Module 4: Data Visualization

In this module, you will learn the basics of effective communication with data visualization. You will learn how the human brain is adapted for understanding data through visualization and what effective sensory engagement looks like. This module will also introduce principles to matching data types to graphical forms for effective communication, basic design principles and data visualization ethics. Students will learn to redesign data visualizations with this new knowledge. Most importantly this section is about conveying knowledge and information through visual stories, not just dry statistics.

At this end of this module, you will be able to:

- Appreciate the purpose of data visualization
- Develop an effective data visualization design strategy
- Match data and graph types
- Address ethical issues that arise with data visualization
- Convey stories with data visualization

Getting Started
Simple fixes to newborn deaths

Kenya’s high child death rate of 73 per 1,000 children could be greatly reduced if the country were to employ simple and practical solutions to reduce preterm births and complications, which are the leading cause of death among newborns.

The fourth Millennium Development Goal (MDG 4) aims to reduce the 1990 mortality rate among under-five children by two thirds. Child mortality is also closely linked to MDG 5 - to improve maternal health. Since more than one third of all child deaths occur within the first month of life, providing skilled care to mothers during pregnancy, as well as during and after birth, greatly contributes to child survival.

In this module, you will use data to quantify and explain the problem of major development issues, like the rate, causes and solutions for newborn deaths.

Lesson 1: Purpose of Data Visualization
In this lesson, we will review some best practices to follow when considering data visualizations to tell a data story.

Why do we use data visualization?

---

1 http://www.internewskenya.org/dataportal/data/S2

From Evidence to Stories: Thinking Like a Data Journalist
Visual communication has some distinct advantages over verbal or text communication. This section explores how visualization can be more precise and concise than other forms of communication.

**Why do we use data visualization?**

- The brain likes images.
- Makes the data easier to understand and compare
- Grabs the audience’s attention
- Allows the audience to explore the data
- Tells a story about the data

**Exercise: Simplifying Numbers**

Consider these statements and answer the following questions.

**Statement**

- 50% of the population is female
- In 2012, the official poverty rate in Armenia was 32.4%
• In 2012, 19.3% of children in Armenia were stunted and 15.3% were overweight.

• Armenia’s neo-natal mortality rate is 0.1%

• How many in 1,000 babies die before they turn 28 days old?
Lesson 2: Matching Data and Graph Types

In this lesson, we will examine different data types, determine the message behind the data and match the data to an appropriate graphical form.

What to look for in data?

- Trends
- Contrast
- Outliers

3 https://tutorials.infogr.am/

From Evidence to Stories: Thinking Like a Data Journalist
One of the first steps towards developing a public interest story is learning how to interpret and question data - this will help you to define the ‘angle’ for your story. Before we begin working with datasets, let us begin our data literacy journey with learning how to interpret data visualizations.

Importantly, data visualizations help you to compare or connect two or more things, which are useful cues to find angles for data-driven stories. For instance, maps can help compare administrative units (say the population density of two countries); while line graph can help compare time periods (say the population growth over two decades).

<table>
<thead>
<tr>
<th>Visualization</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bar chart</strong></td>
<td>Compare data across categories</td>
</tr>
<tr>
<td><strong>Line graph</strong></td>
<td>Compare data across time</td>
</tr>
<tr>
<td><strong>Scatter plot</strong></td>
<td>Compare interactions between two variables</td>
</tr>
<tr>
<td><strong>Maps</strong></td>
<td>Compare data across geographical units</td>
</tr>
<tr>
<td><strong>Pictogram</strong></td>
<td>Compare aspects of data through human depictions</td>
</tr>
</tbody>
</table>
Let’s look at the following types of commonly-used data visualizations: bar chart, line graph, pie charts, maps, and pictograms.
Interpreting Data Visualizations: Bar Chart

Bar charts are used to compare values across categories. In this example\(^5\), total number of refugees is compared across countries of their origin.

**Questions:**

- Which countries produce the most refugees?
- How many times more refugees come from Afghanistan than from Iraq?
- What tells you that you can verify the data in this chart?

Interpreting Data Visualizations: Line Graph

Line graphs can be used to compare data over time. In this example\(^6\), predicted population of continents is compared across ten decades.

Questions:

- Which continent has the biggest population now?
- Which continent will see the biggest population growth?
- What makes you notice Africa right away?


From Evidence to Stories: Thinking Like a Data Journalist
Interpreting Data Visualizations: Pie Charts

Pie charts are a popular visual form to show parts of a whole, for example, causes of death among young adults. Pie charts always have to add up to 100% (not the top two causes, all the causes). Pie charts should not have more than five slices because if there are more, the sizes are difficult to compare. They should also be different sizes.

Questions:

- What does the chart on the right measure?
- Which product has had the highest sales?
- What is incorrect about the first chart?

7 http://de.slideshare.net/vis4/making-data-visualizations-a-survival-guide

From Evidence to Stories: Thinking Like a Data Journalist
Maps are a familiar type of data visualization – which are important for public interest stories as they let you compare geographies. There are various types of map visualizations; this example shows the visualization type called “Choropleth”. These maps use differences in shades or patterns inside defined areas in a map to indicate the average value of a variable – for instance population density or per-capita income. This example\(^8\) shows three maps that display indicators for Sudan related to water and sanitation, schooling, and food consumption.

**Questions:**

- What do these different maps tell you?
- Do you see any trends?

---

\(^8\) [http://www.bbc.co.uk/news/world-africa-17126340](http://www.bbc.co.uk/news/world-africa-17126340)

From Evidence to Stories: Thinking Like a Data Journalist
Interpreting Data Visualizations: Pictogram

Pictograms are data visualizations that present information through pictorial resemblances to humans or other objects. This example highlights the impact of malnutrition on children’s health in Kenya.

Questions:

- How many children are too short for their age in Kenya?
- How many children die of causes related to malnutrition?
Exercise: Matching Chart and Data Types

Choose a Chart!

Prevalence of unemployed people in different age groups.

Prevalence of unemployment.

Percentage of female candidates in elections from 1990-2014

Number of cases of tuberculosis in different provinces.

Indicate with an X in the box which chart type you would use for each of the listed data types.
Lesson 3: Design and Color Basics

Basic Design Concepts

- **Simplicity**: Choose a maximum of three colors and fonts and stick with them consistently
- **Hierarchy**: All visualizations need a focal point to guide them through story
- **Brevity**: Keep text short and to the point
- **Creativity**: Incorporate playful design that related to the topic
- **Clarity**: Label clearly, specify units, use a legend when necessary

Five Golden Rules

1. **No 3D Effects**
   Three dimensional graphics distort the data. Whichever part of the data is closest to the reader is distorted to look larger. In the example below, see how a pie chart that is evenly divided into thirds, when 3-D effect is added, distorts the data to look like the slice closest to the reader is larger.
Avoid 3D charts (at all costs)
2. Sort data from largest to smallest.

You want your audience to easily be able to make a visual comparison across categories. To make this easier, sorting your data will order the bars in a bar graph or slices in a pie chart from greatest to least, making the visual comparison much easier.

9 http://de.slideshare.net/vis4/making-data-visualizations-a-survival-guide

From Evidence to Stories: Thinking Like a Data Journalist
3. Choose maximum two colors or shades of the same color for your graphic. Stay away from rainbow colors.

Simple colors in the same color range and a consistent font make your visualizations look more professional and credible. Your headlines, labels and text should be clear and explain the visualization to the audience.

4. Your headline should tell the story!
Before you create a data visualization, you should know what the message is. We create drafts in Excel to ensure that we have a clear headline and message before we start. It is important to be able to tell the audience where the data comes from but it’s even more important to tell them through a strong headline, what the message of the story is.
Module 4: Data Visualization

Distribution of Catholic population


10 From Evidence to Stories: Thinking Like a Data Journalist
Why are you showing this to me?

Samstag, 27. April 13

http://www.theguardian.com/world/interactive/2013/mar/05/world-map-catholic-population

From Evidence to Stories: Thinking Like a Data Journalist
Examples:

Bad headline: Rates of Domestic Violence in Afghanistan

Good Headline:
Bad Headline: **Top 10 Types of Registered Domestic Violence Cases**

Good Headline:

---


From Evidence to Stories: Thinking Like a Data Journalist
Bad Headline: Resolution Methods for Domestic Violence in Afghanistan

From Evidence to Stories: Thinking Like a Data Journalist
5. **Use Direct Labeling**

Having to jump back and forth between a key and a graph is difficult for readers. Whenever possible, use direct labeling on visualizations.
Exercise: Redesign a Data Visualization

Examine the below visualizations and make a list of at least five changes you would make.
1. ____________________________________________________________________________________
   ____________________________________________________________________________________

2. ____________________________________________________________________________________
   ____________________________________________________________________________________

3. ____________________________________________________________________________________
   ____________________________________________________________________________________

4. ____________________________________________________________________________________
   ____________________________________________________________________________________

5. ____________________________________________________________________________________


From Evidence to Stories: Thinking Like a Data Journalist
Module 4: Data Visualization

From Evidence to Stories: Thinking Like a Data Journalist
Exercise: Find a graphic in your local media and redesign it based on these principles.
Lesson 4: Ethics of Data Visualization

Look vs. Function
Take a look at the following three visualization examples, and answer the following questions.

Example 1: Gun violence in America

Example: How recession shaped the Economy


From Evidence to Stories: Thinking Like a Data Journalist
Example 3: Paths to White House

Questions

- What story does the data visualization tell?
- What big picture does the visualization provide?
- What details do you understand from the visualization?
- Is there text to help you understand?

Ethics of Data Visualization

Just as with any other type of source material, just because you have access to detailed data, does not mean that you necessarily want to include all of it in your story. You have to weigh the pros and cons of...

17 http://www.nytimes.com/interactive/2012/11/02/us/politics/paths-to-the-white-house.html?_r=0

From Evidence to Stories: Thinking Like a Data Journalist
different levels of aggregation and always return to asking yourself what is the news value of the data and what is the best way to present the data so that citizens can make better decisions.

Privacy

Example: Gun Registration

This map\textsuperscript{18} was created by a journalist who through an access to information request, obtained the name and address for every US citizen in his area who has a gun license. He put this data onto a dot map. When a reader clicks on a point on the map, the name and address of the person with the gun license appears on the map. This sparked a debate on privacy vs public safety in the media.

\begin{center}
\textbf{The Journal News published a map showing every handgun permit holder in two New York counties}
\end{center}

\begin{center}
\textsuperscript{18} http://www.slate.com/articles/news_and_politics/jurisprudence/2013/01/the_journal_news_gun_map_the_first_amendment_and_state_law_gave_the_new.html
\end{center}
One of the biggest global achievements for global health over the last century has been the widespread introduction of vaccines to save lives. In the United States, a group of vaccine deniers are putting public health at risk. Through this visualization\(^\text{19}\) - which has as a few basic components: years from left to right, a black vertical line representing the introduction of the vaccine, states from top to bottom, and colored squares indicating the number of people who have died from that disease - a viewer can understand at a glance that with the introduction of vaccines, millions of lives have been saved across the country. The visualization is a simple fact-checking exercise to quell public hysteria about a proven scientific fact.

What issues in your community could be addressed with data? How can that data be put into a form that people will relate to, that will help them understand how the data can improve their quality of life? Telling an effective story is key to transforming data into insight and action.

In this lesson, we will look at the objectives of data-driven storytelling and different visual forms that can enhance your story.

\(^{19}\) http://graphics.wsj.com/infectious-diseases-and-vaccines/

From Evidence to Stories: Thinking Like a Data Journalist
Objectives of Data Stories

- Collect information that we can use strategically
- Influence policy
- Inform public debate
- Expose wrong-doing
- Create awareness and understanding of complex issues
- Explore options for solving problems using data
The form a data story takes depends a bit on the audience and your purpose. The audience may be a policy maker who you would like to make a specific decision based on that data, a general public that you are trying to help understand the complexities of an issue, a group of researchers who are weighing factors before developing recommendations. In general, data stories fall into a few general categories:

- Collect information that we can use strategically
- Influence policy
- Inform public debate
- Expose wrong-doing
- Create awareness and understanding of complex issues
- Explore options for solving problems using data

20 https://tutorials.infogr.am/

From Evidence to Stories: Thinking Like a Data Journalist
Let’s look a few common types of data stories and discuss the following aspects: audience, information retention, emotional impact, and call to action.

**Infographics: Example 1**

Infographics are graphic visual representations of information, data or knowledge intended to present information quickly and clearly to illustrate patterns or trends. They enable you to package different kinds of information together to deliver an accessible message about a topic.

Take a look at this infographic21 from UNICEF, and try to answer these questions:

- Who is the audience?
- What information are they supposed to retain?
- What creates an emotional impact?
- What should they do after seeing this?


*From Evidence to Stories: Thinking Like a Data Journalist*
Infographics: Example 2

Here is another example of an infographic about child health in Kenya from Internews:

http://www.internewskenya.org/dataportal/data/52

Questions:

- Who is the audience?
- What information are they supposed to retain?
- What creates an emotional impact?
- What should they do after seeing this?
Charts are the most direct way to display data findings but often require a bit of interpretation from the audience. A strong narrative can help readers identify the important findings conveyed by the charts.

Take a look at this chart from Forbes that presents data from UNHCR: http://www.forbes.com/sites/niallmccarthy/2015/06/18/the-countries-with-the-most-refugees-per-1000-inhabitants-infographic/

Questions:

- Who is the audience?
- What information are they supposed to retain?
- What creates an emotional impact?
- What should they do after seeing this?
Here is another example of a chart from a survey company called Gallup that presents data from a poll of sub-Saharan Africans about their dissatisfaction with public needs in their communities: http://www.gallup.com/poll/113872/opinion-briefing-achieving-gains-africa.aspx

**Questions:**

- Who is the audience?
- What information are they supposed to retain?
- What creates an emotional impact?
- What should they do after seeing this?
Maps: Example 1

Maps are representations of geographical data and enable easy comparisons between data in different regions. Take a look at this interactive map that presents data about outbreaks of vaccine-preventable diseases: http://www.cfr.org/interactives/GH_Vaccine_Map/#map

Questions

- Who is the audience?
- What information are they supposed to retain?
- What creates an emotional impact?
- What should they do after seeing this?
Maps: Example 2

Here’s another example of an interactive map that presents data-driven insights about migration in Europe: [http://www.themigrantsfiles.com/](http://www.themigrantsfiles.com/)

Questions:

- Who is the audience?
- What information are they supposed to retain?
- What creates an emotional impact?
- What should they do after seeing this?
‘X’ Issues in ‘X’ Charts: Example 1

Often, the objective of data stories is to explain the different factors affecting an issue. This lends itself well to explaining an issue in a series of simple charts that add complexity to the audiences’ understanding of the topic. This could include a variety of bar, line and pie charts or maps that each add a layer of meaning to the topic.

Take a look at this a news story that uses various charts and maps to investigate the Ebola crisis:

**Questions:**

- Who is the audience?
- What information are they supposed to retain?
- What creates an emotional impact?
- What should they do after seeing this?

From Evidence to Stories: Thinking Like a Data Journalist
‘X’ issues in ‘X’ Charts: Example 2

Here’s is another example from the World Bank, where data from various sources is presented in a series of seven charts to summarize the global state of gender:

Questions:

- Who is the audience?
- What information are they supposed to retain?
- What creates an emotional impact?
- What should they do after seeing this?
News Apps: Example 1

In an effort to personalize data, more CSOs and media outlets are creating online applications that allow users to explore the data themselves. So, for example, a site may allow you to enter your height and weight and provide a personal recommendation for the number of calories you should consume a day. This allows people to personalize the data experience to their own lives.


Questions:

- Who is the audience?
- What information are they supposed to retain?
- What creates an emotional impact?
- What should they do after seeing this?
News Apps: Example 2

Take a look at this online news application that helps a user interact with data from their perspective. ProPublica created this application using data about death and complication rates for surgeons in the USA: https://projects.propublica.org/surgeons/

Questions

- Who is the audience?
- What information are they supposed to retain?
- What should they do after seeing this?
Exercise: Writing Headlines

What have the Millennium Development Goals (MDGs) achieved? The following set of visualizations explores the progress that has been made towards achieving the MDGs, which had a target date of 2015. From this broad overview of data\(^{22}\), we can identify the most interesting and relevant findings to our audience and write headlines that draw attention to progress, or the lack thereof.

**Millennium Development Goals**

Using the skills learned in the last lesson, write a one-sentence headline for each of the following MDG graphs.

---

1. To eradicate extreme poverty and hunger

**MDG 1**

People living on less than $1.25 a day

- **1990:** $1.25 less
- **2015:** $1.25 more

47% 14%

**Global number of extreme poor (million)**

- **1990:** 1,926
- **1999:** 1,751
- **2015:** 836

**Proportion of undernourished people**

- **1990-92:** 23.3%
- **2014-16:** 12.9%

One sentence headline:
2. To achieve universal primary education

**MDG 2**

Global out-of-school children of primary school age

<table>
<thead>
<tr>
<th>Year</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>100 million</td>
</tr>
<tr>
<td>2015</td>
<td>57 million</td>
</tr>
</tbody>
</table>

Primary school net enrolment rate in sub-Saharan Africa

<table>
<thead>
<tr>
<th>Year</th>
<th>Net Enrolment Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>52%</td>
</tr>
<tr>
<td>2000</td>
<td>60%</td>
</tr>
<tr>
<td>2015</td>
<td>80%</td>
</tr>
</tbody>
</table>

*One sentence headline:*
3. To promote gender equality

**MDG 3**
Primary school enrolment ratio in southern Asia

<table>
<thead>
<tr>
<th></th>
<th>1990</th>
<th>2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>Girls</td>
<td>74</td>
<td>103</td>
</tr>
<tr>
<td>Boys</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

90% of countries have more women in parliament since 1995

One sentence headline:
4. To reduce child mortality

**MDG 4**

Global number of deaths of children under five (million)

<table>
<thead>
<tr>
<th>Year</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>12.7</td>
</tr>
<tr>
<td>2015</td>
<td>6</td>
</tr>
</tbody>
</table>

Global measles vaccine coverage

<table>
<thead>
<tr>
<th>Year</th>
<th>Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>73%</td>
</tr>
<tr>
<td>2013</td>
<td>84%</td>
</tr>
</tbody>
</table>

One sentence headline:
5. To improve maternal health

One sentence headline:
6. To combat HIV/AIDS, malaria, and other diseases

One sentence headline:
7. To ensure environmental sustainability

**MDG 7**
Access to piped drinking water since 1990 (billion)

<table>
<thead>
<tr>
<th>Year</th>
<th>Access (billion)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>2.3 billion</td>
</tr>
<tr>
<td>2015</td>
<td>4.2 billion</td>
</tr>
</tbody>
</table>

Terrestrial and marine protected areas in Latin America and the Caribbean

<table>
<thead>
<tr>
<th>Year</th>
<th>1990</th>
<th>2015</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>8.8%</td>
<td>23.4%</td>
</tr>
</tbody>
</table>

One sentence headline:
8. To develop a global partnership for development

**MDG 8**

Official development assistance

- 1990: 81 bn
- 2014: 135 bn

**Global internet penetration**

- 2000: 6%
- 2015: 43%

One sentence headline: