PRACTICAL GUIDE TO DATA STORYTELLING

in Voluntary National Reviews and SDG Reporting

United Nations Statistics Division

October 2022

CONTENTS

I. Introduction

Voluntary National Reviews (VNRs) aim to facilitate the sharing of experiences, including successes, challenges and lessons learned of SDG progress, with a view to accelerating the implementation of the 2030 Agenda.[1] VNRs provide a snapshot of the progress made by countries on reaching the Sustainable Development Goals (SDGs) and targets and are a critical peer-learning and lesson-sharing process in global and national SDG reporting.

VNRs and SDG reporting should be centered around data, meaning the preparers of VNRs and SDG materials need to collect, analyze, and decipher a large amount of data. They also need to communicate their findings from the data to a broad audience that has varying levels of data literary—from rudimentary to mastery, as well as a myriad of interests in the data and findings.

Statisticians and data scientists are typically skilled in data analysis and interpretation, yet many are not equally fluent in communicating their findings in a way that is not only informative but also engaging.

Having recognized the challenge that individuals and institutions that oversee or conduct VNRs and SDG reporting are facing in getting data messages across, the United Nations Statistic Division (UNSD) prepared *A Practical Guide to Data Storytelling in Voluntary National Reviews and SDG reporting*, to help countries and national statistics offices better tell data stories for their audience.

About this document

This practical guide is intended to help statisticians, data scientists, policy makers, analysts, managers, and communicators tell convincing data stories in VNRs and SDG reporting, with a focus on understanding the context and constructing the narrative. It contains examples drawn from VNRs and SDG reporting as well as other external sources.

Cultures differ, and data storytelling practices and approaches vary from country to country. This practical guide serves as a generic reference that can be tailored to suit local contexts and situations.

> *The Guide was developed by consultant Hong Xiao. Yongyi Min and Heather Page, UNSD, also contributed to the preparation of the document. Please provide any comments or questions on the Guide to Heather Page at pageh@un.org.*
>
> *Design by Dewi Glanville.*

---

[1] https://sustainabledevelopment.un.org/vnrs/

II. Understand data storytelling

Data are more than numbers. They are a snapshot of real life and information condensed in a numeric way. People use data widely to communicate their ideas, support their arguments and inform decision making. Many statisticians and data analysts incorporate data and statistics in reports and presentations as text, bullet points, charts, and graphs, hoping that the audience will digest the numbers in the same way as they do.

However, the data points that are meaningful to the authors may lack the same punch for the audience. For statisticians and data scientists, simply showing data without helping the audience understand the meaning of the data runs the risk of throwing too much analytical work at the audience. As a result, the audience may misinterpret the data or even lose interest in the data altogether.

Do data speak for itself?

Most of the time, data do not speak for themselves. They give the 'what', but only humans know the 'so what' and the 'why'. In essence, numbers are boring, people are interesting.[2] Statisticians and data scientists need to help the audience look at the numbers but also understand the lives behind the numbers. As such, in addition to the quantitative work of collecting and analyzing data, an equally—if not more—important part of the job for a statistician or a data scientist is qualitative: asking questions, creating narratives from the data, and telling the stories behind the data.

The science of storytelling

Storytelling has been an integral part of humanity for thousands of years. Stories are appealing and enjoyable. They generate interest and entertain audiences in the forms of a movie, a novel, or even a description of a real-life experience. But why do people feel so engaged by a story?

Because the human mind is a story processor, not a logic processor.[3] Neuroscientists have found that emotion—not logic—helps human brains navigate options and arrive at a decision.[4] Researchers find that a story not only activates the language processing parts of the audiences' brain but also other areas in the brain that help the audience experience the events of the story.[5] Compared to facts and numbers, stories stimulate more areas of the brain and change how the audience feels, perceives, and experiences in real life.

---

[2] Hans Rosling, Ted Talk Numbers are boring, people are interesting

[3] Jonathan Haidt, "The Righteous Mind: Why Good People Are Divided by Politics and Religion", Vintage; Illustrated edition, February 12, 2013.

[4] https://www.youtube.com/watch?v=IifXMd26gWE

[5] https://www.nytimes.com/2012/03/18/opinion/sunday/the-neuroscience-of-your-brain-on-fiction.html?pagewanted=all

These findings pay dividends for data storytellers. If statisticians and data scientists can package their numbers and insights into a data story, they will build a bridge for the data to reach the influential, emotional side of their audiences' brain, and help the audience not only see the statistics but also feel the stories behind them. Consequently, these data stories will[6]:

- engage the audiences' brain at intuitive, emotional, and rational levels;
- create a powerful connection between the storyteller and the audience; and
- move the audience to feel and act on the narrative of the story.

---

[6] Nancy Duarte, "Data Story – Explain Data and Inspire Action Through Story", IDEAPRESS publishing, 2019.

Components of data storytelling

Data storytelling is a methodology for communicating information, tailored to a specific audience, with a compelling narrative. It is considered the last mile of your data analysis journey and arguably the most important aspect.[7] It is a multidisciplinary process that combines the results of data analysis with compelling narratives and presents the combination via text and visuals to inform, engage, and influence the audience.[8] It incorporates knowledge and skills from data analytics, design, and communication.

Data storytelling consists of four key components[9], namely, **context, data, narrative, and design (visualisation)**.

Chart 1. Components of Data Storytelling
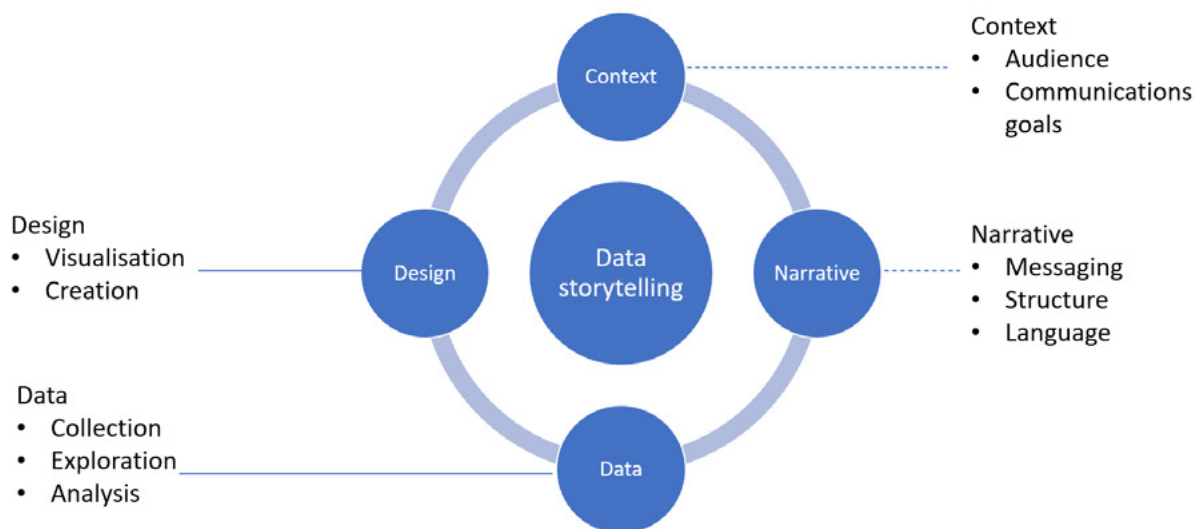
Among the four components, data analysis and data visualization have gained traction and are widely recognized for their critical roles in data storytelling. By contrast, understanding the context and developing a narrative are underwhelmingly valued and deserve more attention from data storytellers. This guide is focused on understanding the context and developing a compelling narrative in SDG data storytelling.

Once the initial data exploration stage ends, data storytellers should decide on what in the data matters to the audience and aim to travel the last mile of sharing the insights with the audience through text, visuals, or both. These are critical processes that will complete the data storytelling cycle and provide a guided narrative for the audience to explore and understand the data and its meanings.

---

[7] https://www.nugit.co/what-is-data-storytelling/

[8] Venngage Data Storytelling Benchmark Report 2021

[9] Venngage Data Storytelling Benchmark Report 2021

What is a data story?

Data stories come in many shapes, forms, and sizes. A compelling data story can be presented using an infographic, an article, a presentation, a webpage, a chart with a proper headline or a comprehensive report. Data stories can be roughly categorized into two groups-- informative and persuasive. Informative data stories are aimed to inform the audience of a finding or an issue and are commonly seen in the media.[10] Persuasive data stories are very commonly seen in the marketing industry in the forms of a campaign or a client presentation, aimed to generate sales.

SDG data stories—such as the SDG reports issued annually by the United Nations Statistics Division (UNSD) and the Voluntary National Reviews (VNRs) that the United Nations Member States produce periodically—would fall mostly under the 'informative' umbrella. However, development agencies also use SDG data stories in fundraising campaigns or programme documents, aimed to demonstrate programme progress and generate more funding.

What makes a good data story in VNRs and SDG reporting?

Telling good data stories in VNRs and SDG reporting is challenging because the SDG content is dense and heavy by nature. In VNRs and SDG reporting, data storytellers need to analyse and pull insights out of potentially enormous amounts of data and report SDG progress across countries and regions on over 230 global SDG indicators. Not to mention the large number of technical terms that are commonly used in SDG discussions. Audiences of the SDG materials need to jump over these common hurdles to understand data in VNRs and SDG reporting.

As such, SDG data story tellers are responsible for making it easier for audiences by building a bridge to their brain via data stories, so that the audience can understand SDG data and their meaning more easily. To do so, storytellers need to better understand the context that they operate in and their audiences' needs and wants, so that they can develop a convincing and compelling data story to better narrate their data and insights through strategically selected text and visuals.

In general, **a good data story in VNRs and SDG reporting should:**

- **be relevant for the audience**
  The content (including text and visuals) needs to match the audiences' current level of data literacy and knowledge on the topic. A good data story shares what the data says and highlights for the audience what matters in the data.

---

[10] How to Tell an Effective Data Story, InformationWeek.com, January 09, 2018

- **be based on good data**[11]

  The data should be from a reputable source and / or collected in a way that appropriately represents what is needed to tell an objective data story.

- **deliver a clear, compelling narrative**

  The story should have a robust storyline, coupled with a clear, powerful title (or headline) and a well-constructed key message that brings out the meaning of the data for the audience. It is important to use easy-to-understand language so the audience will not be overwhelmed by exceedingly difficult language choices.

- **use effective and suitable visuals**[12]

  All visuals including charts, graphs, icons, and photos should serve the sole purpose of helping the audience understand what the data means. The visuals need to be appropriate for the data, well-labeled, legible, and not misleading.

Data storytelling is a powerful tool to engage, influence, inform and inspire the audience. Given the wide scope and complexity of VNRs and SDG reporting, data storytellers are encouraged to leverage storytelling techniques to package the data and insights into a memorable data story that is relevant for the audience and make the story stick with the audience.

---

[11] For more information on this topic, please refer to the *Practical Guide for Evidence-Based Voluntary National Reviews, UNSD*.

[12] For more information and tools for data visualization in SDG reporting, please refer to the SDG & COVID-19 Data Visualization Toolkit.

III. Analyse the Context

Statistics and data stories need context to be alive. Before making any move on data storytelling in VNRs and SDG reporting, statisticians and data scientists need to answer two fundamental questions about the context in which they are operating:

1. What goals should be achieved from telling the data story?
2. Who are the audiences?

## Communications goals of VNRs and SDG reporting

VNRs and SDG reporting take place annually or every few years. Therefore, data storytellers are blessed with the opportunity of reporting on latest developments towards realizing SDG Goals and meeting SDG targets over the reporting period, rendering their reporting newsworthy and highly anticipated. Therefore, **the immediate goals of VNRs and SDG reporting** are to inform and engage the audience by:

• presenting latest progress and trends related to SDG Goals and targets
• sharing new insights and observations related to SDG Goals and targets

In the meantime, data storytellers are often challenged by the need to find new angles, new perspectives, and new stories from the SDG data that may not change much from year to year. If a year-on-year perspective is difficult to present, data storytellers could explore whether a comparative narrative would surface if data from multiple years are brought into play. Sometimes, no change or no or limited progress or progress of a focused population group, such as women or youth, could also be news. Taking a longer-term view on data echoes with the **longer-term communication goals of VNRs and SDG reporting**, namely,

• influencing perceptions on progress made around SDG Goals and targets
• generating actions towards achieving SDG Goals and targets

Chart 2. Communication Objectives of SDG Data Storytelling
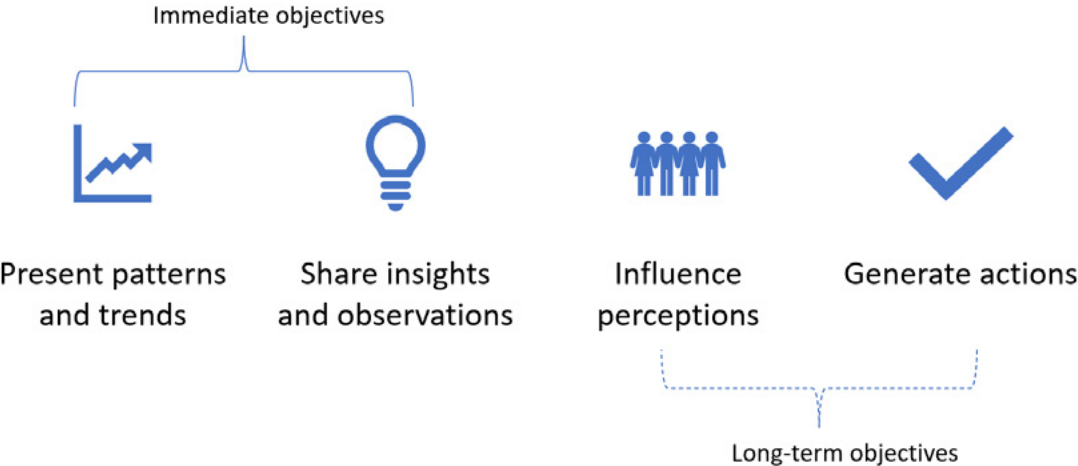
Audience analysis

As products are manufactured to be consumed, data stories are developed for an audience. Data storytellers need to begin with the audience in mind. Due to the extraordinary depth and breadth of coverage that VNRs and SDG reporting have, these reports should not be aimed at serving a single group of audience, but a broad range of organizations and individuals that are from various backgrounds and with varying degree of interest in the data story. The audiences include, but are not limited to:

- Media
- Government agencies
- International organisations
- General public
- Civil society
- Businesses
- Academia
- Other stakeholders

Data storytellers are encouraged to conduct an audience analysis using methods such as surveys, interviews, focus groups, and web analytics to profile their audiences so that the storytelling can be tailored to meet audience needs. It is particularly important to understand the data literacy level of the audience and what data stories might be relevant for them. If there are multiple audiences with varying levels of data literacy, it is advisable to provide several highlights via easy-to-understand text and visuals and make resources such as metadata available for more experienced audiences to pursue further.

The following questions can be used in an audience analysis:

- Who are the audiences?
- How is the data story relevant to the audiences?
- Why do the audiences need the information?
- How will the audiences use the information?
- What role(s) do the audiences play (decision making, influencing, or gathering information)?
- What data are available that would support the storytelling needed by the audiences?
- What do you need your audiences to do with the data story?
- How experienced are the audiences with data and the data story (laymen versus fluent)?
- What do the audiences care about in the data story (microdata, metadata, key figures, key messages, explanatory texts, guidance on how to understand statistics, insights for sharing, or in-depth analyses)?

Some national statistical agencies venture to profile their audiences in greater detail by categorizing them into specific personas such as expert analysts, information foragers, inquiring citizens, technical users, and policy influencers.[13] By defining each persona's data preference, behavioral patterns, and other traits, data storytellers can tailor-make data products for each of the personas.

All in all, the more data storytellers know and understand their audiences, the more likely they will be successful in telling the data story for their intended audiences.

---

[13] https://style.ons.gov.uk/category/writing-for-the-web/personas/.
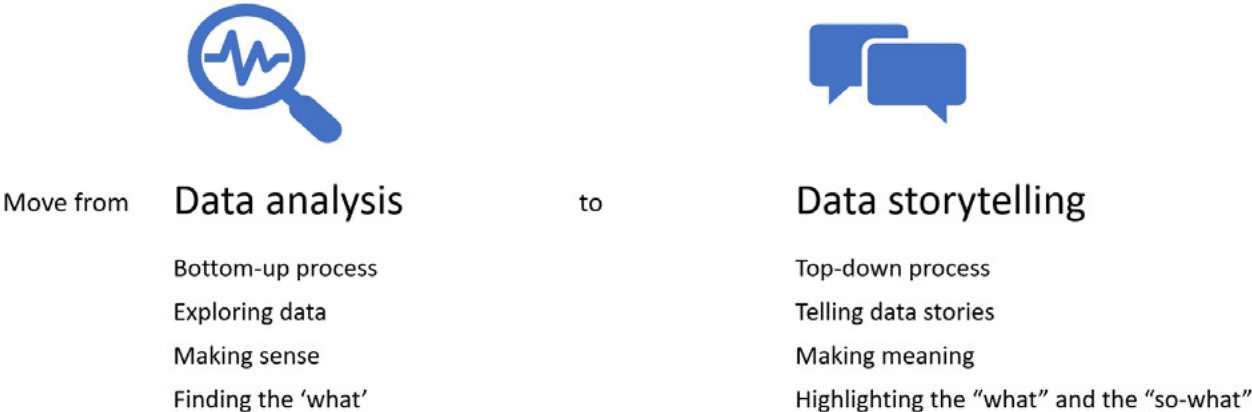
IV. Construct the story

After analyzing the data, the context, and the audiences, data storytellers can start tying all the information together and give the audience a framework—the story—to follow through. This marks a critical turning point where a statistician or a data scientist transitions from a data analyst to a data storyteller or a data communicator.

Mindset change: From a data analyst to a data storyteller

Statisticians and data scientists wear two hats—'**data analyst**' and '**data storyteller**'—in VNR and SDG reporting. As a **data analyst**, they work on collecting data, processing the data and conducting the analyses. It is a bottom-up process where a data analyst moves from sorting out the details to forming higher-level insights. As a **data communicator or a storyteller**, statisticians and data scientists put together the narrative—the storyline, the title, the key message(s), the structure, and the language—and tell the data story for the audience. It is a top-down process where data storytellers present the core statistical finding(s) and explain why and how they matter.

Chart 3. Mindset Change: From Data Analyst to Data Storyteller

Move from   Data analysis   to   Data storytelling

Bottom-up process   Top-down process
Exploring data   Telling data stories
Making sense   Making meaning
Finding the 'what'   Highlighting the "what" and the "so-what"

Being aware of the differences between these two roles, data storytellers can better tell data stories for their audience and help the latter grasp the meaning of the data. Here are two examples that show the difference between telling a data story from the perspective of a data analyst and that of a data storyteller.
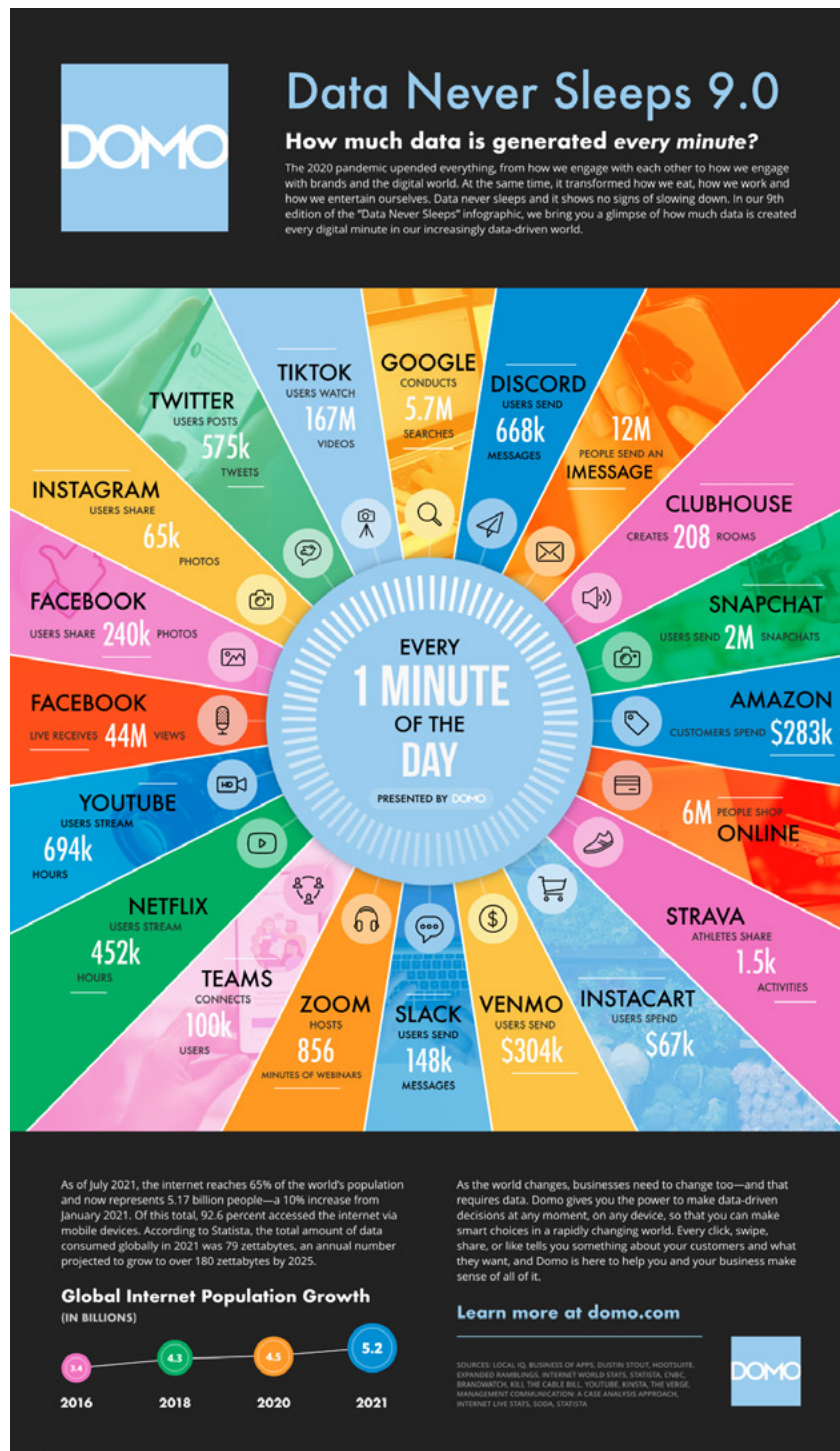
Source: Data Never Sleeps 9.0 Infographics from Domo

Each year, cloud-based operating system provider Domo releases an infographic titled *Data Never Sleeps* to demonstrate how netizens generate, use, and view data in a minute across high-traffic online platforms and applications such as Instagram, YouTube, Netflix, and Amazon. The infographic casts a glimpse into the volume and variety of data generated and distributed over the internet.

Despite the importance of the topic and the huge amount of information presented in the infographic, the *Data Never Sleeps 9.0*, released in 2021, instead of illuminating the audience on the topic, unfortunately becomes one of the so-called "data dumps"—a data presentation packed with more facts and figures than a human being can absorb in a single sitting[14]. The infographic neither brings out a clear message nor a compelling story to frame the topic. It fails to highlight the implication of the data and explain the 'so-what' about the data. As a result, the audience is left to wonder on their own why they should care about the infographic and how they can relate to the data.
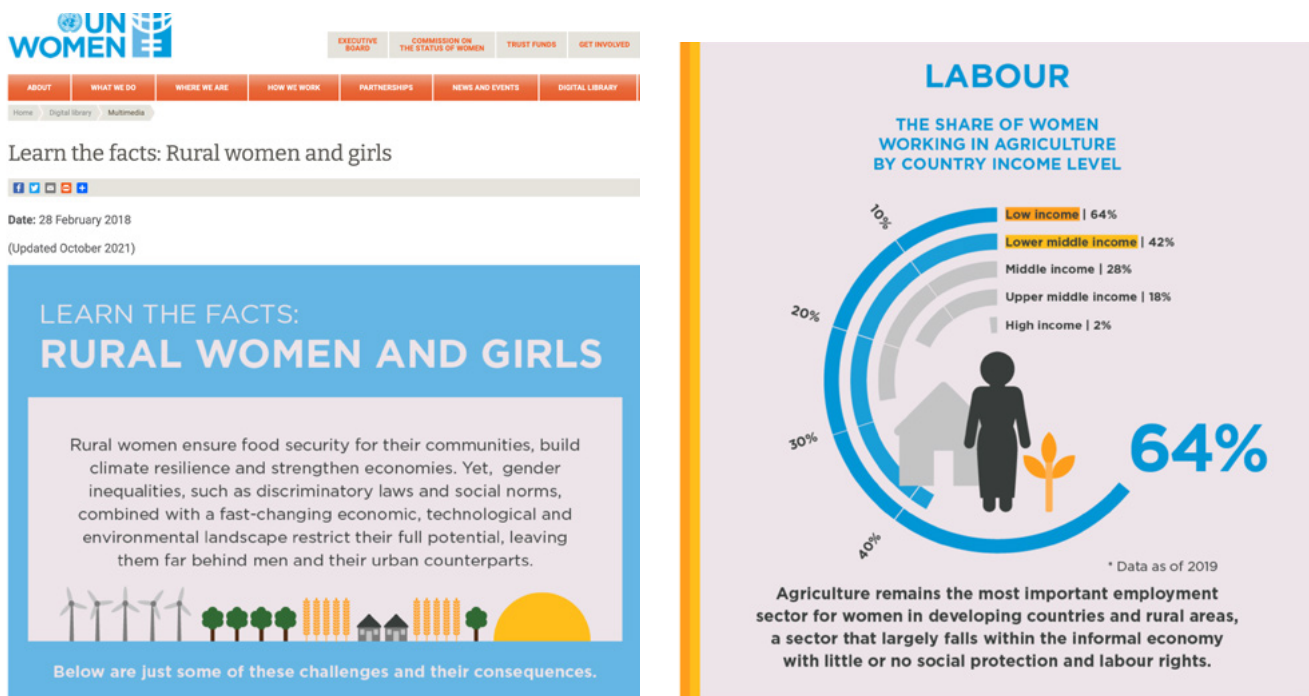
Source: Learn the facts: Rural women and girls | UN Women

In contrast, in the two sections taken from UN Women's infographics (above) tilted "Learn the facts: Rural women and girls", data storytellers communicate their findings in a coherent data story by sharing insights and explaining the meaning of the data, instead of dragging the audience through the analysis process every step of the way like what is seen in the previous example.

In the UN Women infographics, the storytellers do not start their story by explaining the concepts and how the data are collected and analysed—not because such information are not important but because they are not the primary focus of the audience of the infographics. Such information can be conveniently included as references for audiences to

---

[14] More Data Stories, Fewer Data Dumps, Forbes, November 30, 2020

explore at their own leisure. Instead, storytellers open the narrative with a eye-catchingkey message—gender inequalities, among other things, restrict rural women's and girls' full potential and leave them behind men and their urban counterparts. The audiences are well informed of the significance of the issue and are now ready for further insights and details on the topic.

Subsequently the storytellers support the key message with specific data stories. The right section in the above table presents one of these data stories, on the employment of rural women and girls. It first highlights the key statistical finding—or the 'what'—on the share of women working in agriculture by country income level. Then the data storytellers proceed to spell out the specific social and economic impacts that the agriculture work has on rural women and girls in low-income and lower-middle-income countries. In other words, data storytellers help their audiences to see both the 'what' and the 'so what' in the data story.

The comparison between the two examples shows that telling data stories from the audiences' perspective by putting the key message upfront and explaining the meanings of the data will go a long way towards engaging the audience and getting the key message across. Statisticians and data scientists will benefit greatly from putting their audiences at the centre of the data storytelling.

How to make VNRs and SDG materials friendlier to the audience?

There are two golden rules on how to make the VNRs and SDG materials more accessible to the audience.

      1. Keep it simple
      2. Share insights

To **keep the data stories simple**, data storytellers need to create a clean, crisp **storyline** with a **key message** supported by several **core arguments**. It would be ideal to use no more than three arguments because too many arguments will make it hard for the audience to follow along. A powerful **title or headline** will help draw people's attention. Of course, using fewer big words, jargon, and acronyms will make the data stories more accessible to the audience, particularly those who are less experienced with data or not familiar with the topic of the data story. Finally, intentional **visuals** should be used strategically to support the narrative and help the audience understand the data story.

Also, **share your insights about the data**. Insights about data are not the data storytellers' personal opinions. Instead, they show why and how the statistical findings presented in the data story are important. If there is too much to tell or too many insights to share, storytellers may step back and contemplate on whether there is an overarching story that can weave all insights together. Writing down what you want to convey about the "what" and the "so what" in 1-2 full sentences will help data storytellers organize their thoughts and clearly define the data story.

Let's look at each of these elements in greater detail.

Storyline

The first step to telling a good data story in VNRs and SDG reporting is to identify a story worth telling. Storytellers need to know what they want to tell the audience and find a compelling story to supply the audience with context, insight, interpretation—all the things that make data meaningful and analytics relevant and interesting.[15] Storytellers need to identify a theme or a perspective from the data and develop it into a robust storyline. Here are several commonly used approaches to locating a story in data.[16]

- **Identify trends and patterns**
  Patterns and trends indicate how and in what direction the SDG progress is made. Identifying new or evolving trends and patterns help the audience pinpoint where actions and resources are needed to further drive SDG progress.

---

[15] https://www2.deloitte.com/us/en/insights/topics/analytics/data-driven-storytelling.html

[16] https://www.lucidchart.com/blog/how-to-tell-a-story-with-data

- **Draw comparisons**

  Comparisons can help the audience understand how data relates to one another and why.
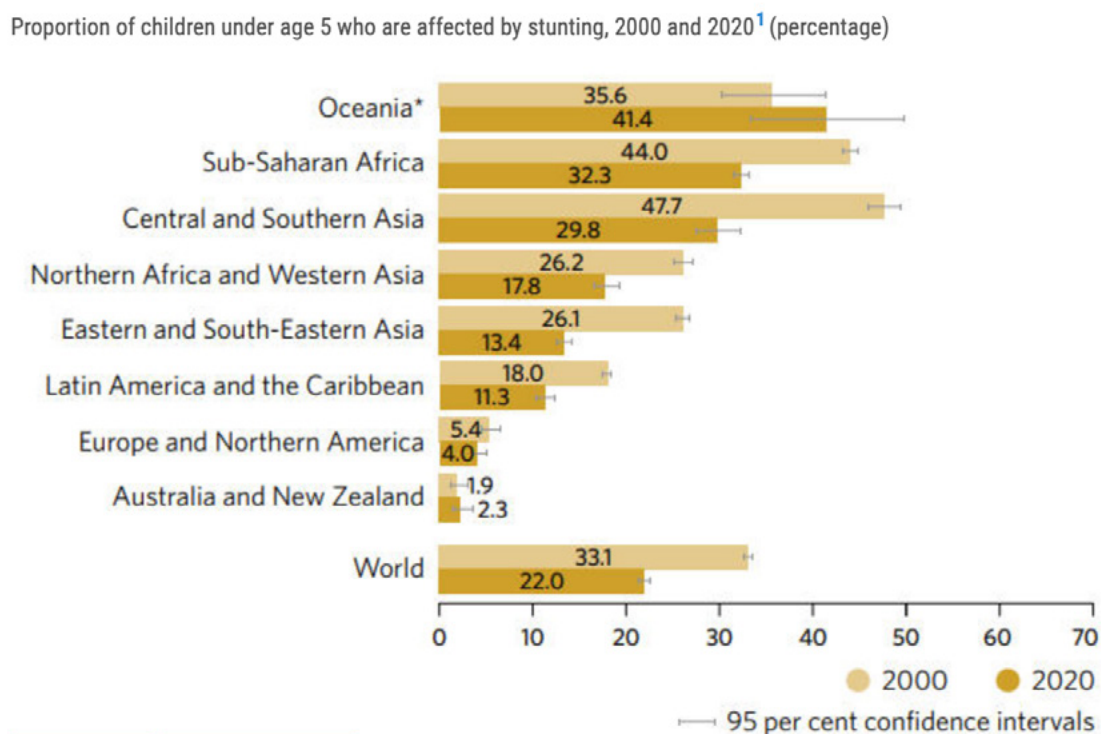
- **Look for correlations**

  Are data points connected in some way? Are there any interesting or surprising correlations? These relationships can develop into a compelling data story.

- **Check on outliers**

  Outliers are any data that act unusually. They can be very useful in highlighting an insight. Look for outliers and ask why. Why is the data behaving that way? What is the cause? This may lead to a very interesting storyline.

Here is an example that shows how SDG data storytellers may develop multiple storylines from one set of data. The following data chart is taken from the United Nations SDG Report 2021, showing the proportion of children under age 5 who are affected by stunting from 2000 and 2020.

Proportion of children under age 5 who are affected by stunting, 2000 and 2020[1] (percentage)

Region	2000	2020
Oceania*	35.6	41.4
Sub-Saharan Africa	44.0	32.3
Central and Southern Asia	47.7	29.8
Northern Africa and Western Asia	26.2	17.8
Eastern and South-Eastern Asia	26.1	13.4
Latin America and the Caribbean	18.0	11.3
Europe and Northern America	5.4	4.0
Australia and New Zealand	1.9	2.3
World	33.1	22.0

● 2000   ● 2020

⊢——⊣ 95 per cent confidence intervals

Show Data   Download to CSV

* Excluding Australia for New Zealand.

Source: https://unstats.un.org/sdgs/report/2021/Goal-02/

Using the storyline-finding approaches highlighted above, storytellers could generate the following storylines from the dataset.

- **Identify trends and patterns:** Great progress has been made globally on reducing stunting between 2000 and 2020.

- **Draw comparisons:** Oceania, Sub-Saharan Africa, and Central and Southern Asia lag the rest of the world on reducing stunting in the past two decades. These three regions may need stepped up efforts to reduce stunting.

- **Look for correlations:** The underwhelming performance of Oceania, Sub-Saharan Africa, and Central and Southern Asia in reducing stunting indicates a possible correlation between stunting and the level of economic development in the regions. Maybe more targeted campaigns and programming will help these regions tackle stunting.

- **Check on outliners:** Oceania is the only region that has higher stunting statistics in 2020 than in 2000. Further study is needed to find to find the reason and asolution.

This exercise reveals how several different stories could stem from the same set of data. Data storytellers should test as many options as possible and select the storyline that is the most relevant for the targeted audiences.

Title (headline)

A powerful title or headline captures the audience's attention and reduces the cognitive load on the audience. The title concisely summarizes the core insight generated from the data. Normally a good data story title is SMART—the acronym of specific, measurable, action-oriented, relevant, and time-bound. Although a title does not have to contain all five elements to be good, the SMART rule provides a framework that data storytellers can use while drafting the title for their stories.

Chart 4. SMART Titles

Specific   Measurable   Action-oriented   Relevant   Time-bound
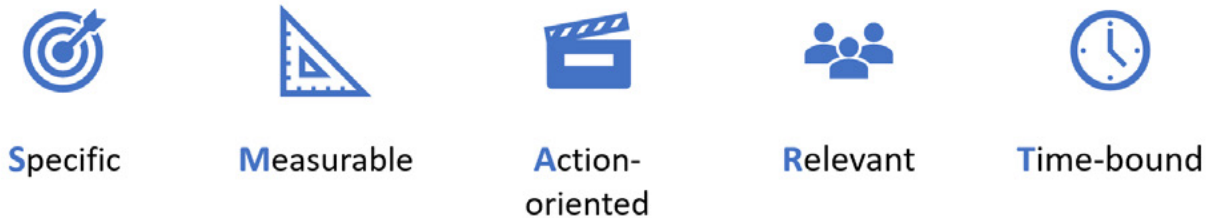
Here are some examples of **SMART** titles:

- **In 2020, the destruction of seized, found, or surrendered weapons increased by 15%** (Source: Voluntary National Review -- URUGUAY 2021, Page 98)

- **A decade of progress in reproductive, maternal and child health could be stalled or reversed by the pandemic** (Source: UN SDG Report 2021 Goal 3)

- **Global warming of 1.5°C and 2°C will be exceeded during the 21st century** (Source: IPCC | Climate Change 2021: The Physical Science Basis)

- **Pfizer Vaccine Has 23% Efficacy Vs. Omicron in South African Lab Study** (Time, December 12, 2021)
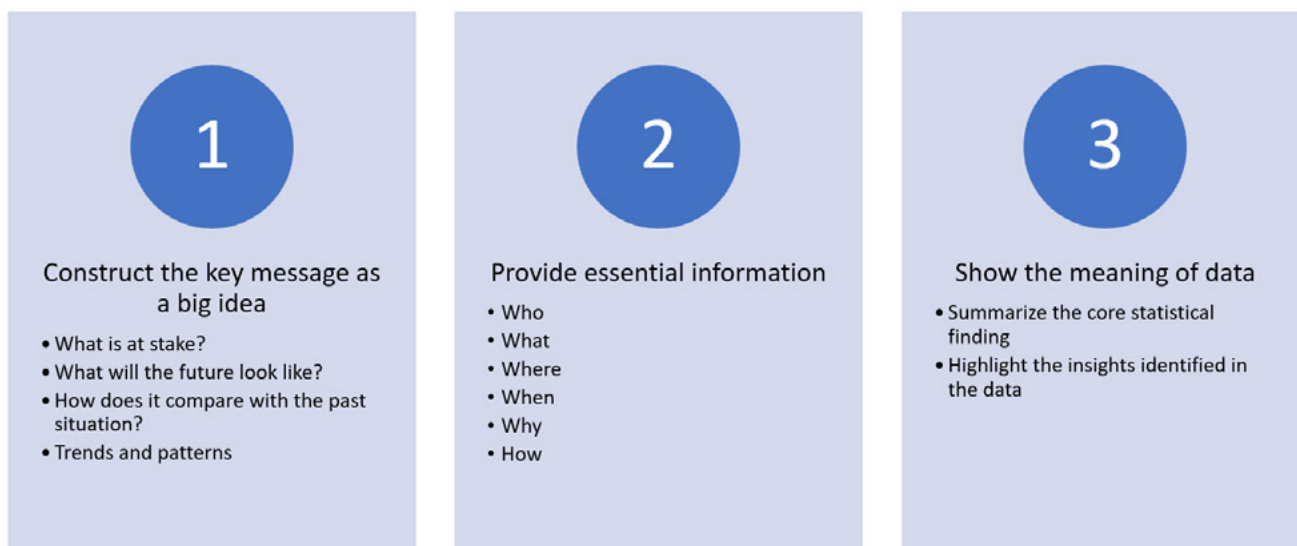
Key message

The key message of a data story tells the audience the key statistical findings from the data and explains why it is important to the audience. It helps 'humanize' the data and enable the audience to understand the data and its meaning. By doing so, the data stories will stick with the audience.

How to write a powerful key message:

The key message is often the main points about issues or themes, rather than a mere description of data. While drafting the key message of a data story, data storytellers need to crystalize the meaning of the statistics, focusing on how the findings affect people. Here are some quick tips on writing concise key messages for data stories (Chart 5).

Chart 5. How to Write a Key Message for Data Stories

1

Construct the key message as a big idea

- What is at stake?
- What will the future look like?
- How does it compare with the past situation?
- Trends and patterns

2

Provide essential information

- Who
- What
- Where
- When
- Why
- How

3

Show the meaning of data

- Summarize the core statistical finding
- Highlight the insights identified in the data

Let's analyze two well-written key messages. The first one is taken from Norway's VNR report 2021.

Norway has a well-developed system of universal health coverage and access to quality health care for all. Life expectancy is high and increasing; in 2020 it was 85 years for women and 81.5 years for men. However, non-communicable diseases and social inequities are still public health challenges. Women and men with a short education (primary/lower secondary level) can expect to live about five years less than those with a long education.

Source: Voluntary National Review 2021 Norway, Page 60.

- **Big picture: Norway has a good health care coverage for all citizens.**
- **Key statistical findings:** Life expectancy is high and increasing; in 2020 it was 85 years for women and 81.5 years for men.
- **Insights:** Non-communicable diseases and social inequalities remain public health challenges. Lower level of education many negatively affect life expectancies.

Overall, this short paragraph describes the trends and patterns about health care coverage and access in Norway. It consists of core statistical findings on life expectancies as well as insights on the public health bottlenecks and challenges that need to be tackled. The statistics on life expectancy support the narrative and are relevant for the audience.

The second example is a key message taken from the 2021 Annual Report of the Global Prepared Monitoring Board (GPMB).[17]

If the first year of the COVID-19 pandemic was defined by a collective failure to take preparedness seriously and act rapidly on the basis of science, the second has been marked by profound inequalities and a failure of leaders to understand our interconnectedness and act accordingly. In the most glaring example, as of 20th October, 63% of those living in high-income countries had received at least one dose of COVID-19 vaccine. In low-income countries, only 4.5% had received the same.[3] The failure to act in the interests of all has prolonged the pandemic for all, as new variants circulate worldwide.

Figure 1 | COVID-19 vaccine doses administered by country income group

For vaccines that require multiple doses, each individual dose is counted. As the same person may receive more than one dose, the number of doses can be higher than the number of people in the population.
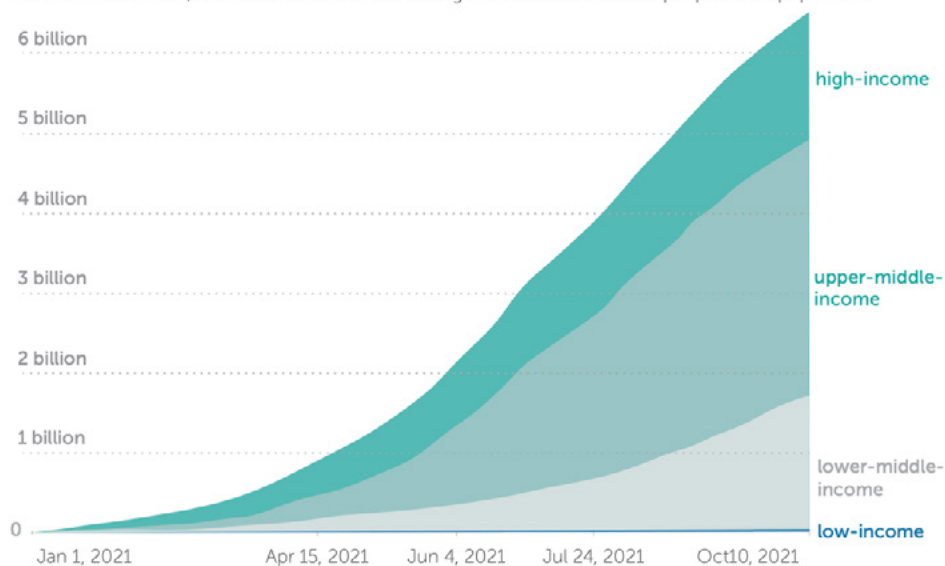
The large majority of vaccine doses have been administered in high- and upper-middle-income countries. Very few have reached low-income countries. Source: Our World in Data, World Bank.

---

[17] Global Preparedness Monitoring Board Annual Report 2021, Page 11

- **Insight:** "... the second (year of the COVID-19 pandemic) has been marked by profound inequalities and a failure of leaders to understand our interconnectedness and act accordingly."
- **Statistical finding:** 63% of those living in high-income countries had received at least one dose of COVID-19 vaccine. In low-income countries, only 4.5% had received the same.

The key message consists of a powerful observation on the world's fight against COVID-19 in 2021 and relevant statistical findings (data on vaccinations in countries at different income levels) to support the observation. The chart visualizes the statistical findings and makes it easier for the audience to follow the narrative.

Use a topic sentence in paragraphs

A key message may take the form of a paragraph consisting of 2-3 sentences. Alternatively, it may appear as a topic sentence in a paragraph that operates as a standalone data story. Topic sentences in paragraphs will help the audience move between topics without losing sight of the key messages and core insights.

Normally the topic sentence is the first and most important sentence in a paragraph. It states the subject of the whole paragraph or gives the overview of the sentences to follow. The supporting sentences after the topic sentence help to develop and support the main idea. These sentences give specific details such as data and facts related to the topic sentence. If the idea discussed in the paragraph is complex or the supporting sentences contain lots of details, the storytellers may also use a final or concluding sentence to restate the main idea to bring the audience back to the key message of the paragraph.

Here is a VNR example on how data storytellers use a topic sentence at the beginning of each paragraph to give the audience a quick snapshot of the data story that is about to follow. It is also helpful to put the topic sentences in bold text—as visual cues for the audience.

The causes of undernutrition have been inadequately addressed. These are primarily poor infant and young child feeding practices (including low rates of exclusive breastfeeding for children under the age of 6 months), cessation of breastfeeding before the recommended 2 years of age, poor maternal nutrition, high adolescent birth rates in some communities, and poor sanitation and hygiene practices, including open defecation. The percentage of children below the age of six months who were exclusively breastfed remains low, although slightly increased from 40.4% in 2012 to 44.9% in 2018.

The COVID-19 pandemic is likely to worsen the nutrition and food security situation across the country. At the global level, due to COVID-19, it is estimated that there could be a 14.3% increase in the global prevalence of moderate or severe wasting among children under five years as this relates to a prediction in mortality rates and country specific losses in GNI per capita. Although there is no official predicted data for Lao PDR, the trend will be similar to those of the global picture. Making progress toward the target of reducing undernutrition is the most challenging task for the country. Stunting is declining but still affects nearly 33% of children under 5 years and wasting, 9%, which is rising. COVID-19 will increase wasting by 14.3%. Children from rural areas, poorer households, ethnic groups and whose mothers are not educated were more likely to be stunted. Food insecurity is unevenly distributed across districts and low-income households and farm laborers were most likely to be affected by COVID-19.

Source: Lao People's Democratic Republic Voluntary National Review on the Implementation of the 2030 Agenda for Sustainable Development 2021, Pages 32

To summarize, an effective topic sentence:

- is the key message of a paragraph,
- is clear, concise, and specific,
- informs the audience of the subject to be discussed in the paragraph,
- motivates the audience to continue reading, and
- allows the audience to skim through the first sentences of every paragraph to get a quick overview of the entire story, article, or report.
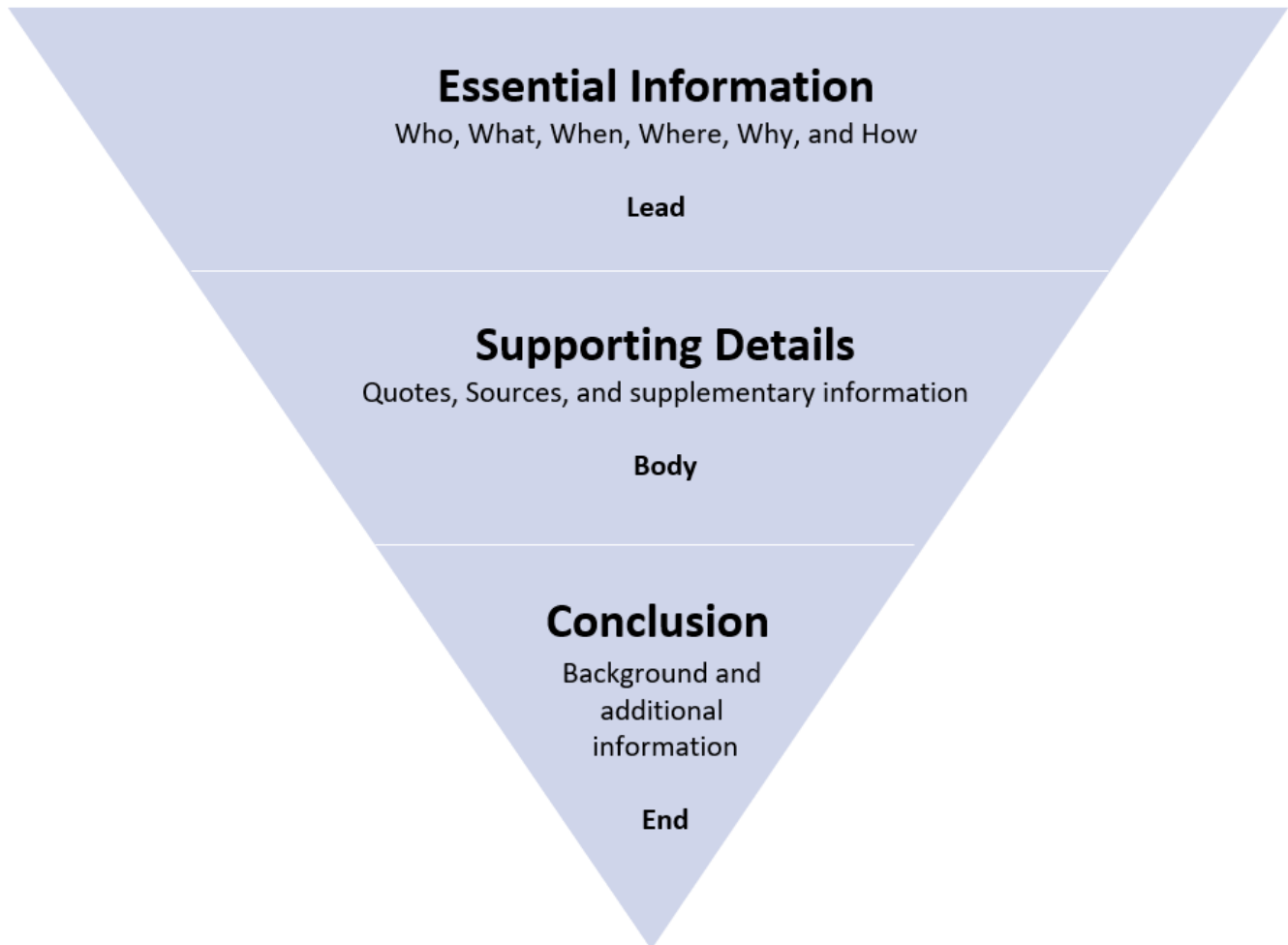
Structure

The structure of a data story is its narrative flow or the order in which the audience experiences the story. An effective structure properly sets the stage, and then guides the audience to find and understand the insights from the data. Following are two structures that a data storyteller may use in constructing a data story. Both are in the shape of a pyramid.

Inverted Pyramid Model

The Inverted Pyramid Model (Chart 6) is widely used by journalists. Simply, the story starts with the conclusion on top. It is called the 'lead'. A lead is followed by supporting details such as quotes, sources, and any supplementary information. At the bottom of the structure lies the background and additional information that the readers might want to know.

Chart 6. Inverted Pyramid Model

Essential Information
Who, What, When, Where, Why, and How

Lead

Supporting Details
Quotes, Sources, and supplementary information

Body

Conclusion
Background and
additional
information

End

The lead is the most important part of the inverted pyramid structure. It is the key message of the story and places the key findings in context. It contains all essential information that can answer questions of who, what, when, where, why, and how.

The inverted pyramid structure was originally created for the benefit of editors in traditional media because if the space is limited, the editors may start chopping things down from the bottom layer of the pyramid as the information there is less important compared to that in higher layers.
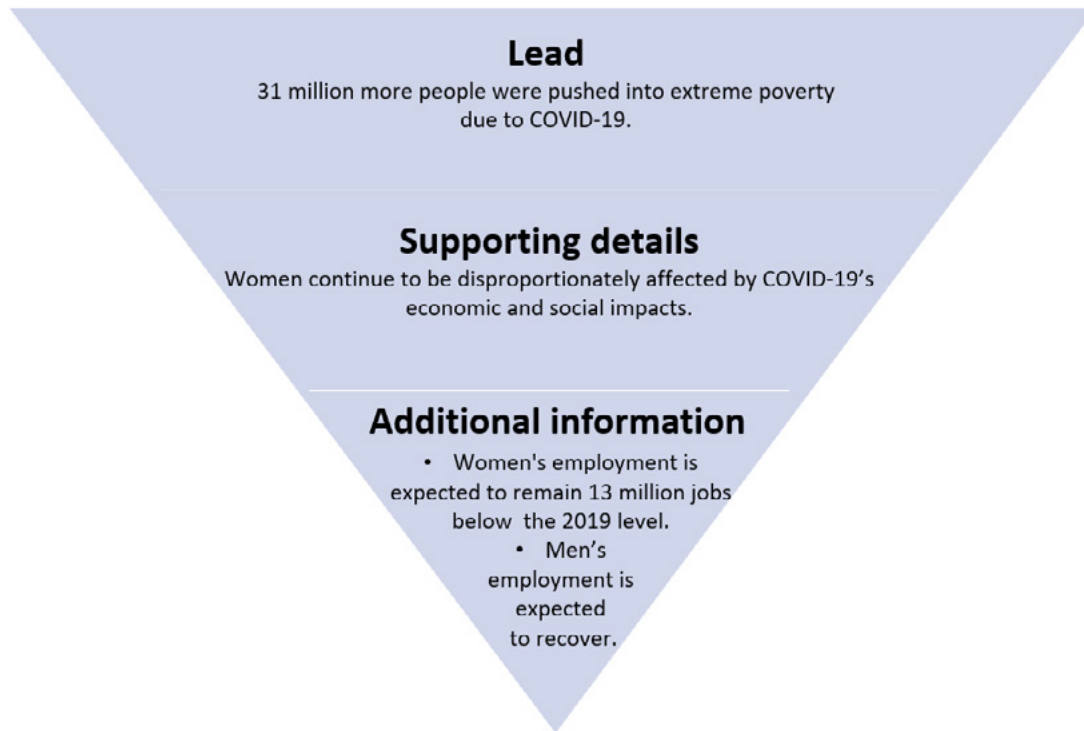
Below is a data story taken from Bill & Melinda Gates Foundation's Goalkeepers Report 2021. It uses text and animated visuals to discuss COVID-19 pandemic's impact on poverty.

Millions More in Extreme Poverty

For many, the economic impacts of the pandemic continue to be severe and enduring. We know we may seem like unlikely messengers on this topic—we're two of the most fortunate people on the planet. And the pandemic has made that even more clear. People like us have weathered the pandemic in good shape, while those who are most vulnerable have been hit the hardest and will likely be the slowest to recover. An additional 31 million people around the world have been pushed into extreme poverty as a result of COVID-19. Although men are 70% more likely to die from COVID-19, women continue to be disproportionately affected by the economic and social impacts of the pandemic: This year, women's employment globally is expected to remain 13 million jobs below the 2019 level—while men's employment is largely expected to recover to pre-pandemic rates.

1.8M
1.5M
26.6M
4.0M
POVERTY LINE

This graphic does not include regions that achieved a total net reduction of 2.6 million people in extreme poverty.

Legend:
Sub-Saharan Africa | South Asia | Latin America and the Caribbean
North Africa and Middle East

POVERTY LINE
4.0M

This graphic does not include regions that achieved a total net reduction of 2.6 million people in extreme poverty.

Legend:
Sub-Saharan Africa | South Asia | Latin America and the Caribbean
North Africa and Middle East

Source: Gates Foundation Goalkeepers Report 2021

In this example, storytellers use an inverted pyramid structure to narrate the story. It opens with the lead or the key message that the COVID-19 pandemic pushed 31 million more people into extreme poverty. Subsequently, the storytellers move to highlight gender inequality and states that women bear the brunt socially and economically. Additional data on women's employment recovery in comparison to that of men are used to close the story. Since this report is presented online, data storytellers use animated visuals to show the core statistical finding (31 million people pushed below the poverty line). The colored legends demonstrate the composition of the 31 million people by region. The narrative flow can be seen clearly in the following inverted pyramid (Chart 7).

Chart 7. Gates Foundation Example Analysed in an
Inverted Pyramid Structure

Lead
31 million more people were pushed into extreme poverty
due to COVID-19.

Supporting details
Women continue to be disproportionately affected by COVID-19's
economic and social impacts.

Additional information
• Women's employment is
expected to remain 13 million jobs
below the 2019 level.
• Men's
employment is
expected
to recover.

By putting the conclusion or the essential information on top, data storytellers allow the audience to

- quickly grasp the essence of the story by reading the 'lead', and
- explore further details if need.

Additionally, putting the essential information upfront carries a special benefit for data stories posted on a webpage. Specifically, having most of the relevant keywords in the data story condensed into the 'lead" will greatly boost the Search Engine Optimization (SEO) for the data story. It will help the data story appear organically at the top of the keyword-driven search results.

Pyramid Model[18]

The Pyramid Model is another storytelling structure. It is a tool for communication and a tool for thinking. It is well used in the corporate world, particularly in the management consulting industry where consultants pack their findings and recommendations into reports and presentations developed using the Pyramid Model. The simplest form of the Pyramid Model (Chart 8) consists of a key message in the top layer of the pyramid, core arguments that support the key message in the middle layer, and additional relevant details that support the core arguments in the bottom layer.

---

[18] Barbara Minto, "The Pyramid Principle", Prentice Hall; 3rd edition, January 1, 2010

Chart 8. Pyramid Model
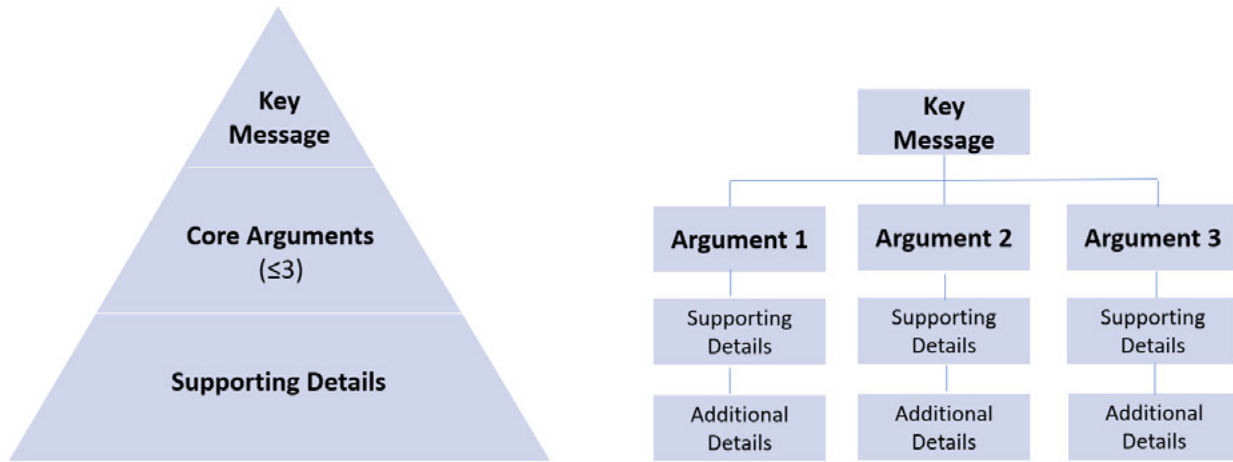
The following data story is taken from an online article about sustainable fishing published by the Marine Stewardship Council (MSC)[19]. The story presents some fundamental data and facts about the ocean. Its narrative is structured using the pyramid model.

The ocean is vast

The ocean is so big that **most of it remains unexplored**. Scientists don't know exactly how many species of plants and animals live in the ocean, but some think that more than 90% of them have yet to be discovered.

There's one thing that everyone is agreed on, though: **the ocean is home to an extraordinary variety of plants and animals.**

20%
OF THE OCEAN
HAS BEEN
EXPLORED

---

[19] MSC is an international non-profit that uses a labelling system and fishery certification program to recognise and reward sustainable fishing practices.
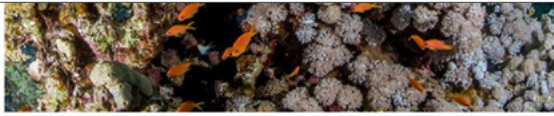
From coral reefs to polar seas, the ocean bursts with life and diversity.

And **much of this life is essential to sustaining us**. Around the world, more than a billion people rely on fish for their main source of protein, while around 1 in 10 depends on fishing for their livelihood.

91%
OF SPECIES
ARE YET TO BE
DESCRIBED

We rely on the ocean's wild bounty so much that **fish is the most traded food in the world – above tea, coffee, bananas** and **sugar.**
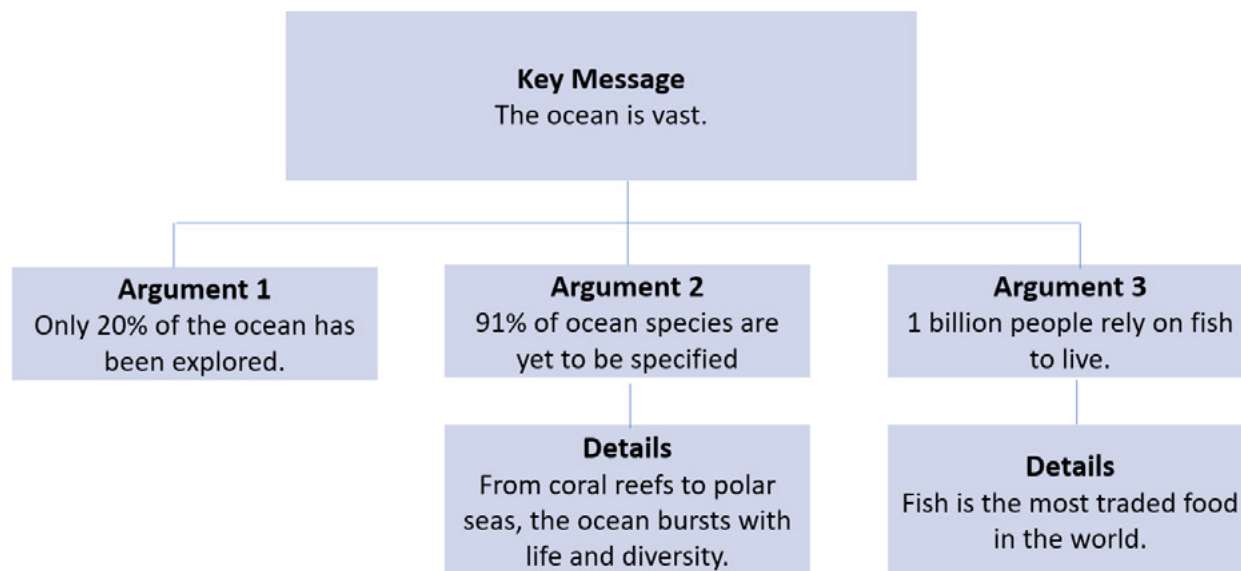
1BN
PEOPLE RELY
ON FISH
TO LIVE

Source: Dad fishes for the future, MSC

The data story is centred on the key message that the ocean is vast. The data storytellers use three arguments to support the key message, namely, the ocean is underexplored, a predominant majority of ocean species are not yet specified, and much of the ocean lives are essential to sustaining human beings. These arguments are supported by relevant data and details. The analysis of this data story using the pyramid structure is shown in Chart 9.

Chart 9. MSC Story Analysed in a Pyramid Structure

Key Message
The ocean is vast.

Argument 1
Only 20% of the ocean has been explored.

Argument 2
91% of ocean species are yet to be specified

Argument 3
1 billion people rely on fish to live.

Details
From coral reefs to polar seas, the ocean bursts with life and diversity.

Details
Fish is the most traded food in the world.

No matter which pyramid structure is used, it is important for data story tellers to leverage these structures and comb through the insights, facts and numbers to organize them into a coherent data story that can be easily digested by the audience. This thinking process should happen before any data visualization is done. After all, all visual elements are part of the overall narrative. They are created to help the audience understand the data story, not the other way around.

Language

As VNRs and SDG reporting are targeted for a wide range of audiences, data storytellers should make conscious language and style choices so that even a layperson with no prior statistical knowledge can easily understand the story. With this said, data storytellers can always make statistical information, such as a statistical annex, technical notes, and/or metadata, available as supplementary resources for the expert audience.

General writing tips

Following are some general writing tips that can make the content heavy VNRs and SDG materials easier to read, follow, and understand for audiences:

- Put key messages first.
- Break the text into small chunks (One paragraph - one argument or one idea).
- Start a paragraph with a topic sentence.
- Write short paragraphs (no more than three sentences in one paragraph).
- Write short sentences (15-25 words per sentence on average).

- Use visual cues such as headings, subheadings, bold text, and bullet points to guide the audience.
- Use active voice.
- Use appropriate verbs.
- Minimize the use of jargon. If jargon is necessary, explain it the first time it appears.
- Minimize the use of long words. Use more common alternatives.
- Minimize the use of acronyms. If acronyms are necessary, explain them the first time they appear.
- Avoid redundancy.

Print writing versus web writing

Data storytellers should note that print writing and web writing are different.[20] They have different styles and are suitable for carrying different content.

Chart 10. Print Content Versus Web Content

Print content is	Web content is
passive	active
linear	interactive
Slow	fast
detailed	concise
complete	Complete or fragmented

In the print world, you can present as many details as possible and the readers will move from page to page, following the structure that is created for them. However, on the web, the audiences log onto a webpage hoping to find the information that they need as quickly as possible. The web audience will scan the webpage roughly in the order of writing the letter 'F', first on top from left to right horizontally and then the left side of the webpage from the top to the bottom vertically. In this case, they need to be served with the most important information quickly. Chart 11 shows some tips on how to repurpose your print materials for the web.[21]

---

[20] Writing Style for Print vs. Web

[21] Repurpose print material for the web | plainlanguage.gov

Chart 11. How to Repurpose Print Content for the Web

🖥	Keep the most important message at the top of the web page.
⛭	Chunk the content into logical sections.
📰	Use headings, subheadings, and bold text to help the audience navigate the content.
📋	Highlight key facts in bulleted lists.
📊	Use intentional, creative visuals to help explain complex content.

These rules can also be used to make the print materials sharper and crisper. For example, it is advisable to put the key message and core insights upfront in an inverted pyramid structure or a pyramid structure as explained earlier in this guide. These rules could also be used in developing content for social media with creative adaptations. For example, storytellers can chunk the content into post-size units for social media and broadcast one unit per post. In both contexts, storytellers can provide visual cues to help the audience understand the meaning of the data.

V. Think beyond storytelling

In conclusion, a statistician or a data scientist needs to develop the mindset of a storyteller and practice the storytelling techniques covered in this guide as much as possible. Here are a few guiding principles that a data storyteller should follow.

Start with the end in mind /set the objective

A data story needs to present both the core statistical finding(s) and its meaning to the audience. In other words, the audience needs to know 'what' and 'so-what' after reading the data story. Simply showing data is not enough and is not considered storytelling.

Develop the narrative before visualizing data

Step back and find what you really want to say. Then formulate the storyline and structure the story so the audience can follow and understand the story easily. Don't rush to create charts and visualizations. Visualization is not data storytelling, but rather a component of data storytelling.

Connect with the audience

Meet the audience where they are with a compelling narrative and intentional visualizations. Make the data story friendly by keeping it simple and showing the data and the insight clearly in easy-to-understand style, language, and visuals. Make sure even a layperson audience finds the data story relevant and can understand the story easily.

Advocate for the data story

A good data story still needs marketing. Develop communications strategies early on to promote the data story in the VNRs and SDG reporting and ensure it gets the attention it deserves. Use the established data visualization toolkit developed by UNSD to create and leverage spin-off communications assets such as infographics, videos, newsletters, social media posts, etc. to promote the data story. Be creative.

VI. Further readings

Data storytelling

- Nancy Duarte, "Data Story: Explain Data and Inspire Action Through Story", IDEAPRESS Publishing, 2019; Duarte DataStory®

- Cole Nussbaumer Knaflic, "Storytelling with data", Wiley, 2015, https://www.storytellingwithdata.com/

- Will Storr, "Science of Storytelling: Why Stories Make Us Human and How to Tell Them Better", Harry N. Abrams, March 10, 2020

- https://narrativescience.com/resource/webinar/how-to-tell-a-compelling-data-story-ft-zach-mazzoncini-data-storytelling-virtual-summit/

- Data Storytelling in Marketing: Venn Benchmark Report 2021

- Data Storytelling: The Essential Data Science Skill Everyone Needs

- https://www.nugit.co/what-is-data-storytelling/

- https://narrativescience.com/data-storytelling/

- UNECE's Making Data Meaningful series very relevant, including the Guide to Writing Stories About Numbers: https://unece.org/statistics/making-data-meaningful

Context and audience

- Purdue Online Writing lab

- https://style.ons.gov.uk/category/writing-for-the-web/personas/

Writing

- Barbara Minto, "The Pyramid Principle", Prentice Hall; 3rd edition, January 1, 2010, http://www.barbaraminto.com/

- Style.ONS: A guide to writing about statistics

- https://datajournalism.com/

- https://www.theguardian.com/membership/datablog/2021/sep/13/numbers-you-can-tell-stories-with-a-decade-of-guardian-data-journalism

- https://training.npr.org/2016/10/12/leads-are-hard-heres-how-to-write-a-good-one/

- https://www.clearvoice.com/blog/the-dos-donts-of-writing-a-good-lead-according-to-dear-megan/

- Federal plain language guidelines | plainlanguage.gov

Data visualization

- Nathan Yau, "Data Points: Data Visualization Than Means Something", Wiley, 2013, https://flowingdata.com/

- Cole Nussbaumer Knaflic, "Storytelling with data", Wiley, 2015, https://www.storytellingwithdata.com/

- https://www.juiceanalytics.com/writing/20-best-data-storytelling-examples