STRENGTHENING ANALYSIS AND USE OF ROUTINE FACILITY DATA FOR MATERNAL, NEWBORN, CHILD AND ADOLESCENT HEALTH

Principles and approaches for analysis, visualization and interpretation of routine health facility data for MNCAH

World Health Organization

Objectives of the session

- Review key health data terms and concepts
- Introduce common principles and approaches to analysis of routine health facility data
- Understand the basic concepts of data visualization
- Explore tips for data visualization
- Introduce data interpretation

From data to information

What do we mean by analysis?

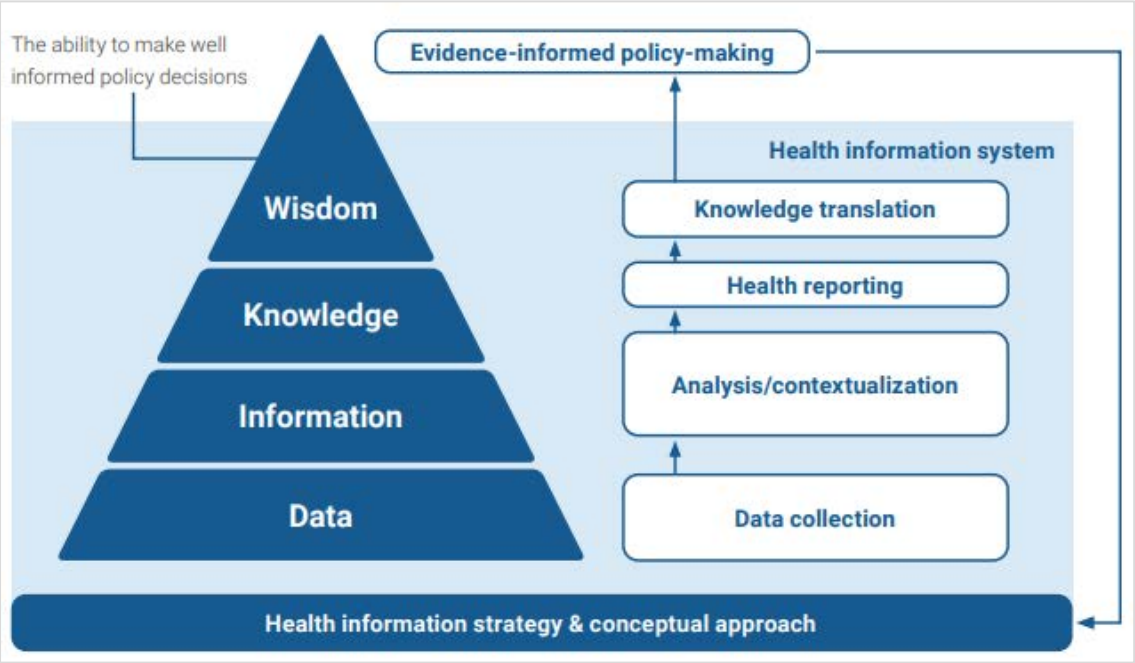
Data refers to raw, unprocessed numbers, measurements, or text

Analysis

Information refers to data that are processed, organized, structured, or presented in a specific context.

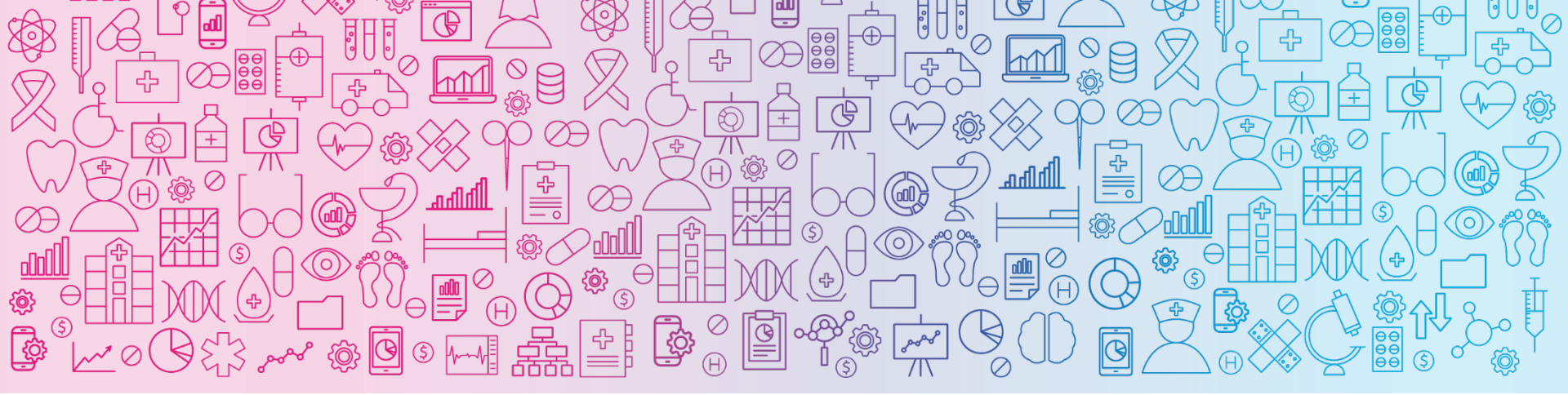
Analysis…

- …does not only imply using statistical software or performing complex calculations.
- …is the process of transforming data into useful information.
- …is examining data in the context of questions that you need answered.



From WHO EURO Support tool to strengthen health information systems: Guidance for health information system assessment and strategy development

Adapted from MEASURE Evaluation: Introduction to Basic Data Analysis and Interpretation for Health Programs: A Training Tool Kit



KEY HEALTH DATA TERMS AND CONCEPTS

World Health Organization

Definitions of key public health measurement concepts

Coverage Measure of the extent to which the services provided cover the potential need for these services in a population

Expressed as a percentage in which:

- the numerator is the number of service units provided, multiplied by 100, and
- the denominator is the target population in need of the services

Correlation When two factors (or variables) are related, but one does not necessarily cause the other

Trend Pattern of changes (i.e. in a health condition, outcome, intervention coverage) over a period of time

Association Statistical relationship between two or more events, characteristics, or other variables

Causation A factor (characteristic, behavior, event, etc.) that directly influences the occurrence of disease. A reduction of the factor in the population should lead to a reduction in the occurrence of disease

Sources:
WHO: Toolkit for analysis and use of routine health facility data- General Principles
US CDC: https://www.cdc.gov/csels/dsepd/ss1978/lesson1/section11.html

Statistical terms and concepts: Ratio

Term	Definition	Calculation/Notes
Ratio	Comparison of values, showing their size in relation to each other	• Calculated by dividing the first value by the second (a/b) • Can be written in various ways, such as: a to b, a per b, a:b • The numerator is not contained in the denominator

EXAMPLE

In district X, there are 600 nurses and 200 clinics. What is the ratio of nurses to clinics?

- 600/200 = 3 nurses per clinic
  - A ratio of 3:1 or 3 nurses to 1 clinic

Sources:
WHO: Toolkit for analysis and use of routine health facility data- General Principles
MEASURE Evaluation: Introduction to Basic Data Analysis and Interpretation for Health Programs: A Training Tool Kit

Statistical terms and concepts: Proportion

World Health Organization

Term	Definition	Calculation/Notes
Proportion	Number of events or cases that occur in a defined population, expressed as a fraction, decimal or percentage	• Used to compare part of the whole • All events or persons in the numerator are also included in the denominator

EXAMPLE

If a health facility sees 20 female children and 30 male children in one day, what is the proportion of female children seen that day?

- 20 females + 30 males = 50 children
- 20 female children / 50 total children
- Reduced: 2/5 of children seen that day were females

Note on percentages

- A percentage is a way to express proportion (proportion multiplied by 100).
- From the example above, where 2/5 children seen in the health facility that day were female, 40% of children seen were females (2/5 = 0.4 x 100 = 40%)

Sources:
WHO: Toolkit for analysis and use of routine health facility data- General Principles
MEASURE Evaluation: Introduction to Basic Data Analysis and Interpretation for Health Programs: A Training Tool Kit

Statistical terms and concepts: Rate

Term	Definition	Calculation/Notes
Rate	The frequency with which an event or case occurs in a defined population over a specified period of time	• Often expressed as events per 1000 population per year – Some rates are expressed per 10 000 or per 100 000 population • All events or persons in the numerator are also included in the denominator • Numerator and denominator must be from the same period

EXAMPLE

Infant mortality rate: # of deaths / population at risk in the same time period x 1000

- 75 infants died out of 4000 infants born that year
- 75/4000 = 0.0187 = 18.7
- 19 infants died per 1000 live births

Sources:
WHO: Toolkit for analysis and use of routine health facility data- General Principles
MEASURE Evaluation: Introduction to Basic Data Analysis and Interpretation for Health Programs: A Training Tool Kit

Review understanding of statistical terms and concepts

For each of the following examples, identify which is a **count, ratio, proportion or rate**

- 31 hospitalised cases of injuries among adolescents per 100 000 per year deaths **--> Rate**
- 512,020 live births in 2022 **--> Count**
- 52% of newborns born in a health facility are put to the breast within one hour of birth **--> Proportion**
- 4 nurses per 1 000 patients **--> Ratio**

Health data terms: measurements of central tendency

Term	Definition	Notes and examples
Mean	• Average of the dataset • Sum of all the values in a set divided by the number of values in the set	• Sensitive to extreme values: An extreme value in the set can result in a very high and unrepresentative mean – Example: (22+18+30+19+37+33) = 159 ÷ 6 = 26.5
Median	When values in a set are ranked from smallest to largest, the median is the value in the middle of the list	• Not as sensitive to extreme values • When there is an odd number of values, the median is the middle value – Example: Median of 2,7,4 = 4 • When there is an even number of values, the median is the average of the two middle values – Example: Median of 2,7,4,12 --> (4+7) /2 = 5.5
Mode	The value that occurs the most often in a set of data; determined by taking the number of times each value occurs and identifying the most common	• If no number is repeated in a list, there is no mode • Also a distribution can have more than one mode, e.g. bi-modal • Not commonly used in statistical analysis but may be used to described, for example, the day of the week when most outpatient consultations occur

Sources:
WHO: Toolkit for analysis and use of routine health facility data- General Principles
MEASURE Evaluation: Introduction to Basic Data Analysis and Interpretation for Health Programs: A Training Tool Kit

Review calculating the mean and the median

Month	Number of outpatient visits to a facility by adolescents
January	121
February	108
March	109
April	320
May	132
June	125

MEAN

- Add all observed values: (121 + 108 + 109 + 320 + 132 + 125) = 915
- Divide the sum by the number of observations = 915/6

Mean = 152.5

MEDIAN

- Arrange all observed values from lowest to highest: 108 109 121 125 132 320
- Find the middle position of the distribution. Since there is an even number of observations (6), use the average of the two middle values (121 + 125)/2

Median = 123

PLANNING FOR DATA ANALYSIS

World Health Organization

Step 1: Understand why you want to analyse the data

- Make sure you understand the reason why analysis is required
- Know the target audience and align
- Make a data analysis plan

How do we make sense of all of this data?!

Adapted content from original slide from:

Example: Analysis to answer programmatic questions

Analysis is looking at the data in light of the questions you need to answer.

How would you analyze data to determine: ***"Is the MNCAH programme meeting its objectives?"***

- **Question:** Is the MNCAH programme meeting its objectives?
- **Analysis**: Compare current MNCAH programme performance against programme targets to learn how far you are from achieving the targets
- **Interpretation**: Why have you achieved or not achieved a target, and what does this mean for the MNCAH programme?

→ **Answering the question may require more information**

Step 2: Review available data and indicators

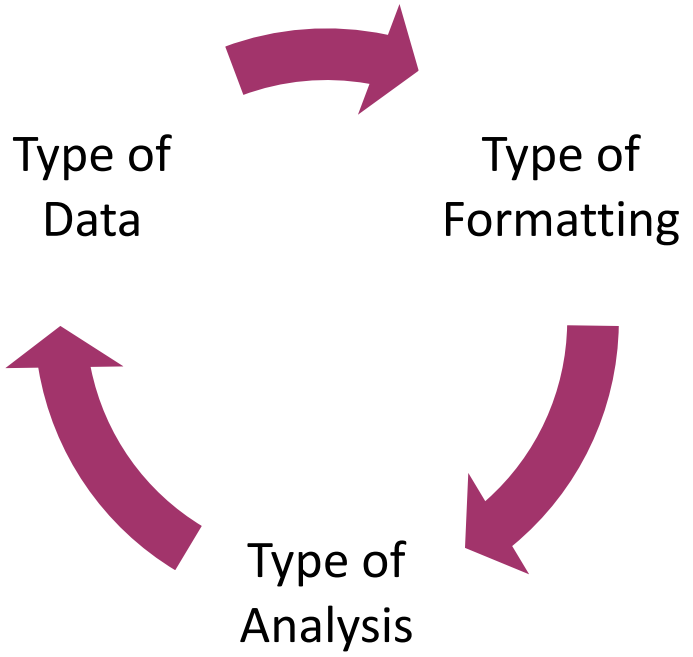
It is critical to identify and understand the indicators of interest for analysis

- Identify/review MNCAH indicators
  - Purpose of data analysis is critical in indicator identification and review
- Make sure indicators are understood
  - What is being measured? Which numerators and denominators used to calculate the indicator? What are their data sources?
- Identify targets and/or baseline values, where they exist
  - For comparison and assessment of progress
- Identify the denominator data
  - Estimate the denominator where necessary

Adapted content from original slide from:

Step 3: Prepare data and plan for analysis

- Do not just dive in!
  - Do proper data preparation and make sure you understand the data (e.g. numerator and denominator, data sources, quality issues)
- Download, organize and consolidate data
  - What type and in what format the data come from can determine the type of analysis that is feasible
- Assess data for the common data quality problems
  - Do data quality adjustment where necessary
    - Document and note any adjustments made to data and make sure to keep the original unadjusted data



Adapted content from original slide from: World Health Organization Regional Office for Africa

World Health Organization

Understand the types of data you will analyse

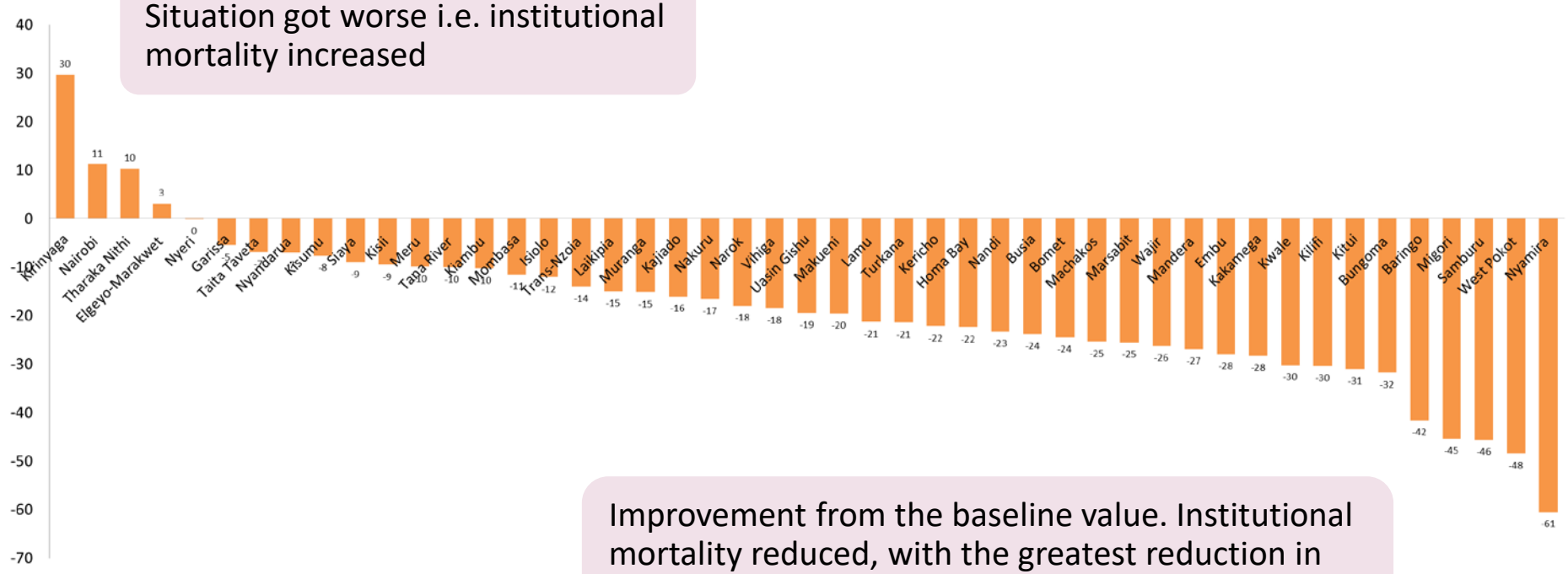
Most data reported via routine health management information systems are quantitative

- Counts/frequencies/tallies
  - E.g. number of children seeking care in a facility for acute respiratory infection
- Calculated indicators: percentages, rates, ratios
  - E.g. Proportion of newborns receiving postnatal care within 2 days in facilities

Type and level of analysis

- Choose a level of analysis that is appropriate
- The type of analysis you choose is often dictated by the type of data you have (i.e. categorical, continuous, etc.)

Example: Assessment of progress from baseline value



Situation got worse i.e. institutional mortality increased

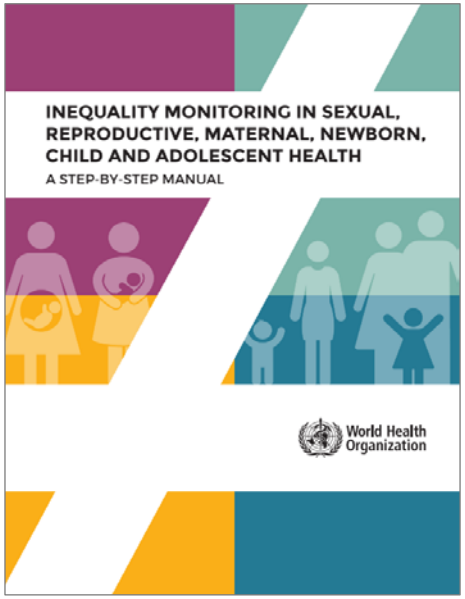
Improvement from the baseline value. Institutional mortality reduced, with the greatest reduction in Nyamira county followed by West Pokot

Slide from:

World Health Organization
REGIONAL OFFICE FOR Africa

Equity analysis

- Equity analysis refers to examination of performance of an indicator by population groups such as place of residence, wealth status, education level, age, sex, etc.
- The aim is to identify populations that are disadvantaged or underserved to guide efforts and resources to address inequalities between them through policies and programmes
- Equity analysis involves:
  - Disaggregation of results by population groups
  - Examination of inequality gaps between the population groups
  - Examination of trends in inequality over time and between the groups
- Equity analysis in analytical reviews could answer the following questions:
  - How does the performance of an indicator differ by population groups?
  - How wide is the inequality gap between the groups?
  - Are there populations that consistently lag behind?
- Note that ability to analyse data through an equity lens is contingent on which equity variables are incorporated into the data collection tools

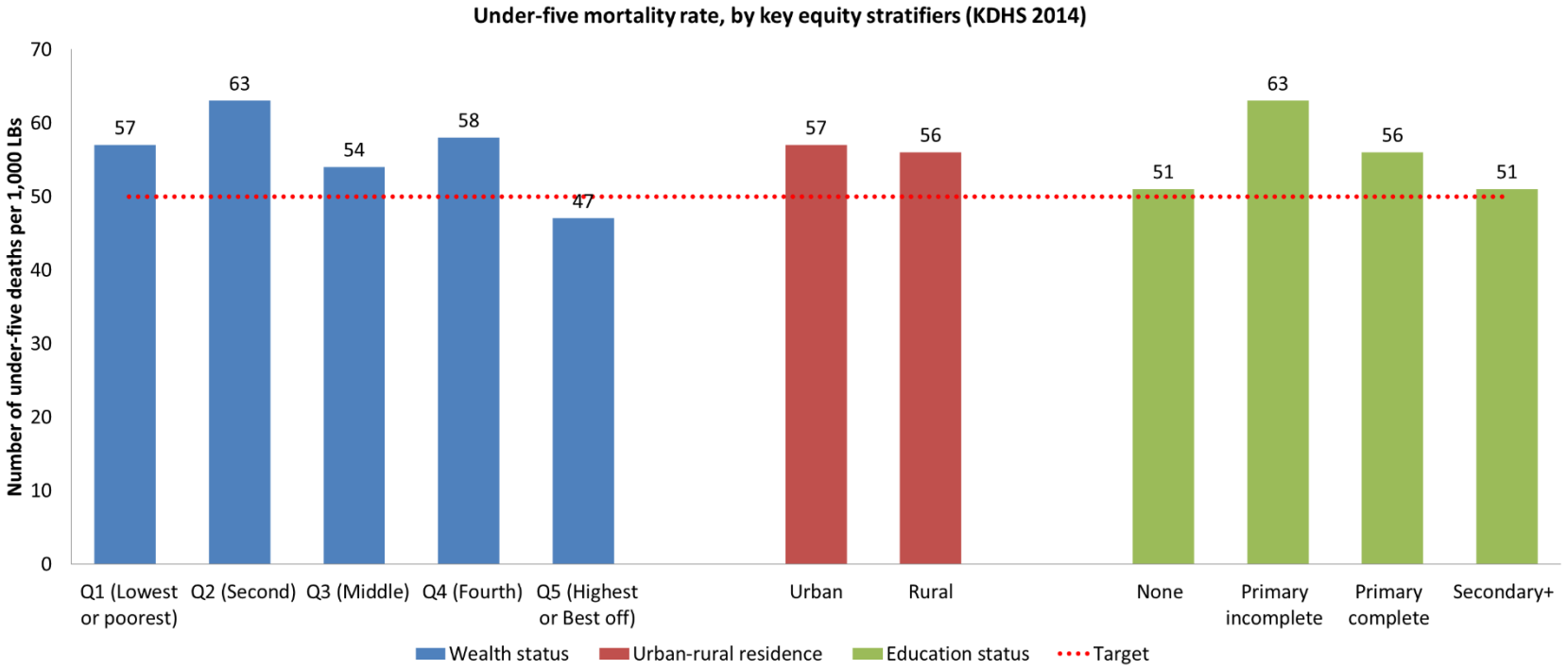


WHO (2022) : https://www.who.int/publications/i/item/9789240042438

Adapted content from original slide from: World Health Organization REGIONAL OFFICE FOR Africa

Examples of equity analyses (1/2)

Disaggregation of results by wealth quintile, residence type, and education level



Key interpretations

- No clear pattern by wealth status but general trend suggests that under-5 mortality rate (U5MR) reduces with improvement in wealth status
- U5MR in urban and rural areas almost the same
- No clear pattern by education status but the general trend suggests decline in U5MR with increase in education level

Slide from:

Examples of equity analyses (2/2)

Disaggregation of data by province, wealth quintile, residence type, and education level over time

Under-five mortality rate

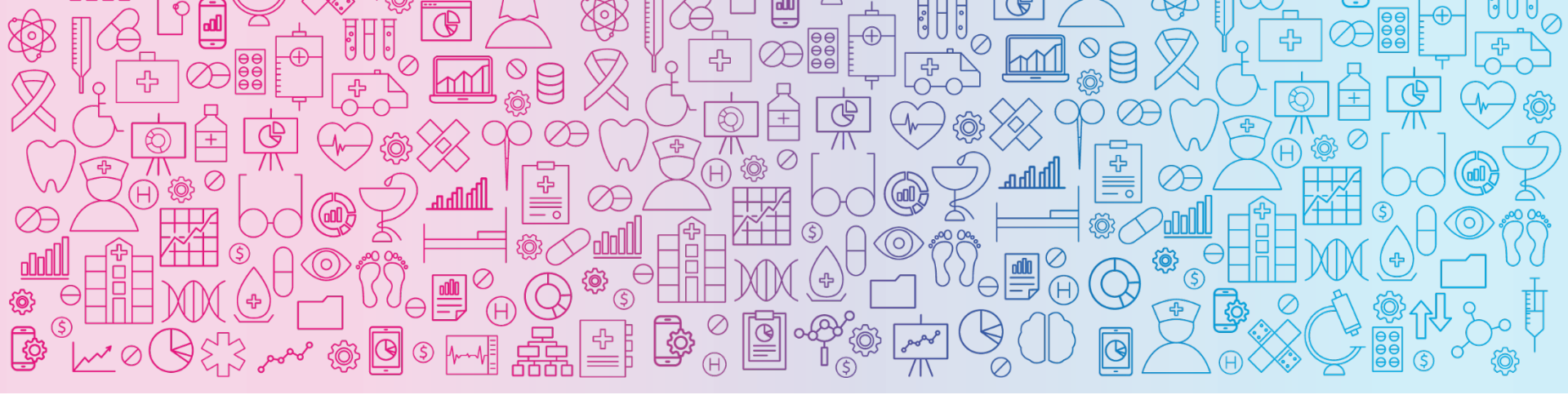
	KDHS 2023	KDHS 2008-2009	KDHS 2014
Province			
Nyanza	206	149	82
Nairobi	95	64	72
Western	144	121	64
Coast	116	87	57
Rift valley	77	59	45
Eastern	84	52	45
North Eastern	163	80	44
Central	54	51	42
Vealth status			
Q1 (Lowest or poorest)	149	34	57
Q2 (Second)	109	40	63
Q3 (Middle)	121	26	54
Q4 (Fourth)	77	12	58
Q5 (Highest or Best off)	91	13	47
Urban-rural residence			
Urban	93	74	57
Rural	117	86	56
Education status			
None	127	86	51
Primary incomplete	145	112	63
Primary complete	98	68	56
Secondary+	63	59	51

Key interpretation

Under-five mortality rate is consistently higher in:

- Nyanza province than elsewhere; followed by Western province
- The first and second wealth quintiles than in others
- Rural than in urban areas
- Among children whose mothers had incomplete primary education compared to other education levels

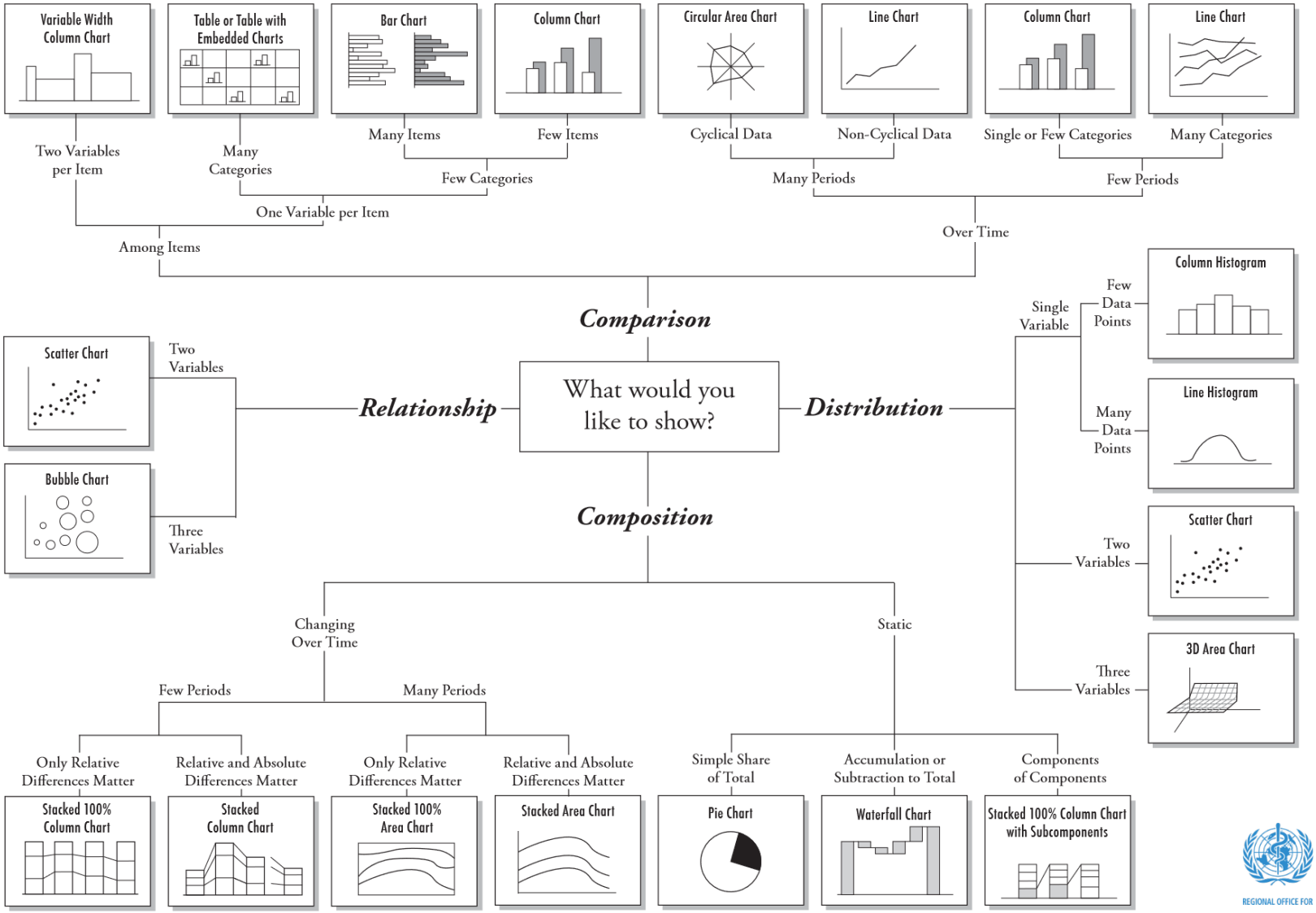
Slide from:



DATA VISUALIZATION AND INTERPRETATION

World Health Organization

Which visualization best enables clear interpretation of your analysis?



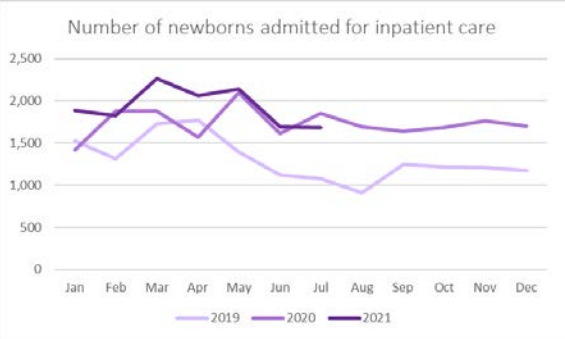
Slide from:

World Health Organization
REGIONAL OFFICE FOR Africa

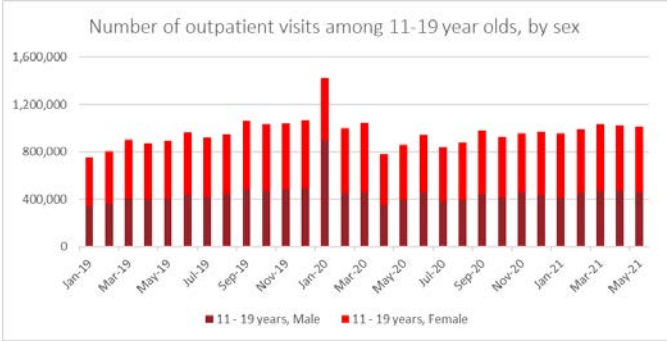
What are common types of data visualizations and when should they be used?

Bar charts compare categories of data

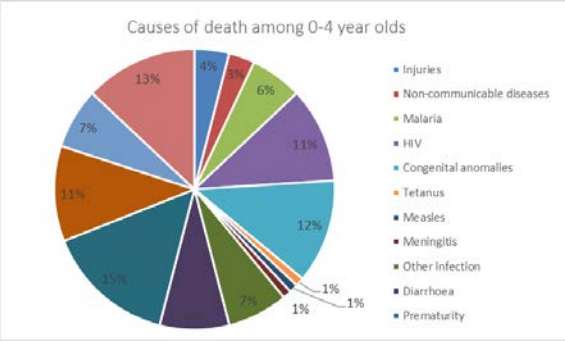
Line graphs display trends over time



Stacked bar charts represent components of a whole and compare wholes (or multiple values)



Pie charts show percentages or proportional share of a total (100%)



Data can also be summarized in a **table**

Percentage of deliveries by caesarean section												
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2019	11%	12%	13%	12%	12%	12%	12%	11%	11%	11%	14%	12%
2020	11%	11%	11%	13%	12%	12%	14%	11%	11%	12%	11%	11%
2021	11%	11%	12%	11%	10%	11%	11%					

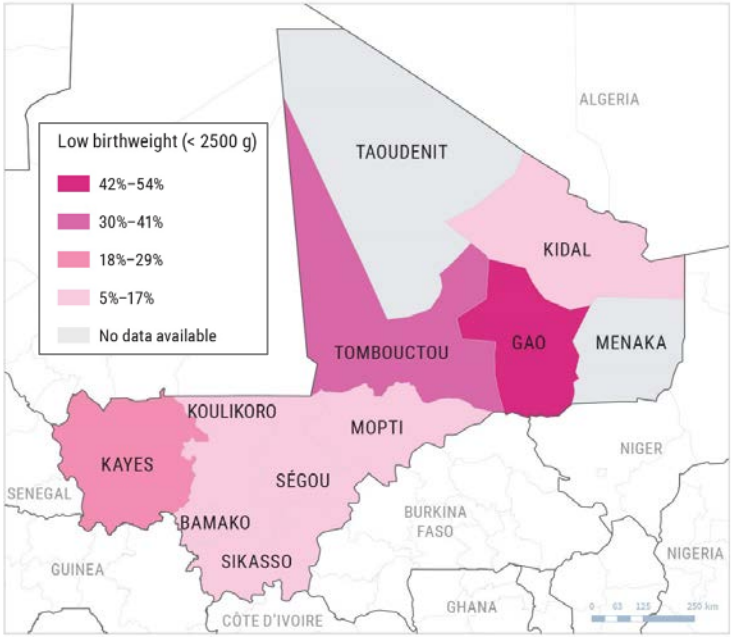
Reference for definitions: MEASURE Evaluation
Introduction to Basic Data Analysis and Interpretation for Health Programs: A Training Tool Kit

Thematic maps

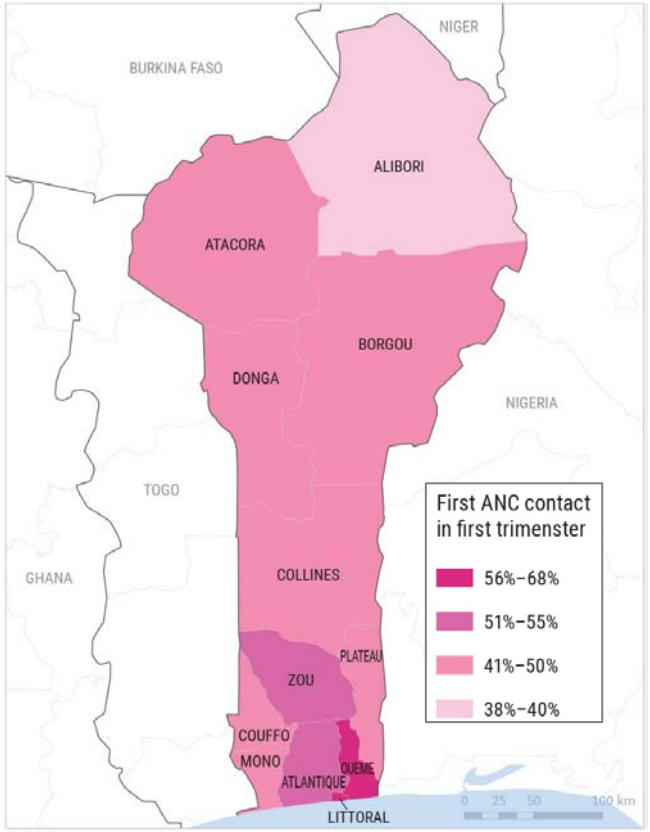
In these examples, **maps** display performance of key indicators across various subnational areas or facilities, allowing for easy identification of differences in performance of indicators between areas

- Note, only a single time point may be displayed (i.e. for a year, quarter, month, etc.)

Proportion of live births that weigh less than 2500 g, by region, Mali, 2021 (*Mock data*)



Annual proportion of antenatal clients with 1st antenatal care contact in the first trimester, by region, Benin, 2021 (*Mock data*)

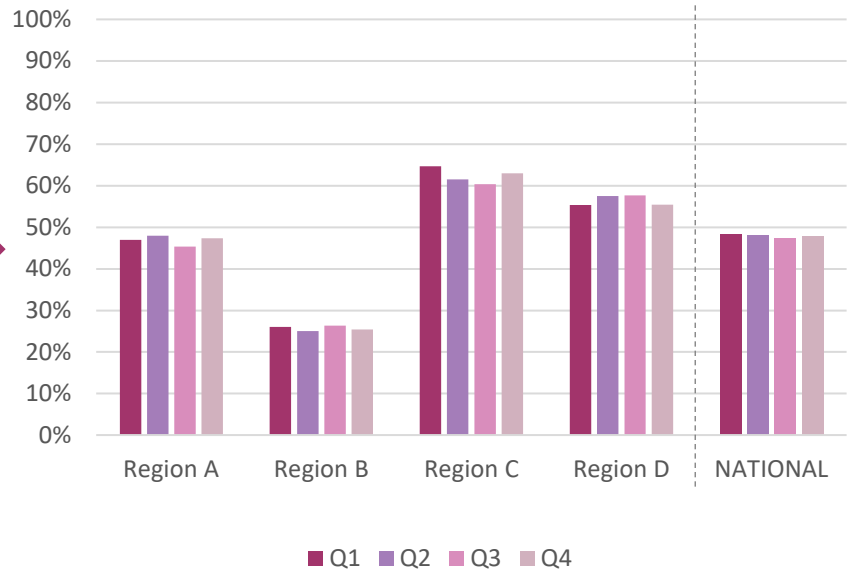


What would you like to show?

Proportion of newborns breastfed within first hour of birth

	Region A	Region B	Region C	Region D	NATIONAL
Q1	47%	26%	65%	55%	**48%**
Q2	48%	25%	62%	58%	**48%**
Q3	45%	26%	60%	58%	**47%**
Q4	47%	25%	63%	55%	**48%**

Proportion of newborns born in a facility put to the breast within one hour of birth, quarterly, by region
(*Mock data*)



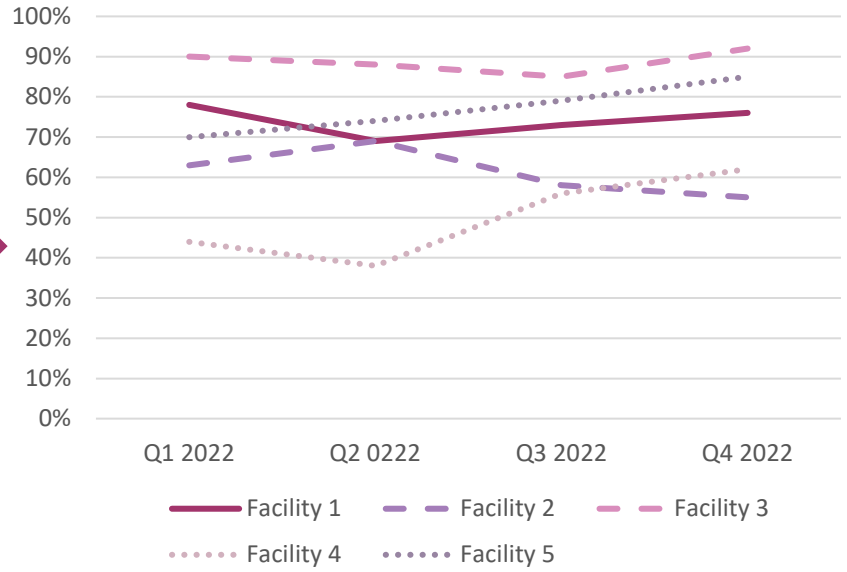
In this example, a **bar chart** is used to show a comparison of early initiation of breastfeeding over time across four regions (and the national average)

What would you like to show?

Treatment for cases of childhood diarrhoea in facilities over time

	Q1 2022	Q2 2022	Q3 2022	Q4 2022
Facility 1	78%	69%	73%	76%
Facility 2	63%	69%	58%	55%
Facility 3	90%	88%	85%	92%
Facility 4	44%	38%	56%	62%
Facility 5	70%	74%	79%	85%

Proportion of cases of diarrhoea among children 0-9 years treated in facilities with oral rehydration salts and Zinc, quarterly by facility (*Mock data*)



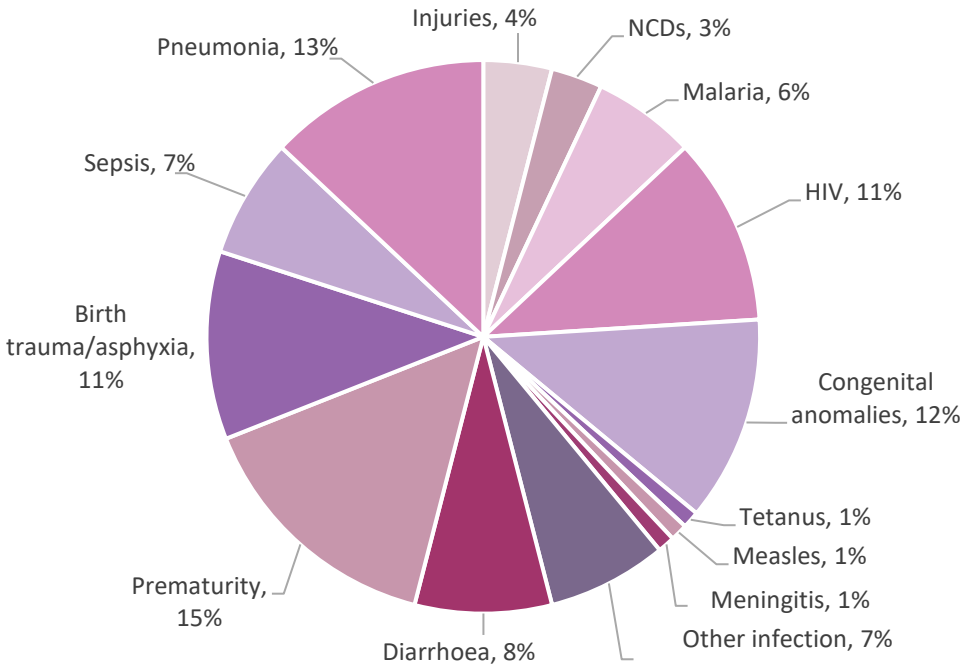
In this example, a **line graph** is used to show trends in treatment of childhood case of diarrhoea over time by facility

What would you like to show?

Causes of death for children under-five

Cause of death	%
Injuries	4%
NCDs	3%
Malaria	6%
HIV	11%
Congenital anomalies	12%
Tetanus	1%
Measles	1%
Meningitis	1%
Other infection	7%
Diarrhoea	8%
Prematurity	15%
Birth trauma/asphyxia	11%
Sepsis	7%
Pneumonia	13%
Total (all causes)	100%

Distribution of deaths in children 0-4 years by cause, 2022
(*Mock data*)



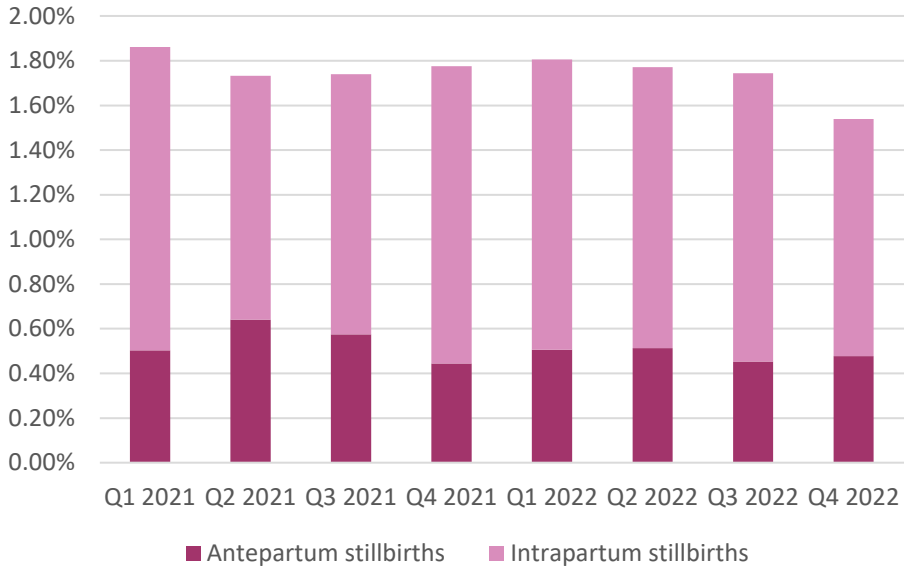
In this example, a **pie chart** is used to show the distribution of causes of death among children 0-4 years of age

What would you like to show?

Stillbirths in facilities as a proportion of all births in facilities

	Antepartum stillbirths	Intrapartum stillbirths
Q1 2021	0.50%	1.36%
Q2 2021	0.64%	1.09%
Q3 2021	0.57%	1.17%
Q4 2021	0.44%	1.33%
Q1 2022	0.51%	1.30%
Q2 2022	0.51%	1.26%
Q3 2022	0.45%	1.29%
Q4 2022	0.48%	1.06%

Stillbirths in facilities as a proportion of all births in facilities, quarterly 2021-2022 (*Mock data*)



In this example, a **stacked bar chart** shows the proportion of antepartum and intrapartum stillbirths among all stillbirths in facilities over a two year period

Displaying several indicators together

Presenting multiple indicators together in a single visual can offer a broader picture of MNCAH or highlight the performance of one indicator or state/region/district (or disaggregation (e.g. sex, age group)) in comparison to others.

- **When more than one indicator is presented on the same chart as a proportion, all indicators must use the same denominator.**
- Dashboards or scorecards combining indicators, independent of scales are not subject to this rule to the same extent, but should include clearly labeled, related indicators and consistent time periods.

Text from: WHO Analysing and using routine data to monitor the effects of COVID-19 on essential health services - Practical guide for national and subnational decision-makers

Analysis of a core set of indicators – displaying a key set of indicators in one visual

World Health Organization

In the figure here, four indicators are presented to show the change in several districts from a common baseline. These show a decline in ANC, facility births, measles vaccination and treatment for children with malaria, thereby providing a snapshot of the performance of key MNCAH tracer indicators.

Changes in RMNCH indicators from baseline to current reporting period

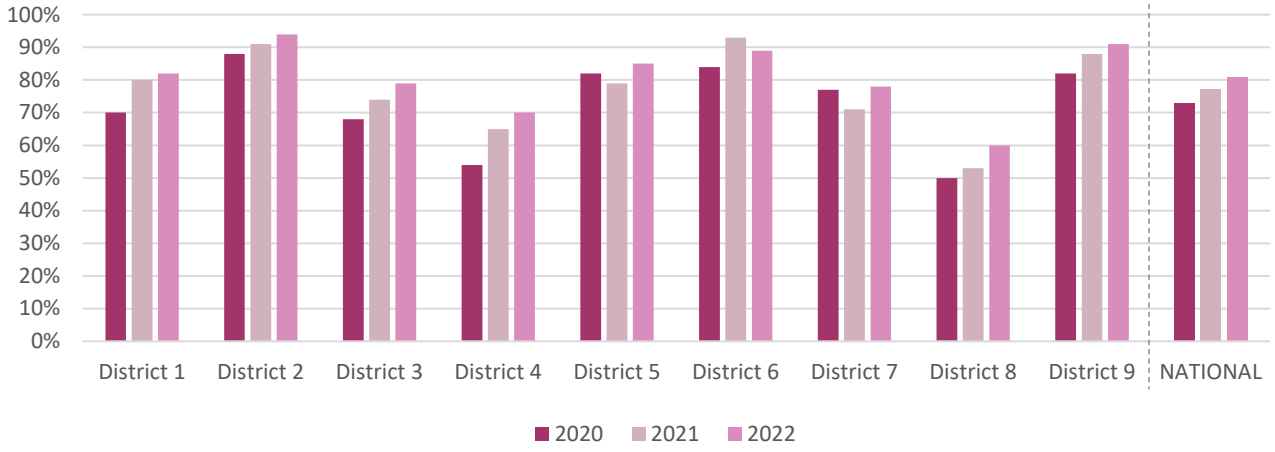
Region	District	ANC sevices	Facility births	MCV1	Treatment for children with malaria
Region 1	District 1	−20%	−15%	−22%	−25%
	District 2	−40%	−50%	−30%	−48%
	District 3	−30%	−25%	−35%	−50%
	District 4	−22%	−20%	−25%	−40%
Region 2	District 5	−10%	−11%	−8%	−15%
	District 6	−7%	−5%	7%	−12%
	District 7	−9%	−9%	−7%	−15%
	District 8	−11%	−8%	−12%	−10%
Region 3	District 9	−31%	−27%	−15%	−47%
	District 10	−25%	−30%	−28%	−45%
	District 11	−30%	−35%	−32%	−47%
	District 12	−35%	−40%	−30%	−38%
National		−25%	−28%	−22%	−33%

ANC: antenatal care; MCV1: measles-containing vaccine first dose.

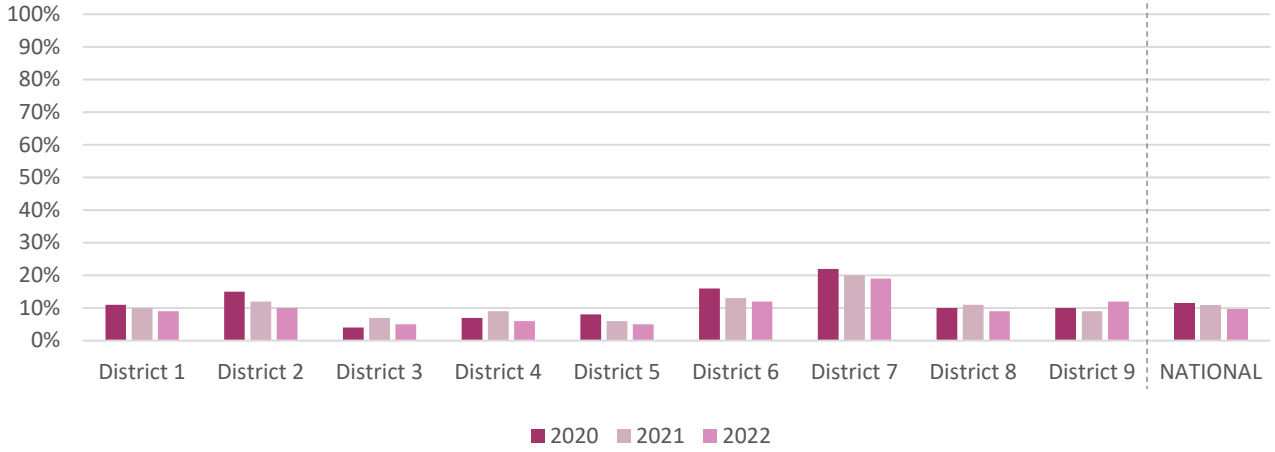
Text and image from: WHO Analysing and using routine data to monitor the effects of COVID-19 on essential health services - Practical guide for national and subnational decision-makers

Analysis of related indicators

Proportion of births in health facilities, by district, annual average for 2022 (*Mock data*)



Proportion of deliveries by caesarean section in health facilities, by district, annual average for 2022 (*Mock data*)



As these two indicators use different denominators, they are not displayed on the same chart, however, it is helpful to review proportion of deliveries by caesarean section with the related indicator, proportion of births in facilities, to provide context for interpretation.

Recommendations for improving accurate interpretation of data visualizations

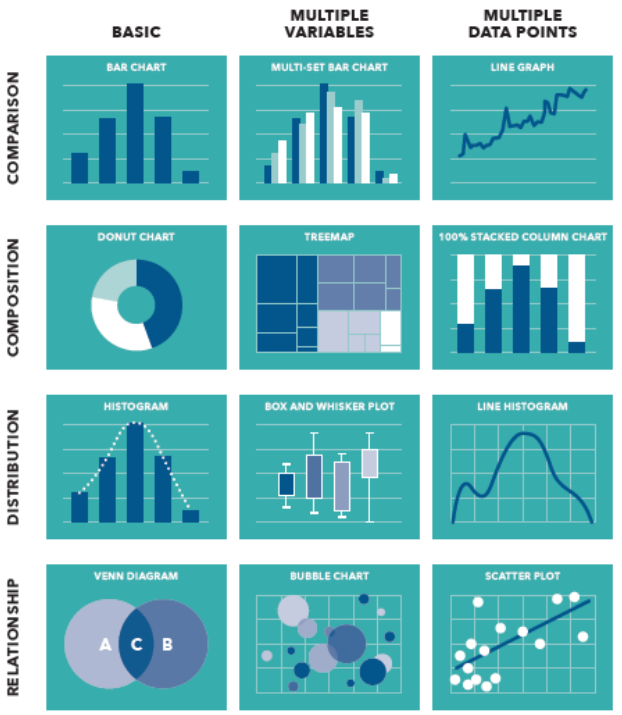
World Health Organization

A few recommendations

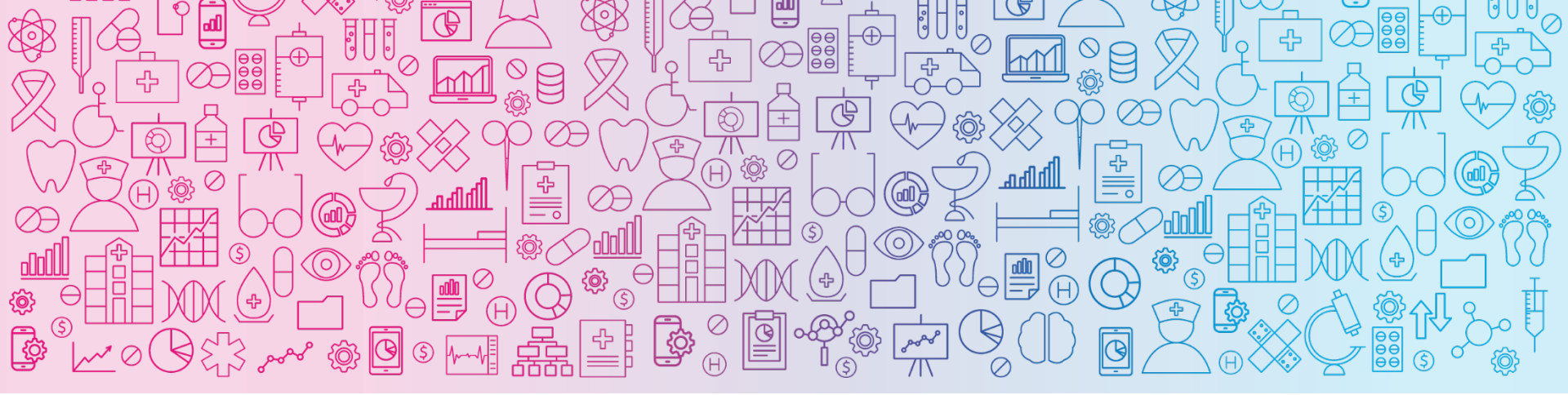
- Make sure your visualization has a title
- Label the components (units, variables, etc.) of your graphic
- Record the data source with date
- Provide brief narrative explanation of the visualization

Tips to better enable visualizations for clear communication

- Test several options and compare
  - Best graph and disaggregation to see patterns related to question
  - How can the data visualization be most easily understood by diverse audiences?
- Customize visualizations for specific audience
  - What best conveys your information for a specific person/audience (e.g. programme manager, health worker) or purpose (e.g. advocacy materials, technical report)?
- Annotate context information to aid interpretation
  - Circles, arrows, text, benchmark lines



Sources for adapted content and visualization:
WHO-EURO Effective Communication of Immunization Data
MEASURE Evaluation Introduction to Basic Data Analysis and Interpretation for Health Programs: A Training Tool Kit



DATA INTERPRETATION

World Health Organization

Interpreting data

What do we mean by interpretation?

- Process of making sense of information
- Adds meaning to information by making connections and comparisons and exploring causes and consequences

Understanding the limitations of your data is critical for interpretation

- The type and source of the data can limit what the data can and cannot tell you
  - For example, the total count of antenatal contacts across all health facilities in a country does not imply population-level coverage of antenatal care
- Transparency about the quality and any potential bias of your data should be included in the interpretation
  - For example, changes in monthly numeric counts of antenatal contacts may be due to different numbers of facilities reporting each month, which should be considered when examining trends in the data over time

Adapted content from:
Source: MEASURE Evaluation Introduction to Basic Data Analysis and Interpretation for Health Programs: A Training Tool Kit

Interpreting data

World Health Organization

Relevance of finding → Reasons for finding → Consider other data → Conduct further research

What are the data and information telling us? What do the findings mean for the MNCAH programme?

- Have we met our programme targets?
- Are levels and trends of MNCAH health service delivery indicators higher, lower or about the same as in previous reporting periods?

Why are we seeing these levels and trends for these MNCAH indicators?

- What are some possible reasons for changes in the performance of MNCAH indicators?
- Why may there have been a change in intervention coverage or in health outcomes?

What other information might be needed to understand the data?

- Are there other existing data sources that should be explored (e.g. administrative records (human resource or health financing data), health facility assessments)?
- Are there other people that should be consulted to help explain the data?

How can I answer any remaining questions about what the data are telling me?

- Do I need to gather new information?
- Are there additional analyses that can be done with existing data sources?
- Are there other actions I should take (i.e. supervisory visit, special survey, stakeholder meeting, etc.)?

Adapted content from:
Source: MEASURE Evaluation Introduction to Basic Data Analysis and Interpretation for Health Programs: A Training Tool Kit

What is the meaning of this visualization?

Proportion of newborns breastfed within first hour of birth

Key interpretation considerations

What do the data tell me?

The reported proportion of newborns born alive in facilities who were put to the breast within one hour of birth was consistent across all quarters of 2022 in each given region.

The reported proportion of newborns born alive in facilities who were put to the breast within one hour of birth in Region B was lower than in other regions and lower than the national average.

Proportion of newborns born in a facility put to the breast within one hour of birth, quarterly (2022), by region (*Mock data*)

What do they <u>not</u> tell me?

- Early initiation of breastfeeding at population level
- Trends in early initiation of breastfeeding in facilities previous years
- Reasons for differing proportions of intervention coverage in each region
- Reporting completeness

What is the meaning of this visualization?

Treatment for cases of childhood diarrhoea in facilities over time

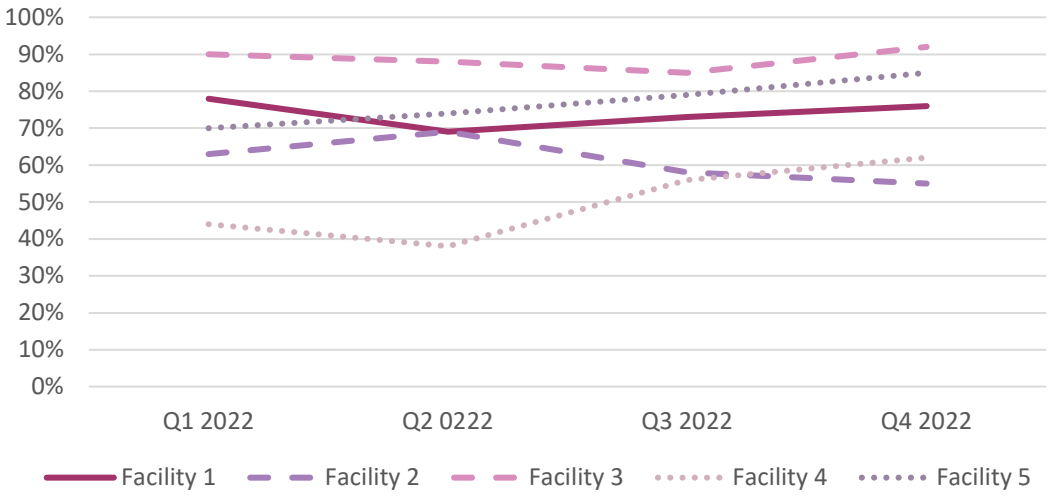
Key interpretation considerations

What do the data tell me?

The reported proportion of cases of diarrhoea among children 0-9 years treated with ORS and Zinc in five health facilities for each quarter of 2022.

In Facility 4, the proportion of childhood cases of diarrhoea treated with ORS and Zinc is lower than in other facilities for Q1-Q2 2022.

Proportion of cases of diarrhoea among children 0-9 years treated in facilities with oral rehydration salts and Zinc, quarterly by facility
(*Mock data*)



What do they <u>not</u> tell me?

- Prevalence of childhood diarrhoea in the population
- Treatment for diarrhoea outside of facilities
- Treatment of diarrhoea in or outside of facilities with other treatment types (e.g. just ORS, etc.)
- Reporting completeness
- Stock levels of ORS and/or Zinc

What is the meaning of this visualization?

Causes of death for children under-five

Key interpretation considerations

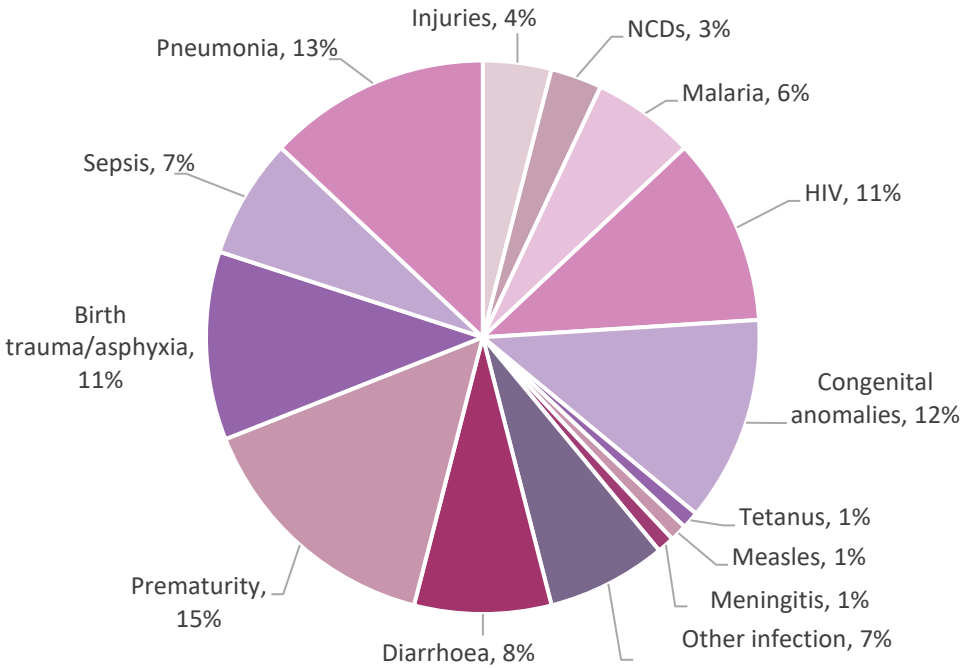
What do the data tell me?

The reported top 5 most common causes of death for children under-five in 2022 were: pre-maturity, pneumonia, congenital birth trauma/asphysxia, HIV.

What do they not tell me?

- Trends in causes of death from previous time periods
- The reasons for the distribution of causes of death

Distribution of deaths in children 0-4 years by cause, 2022
(*Mock data*)



What is the meaning of this visualization?

Stillbirths in facilities as a proportion of all births in facilities

Key interpretation considerations

What do the data tell me?

The reported proportion of stillbirths in facilities out of all births in facilities was under 2% for all quarters of 2021 and 2022.

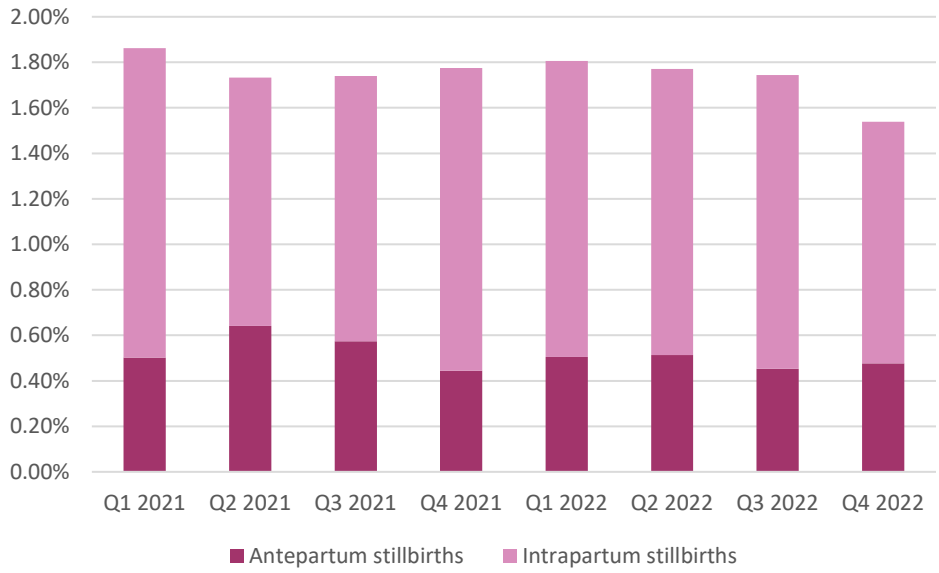
There was a slight decrease in the proportion of stillbirths in Q4 2022 in comparison to previous reporting periods.

Most stillbirths in facilities were intrapartum stillbirths.

What do they not tell me?

- Reporting completeness (i.e. for Q4 2022)
- Stillbirth rates outside of facilities
- Trends in stillbirths prior to 2021

Stillbirths in facilities as a proportion of all births in facilities, quarterly 2021-2022 (*Mock data*)



Considerations for data interpretation

World Health Organization

Make sure your data are standardized

- Check that the data share consistent denominators, units, and population bases, particularly when comparing data across different groups or time periods.

Consider the impact of data quality

- Review data quality metrics where available and include these in the interpretation. Be honest about limitations in the data.

Correlation does not mean causation

- Other factors may influence the relationship. Lower health outcomes in rural areas do not imply that being in a rural area directly caused those outcomes. Factors like population density, access to health services, health workforce density, etc. can also have an impact.

Consider the generalizability of your data to the entire target population

- Be cautious about drawing broad conclusions based on a limited or biased sample size. For example, RHIS data often only reflect services provided in (public sector) facilities, which may not represent the entire target population.
- Be clear on what your data can day and also what they cannot say.

Review data or information supporting various factors/outcomes

- Selectively choosing data (or data sources) to support a conclusion and ignoring contradictory evidence may lead to biased interpretation.

Exercise

- Complete exercises under **Analysis, visualization and interpretation of MNCAH data** in *Companion exercises to strengthen analysis and use of health facility data for MNCAH*.

- There are two parts to the exercise:
  - Part 1: Key health data terms and concepts
  - Part 2: Triangulation, analysis and interpretation of MNCAH data: case study
- Review responses in plenary.