lines, as shown on the table. 63.4% of women age 15-49 in Lusaka know that HIV can be prevented by using condoms and limiting sex to one uninfected partner.

Table 13.2 Knowledge of HIV prevention methods ¹

Percentage of women and men age 15-49 who, in response to prompted questions, say that people can reduce the risk of getting being infected with HIV by using condoms every time they have sexual intercourse, by having one sex partner who is not infected and has no other partners, and by abstaining from sexual intercourse, by background characteristics, Zambia 2007

Background characteristic	Women				Number of women
	Using condoms ¹	Limiting sexual intercourse to one uninfected partner ²	Using condoms and limiting sexual intercourse to one uninfected partner ^{1,2}	Abstaining from sexual intercourse	
Age					
15-24	71.0	86.6	65.4	82.5	2,944
..15-19	66.5	83.0	60.2	80.2	1,574
..20-24	76.3	90.7	71.3	85.3	1,370
25-29	76.0	90.8	71.8	85.3	1,363
30-39	75.5	92.6	72.3	88.0	1,803
40-49	71.7	91.8	69.2	85.7	1,036
Marital status					
Never married	67.9	84.4	61.9	80.8	1,856
Married/Living together	75.3	91.8	71.9	86.2	4,402
Divorced/Separated/Widowed	73.7	90.3	68.8	86.9	888
Residence					
Urban	73.1	89.9	68.2	84.0	3,009
Rural	73.3	89.6	69.4	85.6	4,137
Province					
Central	71.9	90.6	69.5	88.8	659
Copperbelt	72.9	89.7	69.0	79.8	1,264
Eastern	79.1	94.9	78.1	93.8	971
Luapula	69.8	91.6	66.2	79.3	530
Lusaka	69.1	89.1	63.4	85.2	1,172
Northern	66.4	79.0	56.2	78.8	966
North-Western	74.7	91.4	71.6	83.7	365
Southern	75.9	91.0	71.9	84.9	727
Western	85.7	95.4	84.5	93.5	492
Education					
No education	66.9	87.8	63.6	83.4	744
Primary	72.6	88.9	68.3	84.4	3,891
Secondary	75.0	90.6	70.1	85.2	2,140
More than secondary	81.8	96.6	79.0	91.1	371
Wealth quintile					
Lowest	74.7	89.7	71.0	88.4	1,240
Second	71.6	88.5	68.0	84.4	1,283
Middle	72.7	89.1	68.1	83.3	1,280
Fourth	72.2	89.0	67.6	83.4	1,567
Highest	74.5	91.5	69.9	85.3	1,776
Total 15-49	73.2	89.7	68.9	84.9	7,146

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

- a) What percentage of women age 15-24 who are aware that HIV can be prevented by using condoms AND limiting sex to one partner?
- b) In what province(s) are women most aware that HIV can be prevented by abstaining from sex?
- c) Are urban or rural residents more likely to know that HIV can be prevented by using condoms?

a) 65%; b) Western and Eastern-94% ; c) It is the same (73%).

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 about two separate groups of men a) all men age 15-49 and b) only men in that age group who reported that they paid for sex in the past 12 months.

Step 2: Identify the two panels. First identify the columns that refer to all men (a) and then isolate the columns that refer only to the men who reported having paid for sex (b).

Step 3: Look at the first panel. What percentage of all men age 15-49 have paid for sex in the past year? It's 4.7%.

Now look at the second panel. How many men are included in this group? Only 283, or 4.7% of 5,995 men who were asked about paying for sex. The second panel is a subgroup of the first.

Step 4: There are very few men who report having paid for sex in the past year. Once these men are further divided into the background characteristics categories, there may be too few cases for the percentages to be reliable.

For example, look to see the percentage of men in Western who used a condom at last paid intercourse: 29.3%. This percentage is in parentheses because it is based on fewer than 50 men (unweighted) in this category. This means use this number with caution—it may not be accurate. (For more information on weighted and unweighted numbers, see example 4.)

If a number is replaced by an asterisk, it is because there are fewer than 25 people in that category, and results for this group are not reported. The subgroup is too small, and therefore, those data are not reliable.

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

Table 13.9 Payment for sexual intercourse and condom use at last paid sexual intercourse: Men

Percentage of men age 15-49 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, Zambia 2007

Background characteristic	Payment for sexual intercourse in the past 12 months		Condom use at last paid sexual intercourse	
	Percentage who paid for sexual intercourse	Number of men	Percentage reporting condom	Number of men who paid for sexual intercourse in the past 12 months
Age				
15-24	5.0	2,482	49.7	124
..15-19	4.3	1,416	46.1	61
..20-24	5.9	1,066	53.1	63
25-29	6.4	977	63.3	62
30-39	4.6	1,671	54.6	77
40-49	2.3	865	(65.8)	20
Marital status				
Never married	5.7	2,553	52.6	146
Married or living together	3.3	3,168	58.5	105
Divorced/separated/widowed	11.8	274	55.5	32
Residence				
Urban	4.3	2,601	62.0	113
Rural	5.0	3,395	50.6	170
Province				
Central	7.3	559	(55.7)	41
Copperbelt	2.8	1,140	*	32
Eastern	5.3	795	(45.8)	42
Luapula	1.3	387		5
Lusaka	6.1	1,072	74.6	66
Northern	2.2	805	*	18
North-Western	9.8	303	53.3	30
Southern	5.7	621	(57.2)	35
Western	4.8	315	(29.3)	15
Education				
No education	3.4	267	*	9
Primary	5.4	2,775	51.3	151
Secondary	4.6	2,512	59.8	116
More than secondary	1.7	441	*	8
Wealth quintile				
Lowest	5.2	1,114	47.5	58
Second	4.1	869	53.2	35
Middle	5.5	1,097	40.5	61
Fourth	6.2	1,381	63.2	86
Highest	2.8	1,534	71.5	43
Total 15-49	4.7	5,995	55.1	283

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.

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

a) In what age group is payment for sex the most common?

b) Among the men with primary education who paid for sex, what percentage used a condom the last time they had paid sex? Can you use this answer with confidence? Why or why not?

a) 25-29 (6.4%); b) 51.3%. There is no parentheses or asterisk, and there are 151 men in this group (more than 50), so you can use this figure with confidence.

Example 3: HIV Prevalence by Background Characteristics

Comparing data and understanding patterns

Step 1: Read the title and subtitle. In this case, the table is about HIV prevalence among women age 15-49 in Zambia.

Step 2: Scan the column headings (the top horizontal row). In this case, there is only one variable, percentage of women who are HIV-positive.

The third column (number tested) indicates how many women in each category were tested for HIV in the survey. In this case, 2,317 women living in urban areas were tested for HIV in the survey. In total, 5,502 women were tested.

Step 3: Scan the row headings (the first vertical column). These show the different categories the data are divided into based on the population characteristics. In this table, HIV prevalence is presented by urban/rural residence, province of residence, religion, educational level, employment status, and wealth. The data in these categories will help you understand how HIV prevalence varies throughout the country.

Step 4: Answer the following questions to understand how HIV prevalence is spread throughout the population:

- Where is HIV the highest and the lowest?
- Look for patterns: do results go up or down? Does prevalence vary within specific populations? For example, is there a clear pattern of HIV prevalence by wealth? by education?
- Compare different groups: do urban residents have a different HIV prevalence than rural residents?

Step 5: What does all this mean? Once you see that urban residents and residents of Lusaka and Central have the highest HIV prevalence, how can this knowledge affect your work? Program planners and policymakers can use this detailed information to inform their activities. Perhaps they should focus prevention campaigns in those areas and make sure that there are health centers prepared to deal with larger numbers of HIV-positive women in those areas.

What other trends exist? In Zambia, HIV prevalence is highest among the wealthiest women. What does this mean for the program planner? It means that outreach should not be targeted only at the poorest women in the population. Wealthy women are more likely to carry HIV infection than poorer women. Programs should be designed to address them. There is also a clear pattern among educational levels, but there is no educational group that is safe from HIV.

Table 14.5 HIV prevalence by socioeconomic characteristics

Percentage HIV positive among women and men age 15-49 who were tested, by socioeconomic characteristics, Zambia 2007

Background characteristic	Women	
	Percentage HIV positive	Number
Residence		
Urban	23.1	2,317
Rural	11.0	3,185
Province		
Central	22.0	507
Copperbelt	21.6	973
Eastern	11.0	748
Luapula	11.5	408
Lusaka	22.4	902
Northern	7.7	744
North-Western	9.1	281
Southern	15.8	560
Western	16.1	379
Religion		
Catholic	14.2	1,106
Protestant	16.7	4,290
Muslim	(1.9)	31
Other	15.9	69
Education		
No education	10.8	549
Primary	15.8	2,937
Secondary	17.4	1,726
More than secondary	21.3	290
Employment (last 12 months)		
Not employed	14.1	2,523
Employed	17.7	2,978
Wealth quintile		
Lowest	8.8	939
Second	9.6	987
Middle	13.3	973
Fourth	22.9	1,193
Highest	21.6	1,410
Total 15-49	16.1	5,502

Notes: Total includes 19 cases with missing information on religion and 2 cases with missing information on employment. Figures in parentheses are based on 25-49 unweighted cases.

Example 4: Understanding Samples and Weighting in DHS Tables

A sample is a group of people that has been selected for a survey. The sample is supposed to represent the entire population that you would like to learn about. Most countries want to collect data and report information for the entire country as well as for a country's provinces.

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

For example, let's say that you have enough money to interview about 7,000 women for a survey that should be representative of both the provinces and the entire country (as in the Zambia table to the right). In Zambia, the provinces are not evenly distributed—some provinces are very heavily populated, while others have very small populations.

A sampling statistician can determine how many women should be interviewed in each province in order to get reliable statistics for the specific indicators. In the case of Zambia, the **blue column (1)** shows the actual (unweighted) number of women selected and interviewed in each province, ranging from 672 in Central to 940 in Eastern. Now, there are enough interviews to get reliable results in each province.

But now there is a new challenge. With this distribution of interviews throughout the provinces, some provinces are overrepresented and some provinces are underrepresented. For example, Western's population in 2000 was about 765,000 people, which only represents roughly 7% of the entire Zambian population. On the other hand, Copperbelt's population in 2000 was 1,500,000, or approximately 18% of the Zambian population. But based on the blue column, our survey has interviewed almost as many women in Western as in Copperbelt! Does this represent the population of the country?

In order to get statistics that are representative of the entire country, the distribution of the women in our sample needs to resemble the distribution of the women in the country. Women from a small province, like Western, should only contribute a small amount to the national total. Likewise, women from a largely populated province, like Copperbelt, should contribute more. So, DHS statisticians adjust or "weight" the numbers of women from each province so that each province'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 7,146 women has not changed, but the distribution of the women in the provinces has been changed to represent their contribution to the total population size. The weighted numbers suggest the true relationship between the populations of Western and Copperbelt: the population of Copperbelt is twice as large as Western.

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 Zambia, you would see that women in each province are contributing to the total sample with the same weight that they contribute to the population of the country. The weighted number of women in the survey now accurately represents how many women live in Copperbelt (18% of the Zambian population) and how few women live in Western (7% of the population).

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 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.

Table 3.1 Background characteristics of respondents
Percent distribution of women and men age 15–49 by selected background characteristics, Zambia 2007

Background characteristic	Women		
	Weighted percent	Weighted	Unweighted
Province			
Central	9.2 ³	659 ²	672 ¹
Copperbelt	17.7	1,264	829
Eastern	13.6	971	940
Luapula	7.4	530	704
Lusaka	16.4	1,172	939
Northern	13.5	966	783
North-Western			
Western	5.1	365	685
Southern	10.2	727	822
Western	6.9	492	772
Total 15–49	100.0	7,146	7,146