Open Data Capacity Development
Module 4: Storytelling with Open Data
Training Syllabus

This training module was developed under the guidance of the United Nations Statistical Division as part of the Data4Now Initiative.

Introduction

This syllabus is intended to guide trainers who are presenting a course on open data for staff in NSOs and other agencies of the national statistical system who are responsible for communicating key messages and insights available in open data and databases of indicators. It may also be useful to other advanced users of NSO data, particularly data intermediaries, who wish to explore the data and convey new messages of their own.

The training module is part of a larger program on the development of open data capacity in official statistics agencies. It is accompanied by a PowerPoint presentation that can be used for group presentations or for individual learning.

What do I need to know before using this module?

This module provides a general introduction to using open data in the context of national statistical agencies to develop compelling, data-driven narratives. It assumes a basic knowledge of the types of data produced by official statistical agencies and the principles of open data.

Learning objectives

The expected learning outcomes for this module include:

- How to develop and convey effective narratives that combine text and visuals to inform and influence an audience.
- How to understand their audience and their data needs.
- How to create effective data visualizations.

Note to trainers: Depending on the pace of the trainer and trainees, it is expected that this module can be delivered in 2 to 3 hours.

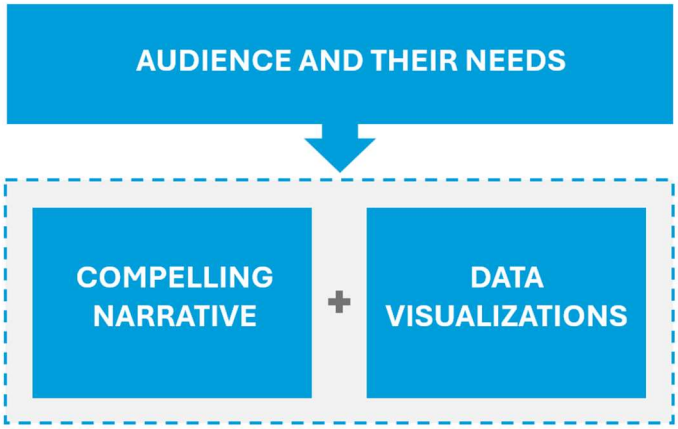
Table of Contents

1. What does it mean to tell a story with open data?

Can open data speak for themselves? No, a table of numbers cannot convey an impactful message on their own. Open data can make a wide range of well-documented information available, coming with assurances that the data can be freely used and redistributed. However, data can only describe the 'what,' and humans must supply the 'why' and the 'so-what'. Someone finding open data will seek to tell their own story but may be limited by their knowledge and biases. The storyteller helps others find the stories in the data.

According to the *Venngage Data Storytelling Benchmark Report 2021*, data storytelling is "a multidisciplinary process that combines the results of data analysis with compelling narratives and presents the combination using text and visual representations of the data to inform, engage, and influence the audience." Based on this definition, the process of storytelling has three core elements: understanding an audience and their needs; conveying the results of data analysis through compelling narratives; and using visual representations to inform, engage, and influence.

Figure 1: Storytelling with open data.



This module starts with an overview of effective data visualization to illustrate and elevate a narrative. A section on crafting compelling narratives based on open data follows. Some instances of storytelling with open data may rely more on narratives, as in the case of a podcast, or in the case of an infographic, on visuals, but an understanding of the audience and what they need to know is fundamental to effective storytelling with open data. The module ends with an in-depth discussion of understanding an audience and their needs.

2. Developing effective data visualizations

Data visualization is a powerful way to illustrate the narrative that we wish to convey using open data. To visualize data, the storyteller should start with an understanding of the

audience and consider what data will be relevant to them and support the story being told. And an effective visualization present only the information (based on high-quality data) that is relevant to the key message, leaving out irrelevant information for the sake of clarity.

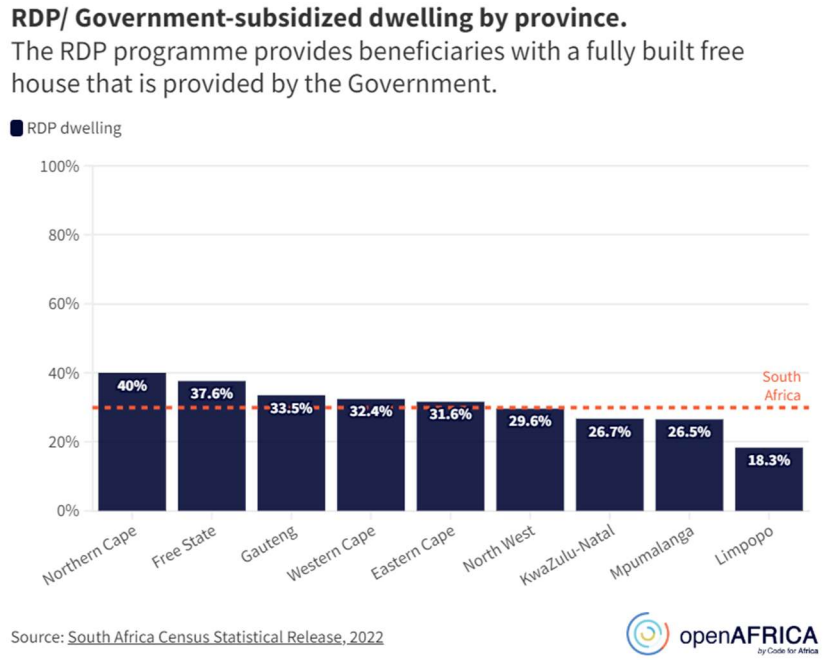
This section begin by analyzing examples of data visualizations accompanying a narrative and discusses how each may be effective or not. With a clear data narrative in mind, the next step is to create effective data visualizations to bring the story to life. Not every visualization may be fit for purpose. Selecting the right charts to fit the data and ensuring that these charts are simple, clear, and accurate will ensure that the data and their messages are properly conveyed and build trust in the audience.

2.1 What makes a good illustration?

There are many ways that a narrative can be effectively illustrated through data visualization. The context and needs of an audience will be the main driver of how the storyteller decides to visualize their data. Through a good visualization, users are helped to understand the data, use the data themselves, and encouraged to explore other open data made available through national statistics offices or other sources.

Code for Africa provides a strong example of storytelling using open data through their article, "Visualising South Africa's society through the 2022 Census," which paints a picture of South Africa's evolving social landscape. Under the housing section, the text describes a shortage of affordable housing and illustrates the availability and distribution of government subsidized dwellings using the data visualization below:

Figure 2: Illustrating affordable housing issues.



Source: Code for Africa.

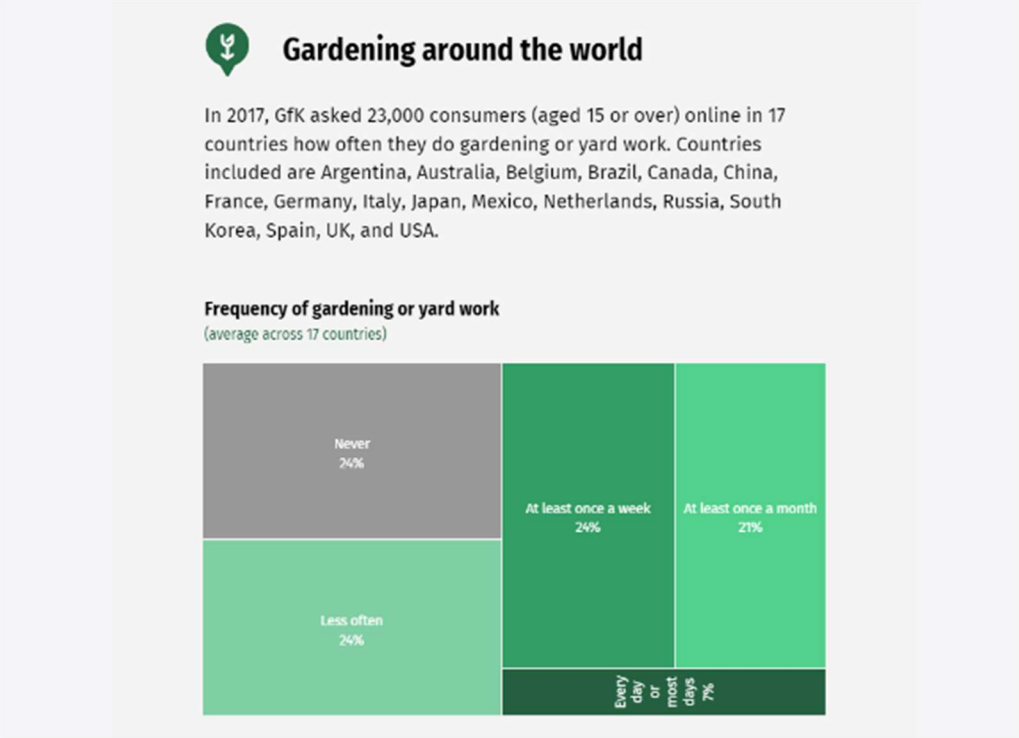
This graphic is clear and simple and effectively illustrates the data in a straightforward manner. Its limited color pallet keeps the attention on the data itself. Its use of a y-axis spanning 0 to 100 percent accurately conveys the percentage of housing that has been subsidized without over-emphasizing any differences while providing a line to show the average across the country.

However, there are ways in which this graph could be further improved and stand on its own. The denominator could be more clearly stated to be all housing, and the timeframe could be explicitly included in the title. However, these details are included in the narrative accompanying the visualization.

A storyteller will need to consider the context and purpose for which they have created a visualization. Below is an analysis of three graphics considering whether they are suited to their purposes.

The GfK published data on gardening and the two visualizations below serve as examples of two extremes. The first example shows how a presentation of data can fail to convey a story. It does not have a key message. It does not have a logical flow. This infographic is an unappealing collage of facts that cannot be considered a data story.

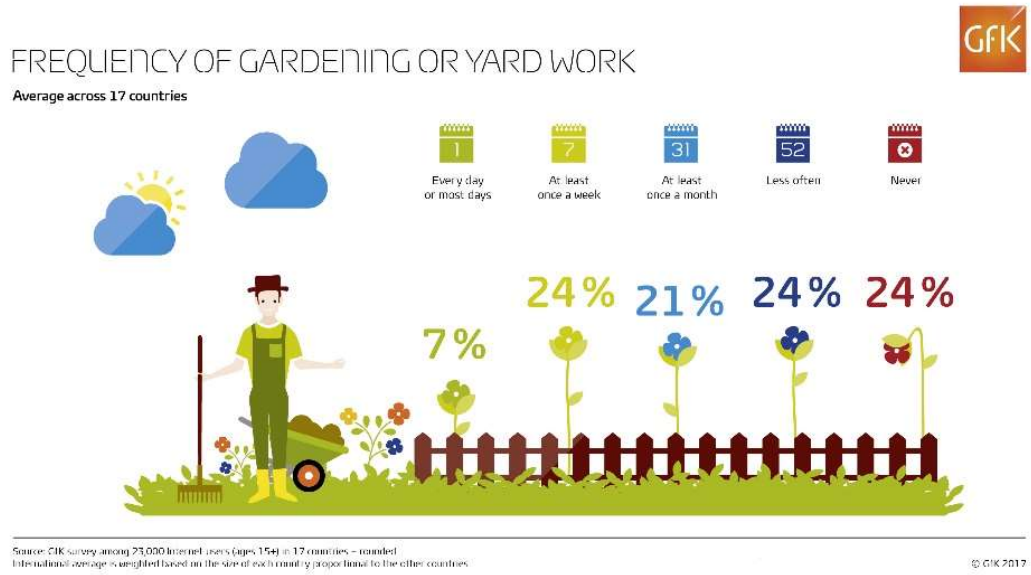
Figure 3: A soulless infographic of gardening around the world.



Source: Infogram.com

The next example below conveys the same data with far more interesting visual elements, but it may lack clarity. While it a far more interesting and visually engaging graphic, it could go further to provide a clear narrative. It paints a literal picture, showing the frequency of gardening work in a garden. It uses calendar icons to visualize the frequencies. The bar chart of flowers conveys the theme of gardening, and the inverted dying flower showing 'never' hints at the consequence of not gardening. However, the other flowers are all the same design, simply with different colors that don't convey a consistent frequency.

Figure 4: A visually overwhelming infographic of gardening around the world.



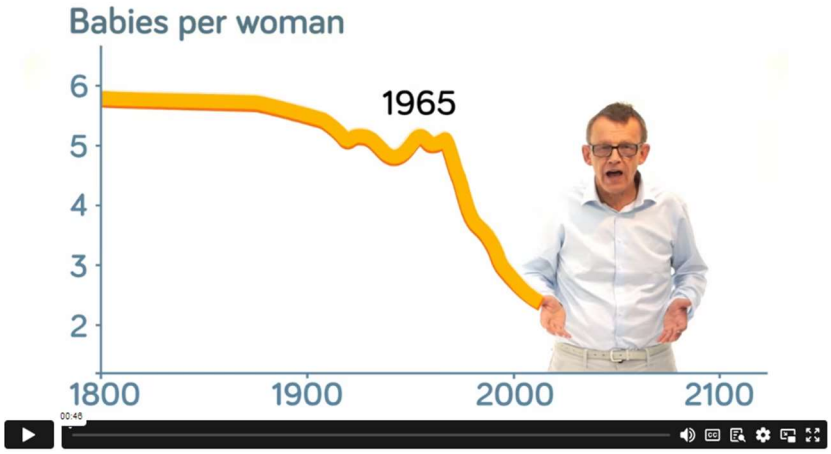
Source: GfK

Each data storyteller will need to decide how to best convey their message using data visualizations, and there is no one-size-fits-all for this. It will depend on the storyteller's judgement about the audience or format. If it is clear and engaging, it will be effective.

A famous example of effective storytelling with data can be found in the animated data visualizations of Hans Rosling. His data storytelling involves a verbal narration with graphics that change as his narrative unfolds. In a talk he gave at TEDxSingapore2015, he asked, "What is the biggest change of our time?" To answer that, he started with a discussion of the number of babies born per woman. Using only a simple line chart, he conveys a narrative that takes us across momentous changes around the world that impact our lives today.

Below is a screenshot from a short video capturing this story that he told using the data.

Figure 5: What is the biggest change in our time?



Source: Gapminder.org

This example is fit for purpose but not necessarily all purposes. The video emphasizes its narrative with minimalistic graphics, but it also makes a tradeoff between a simple, understandable graphic and presenting the actual data values or knowing whether the data were estimated or interpolated. An effective storyteller will need to make judgement calls about how much information to present and how to best combine visualization and narrative to deliver their message to their audience.

"Numbers are boring, people are interesting. You need to look at the numbers but also understand the lives behind the numbers." – Hans Rosling

2.2 Chart selection

Charts are the means through which data will be visualized. These can be helpful in showing comparisons, changes over time, distributions, and correlations. However, it is crucial to ensure that the chart selected is the most appropriate and effective way to achieve visualization goals.

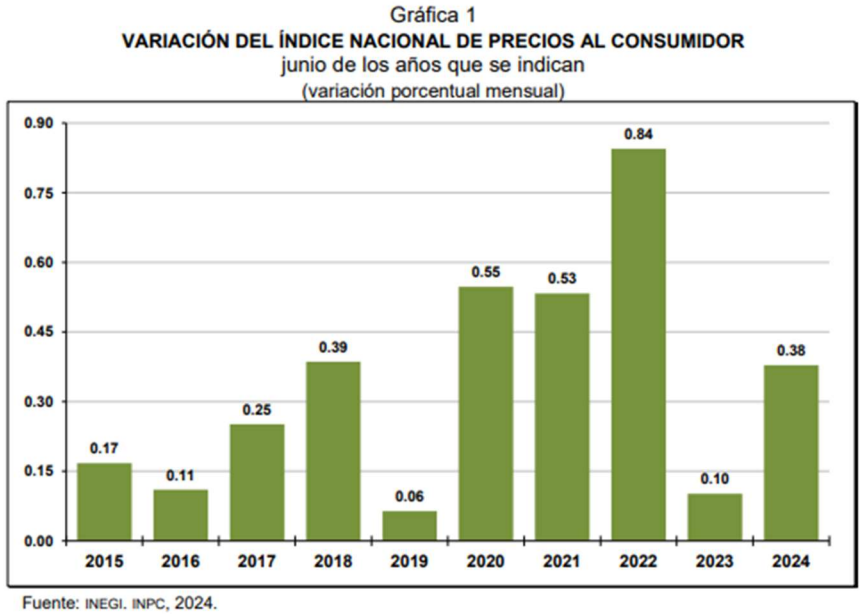
A variety of charts can be used to visualize data. Let's look at when and how to use some of the most common charts.

Bar charts

A bar chart is a versatile means of showing data using rectangular bars, the length of which corresponds to a value. It can visualize comparisons across categories, change over time, and distributions. Bar charts can be horizontal or vertical and can be stacked or clustered.

Below are examples of different bar charts with an explanation of their best practices.

Figure 6: Example of a bar chart visualizing a consumer price index over time.

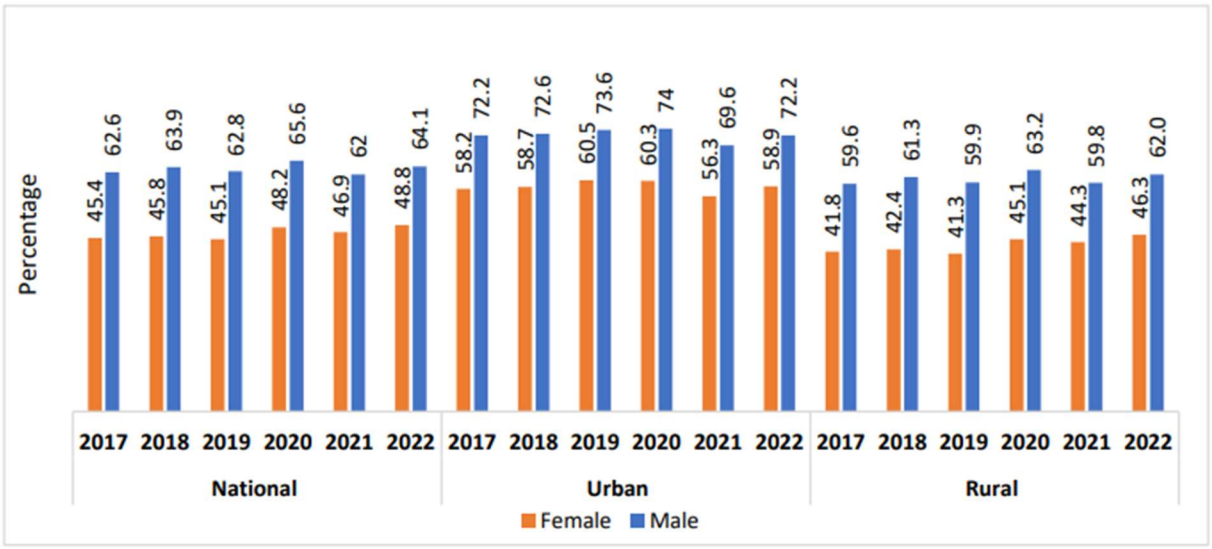


Source: INEGI

The bar chart above is ordered by year, showing fluctuations of the consumer price index over time. However, when a bar chart is presenting categorical data, it may be organized according to the values from smallest to largest, making it easy to quickly identify the highest and lowest performers.

Figure 7: Clustered or grouped bar chart visualizing labor force data by sex and location

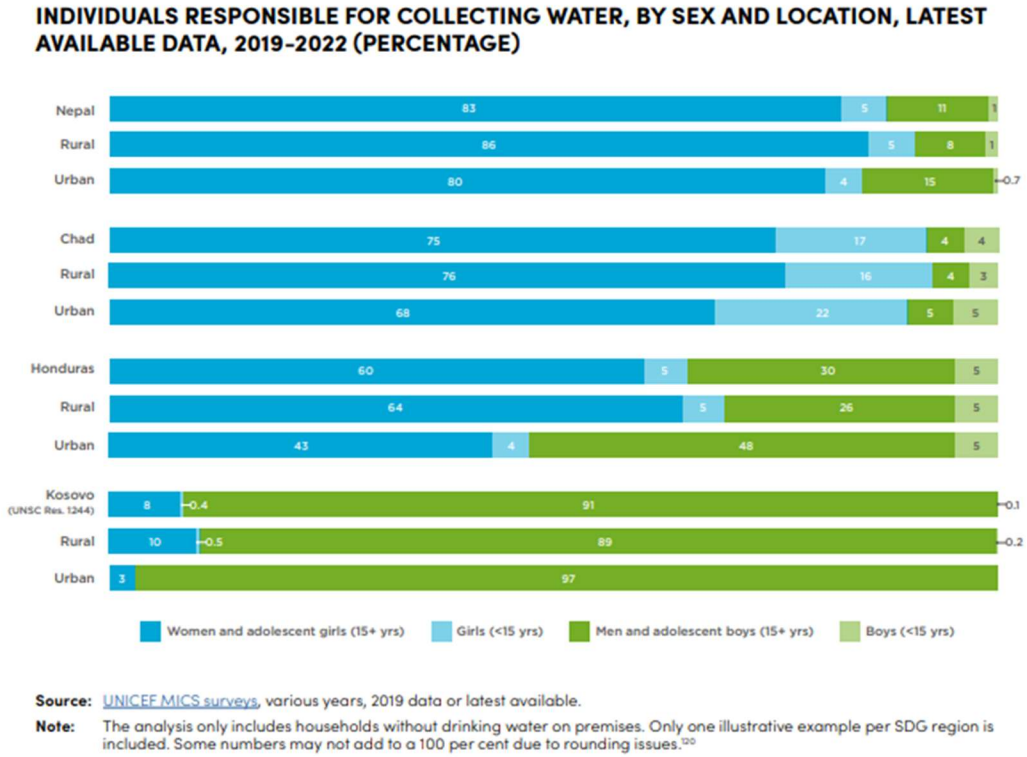
Figure 1:Distribution of the labour force participation rate by area of residence and sex (in %)



Source: National Institute of Statistics of Rwanda (NISR), Labour Force Survey (LFS)2017- 2022

A bar chart is a versatile visualization option that can be configured to show additional dimensions of information. The clustered or grouped bar chart above uses sets of different colored bars to show values by sex, and it further organizes the bars by location. This makes it possible to view three or more dimensions of the data.

Figure 8: Example of a stacked bar chart showing water collection responsibilities by sex and age combinations



Source: UN Women

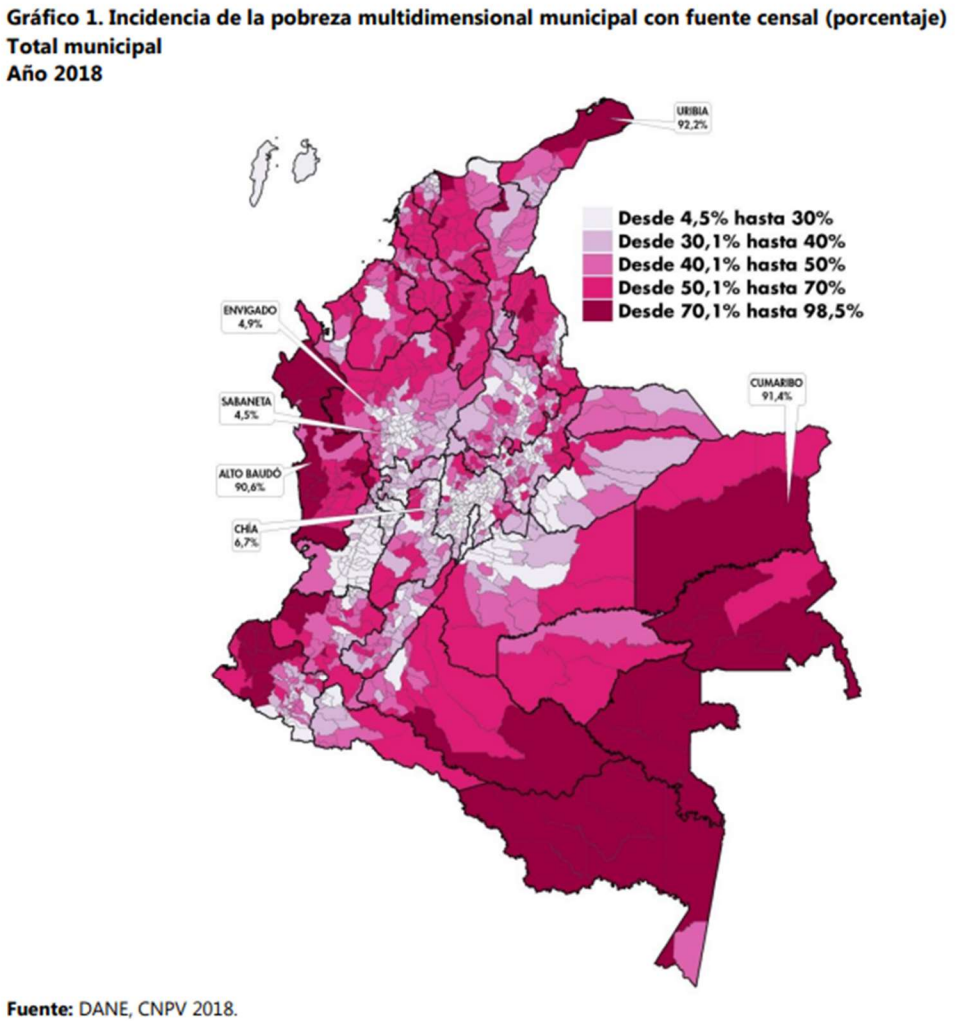
This stacked bar chart offers a succinct way of demonstrating not just the total frequency but also how the total is divided into different components. This chart adds clustering to provide additional organization of the chart by country.

Maps

A map is an ideal method of showing the geographic distribution of the values of an indicator. A map can provide a granular visualization of data at subnational levels or show broad international distributions. Different colors or shades of a color are used to show the changes in a value. The challenge of using maps is that there are cases when some geographic areas are quite small and the important values relevant to that area are difficult to read.

Below is an example of a map visualizing poverty data at the municipal level in Colombia. This is a strong example as small, yet significant areas are highlighted using call-out boxes and it should be possible to view it either in color or black and white due to the strong variations in hues.

Figure 9: Example of a map visualizing poverty data by municipality

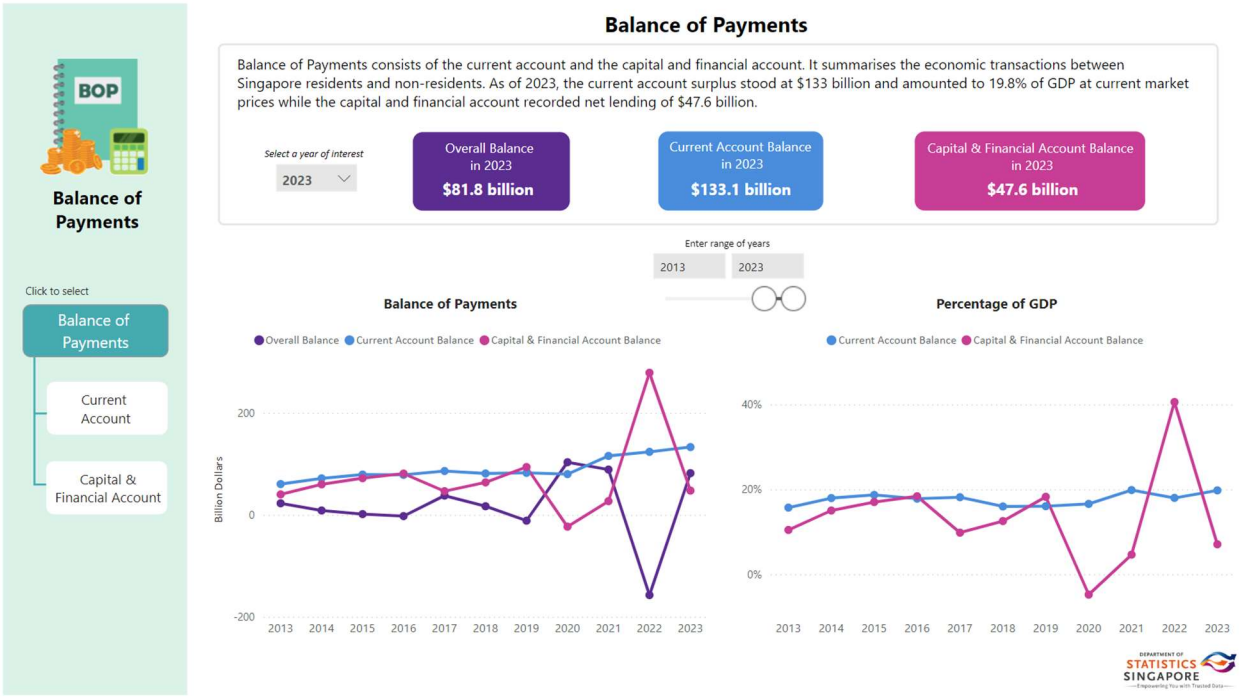


Source: DANE

Line graph

A line graph connects points of information using a line. This type of chart is only appropriate for showing changes over time. They must never be used to illustrate differences between groups or categories as it gives a false impression of progression. The figure below shows an example of line graphs used in a balance of payments dashboard created by the Department of Statistics Singapore.

Figure 10: Line graphs showing balance of payments for Singapore from 2013 to 2023.

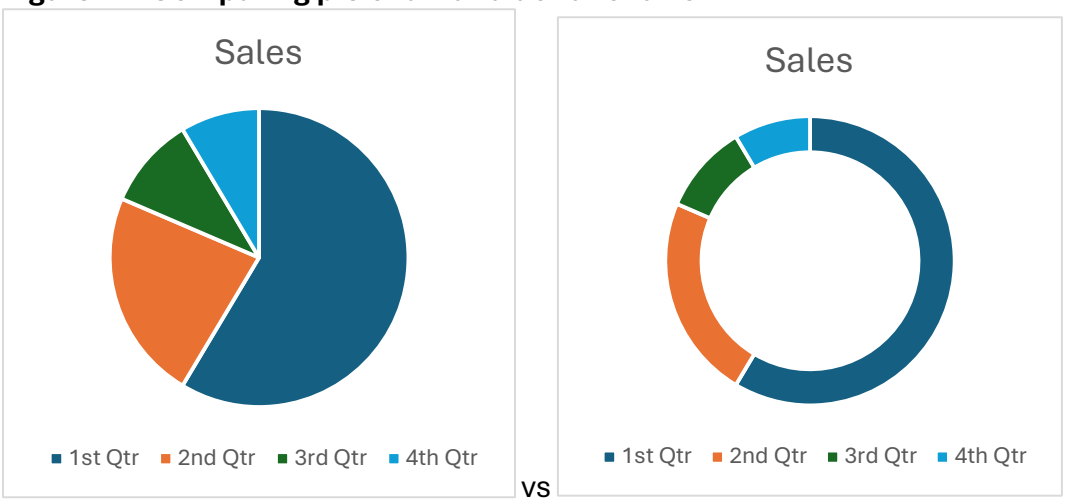


Source: Department of Statistics Singapore.

Donut and pie charts

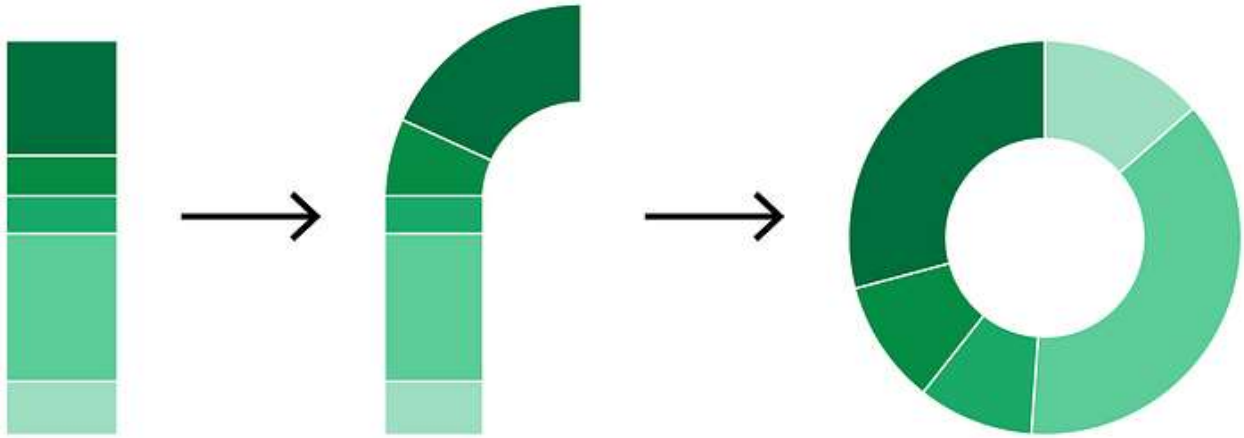
Pie charts are a circle divided into pieces that added up reflect parts of a whole. Donut charts are the same except there is a hole in the center. Many statisticians caution against using either of these charts as it makes it difficult to clearly see the quantities involved.

Figure 11: Comparing pie chart and donut charts



In the case of a donut chart, the hole in the center makes it look like a bar chart that has been wrapped into the shape of a circle. A pie chart focuses the attention on area while a donut chart focuses on length. These charts are used to highlight relative shares or components of a whole. If the goal is to show the disaggregation of an indicator with only a few elements, rather than using a stacked bar chart with a single bar, a donut chart may be useful, but care should be taken to ensure as much clarity as possible.

Figure 12: A donut chart is a wrapped stacked bar chart

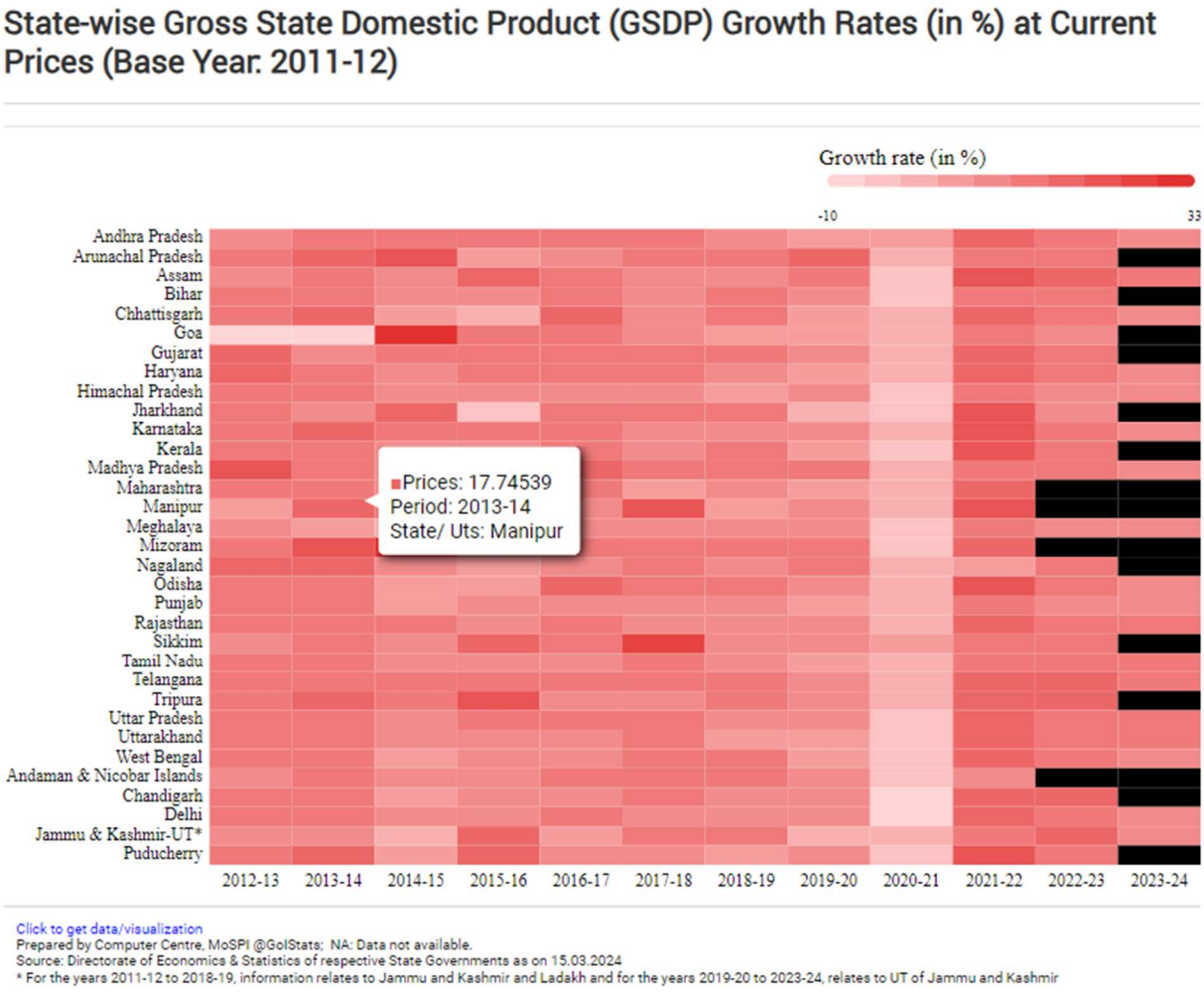


Source: Andrea Robertson

Tables and heatmaps

A table can be added to present data directly. It is often effective to provide a table along with a data visualization to allow an audience to easily find a specific data point. However, it is also possible to use a table as a visualization itself by using colors to highlight the range of the numbers, turning the table into a heatmap.

Figure 13: Example of an interactive heatmap showing India's GSDP by state over time



Source: Government of India, Ministry of Statistics and Programme Implementation

2.3 Creating responsible graphics

Official statisticians will need to take steps to ensure that data visualizations convey information as clearly and accurately as possible. They can promote accessibility through an intuitive and visible color palette, using appropriate scales and axes to avoid misleading an audience, and avoiding distortion and clutter to best convey information.

Accessibility

Ensuring the accessibility of a data visualization will allow people with disabilities to understand and interpret the information presented. The choice of colors and layout are critical to this. A layout should be simple with a consistent structure to allow for easy

navigation across many spectrums of disability. This will involve using a limited color palette with a minimalistic layout and avoiding text heavy compositions.

Text should always be visible with sufficient contrast against the background. The WebAIM Contrast Checker is a useful tool to ensure that all users, including those with visual impairment, will be able to clearly read the text. A contrast ratio of at least 4.5:1 against the background is recommended.

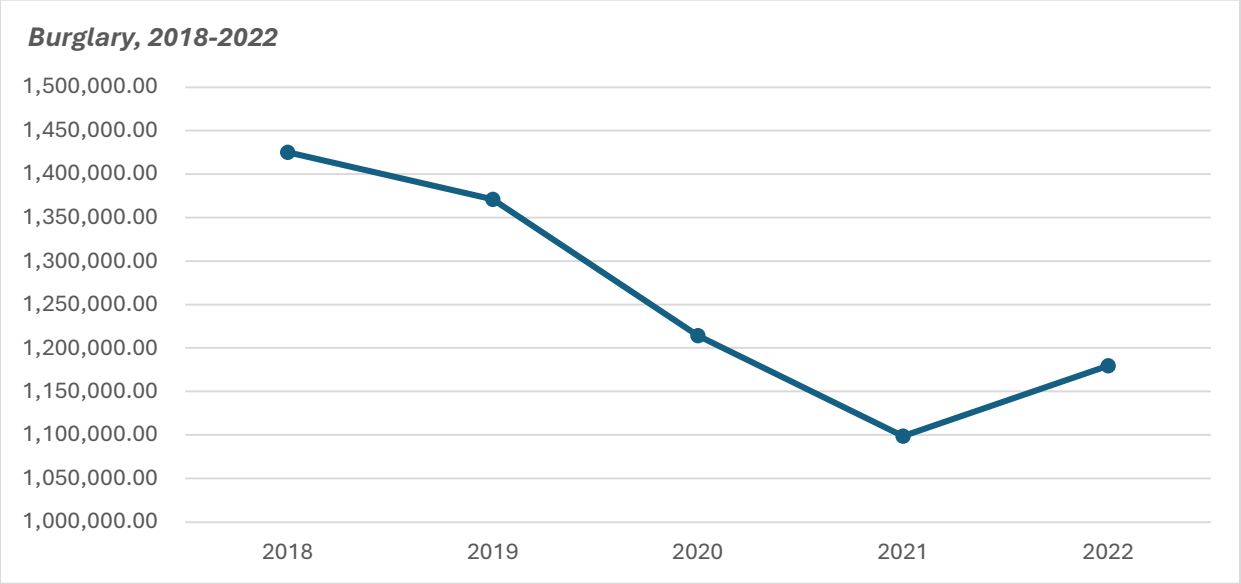
To ensure that a color palette will be visible to those with color blindness, it may be helpful to test selected colors for visibility across the various types of color blindness. David Nichols provides a color palette test through Coloring for Colorblindness that shows how different people may see a set of colors. It may be helpful to use other elements such as shades, shapes, and patterns in a visualization to clearly convey differences between categories.

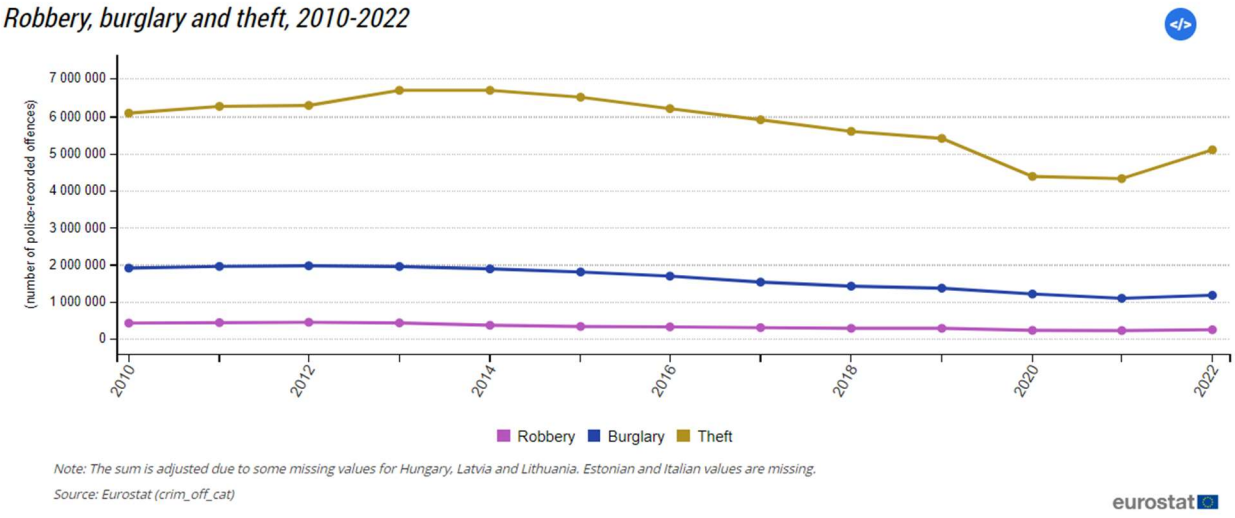
How charts can be misleading

One of the most important ways to ensure that information is clearly conveyed is to set axis values that accurately reflect the scale of the information being conveyed. This will ultimately be a judgement call on the part of the creator who will need to avoid exaggerating or diminishing an effect that needs to be demonstrated.

Below is an example of a graphic that shows an exaggerated change in burglary over five years. The exaggeration is made clear when it is viewed as part of a larger dataset showing robbery, burglary, and theft over a longer period of time.

Figure 16: How a limited axis can be misleading



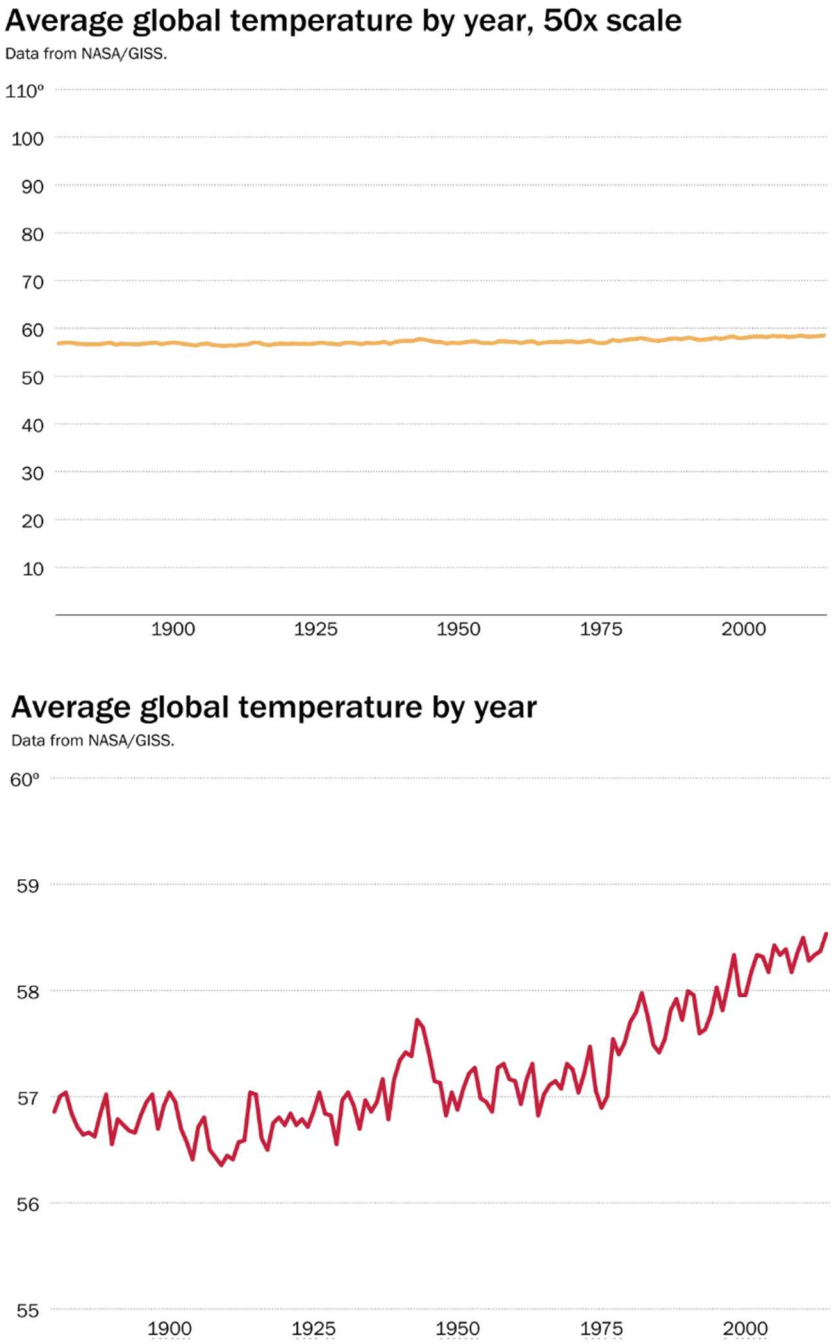


Source: Eurostat.

Starting the vertical axis at near the minimum range or stopping it near maximum is a common method of exaggerating a change. It gives an untrained eye the impression that there has been much variation over time.

In contrast, the graphs below show how important fluctuations can be hidden through an overly extended axis.

Figure 17: How an extended axis can be misleading



Source: Philip Bump, Washington Post.

Maintaining clarity through simplicity

To ensure accuracy and clarity, it is important to focus only on what is most important and avoid overwhelming users with information and visual elements. When creating data

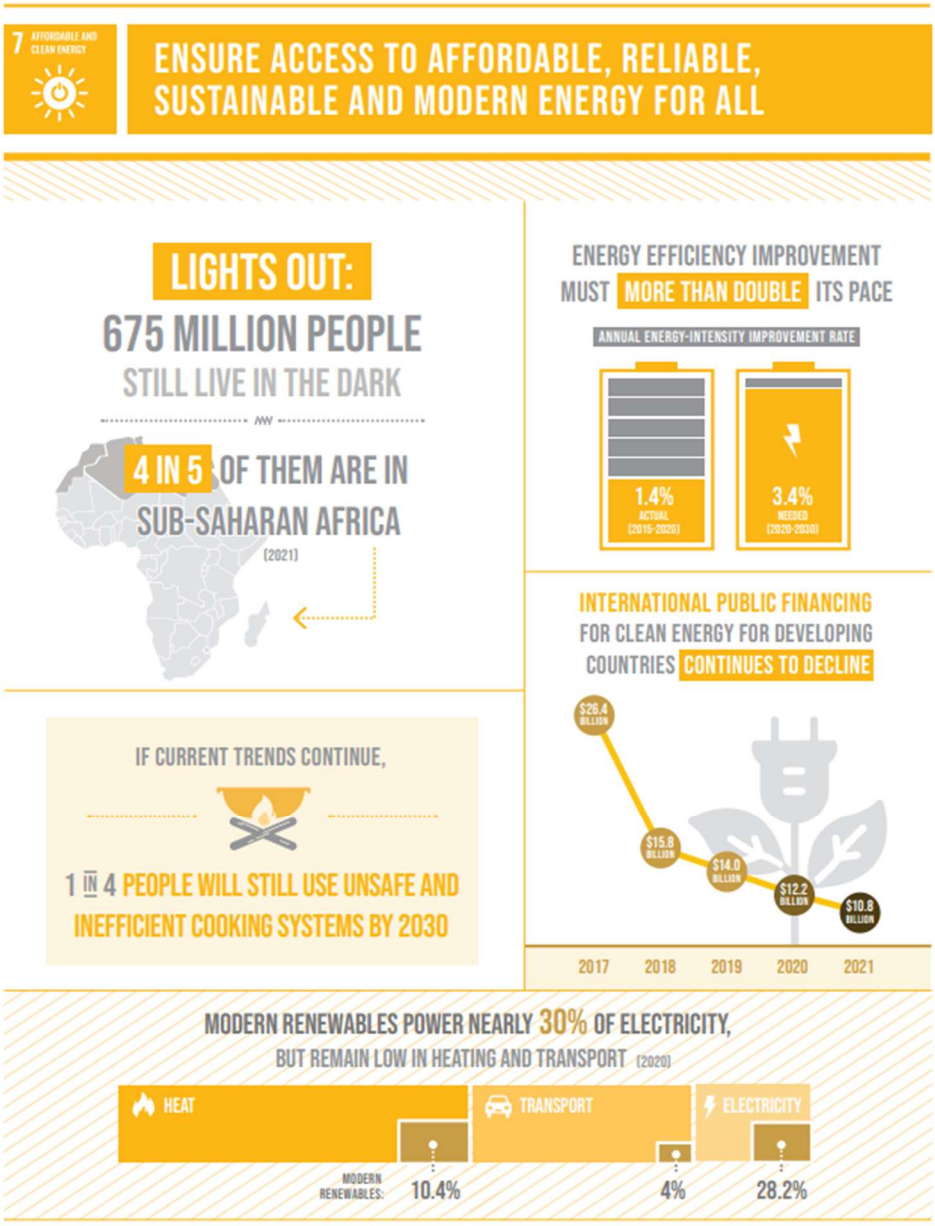
visualizations, less is always more. Overloading a user with information and visual elements leads to confusion.

The following best practices will help achieve clarity:

- Visualize only the most relevant data (focus on the story).
- Limit the color palette.
- Cluster relevant charts, but don't clutter.
- When in doubt, use bar charts.
- Avoid 3D graphics.
- Highlight key messages.

The example below shows an example of an infographic that effectively combines multiple data points with simple visuals and a limited color palette. It divides each indicator into separate sections that make each easy to understand.

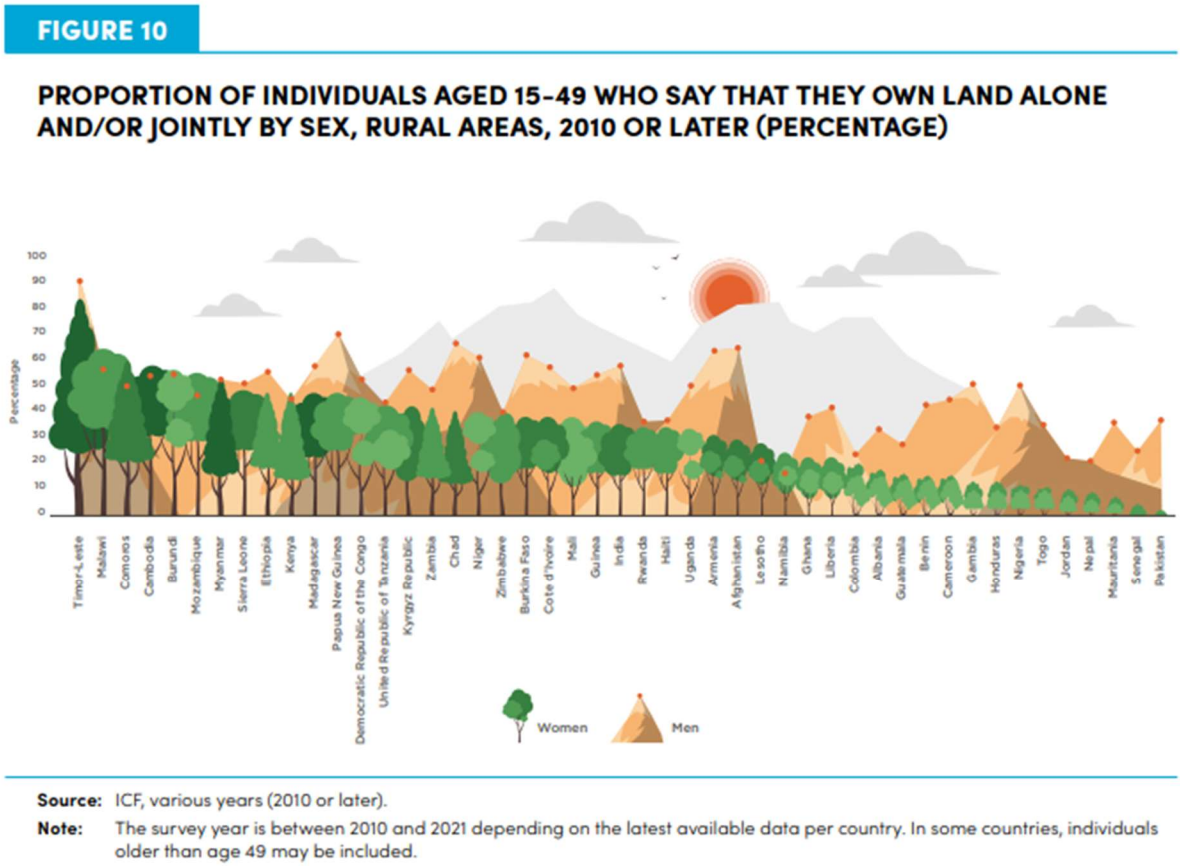
Figure 14: An infographic that maintains simplicity despite conveying an abundance of information.



Source: United Nations.

Avoid complexity for the sake of beauty. The figure below is certainly beautiful, but it sacrifices clarity for the sake of art. It is difficult to see the values represented by mountains and trees. And the orange dots used to show the mountain peaks run the risk of giving the impression of a line graph when the data being shown are percentages across countries.

Figure 15: Example of a beautiful yet overly complex visualization.



Source: UN Women

2.4 Keeping open data open

When open data are used to create data visualizations to convey a story to an audience, it is responsible to ensure that accessibility of the data is maintained. Openness of these data is maintained through the availability of metadata, providing download access to the original data in a non-proprietary and machine-readable format, and maintaining an open data license. These elements can be carried over and continued through data visualizations.

Basic metadata can be included along with a data visualization, citing the original source of the data, when they were uploaded, and including a brief definition of the indicator. At the very least, a link to the metadata associated with the original dataset should be provided. Many sources of open data require that data users identify the original source. If the original data are open and shared in non-proprietary and machine-readable formats such as XLSX, CSV, or JSON, users can inspect and, if they wish, download the original data.

An open license for the data will allow users to re-use and repurpose the data. It would be ideal for the data visualization itself to be shared using CC-BY 4.0 license, which allows others to share or adapt the visualization as long as they provide attribution. This will allow the data visualization to be shared as widely as possible with cascading impacts.

2.5 Creating a final product

As storytelling is a combination of narrative and data visualization, creating a good chart is not the end of the process. Depending on the channel through which the data story will be shared, different final products will be the result. A small social media graphic will present different opportunities than a full-page fact sheet or an in-depth report. NSOs will need to identify the methods used to engage with their audience and key users. These methods or platforms can be:

- Social media: LinkedIn, X/Twitter
- Press releases and bulletins
- Blogs (technical and informal)
- Daily/weekly/monthly/quarterly bulletins
- Fact sheets
- Publications – such as statistical yearbooks, survey reports
- Landing pages on NSO websites: this can be the landing page of thematic statistics, publications, which can provide an overview of data
- Data portals from which data can be downloaded in machine readable formats such as XLSX, CSV or JSON.

Deciding what the best platform to use to convey a story will depend on the audience and the type of material to be presented. A technical audiences may want technical blogs, bulletins, or even the ability to download data, as they may be more focused on the analysis of data or the ability to analyze data.

For more general and broader audiences, NSOs can use social media platforms such as LinkedIn and X to provide more engaging information (whether that is “clickbait” or a “fun” graphic) to target wider audiences—particularly their “followers” from their social media networks. Press releases of statistical products and publications will be of more relevance to media outlets that may cover new data releases. Within each of the platforms, there are varying levels of formality. Press releases are more formal, whereas social media is informal.

3. Crafting a compelling data story

A data story can be told through many different channels, each with their own requirements. A social media post presents different opportunities than a one-page press release or an in-depth report. However, the same basic principles of effective storytelling apply at all levels. A key message must be conveyed in clear, simple terms with engaging language.

3.1 Identifying the key message.

The difference between simply publishing data and storytelling with data is that the storyteller distills an entire world of information into clear messages that an audience wants to know and can easily understand. Thus, the first step is giving in-depth thought into what point the audience should understand and remember.

A narrative is most effective when it is kept short and simple. A compelling story should ask one question and provide a clear and compelling answer. Recall the example of Hans Rosling. He asked, what is the biggest change of our time? And his answer started with a simple, clear message: Women started having fewer children, and this changed human life.

The reality of data is complex and involves many different factors, and it is the role of the storyteller to distill a narrative out of the noise. In an example from the Statistics Canada podcast, Eh Sayers, the storytellers start with a vast array of demographic data and extract a single, thought-provoking question: Who is the 40 millionth Canadian? This is a simple question that allows the storytellers to delve into different types of demographic data as they seek to answer the question. Who are they likely to be? Are they a boy or girl? Do they have siblings? Where are they likely to be born? This single question allows the listener to follow a clear and flowing narrative.

Figure 18: Will the 40 millionth Canadian please stand up?

Source: Statistics Canada

3.2 Understand how users consume information.

Research shows that users want clear, accessible, jargon-free content, written in plain language regardless of their level of knowledge or expertise. However more technical users may wish to access the data directly for further analysis while other users may be content with what has been initially presented.

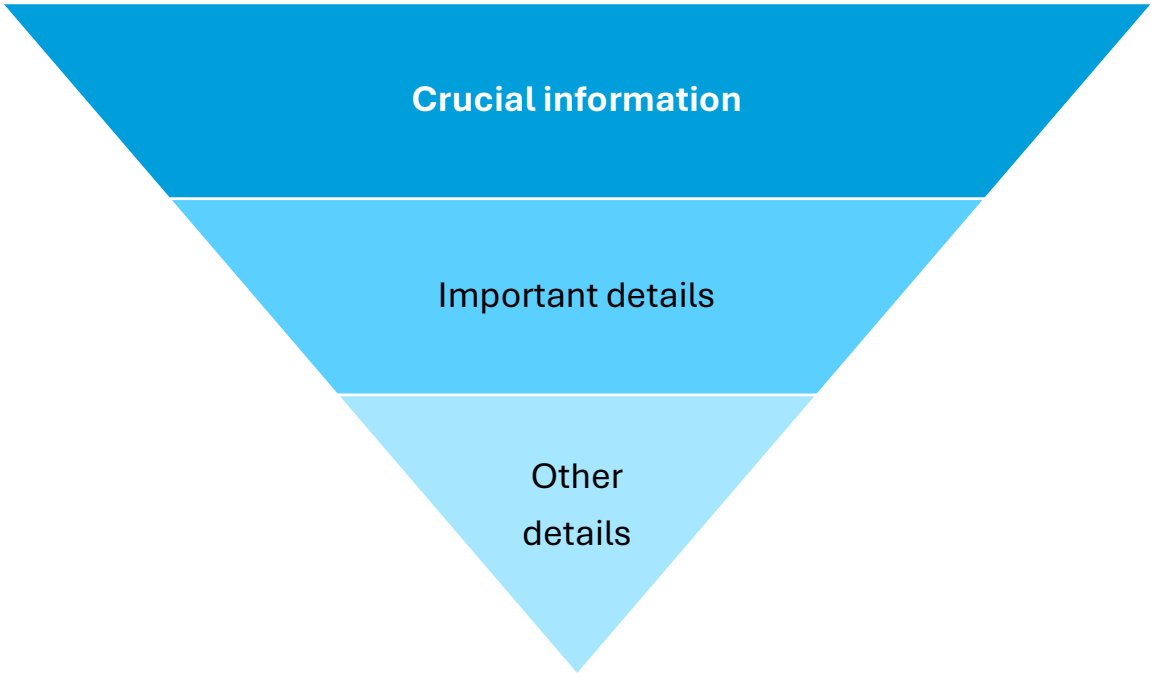
Another characteristic of users is that they want to be able to find the information they need as quickly as possible. If they can't find what they're looking for quickly, they will go somewhere else or not bother. Research shows that the average reader only reads between 20 percent and 28 percent of the text on a web page. It is important to get the point of your content across as concisely as possible.

The most important information in a narrative must be at the start—this is called frontloading. The main information of the content—who, why, what, where, when and how—appears in the first paragraphs so that most users will see it. We use the inverted pyramid to achieve this, placing information in order of importance on the page.

- Most important information
- Important details

- Other general or background information

Figure 19: The inverted pyramid of narrative frontloading.



3.3 Best practices for drafting an effective narrative.

While each narrative will need to consider the availability of data and the needs of an audience, there are a few basic best practices that can serve to make a narrative clear, relevant, and accessible.

- Clear: easy to understand, avoid jargon
- Concise: short and specific to the point, simple but interesting
- Consistent: core messages must build upon each other to create a consistent story
- Tailored: communicate effectively with different audiences by adapting language and depth of information
- Visualized: Telling good stories with the appropriate presentation of data, including visualizations, infographics, and other tools
- Connected: Link data to people

It is important to note that clear and simple messages are not the same as “dumbing down.” Plain language conveys a clear and concise message. It is used with the reader in

mind and with the right tone of voice. It is always possible to use simple language while still respecting an audience.

The sections below cover two key elements that may be included in different types of narratives. A headline message will capture an audience's attention while explanatory text will allow the narrative to unfold in whatever form it takes.

Headline message

Once a key message has been identified, the storyteller must decide how that will be summarized in a single sentence to capture and hold the attention of an audience. This headline message may be used as text of a graphic, in a call-out box in a report, or as the title of a brief. Wherever it appears, it is best to create a sense of urgency by keeping it simple yet powerful and persuasive. It may be helpful to write out several different headlines based on a key message and then to choose the best one after consultation with others.

Below are some examples of headlines that could be used to convey key messages related to extreme poverty in the world. The storyteller starts with a dataset that shows how poverty has changed. After summarizing a few key messages, the objective is to develop headlines that will capture attention and convey a clear message.

Key messages:

- The share of the world's population living in extreme poverty declined to 10 percent in 2015, down from 16 percent in 2010 and 36 percent in 1990. More than a billion people have lifted themselves out of poverty over the last 25 years. However, the pace of change is decelerating.
- Nowcast shows the 2018 rate of extreme poverty at 8.6 percent, and baseline projections suggest that 6 percent of the world's population will still be living in extreme poverty in 2030 if current trends continue.
- Extreme poverty remains stubbornly high in low-income countries and those affected by conflict and political upheaval, particularly in sub-Saharan Africa.

Potential headlines:

- It is less likely that the world will keep its promise to end extreme poverty by 2030
- Conflict and political upheaval have slowed earlier progress to end extreme poverty globally
- Ending poverty by 2030 will require an intense focus on sub-Saharan Africa

Explanatory text

Depending on the storytelling format (whether it is being shared through social media or a full-length report), the extent to which explanatory text can be developed will vary. Some may involve simply a few sentences explaining one data visualization while other stories may allow for multiple paragraphs explaining a sequence of data visualizations.

Here are suggestions for drafting effective explanatory text:

- Make it interesting by linking the data to national policies, goals, and issues that people understand.
- Use headings and bulleted lists to make it easy to scan long sections of text.
- Make the headlines (titles) and first paragraph about the main findings to draw the reader in.
- Keep sentences and paragraphs short.
- Use everyday language (avoid jargon).
- Include definitions or explanation of complex concepts.
- Include clear references to described tables or charts.
- Be careful when describing the relative changes of variables expressed as a percentage – percentage is different than percentage point.
- Stay neutral – look at the different potential angles of what the data show.
- Don't 'table read' by just writing what is already shown in a table– draw out the main findings instead.

Through a strong headline and clear explanatory text with an effective data visualization, a storyteller can connect their audience to the messages of their data analysis.

4. Understanding the audience and their data needs

While the particular story being told may involve a different emphasis on visualization versus narrative, it will always be important to start with an understanding of an audience and what information will be relevant to them. This will allow an effective narrative to be crafted based on available open data and a clear visualization to illustrate it. Everything published should satisfy a user's need. A user's need is what somebody wants to achieve when they visit your website or read your presentation.

If you write with users' needs in mind, it will be easier for people to get the information they need from your content. They will also be more likely to use your data correctly. Different user groups have different needs, but it is also true that most humans are drawn to good stories.

4.1 Why understanding the audience is important.

Understanding your audience will ensure that the narratives and visualizations developed based on open data will be:

- Are relevant, fit for purpose, and help answer people's questions.
- Make a real contribution to our better understanding of the world.
- Instill confidence and deliver maximum impact and value to the user.
- Empower better evidence-based decision-making.

4.2 How to engage with users.

To understand users, identify them and engage with them to obtain their input. There are many approaches to identifying users, such as brainstorming, reviewing previous feedback, public inquiries, using Google Analytics, using social media, talking to people, and advertising.

There are also many channels for engaging with identified users: user satisfaction surveys; social media; toll free lines; radio and TV shows; press conferences, user groups and seminars/webinars; user needs assessments; conferences; bilateral meetings; focus group consultations, roundtable meetings. StatsNorway has used focus groups to gain a deeper understanding of their users while the national statistical office of Malta has used surveys as described in the case study below.

There will be a broad range of users with many levels of expertise. These users may include citizens, civil society organizations, media, business, academics, government analysts, policy makers, ministers and other politicians, international organizations. And within each of these user groups one may find different levels of technical expertise from light to advanced.

While there are many channels for user engagement with a wide range of expertise, these forms of engagement must be matched to the resources and capabilities of the audience. In lower-income countries, traditional methods—such as print media or conferences and presentations—may still be the most effective communication channel.

Case Study from National Statistical Office Malta

The National Statistical Office Malta conducted an NSO User Satisfaction Survey 2022 to help identify their users and to get a sense of who they are. Most of the respondents were identified as light users (63 percent), which meant that they primarily engaged with simple data visualizations and explanations. Intermediate users made up 20 percent of users, and these users sought data in tables and data visualizations and interpretations. Only 17

percent were advanced users who wanted to access the data directly and to adjust tables and download data to create custom visualizations.

The survey also found that the NSO's statistics were used for academic and research purposes by over half of the respondents, followed by market analysis and forecasting (35 percent) and decision-making in business (19 percent).

Regardless of whether a user is light or advanced, storytelling with open data will involve creating clear data visualizations to accompany a compelling narrative. Simply providing raw data to an advanced user may be useful, but it does not involve storytelling. The difference between stories aimed at different audiences will appear in the type of data selected and level of nuance and detail that might be provided.

4.3 How to analyze the audience.

Once an audience has been identified and a dialogue has been established, it is important to consider how their needs will be understood. Below are some key questions to consider as part of analyzing an audience and their needs:

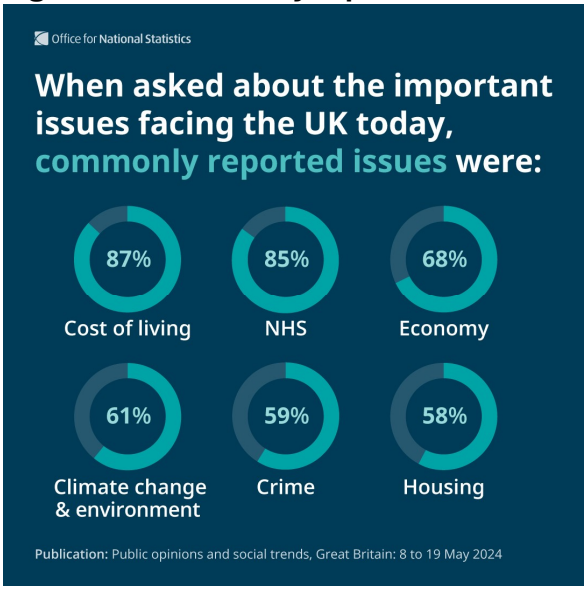
- Who is the audience? Note that there may be more than one audience as people with different interests, background knowledge, time, and technical expertise may care about the same data.
- How is the data story relevant to the audience?
- Why does the audience need the information?
- How will the audiences use the information?
- What role(s) do members of the audience play (decision making, influencing, or information gathering)?
- What data are available that would support the storytelling needed by the audience?
- How experienced is the audience with data and the data story (lay persons or experts)?
- What does the audience care about in the data story? Consider whether they care more about the subject of the data or the data itself.

Case study from the UK's Office for National Statistics

The Policy Evidence and Analysis team at the United Kingdom's Office for National Statistics conducted a survey on public opinions and social trends. While this survey provides knowledge that is relevant to policymakers and decision-makers across a broad range of topics, it is also critical to understanding the core interests of many of their users.

Below is an infographic that was created to convey what respondents said were the most important issues currently facing the UK. This helps flag to a storyteller what data and information will be relevant in this context.

Figure 20: Commonly reported issues facing the UK

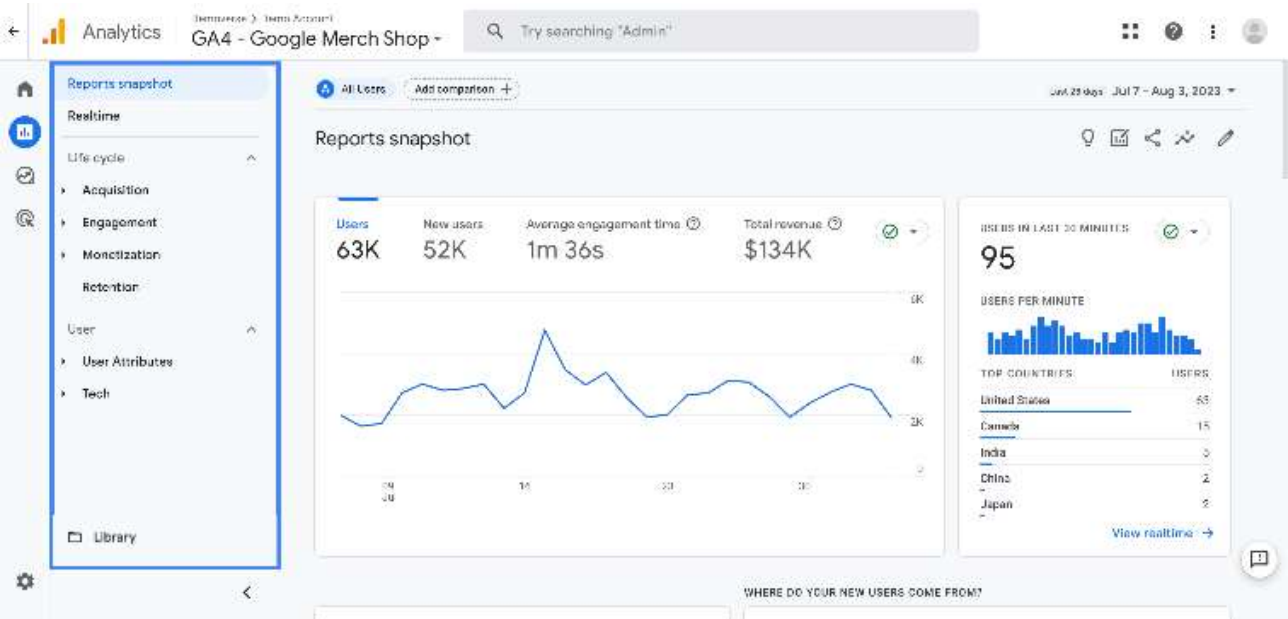


Source: UK Office of National Statistics

4.4 Using web analytics to understand data needs.

Web analytics are an additional means of understanding user needs that allows for real-time activity and behavior to be tracked and observed. Web analytics makes it possible to understand the volume of traffic that an NSO website attracts, user clicks and mouse movement patterns, the return rates of visitors to the website, and user profiles and their geographic origins. And if set up correctly, web analytics can also be used to identify the most frequently accessed datasets. This will show what information is likely to be most relevant to primary audiences. And pair this with what is known about the audience itself to determine how to select or group data for stories. Google analytics is a common option for web analytics as it is free, but many other options are available with different price points and areas of emphasis.

Figure 21: Example of a web analytics dashboard from Google Analytics



Source: Google Analytics Academy

5. Key Takeaways

- Storytelling with open data involves the combination of a narrative with illustrative data visualizations to convey information that is important to an audience.
- Audience:
    - An effective story using open data starts with an understanding of an audience and their data needs.
    - Analyzing an audience and their needs is possible through engagement, which can occur via many channels.
- Narrative:
    - An effective narrative conveys a story in clear, simple terms with engaging language.
    - The most important information in a narrative must be frontloaded at the start followed by other relevant details of decreasing importance.
    - A narrative uses a headline to capture a reader's attention followed by explanatory text to convey a story.
- Data visualization:

- The chart selected to visualize the data will depend on the nature of the data.
- Charts must be made simple, clear, and accurate to accurately convey the messages in a narrative and build audience trust.