

# Module 2: Data analysis key concepts



# Module 2: Learning objectives

- Understand the definition and purpose of data analysis
- Define statistical and M&E key concepts in data analysis

# Data Analysis

- Turning raw **data** into useful **information**
- Purpose is to provide answers to questions being asked at a program site or research questions
- Even the greatest amount and best quality data mean nothing if not properly analyzed – or if not analyzed at all

# Data Analysis

- Analysis does not mean using computer software package
- Analysis is looking at the data in light of the questions you need to answer:
  - How would you analyze data to determine, “Is my program meeting its objectives?”



**USAID**  
FROM THE AMERICAN PEOPLE



**MEASURE**  
Evaluation

# Answering programmatic questions

- **Question:** Is my program meeting its objectives?
- **Analysis:** Compare program targets and actual program performance to learn how far you are from target.
- **Interpretation:** Why you have or have not achieved the target and what this means for your program.
- May require more information.

# Descriptive analysis

- Describes the sample/target population (demographic & clinic characteristics)
- Does not define causality – tells you *what* not *why*
- Example – average number of clients seen per month



**USAID**  
FROM THE AMERICAN PEOPLE



**MEASURE**  
Evaluation

# Basic terminology and concepts

- Statistical terms
  - Ratio
  - Proportion
  - Percentage
  - Rate
  - Mean
  - Median

# Ratio

- Comparison of two numbers expressed as:
  - a to b, a per b, a:b
- Used to express such comparisons as clinicians to patients or beds to clients
- Calculation  $a/b$
- Example – In district X, there are 600 nurses and 200 clinics. What is the ratio of nurses to clinics?

$$\frac{600}{200} = 3 \text{ nurses per clinic, a ratio of } 3:1$$



# Calculating ratios

- In Kwakaba district, there are 160 nurses and 40 clinics
- What is the nurse-to-clinic ratio?

$$\frac{160}{40} = 4$$

4:1 or 4 nurses to 1 clinic

# Proportion

- A ratio in which all individuals in the numerator are also in the denominator.
- Used to compare part of the whole, such as proportion of all clients who are less than 15 years old.
- Example: If 20 of 100 clients on treatment are less than 15 years of age, what is the proportion of young clients in the clinic?
- $20/100 = 1/5$

# Calculating proportions

- Example: If a clinic has 12 female clients and 8 males clients, then the proportion of male clients is  $8/20$  or  $2/5$
- $12+8 = 20$
- $8/20$
- Reduce this (numerator and denominator are multiples of 4) =  $2/5$  of clients = male

# Percentage

- A way to express a proportion (proportion multiplied by 100)
- Expresses a number in relation to the whole
- Example: Males comprise  $\frac{2}{5}$  of the clients or, 40% of the clients are male ( $0.40 \times 100$ )
- Allows us to express a quantity relative to another quantity. Can compare different groups, facilities, countries that may have different denominators



**USAID**  
FROM THE AMERICAN PEOPLE



# Rate

- Measured with respect to another measured quantity during the same time period
- Used to express the frequency of specific events in a certain time period (fertility rate, mortality rate)
  - Numerator and denominator must be from same time period
  - Often expressed as a ratio (per 1,000)

# Mortality rate

- Calculation
- # of deaths ÷ population at risk in same time period x 1,000
- Example – 600 women delivered. Of these, 10 died of obstructed labor
- $10/600 = .016 \times 1,000 = 16.6$

*17 women died of obstructed labor per 1,000 live births*

# Calculating mortality rate

In 2009, Mondello Clinic had 31,155 patients on ART. During that same time period, 1,536 ART clients died.

$$\frac{1,536}{31,155} = .049 \times 1000 = 49$$

*49 clients died  
(mortality rate) per  
1,000 clients on  
ART*



**USAID**  
FROM THE AMERICAN PEOPLE



**MEASURE**  
Evaluation

# Rate of increase

- Calculation
- Total number of increase ÷ time of increase
- Used to calculate monthly, quarterly, yearly increase in health service delivery. Example: increase in # of new clients, commodities distributed
- Example: Condom distribution in Jan. = 200; as of June = 1,100. What is the rate of increase?
- $1,100 - 200 = 900 / 6 = 150$  condoms per mo



**USAID**  
FROM THE AMERICAN PEOPLE

  
**MEASURE**  
Evaluation



# Calculating rate of increase

In Q1, there were 50 new FP users, and in Q2, there were 75. What was the rate of increase from Q1 to Q2?

$$75 - 50 = 25 / 3 = 8.33 = 8 \text{ new clients per month}$$

# Central tendency

Measures of the location of the middle or the center of a distribution of data

- Mean
- Median



**USAID**  
FROM THE AMERICAN PEOPLE



# Mean

- The average of your dataset
- The value obtained by dividing the sum of a set of quantities by the number of quantities in the set.
- Example:  $(22+18+30+19+37+33) = 159 \div 6 = 26.5$
- *The mean is sensitive to extreme values*

# Calculating the mean

- Average number of clients counseled per month

- January: 30

- February: 45

- March: 38

- April: 41

- May: 37

- June: 40

$$(30+45+38+41+37+40) = 231 \div 6 = 38.5$$

Mean or average = 38.5



**USAID**  
FROM THE AMERICAN PEOPLE



**MEASURE**  
Evaluation

# Median

- The middle of a distribution (when numbers are in order: half of the numbers are above the median and half are below the median)
- The median is not as sensitive to extreme values as the mean
- Odd number of numbers, median = the middle number.
  - Median of 2, 4, 7 = 4
- Even number of numbers, median = mean of the two middle numbers.
  - 2, 4, 7, 12 is  $(4+7) / 2 = 5.5$



**USAID**  
FROM THE AMERICAN PEOPLE



**MEASURE**  
Evaluation

# Calculating the median

- Client 1 – 2
- Client 2 – 134
- Client 3 – 67
- Client 4 – 10
- Client 5 – 221
- = 67
- =  $67+134 = 201/2 = 100.5$

# Use the mean or median?

	CD4 count
Client 1	9
Client 2	11
Client 3	100
Client 4	95
Client 5	92
Client 6	206
Client 7	104
Client 8	100
Client 9	101
Client 10	92



**USAID**  
FROM THE AMERICAN PEOPLE

  
**MEASURE**  
Evaluation

# Key messages

- Purpose of analysis is to provide answers to programmatic questions
- Descriptive analyses describe the sample/target population.
- Descriptive analyses do **not** define causality – tell you *what* not *why*



**USAID**  
FROM THE AMERICAN PEOPLE

