

## WEIGHTS IN DHS DATA FILES

- Sample weights are calculated by sampling experts and included in each DHS recode file
- Sample weights compensate for the different probabilities of selection within the sample, and for different levels of non-response.
- Non-response can happen among households, women and men.
- Different weights are calculated for the different units of analysis, or for different survey modules.

UNIT OF ANALYSIS	VARIABLE
Households	hv005
Household members (PR file)	hv005
Women or children	v005
Men	mv005
Women or Men Interviewed for Domestic Violence Module	d005
HIV test results	hiv05

## TWO GOALS OF WEIGHTING

- To make the sample representative of the entire population
  - Units are included in the sample with unequal chance
  - The population is not evenly distributed among different regions and place of residence.
  - Design weights are calculated to retrieve the real population distribution among regions and place of residence
- To account for non-response
  - **After** the survey is finished, design weights are adjusted for non-response among different units of analysis.

## WHICH WEIGHT SHOULD I USE?

Use the weight applicable to the unit of analysis (household, women, men)

- When analyzing **households**, use the **household weight or hv005**
- When analyzing **women**, use the **women's weight or v005**
- When analyzing **men or couples**, use the **men's weight or mv005**
- If analyzing **domestic violence data**, use the **DV weight or d005**
- When analyzing **HIV data**, HIV prevalence or HIV testing data needs the **HIV weight or hiv05**. Variables related to HIV knowledge, attitudes, and behaviors use the women's or men's weight respectively.

# DHS CONVENTIONS FOR PRESENTING DATA IN FINAL REPORTS

- DHS tables show weighted percentages (figures) AND weighted sample sizes (denominators)
- For every table, DHS also checks the unweighted denominator to see if the sample size is too small and thus results are unreliable
- If a figure is based on **<25 unweighted cases**, the figure is not shown and is replaced with a \*
- If a figure is based on **25-49 unweighted cases**, the figure is shown in **(parentheses)** and should be interpreted with caution

Many people have been taught to worry about small sample sizes. However, that refers to the unweighted number of respondents, which are generally not shown in DHS tables. Most DHS tables show the weighted number of households and respondents, and this sometimes may appear to be rather small. But, you have to take into account that the unweighted number can be much larger.

## DHS WEIGHT VALUES CONVENTION

All DHS weights allow for 8 digits, with an implicit 6 decimals

- **Remember: no decimal places in DHS variables**
- If hv005 = 2561000, the actual weight value is 2.561
- If mv005 = 789000, the actual weight value is 0.789



divide by  
1,000,000

## IN SUMMARY...

- Weights restore the representativeness of sampled data
- Weights correct for non-response
- All tables in DHS final reports are weighted unless otherwise noted
- Must use weights when using DHS data!