NATIONAL NUTRITION INFORMATION SYSTEM

TECHNICAL NOTE

Designing Effective Data Visualizations

unicef

World Health Organization

Acknowledgements

The Technical Note on *Designing Effective Data Visualizations* is a component of the full guide on National Nutrition Information Systems (NNIS), a product of the WHO-UNICEF Technical Expert Advisory group on nutrition Monitoring (TEAM), which is supported by the Bill & Melinda Gates Foundation. The TEAM Working Group on Nutrition Information Systems was responsible for the development of this Technical Note, with direct support provided by the Data for Decisions in Nutrition (DataDENT) initiative (www.datadent.org) funded by the Bill & Melinda Gates Foundation.

TEAM acknowledges the contributions of Tricia Aung (DataDENT / Johns Hopkins Bloomberg School of Public Health) as lead writer as well as David Hales (Consultant), Chika Hayashi (UNICEF Headquarters), Rebecca Heidkamp (TEAM Working Group Lead, DataDENT / Johns Hopkins Bloomberg School of Public Health), Louise Mwirigi (UNICEF Headquarters) and Kuntal Saha (WHO Headquarters), who conceptualized and oversaw the production of this document. The Working Group acknowledges the contributions of Julia D'Aloisio (Editor) and Nona Reuter (Designer, UNICEF). The Working Group is also grateful to nutrition colleagues who reviewed the draft and shared specific experiences and insights.

Suggested Citation: N*ational Nutrition Information Systems; Technical Note: Designing Effective Data* Visualizations Geneva: World Health Organization and the United Nations Children's Fund (UNICEF), 2022.

- - - - -

The NNIS Fundamentals Series includes five modules:
Module 1. What is a national nutrition information system?
Module 2. How does a national nutrition information system support a country's nutrition programmes?
Module 3. What is needed to build a useful national nutrition information system?
Module 4. What are the main attributes of a national nutrition information system?
Module 5. What are the main types of data used in a national nutrition information system?

PDF versions of the five modules and different Technical Notes on specific topics can be downloaded from the following website: https://data.unicef.org/resources/nutrition-nnis-guides/

A four-module e-course on national nutrition information systems is available on the same website.

ISBN: 978-92-806-5376-2

November 2022

Please contact:
UNICEF
Data Analytics and Innovation
Division of Data, Analytics, Planning and Monitoring
3 United Nations Plaza
New York, NY 10017, USA

TECHNICAL NOTE

Designing Effective Data Visualizations

INTRODUCTION TO DATA VISUALIZATION

National nutrition information systems (NNIS) can be "data rich" but "information poor" without a thoughtful strategy on how to transform data into actionable evidence.[1] One strategy is data visualization, which involves the use of graphic displays of data to tell a story.

Data visualization is a strategy for communicating and making sense of data.[2] Data visualizations are most effective when they are fit to purpose and easily understood by their target audiences.[3] This technical note presents tips for selecting and producing effective data visualizations for a wide range of purposes and audiences. It complements the NNIS Technical Note titled "The Power of Nutrition Dashboards, Profiles and Scorecards", which provides detail on assembling more specific types of performance monitoring and accountability outputs that use data visualization (dashboards, profiles, scorecards).

Data visualization should be a routine component of data analysis, given its ability to make data more understandable to diverse audiences. Research shows that different ways of visualizing data can influence comprehension and interpretation. Most people process visual information more rapidly than information presented as text. Different data visualization techniques can also enhance recall of information in the short-term.[4]

PURPOSE OF DATA VISUALIZATION ACROSS THE DATA VALUE CHAIN

Data visualization can be used for different purposes along the nutrition data value chain.

During *aggregation and processing,* data can be visualized to facilitate data cleaning and quality assessment. For example, visualization is an important step in data exploration to uncover initial insights on data characteristics and identify data gaps and outliers. During *analysis*, visualizations can be used to quickly assess patterns and distributions or to validate model assumptions. Data visualizations for aggregation, processing and analysis are generally used by people with high levels of data literacy; they are not used to communicate to diverse audiences.

The broader application of data visualization is to improve the *dissemination and use of information.* Data visualizations should be designed to communicate clear and compelling messages to the target audience. Ideally, representatives from the target audience are included in the design process. It is also important to consider issues such as the age, education levels, language fluency and technology access of the target audience.

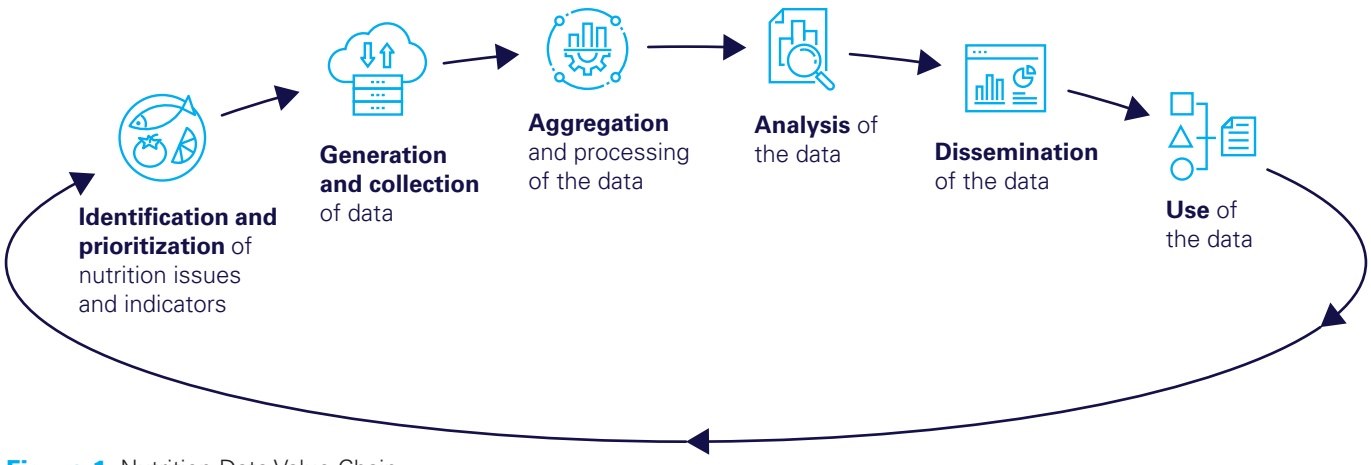


Figure 1. Nutrition Data Value Chain

Table 1. Planning for data visualization: Questions to consider

	Communication goals and key messages	Target audience	Type of output
Questions	What idea or insight do you want the audience to take away and remember from the data visualization?	Who is the data visualization trying to reach? How familiar are they with the data and concepts being communicated? How comfortable are they when it comes to interpreting figures, tables and other data visualizations?	What type of output will the data visualization be used for (e.g., printed brief, projected presentation, mobile-friendly website)? If considering interactive data visualizations, what is the quality of internet and technology among your target audience?
Examples of why this is relevant	• Clear and compelling visualizations are more effective at communicating data. • Formatting choices (e.g., colour, font size, title, and labels) can be leveraged for emphasis.	• Audiences with limited knowledge of the topic area or data types may benefit from annotations in the data visualization that support interpretation. • Individuals have variable comfort levels with reading figures and tables (e.g., some people will reject unfamiliar types of visualizations and/or require verbal explanations to orient them).	• Print quality can influence readability of figures in paper documents. • For a projected presentation, room size, screen size, projector quality and lighting in the room can all affect visibility. • Formatting requirements vary for desktop and mobile devices • Interactive data visualizations may not be appropriate for settings with limited internet connectivity.

Design decisions should be driven by: (1) communication goals and key messages; 2) data literacy and preferences of the target audience; and 3) the type of output the data visualization is being used to communicate (see Table 1).

For more guidance on developing communication goal and key messages, we recommend the WHO Strategic Framework for Effective Communications.

UNDERSTANDING YOUR AUDIENCE THROUGH PERSONAS

Developing personas can be a helpful – and fun – way to better understand your target audience and the ways that data visualization can be tailored to their needs, priorities and preferences.[5] Personas – a technique used in human-centred design – are fictional characters who represent a certain type of end user in your target audience. They are developed and refined based on conversations with actual people from the target audience.[6,7] (More information about personas can be found in the IDEO Open Design Kit.)

The primary objective of the persona is to capture key characteristics that may influence data visualization design choices. Figure 2 shows an example of a persona. The specific format of a persona can be adapted based on your context and priorities.

Data visualization design should account for the fact that this individual: (1) must be able to quickly put together figures with accessible software (e.g., PowerPoint and Excel); (2) must be able to easily explain graphs and would find annotations useful; (3) may require additional help if interpreting a figure that is not a bar or pie chart; and (4) prefers colours that account for colour-blindness.

TWO CRITICAL PRINCIPLES TO GUIDE DATA VISUALIZATION DESIGN

Visualizing ethically

How data are visualized will influence how the audience understands the information. You have an ethical obligation to visualize data accurately. Choosing to visualize only select data or manipulating data to make a point can be misleading – and more importantly, it can compromise users' trust in the data. All data visualizations should list data sources and note any assumptions or omissions made during design.

Figure 2. Persona example to develop a data visualization

<table>
<tr><td rowspan="2"> 

Name: Peter
Position: Policy adviser to the Ministry of Health (MOH)
Primary languages: English (official), Kiswahili (with colleagues)
Primary technology: Smartphone, Laptop
Dissemination preferences: Reports, briefs, PowerPoint
Highest degree: Medical Doctor</td><td rowspan="2">Goal: Visualize Vitamin A coverage data for a presentation to be delivered by the Minister</td><td>Internet connectivity: ●●●○</td></tr>
<tr><td>Technology reliability: ●○○○
Electricity reliability: ●●○○
Familiarity with nutrition data: ●●○○</td></tr>
<tr><td>Daily Responsibilities
• Expected to quickly collate data at the request of the Minister
• Writes speeches and drafts policy guidance for the Minister
• Liaises among Ministry of Health divisions and the Minister</td><td>Relationship with nutrition data:
• Very familiar with reports from Demographic and Health Surveys (DHS) and SMART surveys
• Can directly access and pull data from the District Health Information Software (DHIS2)
• Does not know specifics about how vitamin A data are collected in routine systems versus surveys
• Only uses bar graphs and pie charts in his work</td><td>Additional considerations:
• Both Peter and the Minister have limited nutrition expertise
• Minister wants a rapid briefing of information
• Peter owns an older laptop with Excel that periodically freezes
• Peter is affected by colour-blindness (i.e., he cannot differentiate red from green)</td></tr>
</table>

Avoid information overload

It is unrealistic to expect an audience to remember and act on all presented data. Research and experiences with DHIS2 suggest that individuals can experience information overload when presented with too much information at once.[8] Ultimately, the amount of data included in a single data visualization and the total number of visualizations used should be the minimum required to meet your communication objective.

CHOOSING THE RIGHT TYPE OF VISUALIZATIONS FOR QUANTITATIVE DATA

For quantitative data, visualizations often support two key objectives: (1) to compare numbers (e.g., across time, groups or regions); and (2) to highlight individual numbers. Figure 3 is a decision tree to help guide selection of data visualizations for these purposes. Table 2 details each data visualization type and offers considerations for selecting the most appropriate type based on data visualization research.

The decision tree focuses on a subset of data visualization types that:

- Can be created relatively easily in accessible or free software (e.g., Excel, RStudio, Tableau Public)
- Are more commonly used in global health and nutrition communications (e.g., for DHS and UNICEF reports)
- Are accessible to people with different levels of data literacy (e.g., do not include bubble charts and radar graphs, which can be difficult to interpret)

Research has found that people tend to prefer data visualization styles that are already familiar to them. This does not rule out the possibility of using unfamiliar data visualization types, but unfamiliar data visualization types may require annotation or other supports to guide interpretation.

A common mistake is to prioritize graphics that look interesting or are aesthetically novel. However, a "boring" bar chart may be the visual that resonates best with your audience.

This decision tree draws on resources (chart chooser cheat sheet, cards) developed by Dr. Stephanie Evergreen. Effective data visualization: The right chart for the right data. Sage Publications; 2019 Apr 3.

Figure 3. Data Visualization Decision Tree

This decision tree draws on resources (chart chooser cheat sheet, cards) developed by Stephanie Evergreen Evergreen SD. *Effective data visualization: The right chart for the right data.* Sage Publications; 2019 Apr 3.

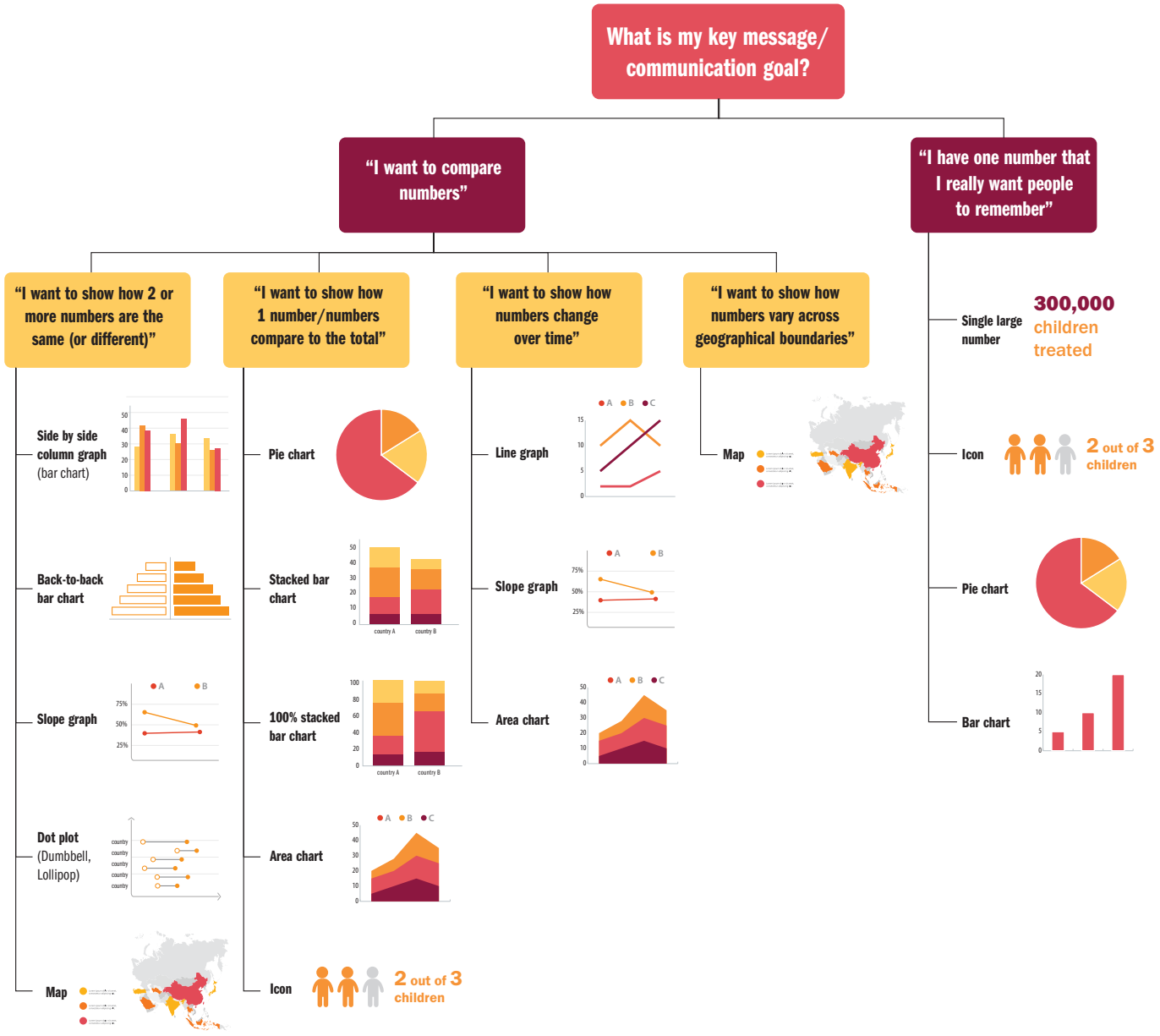


Table 2. Overview of data visualizations featured in the decision tree

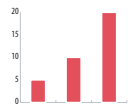
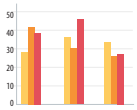
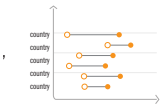
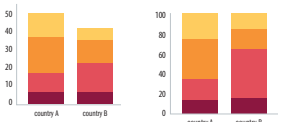
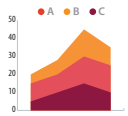
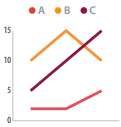
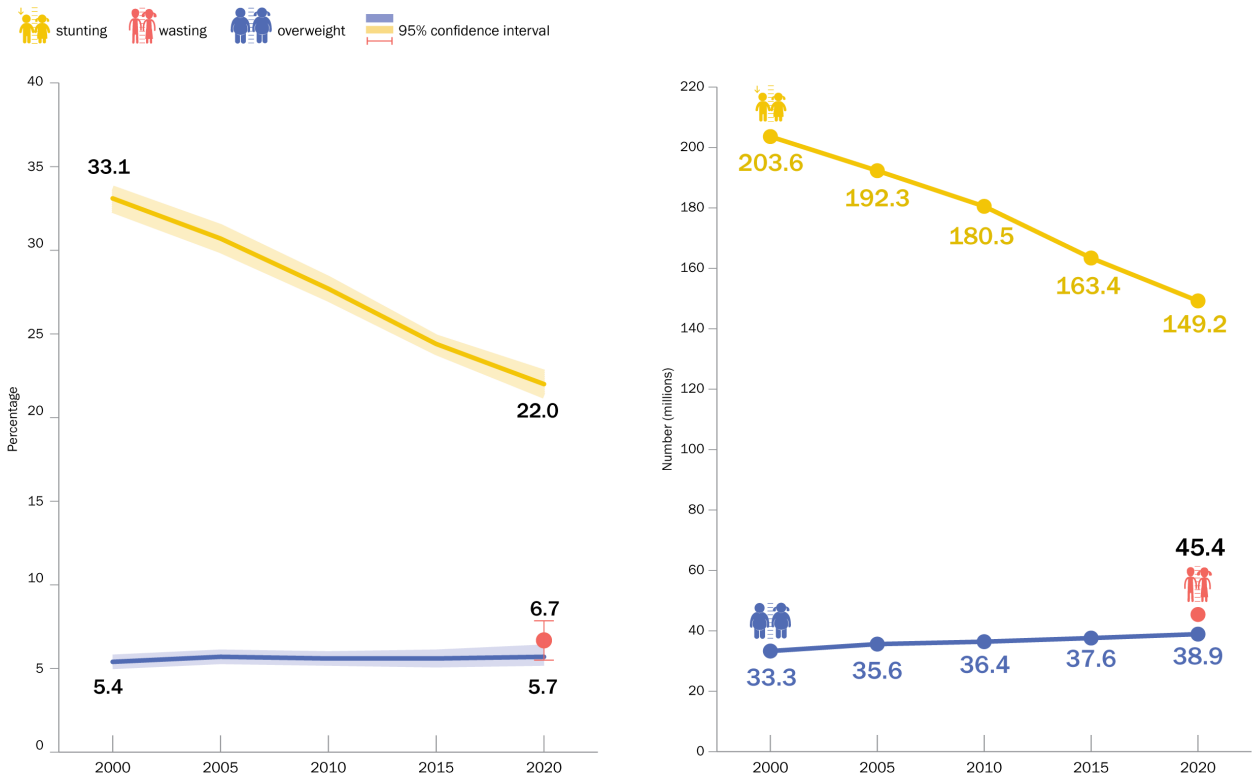
	**Bar charts** are used to compare categories of data. Avoid displaying too many columns, which can impair interpretation. Use a maximum of 3–5 groups and 1–2 columns per group. Bar charts with one column can be used to highlight a single number.
	**Back-to-back bar charts** are used to highlight the distribution of values across two categories. (e.g., a population pyramid). Back-to-back bar charts can be quickly interpreted due to the brain's approach towards looking for symmetry.
	**Slope graphs** are used to compare two groups. Slope graphs attune with the brains easy ability to judge slope/rate of change.
	**Dot plots (also called "equiplots")** are used to emphasize differences among sub-groups. Dot plots compare estimates along a single axis, which is easier for people to interpret rather than estimates on varying axes (e.g., a bar chart).
	**Maps** are mostly used to compare estimates across geographical boundaries. Maps are most effective when geographical boundaries are clearly visible and there is variation in estimates across the areas displayed.
	**Pie charts** are used to show proportion. Recommended when there are 5 or fewer segments and the values across segments are visibly different. Do not use pie charts side by side to make a comparison *(e.g., across two years of data)* as studies suggest it can be difficult to compare them. A 100% stacked bar chart is preferred for this purpose.
	**Stacked bar charts and 100% stacked bar charts** are alternatives to pie charts for presenting proportions.
	**Area charts** are used to illustrate both proportion and change over time.
	**Icons** are used to emphasize single numbers or to visually communicate proportions (e.g. shading in 7 of 10 icons).
	**Line graphs** should nearly always be used to show change over time.
	**Single numbers** in large or unique font or color can be used to highlight one important number.

Figure 4. Use of shading



Communicating statistical uncertainty

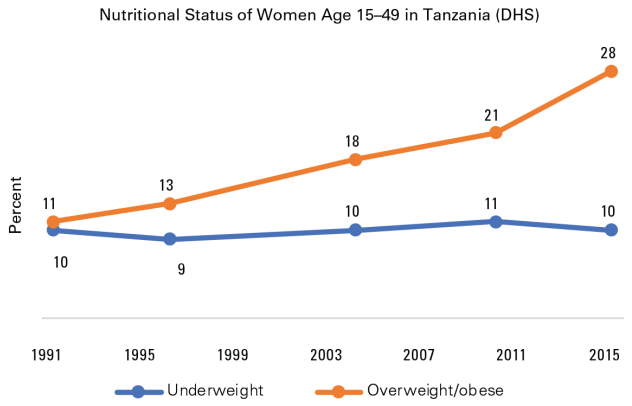
Whether or not to visualize statistical uncertainty (e.g., confidence intervals) around quantitative estimates depends on your target audience, your key messages and your communication goals. Including confidence intervals without guidance can confuse your audience and distract from the message. A 2014 study by Correll and Gleicher concluded that even individuals with technical backgrounds struggle with interpreting statistical significance; as such, it should only be visualized when necessary.[9] The same study found that shaded gradient or "violin plots" to represent statistical uncertainty were interpreted more easily and accurately than bar graphs with error bars and box and whisker plots.[10] In Figure 4, the figure on the left uses shading around stunting estimates because these are modelled, and the figure on the right does not use shading around wasting estimates because these are not modelled.

Any data visualization that includes statistical uncertainty should be accompanied with annotations that support interpretation. These can be incorporated as part of the title, as an annotation within the graphic or in a footnote below the graphic. Alternatives to technical terms "statistical significance" and "confidence intervals" include:

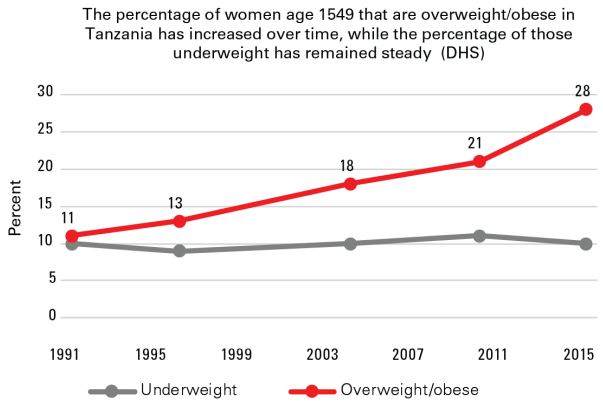
- There is a statistically high chance that the actual number falls within this range
- Statistically, this did not happen by chance
- This is unlikely to be by chance

Figure 5. Formatting to emphasize communication goals

Example A (less preferred)



Example B (more preferred)

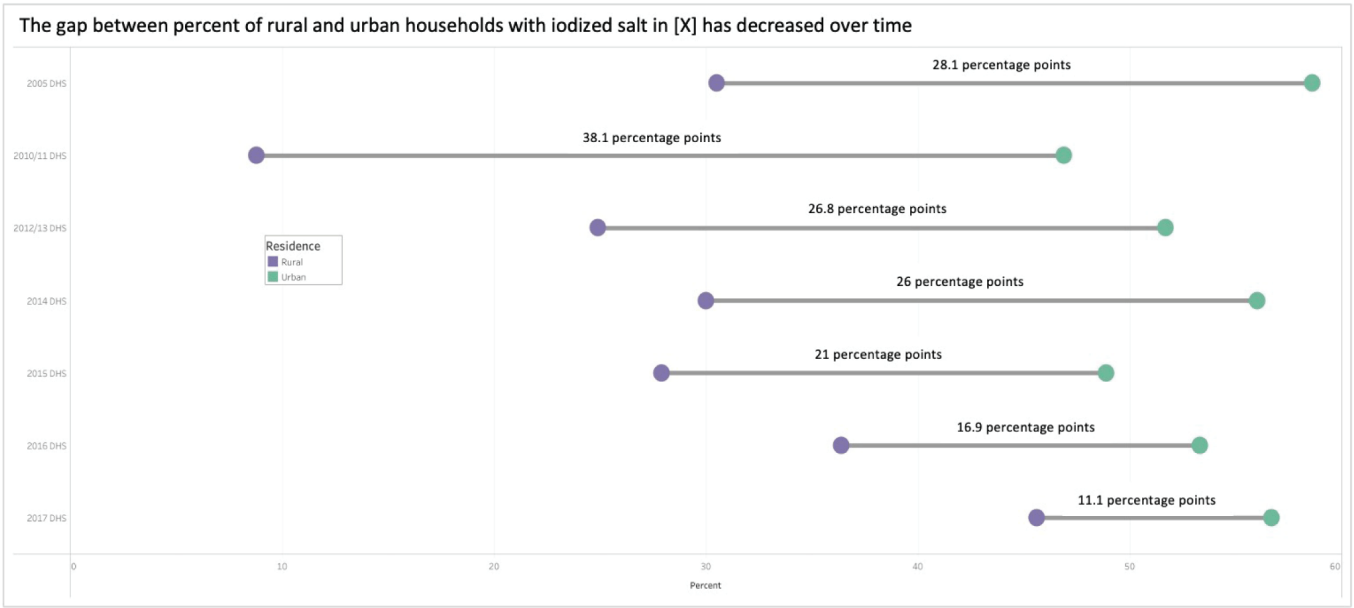


OTHER FORMATTING GUIDANCE

The examples in **Figures 5** and **6** show how formatting can be used to emphasize a communication goal. In **Figure 5**, *Example B* could be considered a stronger option than *Example A.* In *Example A* in Figure 5, the title is general; in *Example B*, the graph title includes the communication goal. Our research suggests that individuals can interpret data visualizations by immediately reading titles, axes and legend labels before focusing on trends. As a result, we recommend that you ensure that these labels are appropriately included and capture the data visualization. In addition, colour and data labels are used more selectively in *Example B* to emphasize the overweight/obese trend line. *Example B* also includes grid lines and a y-axis. These formatting features are often requested by audience members even if they are not always preferred by data visualization experts and designers.

Figure 6 is an example of using annotation to help audiences interpret a potentially unfamiliar data visualization. Dot plots or "equiplots" are increasingly used as an alternative to bar charts to visualize equity. Some people who are new to dot plots struggle to interpret them without assistance. In Figure 6, the plot is annotated so that the communication goal is in the heading and gap in coverage is labelled above each line.

Figure 6. Annotated dumbbell plot to facilitate interpretation



Colour

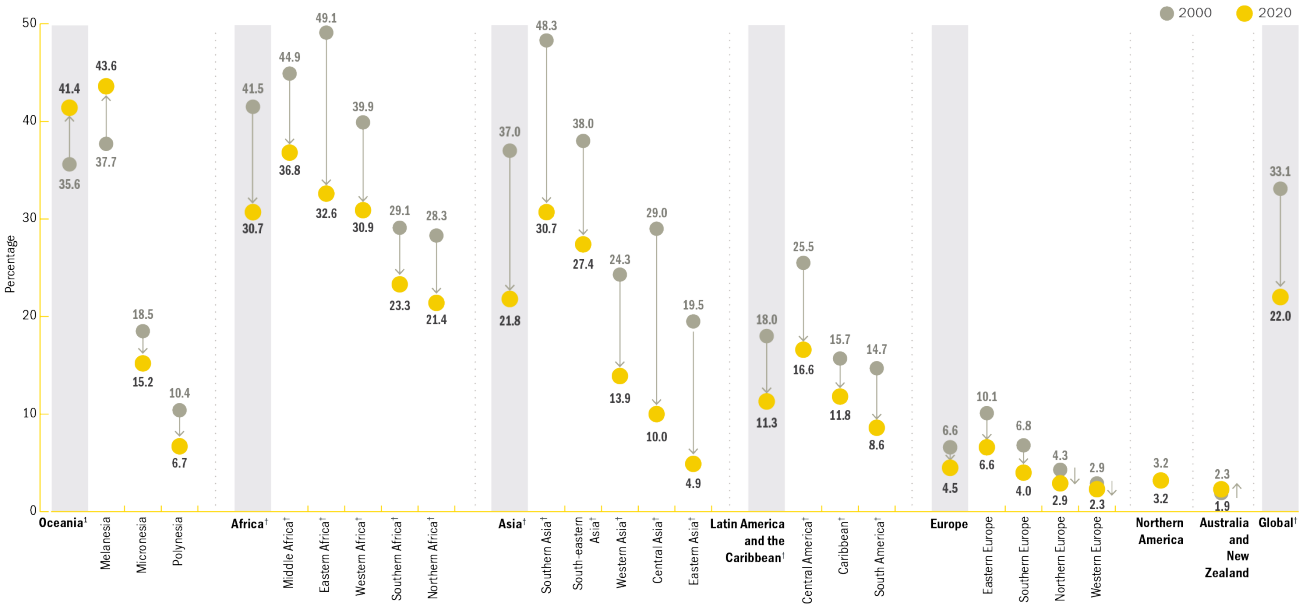
Colour should be used strategically in a data visualization to facilitate interpretation and reinforce communication goals. The use of too many colours can impair interpretation; rather, use colour selectively to highlight specific information.

As much as possible, use a colour-blind-friendly colour palette. An estimated 300 million people globally are colour-blind, with red/green colour-blindness the most common form.[11] Examples of colour-blind friendly colour palettes are available online and can be selected within Tableau and R packages.[12] If you must design using a colour palette that is not colour-blind-friendly, integrate patterns to differentiate elements of the data visualization.

Axes and annotations

All data visualizations should have descriptive titles that align with your communication goals. All axes should be appropriately labelled (e.g., start at 0 when applicable) and spaced consistently. Strategically labelling certain estimates can support interpretation. Some data visualization experts advocate for a high "data to text" ratio – which is a minimalist style that avoids using elements like data labels and gridlines. However, other experts and our own research have debunked perceptions about the usefulness of this trend and advocate for including such annotations to facilitate interpretation.[13] For example, **Figure 7** includes up/down arrows that highlight direction of percentage changes since 2000. **Figures 8** and **9** illustrate how annotations can be used to orient the audience and further support interpretation.

Figure 7. Example showing up and down arrows

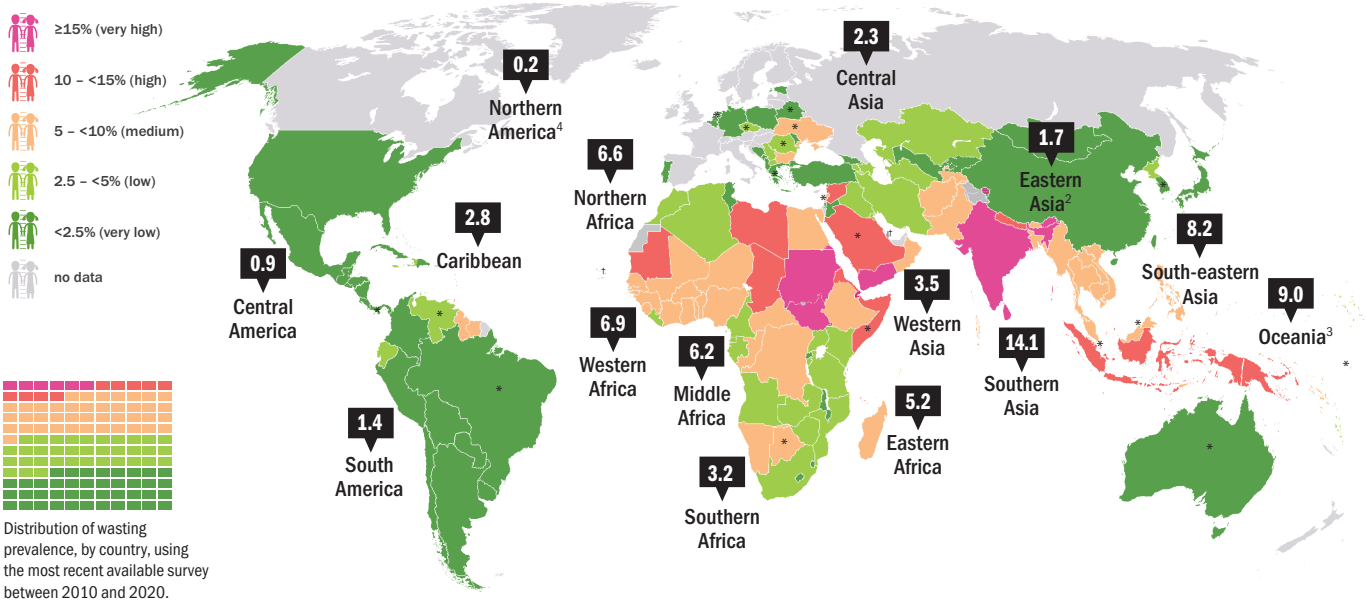


Source: UNICEF, WHO, World Bank Group Joint Malnutrition Estimates, 2021 edition. Note: 1.Household survey data on child height were not collected in 2020 due to physical distancing policies, with the exception of four surveys. These estimates are therefore based almost entirely on data collected before 2020 and do not take into account the impact of the COVID-19 pandemic. However, one of the covariates used in the country stunting model takes the impact of COVID-19 partially into account (see page 3). 2. Oceania excluding Australia and New Zealand. †Represents regions/sub-regions where the change has been statistically significant. See page 14 for the 95% confidence intervals for graphed estimates.

Figure 8. Annotations: Example 1

Southern Asia has the highest wasting prevalence of any sub-region in the world

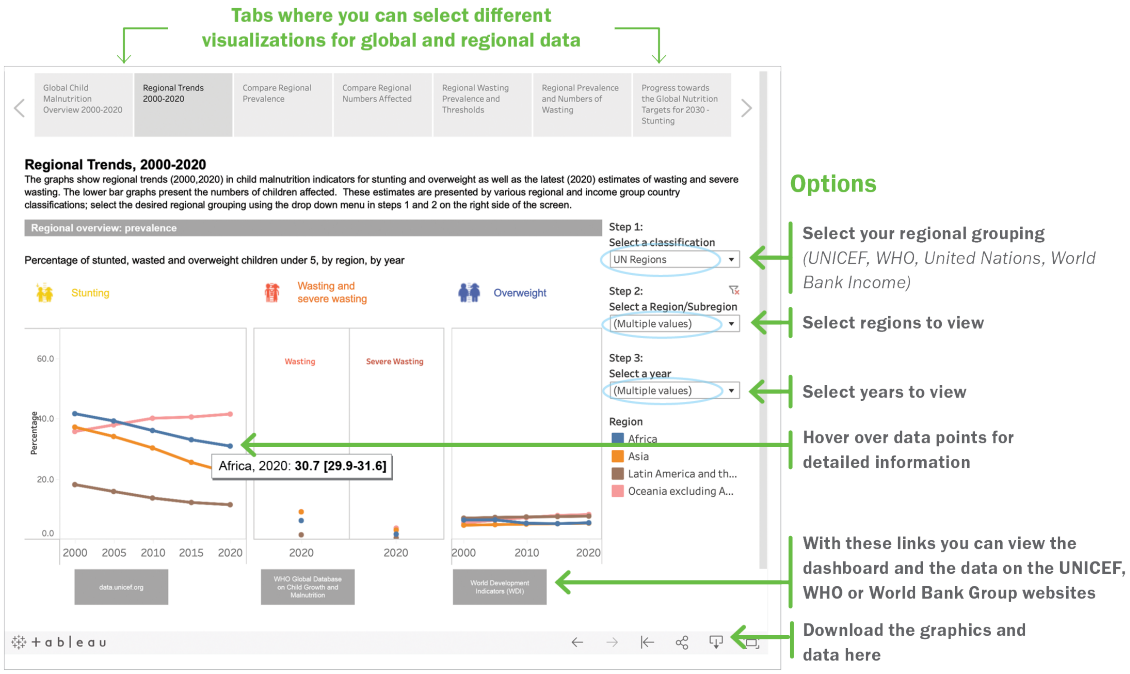
Percentage of children under 5 affected by wasting, by country and United Nations sub-region, 2020[1]



Source: UNICEF, WHO, World Bank Group, Joint Child Malnutrition Estimates, 2021 edition. Note: 1. Country data are the most recent available survey estimates between 2010 and 2020; exceptions where older data are shown (2000–2009) are denoted with an asterisk (*) and where only data prior to 2000 are available the † footnote is used, denoting no recent data. The sub-regional estimates do not account for the impact of COVID-19 given that the collection of household survey data on child height and weight were limited in 2020 due to physical distancing measures with only four national surveys with at least some field work in 2020 included in the JME database; the JME estimates are therefore based almost entirely on data collected before 2020 (see page 3). 2. Eastern Asia excluding Japan. 3. Oceania excluding Australia and New Zealand. 4. The Northern America sub-regional estimate is based on data from only the United States. There is no estimate available for the sub-regions of Europe or Australia and New Zealand due to insufficient population coverage. See section about regional and global estimates on page 27 for an explanation of why trend data are not available for wasting. These maps are stylized and not to scale; they do not reflect a position by UNICEF, WHO or World Bank Group on the legal status of any country or territory or the delimitation of any frontiers.

Figure 9. Annotations: Example 2

Interactive dashboard overview



OPTIMIZING FOR PRINT AND DIGITAL FORMATS

It is essential to design data visualizations that align with the output type and technology available to your target audience. For both print and digital formats, colours should have enough contrast to be visible even if printer and projector quality are not ideal. Light colour shades – particularly yellow –can be problematic for both print and projection. When printing deliverables in black and white, which is less expensive than colour, consider using patterns rather than colours to reduce toner ink. Font size should be adjusted to match the output. Use open source or widely accessible fonts to avoid incompatibility.

Data visualizations are increasingly being published in digital interactive formats. However, interactive data visualizations may not be accessible in settings with unreliable electricity or where internet bandwidth is intermittent, expensive or censored. Furthermore, in many settings, target audiences continue to expect print outputs.

RECOMMENDED RESOURCES

- World Health Organization Analysis and Use of Health Facility Data General Principles, 2018
- The Data for Decisions to Expand Nutrition Transformation (DataDENT) website includes tutorials on how to use Tableau Public to visualize nutrition data
- Evergreen, Stephanie DH. Effective Data Visualization: The Right Chart for the Right Data. SAGE Publications, 2016.
- Chrysantina A, Sæbø JI. Assessing User-Designed Dashboards: A Case for Developing Data Visualization Competency. International Conference on Social Implications of Computers in Developing Countries 2019 May 1 (pp. 448-459). Springer, Cham. [link]
- Aung T, Niyeha D, Heidkamp R. Leveraging data visualization to improve the use of data for global health decision-making. Journal of Global Health. 2019 Dec;9(2). [link]
- IDEO Open Design Kit
- Aung, T. 2019 Data Visualization Society Nightingale Medium. Data Visualization for Audiences in Low & Middle-Income Countries, 2019

ENDNOTES

1 WHO Analysis and Use of Health Facility Data General Principles 2018

2 https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/data-visualization-for-human-perception

3 Aung, T. Data Visualization Society Nightingale Medium. Data Visualization for Audiences in Low & Middle-Income Countries. 2019

4 Evergreen SD. Effective data visualization: The right chart for the right data. Sage Publications; 2019 Apr 3.

5 Holeman I, Kane D. Human-centered design for global health equity. Information technology for development. 2020 Jul 2;26(3):477-505.

6 Aung T, Niyeha D, Heidkamp R. Leveraging data visualization to improve the use of data for global health decision-making. Journal of global health. 2019 Dec;9(2).

7 Neumann C, Dunbar EL, Espino JU, Mtonga TM, Douglas GP. A LMIC-First Manifesto to Developing Electronic Medical Record Systems. Journal of Health Informatics in Africa. 2020;7(2):1-3.

8 Amundsen JI. Dashboards and information overload: Challenges and design guidelines for Health Information Management Systems in developing countries (Master's thesis).

9 Correll M, Gleicher M. Error bars considered harmful: Exploring alternate encodings for mean and error. IEEE transactions on visualization and computer graphics. 2014 Nov 6;20(12):2142-51.

10 ibid

11 See: <https://www.colourblindawareness.org/colour-blindness/>

12 See: <https://www.tableau.com/about/blog/examining-data-viz-rules-dont-use-red-green-together>

13 Aung T, Niyeha D, Shagihilu S, Mpembeni R, Kaganda J, Sheffel A, Heidkamp R. Optimizing data visualization for reproductive, maternal, newborn, child health, and nutrition (RMNCH&N) policymaking: data visualization preferences and interpretation capacity among decision-makers in Tanzania. Global health research and policy. 2019 Dec;4(1):1-4.

November 2022

Published by:
UNICEF
Data Analytics and Innovation
Division of Data, Analytics, Planning and Monitoring
3 United Nations Plaza
New York, NY 10017, USA

unicef

World Health Organization